Staffing for Effective Digital Preservation 2017

An NDSA Report

Results of the 2017 Digital Preservation Staffing Survey

AUTHORS
Winston Atkins, Duke University Libraries (Co-chair)
Cynthia Ghering, Michigan State University
Margaret Kidd, Virginia Commonwealth University Libraries
Carol Kussmann, University of Minnesota Libraries (Co-chair)
Joy M. Perrin, Texas Tech University Libraries
Meg Phillips, National Archives and Records Administration
Sibyl Schaefer, University of California, San Diego
Kathryn Talbot, Kansas State University Libraries
Representing the NDSA Staffing Survey Working Group

TABLE OF CONTENTS

ABOUT THE NATIONAL DIGITAL STEWARDSHIP ALLIANCE 3
EXECUTIVE SUMMARY 4
INTRODUCTION 5
METHODOLOGY 6
THE DATA FILES 6
THE SURVEY CODEBOOK 7
SURVEY CONTENT 8
FINDINGS 8
Background Information 9
Digital Content 10
Digital Preservation Activities 14
Digital Preservation Organization and Staffing 20
Staffing Qualifications and Training 26
Comments about Digital Preservation Program Staffing and Organization 29
ANALYSIS OF STAFFING LEVELS BY AMOUNT OF CONTENT 31
COMPARISON WITH 2012 DATA 36
Survey Response 37
Organization Information 37
Digital Preservation Activities 39
Department Organization and Digital Preservation Comparison 40
Current and Ideal FTE Comparison 41
Perceived Satisfaction in Overall Organization 43
Digital Preservation Specialist Staffing 44
CONCLUSION and NEXT STEPS 45
APPENDIX A: Survey Statistics from 2012 and 2017 48
APPENDIX B: Survey Questions 49
APPENDIX C: Partial List of Respondents 54

ABOUT THE NATIONAL DIGITAL STEWARDSHIP ALLIANCE

Founded in 2010, the National Digital Stewardship Alliance (NDSA) is a consortium of more than 220 partnering institutions that are committed to the long-term preservation of digital information. NDSA's mission is to establish, maintain, and advance the capacity to preserve our nation's digital resources for the benefit of present and future generations. NDSA member institutions represent all sectors, and include universities, consortia, non-profits, professional organizations, commercial enterprises, and government agencies at the federal, state, and local levels.

More information about NDSA is available at http://www.ndsa.org/

Copyright © 2017 by National Digital Stewardship Alliance. This work is licensed under a Creative Commons Attribution 3.0 Unported License.

DOI 10.17605/OSF.IO/3RCQK | ARK c7605/osf.io/3rcqk

EXECUTIVE SUMMARY

In 2017, the National Digital Stewardship Alliance (NDSA) received completed survey responses from 133 institutions engaged in digital preservation, enabling the Staffing Survey Working Group to investigate how these organizations staffed and organized their digital preservation functions, and to identify any changes since the NDSA's 2012 survey[1] on the same topic, published in "Staffing for Effective Digital Preservation: An NDSA Report."[2] We received 78% of the responses from the US, down from 86% in the 2012 survey. Other countries responding were the United Kingdom (8), Germany (4), Switzerland (3), the Netherlands (2), and one each from Australia, Canada, Denmark, Estonia, Luxembourg, Pakistan, Slovenia, Sweden, and Uruguay.

One of the main focuses of the survey is on staffing levels. In response to these questions related to staffing levels, organizations reported an average of 13.6 FTE working in digital preservation activities. However, respondents indicated they would double that to 27.5 FTE in ideal circumstances. They expressed a particular need for more digital archivists, software developers, and cataloger/metadata analysts. Most respondents' organizations (68%) retrained existing staff for at least some digital preservation functions, while 42% also hired experienced digital preservation specialists.

The findings on collection holdings show that the volume of holdings was larger than in the previous survey, but the expected rate of future growth had moderated. Whereas in 2012, 68% expected up to a 49% growth, in 2017, the largest percentage of respondents (73.2%), expected less than 25% growth in the collection. As in the 2012 survey, images and text documents continue to be the majority of digital files being preserved.

The survey also queried respondents about which digital preservation functions and activities were performed in-house and which were outsourced, and what the future plans for these functions and activities were. Most respondents appear to prefer conducting most digital preservation activities in-house. Emulation, which led the list of services considered for outsourcing, was selected by only 35% of respondents.

[1] NDSA Standards and Practices Working Group, "Survey of Staffing Practices and Needs Related to Digital Preservation, 2012," ICPSR34901-v1, Ann Arbor, MI: Inter-university Consortium for Political and Social Research, October 2013, https://doi.org/10.3886/ICPSR34901.v1.

[2] NDSA Standards and Practices Working Group, "Staffing for Effective Digital Preservation, An NDSA Report: Results of a Survey of Organizations Preserving Digital Content," December 2013, http://www.digitalpreservation.gov/documents/NDSA-Staffing-Survey-Report-Final122013.pdf.

In 2017, 46% of respondents were not satisfied with how the digital preservation function was organized within their organization, while 25% believed it was organized properly. This is a significant decrease in satisfaction from 2012, when 43% of respondents agreed or strongly agreed that their digital preservation functions were well-organized. Only 32% of respondents reported that their organization had a dedicated digital preservation department, which is down 1% from organizations responding in 2012.

The 2017 Digital Preservation Staffing Survey provides a useful snapshot of the way digital preservation is accomplished in 2017 and how its practitioners feel about the effectiveness of their current organizational structures. It also builds on the 2012 survey and begins to establish data with which the digital preservation community can identify trends in staffing in the field.

INTRODUCTION

Organizations establishing or scaling up digital preservation programs are faced with many staffing, scoping, and organizational decisions. Some of the questions that need to be answered include: How many staff members are needed and what kinds of skills, education, and experience should they have? What types of positions should the institution create? Should it hire new staff or retrain existing staff? And how should the preservation program be scoped? That is, what functions should be included directly in the program, provided by other parts of the organization, outsourced, or implemented through collaboration with other organizations? What organizational and staffing models work well?

In 2012, several members of the NDSA Standards and Practices Working Group discovered that they shared a common interest in knowing how organizations were staffing and organizing digital preservation programs. They wanted to use the information to improve their own programs and worked together to survey a variety of organizations about their staffing and digital preservation efforts. The 2012 preservation staffing survey and resulting paper, conducted and written by members of the NDSA Standards and Practices Working Group, were designed to shed light on how organizations responsible for digital preservation are addressing these staffing, scoping, and organizational questions.

In late 2016, NDSA members who regularly used data produced from the 2012 survey[3] and report[4] proposed re-surveying organizations preserving digital content. The NDSA

[3] NDSA Standards and Practices Working Group, "Survey of Staffing Practices and Needs Related to Digital Preservation, 2012."

Coordinating Committee approved the project and a new working group was formed from NDSA members from multiple interest groups. The new Staffing Survey Working Group modeled the 2017 survey on the 2012 survey to provide updated data as well as to afford the opportunity to shed light on changes in staffing priorities that have developed over the intervening five years.

METHODOLOGY

We based the 2017 survey on the survey administered in 2012 to assist with long-term analysis. Over the course of three months in the winter of 2017, working group members reviewed the previous survey questions and modified a few of them based on feedback received in 2012 as well as new concerns. We tried to ensure any modifications hewed closely to the spirit of the original question as much as possible to maintain the ability to compare results between the two surveys. The survey was then transferred to Qualtrics, a web-based survey tool subscribed to by Duke University, an NDSA member organization. Qualtrics provided the ability to publish previews of surveys, so working group members were able to review and provide additional feedback on the look and behavior of the survey prior to its launch.

We announced availability of the 2017 Staffing Survey in March 2017 through multiple national and international mailing lists associated with digital preservation and it remained open through the beginning of April. It was open to any institution choosing to respond although one response per organization was requested. The survey was started 168 times and completed 133 times for a 79% completion rate.

Results were analyzed using the Qualtrics software, Excel, and Google sheets during the summer of 2017. We divided the questions thematically, and smaller teams from the working group collaborated on the initial analysis. All working group members participated in the overall analysis and writing of the final report.

THE DATA FILES

This project was conducted under the auspices of the NDSA, and the team felt it was important to demonstrate a commitment to digital preservation by properly archiving its own data and findings. To this end, we uploaded the exported survey response data in CSV

[4] NDSA Standards and Practices Working Group, "Staffing for Effective Digital Preservation, An NDSA Report: Results of a Survey of Organizations Preserving Digital Content."

format to the NDSA's page in the Open Science Framework, a scholarly commons area with the ability to connect the entire research cycle including archiving data and sharing reports. CSV-formatted data provides ready access to the data and should not require specialized software.

We chose to deposit data for 168 respondents, regardless of survey completion. Those who wish to reuse the data will find that Column G records whether a respondent completed the survey. We excluded, and did not deposit, a small number of responses that were obvious false starts for a respondent who submitted a completed survey later. The false starts were identified by their duplicate IP addresses. We excluded instances of potential respondents who opened the survey only to review it. The latter group typically included a note in a free text field, asking us to delete the response.

To keep all responses anonymous, we deleted data from several fields that were either automatically collected by the Qualtrics software or entered as part of a survey responses. These included the respondent's IP address, their approximate latitude and longitude, and their name and contact information. When respondents asked us not to include their institution's name in the list of organizations completing the survey, or did not answer the question, we deleted those institutional names from the database. As much as possible, we reviewed free text fields to limit information that could be used to identify the respondent.

THE SURVEY CODEBOOK

The codebook establishes the context for the survey and its responses. It assumes that future users will have no prior knowledge of the survey or its data, so much of our effort went into ensuring that the Scope of Study and the Survey Overview sections conveyed our mission and survey goals and provided an overview of our approach to inviting responses. The codebook also documents the full text of each question, the possible responses, and the formats in which a respondent could enter their answer. If a respondent either quit the survey or chose not to answer a question, that decision was recorded with an empty cell or the "-99" value. An empty cell typically meant that the question was formatted for text input, while the "-99" value indicated that the unanswered question was seen but not answered, and formatted for a numeric response.

SURVEY CONTENT

The survey posed a series of questions related to the nature of the institution preserving content, the scope of their digital preservation program, and the number and types of positions responsible for carrying out digital preservation work. It also solicited information about the organizational structure of the program: whether there was a dedicated digital preservation department, and if not, whether the responsibility for this work was centralized in another department or decentralized in other parts of the institution. The survey also queried respondents about which digital preservation functions and activities were performed in-house and which were outsourced, and what the future plans for these functions and activities were. Finally, to learn more about not only how organizations were approaching digital preservation but which approaches were perceived as working well, the survey asked for respondents' perceptions of whether the institution was satisfied with their current organizational structure around preservation. We provided an option in the survey to reference or send related organizational charts or position descriptions.

FINDINGS

This section of the report includes the text of the original questions clustered into themes, a presentation of the results for each question, and a short discussion of the implications of the results. A longer discussion of the larger picture of digital preservation staffing provided by the survey results follows in the Analysis section of this report. A more detailed comparison with 2012 results is also in a separate section in this report.

The major themes discussed below include:

- Background Information
- Digital Content
- Digital Preservation Activities
- Digital Preservation Organization and Staffing
- Staffing Qualifications and Training
- Comments about Digital Preservation Program Staffing and Organization

Since a major goal of the survey was to find out how repositories with different characteristics were currently staffed, whether they were satisfied with their current staffing, and what they would like their staffing model to look like in the future, questions found in several of the sections cover both "what we have" and "what we want." For example, both the Digital Preservation Organization and Staffing and Staff Qualifications

and Training sections of this report shed light on the present situation and the desired future.

Background Information

After an introductory section (Q1), the survey started with general background questions to provide an understanding of the types of organizations that were responding to the survey.

- Q2 - (Required): What is the name of your organization?
- Q3 - Can we include the name of your organization in a list of organizations that responded to this survey? Knowing specific responding organizations may be helpful to people interpreting the survey results. If you agree to this we will still make our best effort to protect your individual survey responses so that no one will be able to connect your responses with you or your organization.
- Q4 - (Required): Which of the following most closely describes the type or function of your organization?
- Q5 - In which country is the responding organization located? [133 respondents]
- Q22 - Please provide your contact information if you are willing to respond to follow-up questions.

The name of the organization (Q2) was required to help us verify that there was only one response from each organization. Organizational names were not shared unless respondents specifically gave permission (Q3). Contact information was also requested for follow-up questions if organizations were willing to provide it (Q22), but remains confidential. We did not include it in the deposited data.

The majority (78%) of the 133 respondents were from the United States. However, we also received responses from the United Kingdom of Great Britain and Northern Ireland (8), Germany (4), Switzerland (3), the Netherlands (2), and one each from Australia, Canada, Denmark, Estonia, Luxembourg, Pakistan, Slovenia, Sweden, and Uruguay (Q5).

As shown in figure 1, the highest percentage of respondents represented academic libraries or archives (46%), followed by government entities (11%), museums (8%), “other” (7%), and national, federal, or legal deposit libraries (6%) (Q4). Fewer than 5% of the respondents were from academic institutions or departments (not a library or archives), historical societies, independent libraries or archives, non-profit organizations, public libraries, research data repositories, for-profit corporations, institutional repositories, or universities. Among the institutions that selected “other” were a research library, two

religious institutions, a health system archives, a television and radio archives, two government archives, and a special collections and archives within a university.

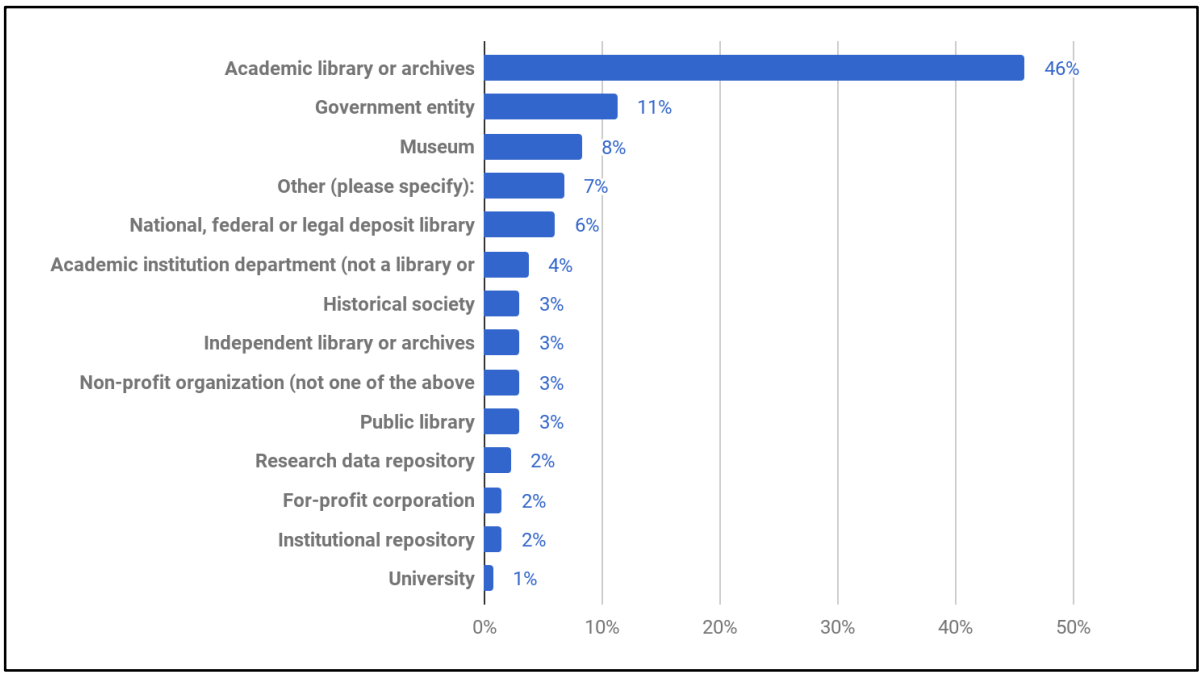

FIGURE 1: Types of organizations represented by respondents

Digital Content

Additional questions designed to elicit information about the scope of preservation efforts focused on the amount of digital content each organization was currently preserving and expected to be preserving in the near future.

- Q6 - How much online or offline storage space are you using for your digital content, not including backup copies? [132 respondents]
- Q7 - What do you expect the percentage of growth to be of your preserved digital content over the next year? Please enter a whole number representing a percentage. [125 respondents]
- Q8 - Roughly how much of each are you preserving, in terms of number of files? (None, A little, Some, A lot) for each type of material listed. [133 respondents]
    - Still/2D images (e.g. TIFF, JPEG)
    - Drawings/vector graphics (e.g. CAD/CAM)
    - Moving images/video
    - Audio recordings
    - Web sites/blogs/social media

- Text/documents (e.g. Word, PDF, TXT)
- Geographic Information Systems (GIS) data
- Spreadsheets or datasets (other than GIS data)
- Databases
- Computer games/software
- Other (please indicate):

Survey results (fig. 2) show that over half of the respondents (58.6%) were preserving 1–50 TB of digital content, 16.5% were preserving 51–100 TB, 14.3% were preserving 101–500 TB, and 8.3% were preserving more than 500 TB.

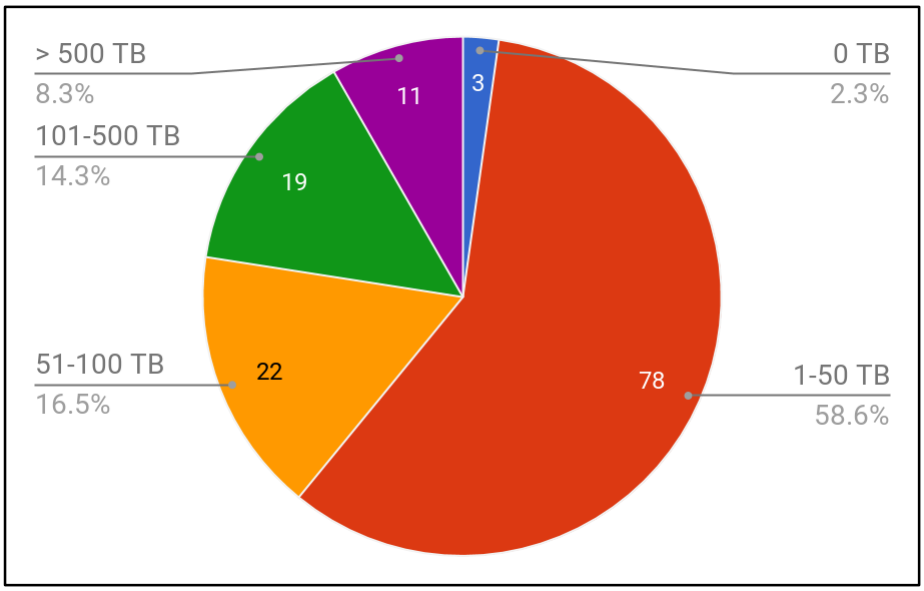

FIGURE 2: Amount of content organizations were preserving

For those that had more than 500 TB of content (for which 8.3% responded), there was a follow-up question to specify how much content they had. Figure 3 displays the results. A full 50% of institutions that had over 500 TB of data had between 500 and 999 TB. After that, 25% of these organizations had 4 petabytes (PB) to 6 PB while only 8.3% had over 6 PB of data.

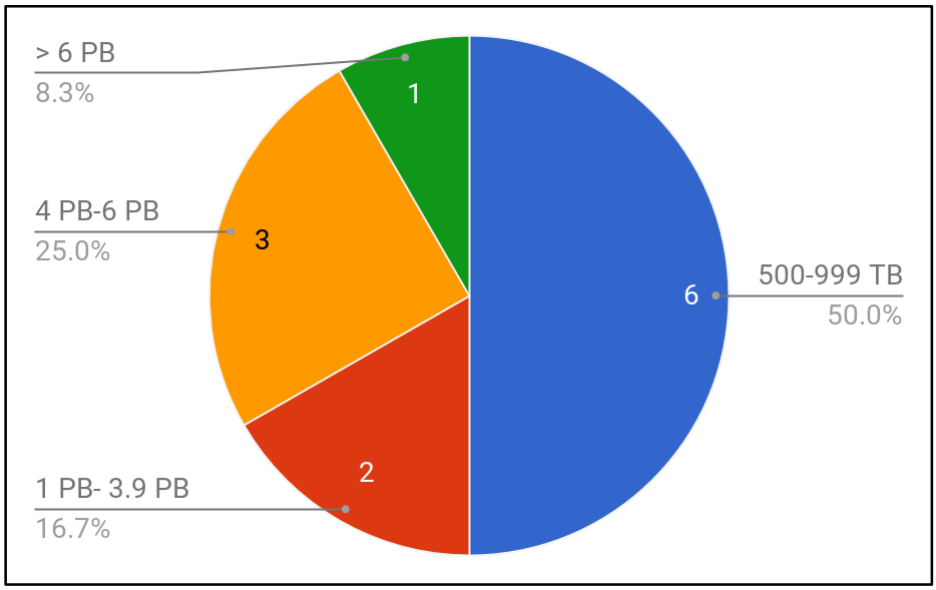

FIGURE 3: Amount over 500 TB

While it is important in interpreting the survey results to know the amount of content organizations were preserving, we also wanted to know the amount of content growth they were expecting in the near future. The survey (Q7) asked for an estimate of the expected percentage growth of preserved digital content over the next year. As figure 4 shows, 73.2% of respondents expected less than 25% growth. The next largest categories are those who think they will experience 26–50% growth and 100% growth, with both being 12.6%. The smallest category are those who think they will experience 51–99% growth. These results show an overall moderate increase in the amount of content growth, rather than overall extreme growth.

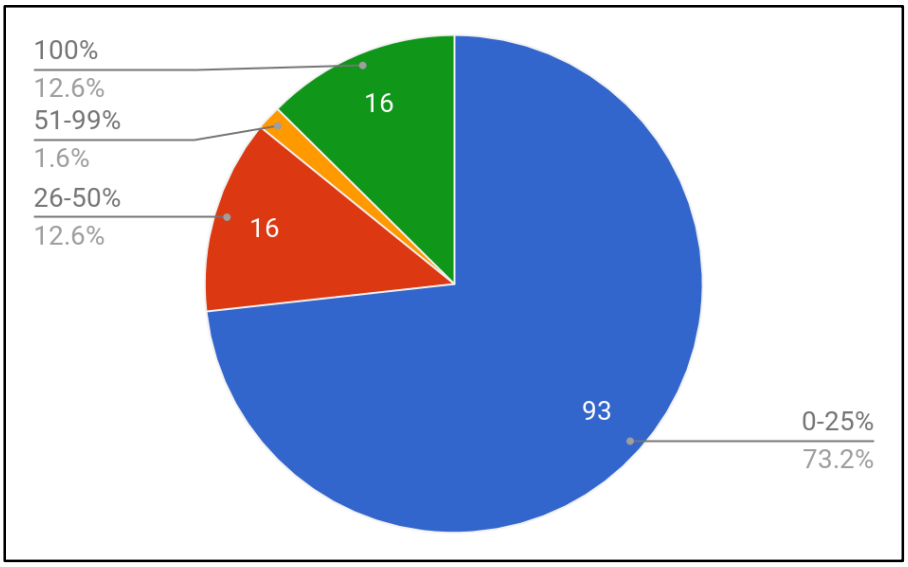

FIGURE 4: Expected percent growth in preserved content

Next, the survey asked respondents to specify and roughly quantify the types of formats they were preserving (Q8). The survey used qualitative responses ("None," "A little," "Some," "A lot") to elicit the overall dominance of types of content within each collection, regardless of size of the collections or the exact number of files or percentage of storage space used for each. Figure 5 shows the results by response count.[5] Images and text documents continue from the 2012 survey to be the majority of digital files being preserved.

[5] The "None" responses are not displayed.

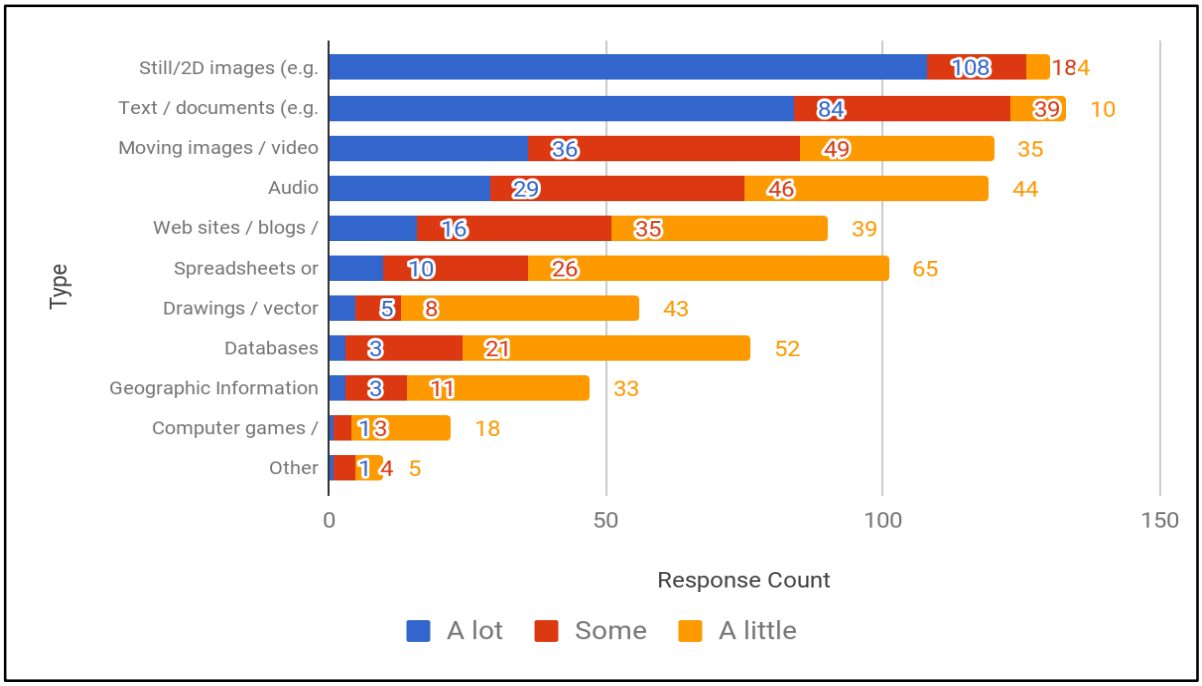

FIGURE 5: Types of content by response count[6]

An additional question was asked for respondents that indicated they had "other" types of files. When defining "other" types of files, respondents mentioned the following file types:

- 3D scans
- Combination of formats not represented including Zip files and others not easily categorized
- Disk images
- Email
- LISTSERVs
- Master data: ontologies, KOS, schemas
- Presentations/Slide Decks
- Project files for exhibitions
- Raw camera files (CR2, NEF)
- SPSS/statistical data

Digital Preservation Activities

To understand how organizations defined digital preservation and how they managed these functions, the survey asked questions about participation in cooperative efforts, the types of activities respondents considered in scope for digital preservation, which of these

[6] Multiple items could be selected by each respondent; "None" responses are not displayed.

were done in-house or outsourced, and what future plans existed for the functions and activities.

- Q9 - Do you participate in any digital preservation consortial or cooperative efforts? [132 respondents]
- Q29 - (shown if the answer to Q9 was yes): What benefits do you gain from your participation? [68 respondents]
- Q28 - In the matrix below, please select all that apply for each activity. Select A for activities that you consider in scope for digital preservation function at your institution, regardless of whether or not you are currently doing the activity. Select B for all of the activities that your organization currently does in-house. Select C for all activities for which your organization currently outsources. Select D for all activities you wish your organization would outsource in the future. (A, B, C, and D can all be checked if applicable to your situation.) [133 respondents]

The responses to Question 9 revealed that 52% of respondents participate in at least one consortium or cooperative network. Those who replied affirmatively were asked to provide the names of the consortia in a free text box (without any prompts). Since there was a wide variety of responses and it was clear that not all respondents interpreted "consortia" to mean the same thing,[7] the survey team did not do a detailed analysis of the particular responses. Nevertheless, the team noted that while the majority of respondents who replied "yes" to Question 9 listed only one consortium, some respondents listed as many as nine. In reviewing the responses some consortia were mentioned very frequently: for example, Digital Preservation Network (DPN) (17), HathiTrust (11), NDSA (8), LOCKSS (7)/CLOCKSS (3), Portico (5), and APTrust (5).

For the follow up question "What benefits do you gain from your participation [in the consortia]?" (Q29) respondents selected networking (68%), training (57%), and storage space (54%) most often. Consulting (35%), access interface (33%), communications/marketing (28%), programming (25%), other (19%), and federated search (16%) were also seen as beneficial (fig. 6).

[7] Because of the differences in the interpretation of "consortia," some members of the same organizations listed them and some did not. For example, several members of the survey team who are members of NDSA did not include NDSA in their response to this question. This indicates that participation in consortia may be higher than what the survey captured.

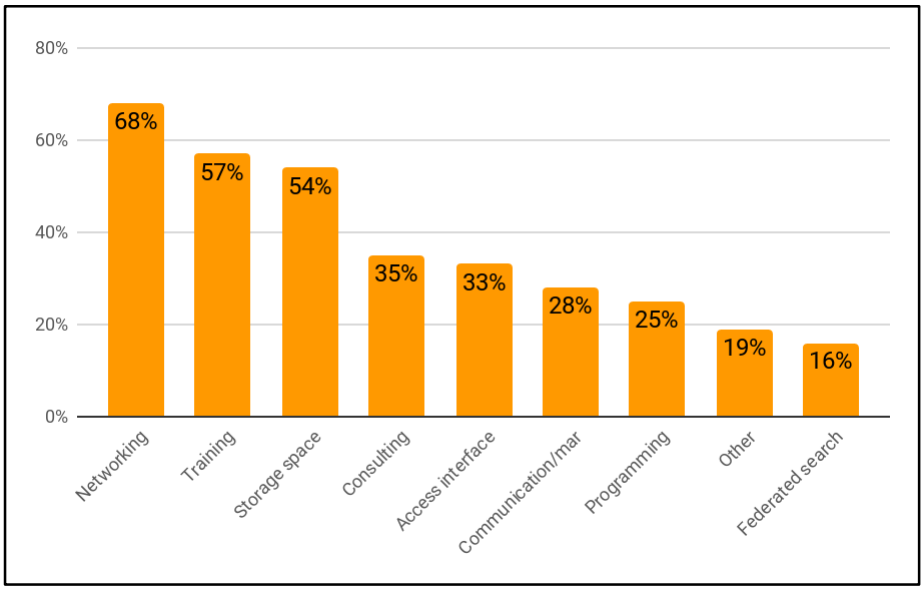

FIGURE 6: Benefits of participation in consortia or cooperative efforts

The remaining question in this section (Q28) was adjusted from the format of the 2012 question because the survey team noted that the previous structure of the question only asked respondents whether the activity was done in-house or outsourced if the respondents had previously identified the activity as "in scope" for digital preservation. That meant that the survey did not gather information on activities that the organizations were performing (either in-house or not) but considered part of other functions. To correct this, the 2017 survey allowed any combination of responses and did not eliminate from further inquiry activities that the organization did not consider in scope for digital preservation.

Asked in table form, respondents were directed to select all that apply for each activity. (Options A, B, C, and D could all be selected.)

- Select A for activities that you consider in scope for digital preservation function at your institution, regardless of whether or not you are currently doing the activity. (This is about how you define digital preservation, not about what you do.)
- Select B for all of the activities that your organization currently does in-house.
- Select C for all activities for which your organization currently outsources.
- Select D for all activities you wish your organization would outsource in the future.

Choices in the table for all parts of Question 28 were:

- Content replication
- Creation of access copies
- Descriptive cataloging
- Development and maintenance of tools

- Development of guidelines for content creators
- Development of preservation policies and strategy
- Digitization
- Emulation
- Fixity checks
- File format identification
- File format validation
- Metadata creation/extraction
- Normalization of files
- Preservation education, training and outreach
- Preservation planning
- Research
- Selection for preservation
- Secure storage management
- Technology watch
- Transformation/migration of formats

The response to part A in Question 28, "activities that you consider in scope for digital preservation function at your institution, regardless of whether or not you are currently doing the activity," revealed that all of the options were considered to be in scope for digital preservation by at least half of the respondents. However, there was considerable range, from preservation planning (88%) at the high-end to emulation (53%) at the low-end, with all but four activities being in scope for over 70% of respondents. Besides emulation, the other three activities considered to be in scope by fewer than 70% of the respondents were descriptive cataloging (68%), tools development/maintenance (65%), and research (64%) (fig. 7).

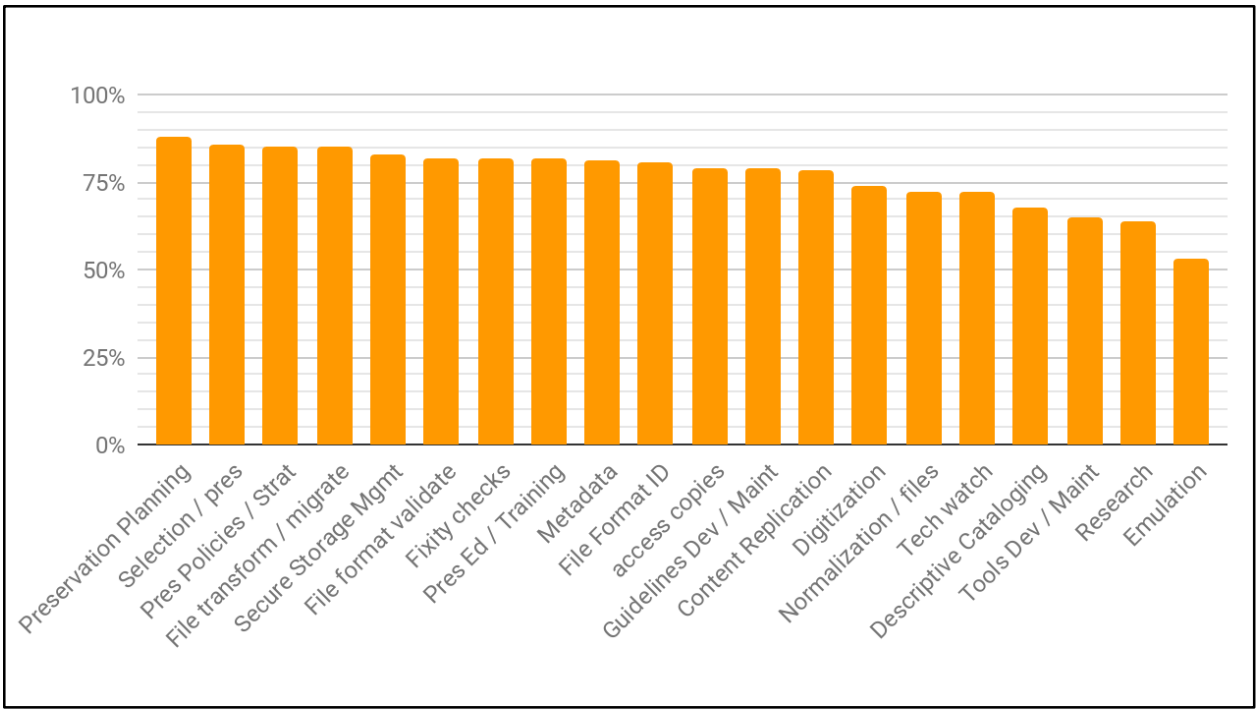

FIGURE 7: In-scope digital preservation activities

Question 28 B and C, "What activities are currently done in-house" and "Which are outsourced" showed that most organizations are currently doing most of the activities the survey mentioned themselves, with the single exception of emulation which is done in-house by only 13% of respondents (fig. 8).

Comparing the responses for emulation currently done (either in-house or outsourced) with the responses to the "in-scope" question, we can see that 53% of organizations consider emulation to be in scope, but a far smaller number are actually pursuing that strategy now (13% are outsourcing and 7% are doing in-house). Presumably, many are planning to explore it more fully in the future.

The *most* commonly outsourced activities were digitization (51%), secure storage management (38%), and tools development/maintenance (32%). The activities respondents are *least* likely to outsource are preservation policies/strategy (4%), selection for preservation (4%), and preservation planning (5%). Since these activities are all high-level management functions closely related to the mission of the organization, it makes sense that those would be kept in-house most often.

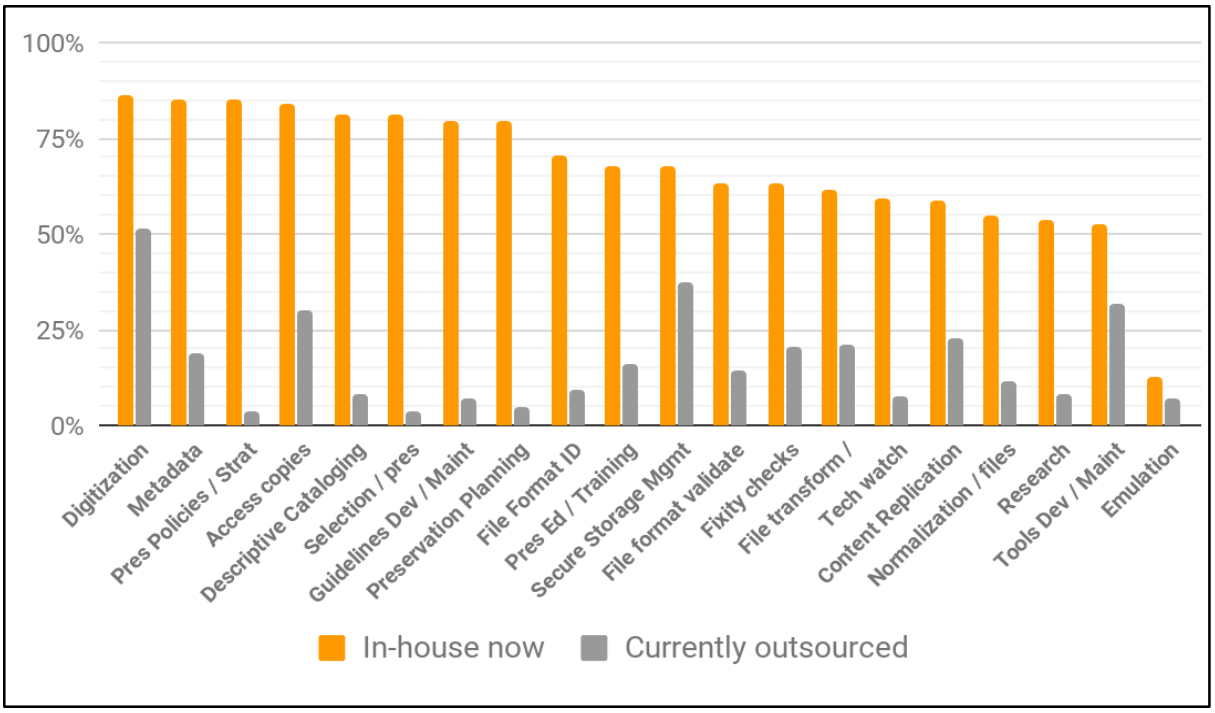

FIGURE 8: Which digital preservation activities are currently done in-house and which are currently outsourced?

The last part of Question 28's multi-part question was "Which of the listed activities would each organization wish to outsource in the future." The most common "yes" responses were for emulation (35%), tools development/maintenance (32%), secure storage management (30%), and digitization (29%), which correspond to the most commonly, currently outsourced activities plus emulation, which is not common now, but in consideration for many organizations (fig. 9). Note that even for emulation, which got the largest number of "yes, we would like to outsource in the future" responses, only 35% of respondents wish to pursue this through outsourcing, which is not very high. By and large, respondents appear to prefer conducting preservation activities in-house, which is an important finding in itself.

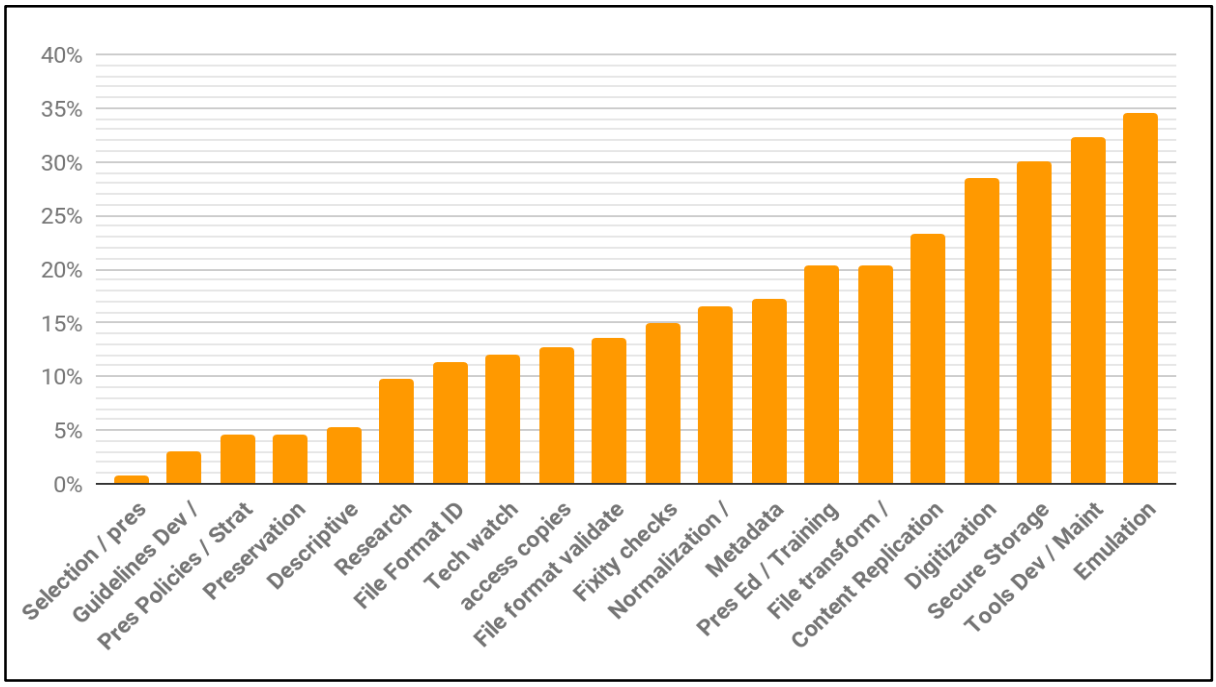

FIGURE 9: Would you wish to outsource this function in the future?

Digital Preservation Organization and Staffing

Organizational Structure

As with the 2012 survey, questions asked which department(s) took the lead in digital preservation, and if respondents were satisfied with the way digital preservation functions were organized within their institutions.

- Q13 - Is there a dedicated digital preservation department within your organization? [133 respondents]
- Q14 - Which department(s) take the lead for digital preservation within your organization? If this is a fairly equally distributed effort choose more than one. [78 respondents]
- Q15 - For each of these positions, how many FTE do you currently have supporting digital preservation (or supporting digital collections, even if the activity is not considered in scope of your DP program) and how many would be ideal? FTE stands for full-time equivalent. For example a 1.0 FTE could mean one person full-time or 2 people half-time; a 0.5 FTE could mean one person half time or two people quarter-time. Please use whole numbers or decimals as appropriate. [130 respondents]
- Q16 - The way our digital preservation function is currently organized (staffing levels, expertise, where they are placed within the larger organization) works well. [133 respondents]

In order to help determine the structure of digital preservation programs, participants were asked to identify if there was a digital preservation department within their organization (Q13). Almost 60% of the 133 respondents said “no” (78), over 30% said “yes” (42) and almost 10% said this response was “not applicable” (13) and provided more information about the structure of their activities (fig. 10).

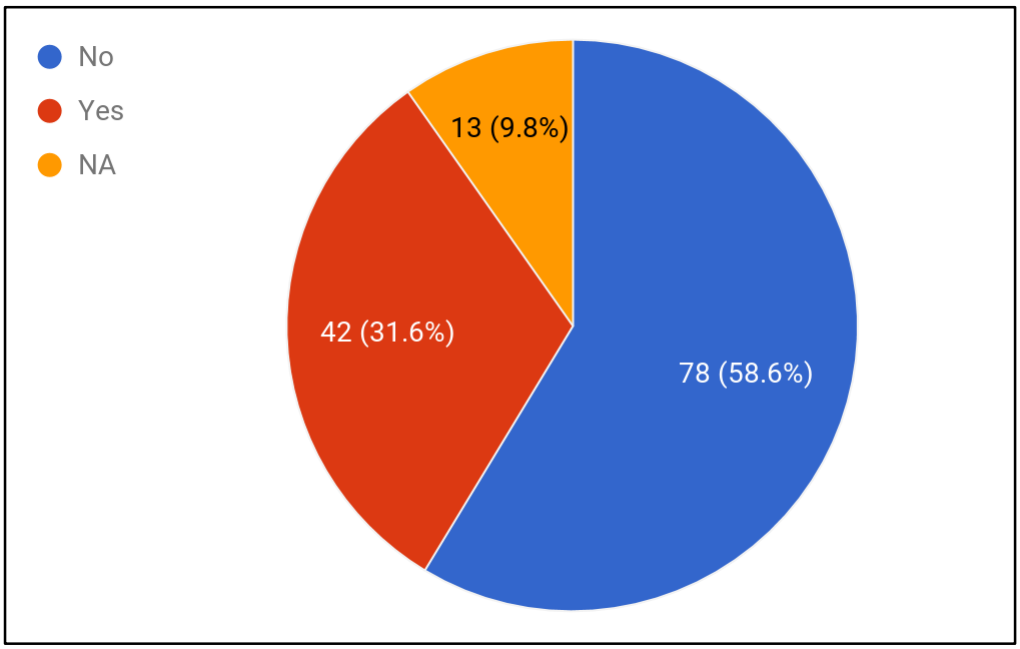

FIGURE 10: Is there a dedicated digital preservation department within your organization?

Thirteen respondents who chose “not applicable” for this answer provided additional information about their organization's approach to digital preservation. These responses were characterized into the following five themes:

- digital preservation is part (but not all) of a department's responsibility (2);
- the overall organization's primary responsibility is digital preservation (2);
- more than one department is responsible for digital preservation (6);
- all responsibilities fall to a single person rather than a department, generally in an organization with a small staff (2);
- the organization is in the process of organizing to include digital preservation (1).

Those that did have a digital preservation department were asked to provide the name of the department. Because department names might identify individual institutions, a word cloud using provided department names was created to share high frequency words (fig.

11). Some high-use words, not surprisingly, include: digital, preservation, archives, services, and collections. Other words of interest include: research, communication, dedicated, scholarship, management, and team.

FIGURE 11: Word cloud of provided department names

In trying to understand which area of an institution might take the lead for digital preservation general categories were provided in Question 14:

- Q14 - Which department(s) take the lead for digital preservation within your organization? If this is a fairly equally distributed effort choose more than one. [78 respondents]

This question asked respondents to characterize the type of department that takes the lead for digital preservation listing the choices of

- A library, archive or other department that stewards the collections;
- Information Technology (IT);
- Preservation department (handling both analog and digital);
- Other, for example the Vault department (Please Indicate).

Respondents were able to select as many as appropriate. Seventy-eight respondents answered this question with 51 selecting only one department (65%), 19 dividing the responsibilities between two areas (25%), and 8 selecting three department areas (10%); indicating that 35% of organizations share the responsibility of digital preservation across multiple departments.

The majority of those responding (56 out of 78) indicated that a library or archive or another department that stewards the collections is responsible for digital preservation, followed by IT (13), other (7), and a preservation department (5) (fig. 12).

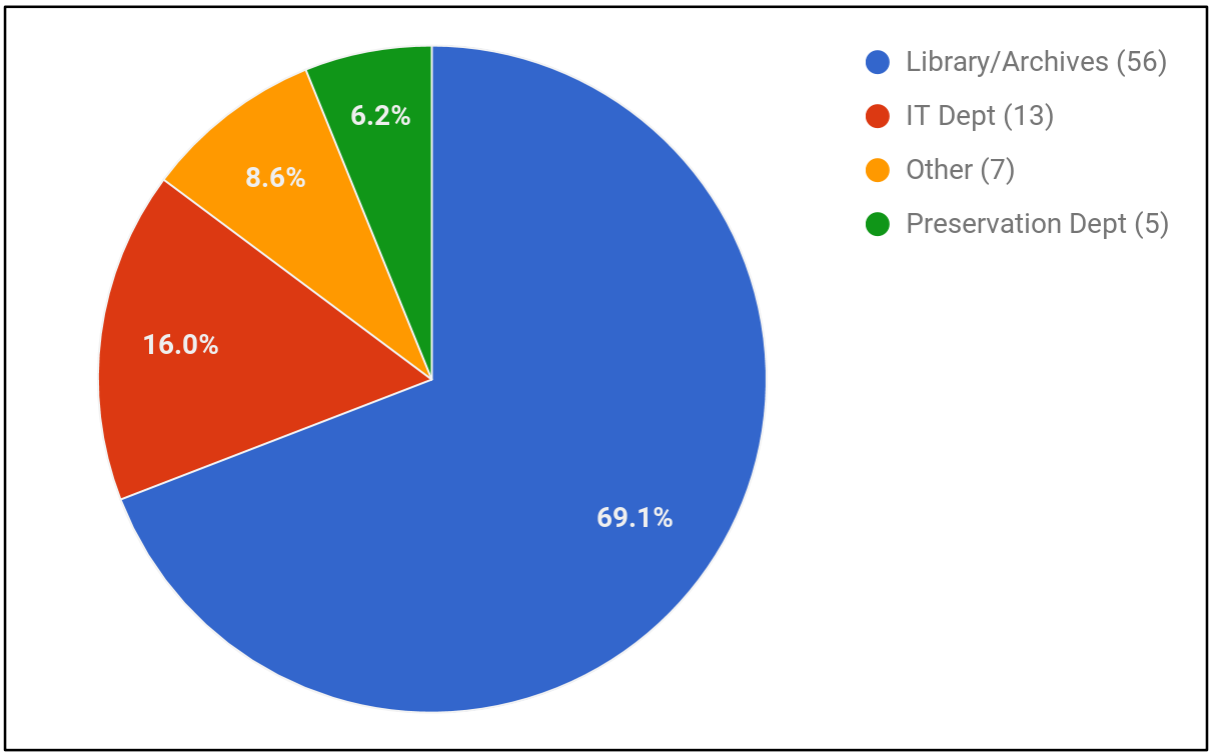

FIGURE 12: Departments responsible for digital preservation

Staffing

The next section of the survey (Q15) addressed staffing for digital preservation activities, asking respondents to quantify their current and ideal staffing levels for twenty specific positions or roles related to digital preservation. The positions listed were the same as those included in the 2012 survey. Respondents were also given the chance to list additional positions or roles in a free text field. To evaluate the matrix of responses, any value recorded as zero or greater was averaged for each position/role. Null responses were not included in the calculations because the intent of the respondent could not be

determined--did they mean zero, or were they registering no opinion? This methodology was consistent with the 2012 survey.

- Q15 - For each of these positions, how many FTE do you currently have supporting digital preservation (or supporting digital collections, even if the activity is not considered in scope of your DP program) and how many would be ideal? [130 respondents]

Totaling the averages for current staffing for all positions, we determined that institutions report an average of 13.6 FTE working on digital preservation activities. Ideally, respondents would like to double that to an average of 27.5 FTE, with a particular need for digital archivists, software developers, and cataloger/metadata analysts. Figure 13 shows respondents' current and perceived future needs based on the results of Question 15.

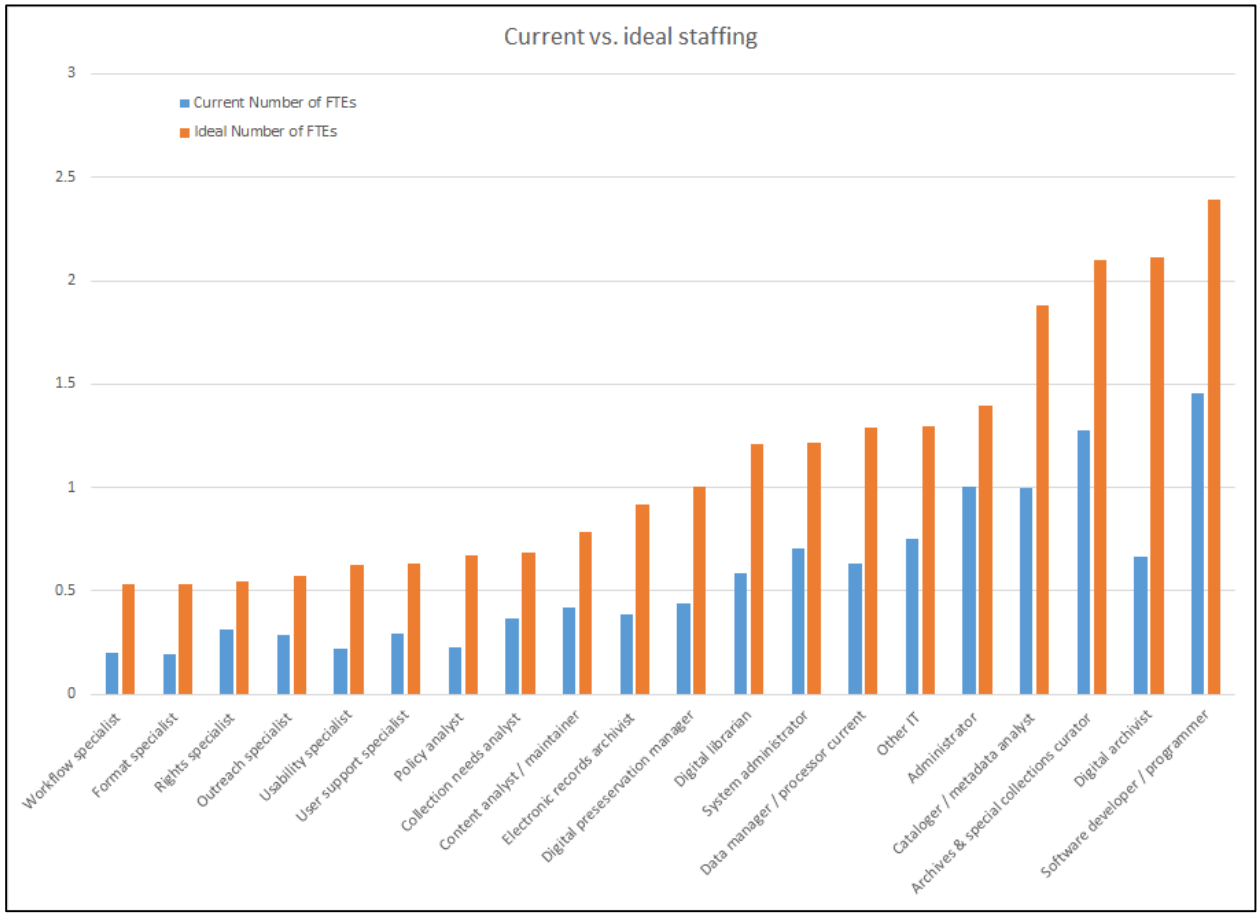

FIGURE 13: Current versus ideal staffing

After asking questions about the location of digital preservation within the organization and the staffing available, a follow-up question was asked to see how satisfied organizations were with the way things were currently structured.

- Q16 - The way our digital preservation function is currently organized (staffing levels, expertise, where they are placed within the larger organization) works well. [133 respondents]

When looking at the responses for the general organization of digital preservation programs respondents were not particularly pleased with their situation. About 46% of respondents indicated a negative view of their current organization, while only 25% were pleased. 29% were unsure, which may indicate either that at times they may feel that things are satisfactory, but other times not as sure, that there are a mixture of good and bad characteristics, or that they were not in a position to evaluate the situation.

Dividing out the responses further, 9% strongly disagree that the current organizational structure works well and 1.5% strongly agree, indicating that there are more negative feelings than positive ones about the overall organization of digital preservation activities (fig. 14).

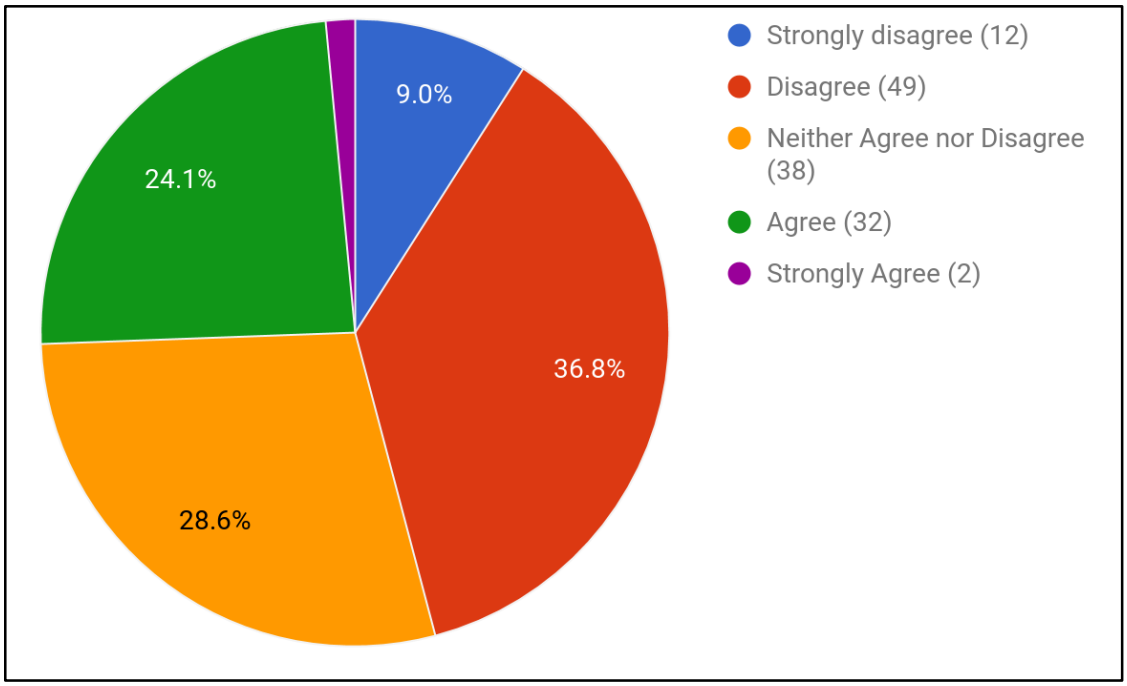

FIGURE 14: The way our digital preservation function is currently organized (staffing levels, expertise, where they are placed within the larger organization) works well.

The survey asked if organizations had organizational charts that they would be willing to share to better understand where and how digital preservation activities fit within the overall structure.

- Q20 - Do you have organizational charts or position descriptions that you'd be willing to share? Any documents you share would provide context to your answers and would be kept private to the NDSA Standards and Practices Working Group unless you give us explicit permission in the future to share more broadly.
- Q21 - If they are online, what are the URL(S)? If they are not on-line, please email this supplementary data to ndsa@diglib.org with the subject line "Staffing Survey."

In response to the first question, 29 organizations answered yes; 10 shared online links; 2 stated they would email them; 2 stated that they were in a foreign language (and were not shared because of this); and 6 emailed information. While interesting, the organizational charts were so specific to each organization that we could extract few generalities. However we appreciate those who took the time to share their documents with us.

Staffing Qualifications and Training

The next set of questions asked whether organizations hired new staff with digital preservation experience or retrained existing staff to manage preservation functions, as well as the importance of particular qualifications when searching for a digital preservation manager. Both questions also appeared on the 2012 survey.

- Q17 - For in-house staff, did you hire experienced digital preservation specialists and/or retrain existing staff? Check all that apply. [130 respondents]
- Q18 - Please rate the importance of each of these items if you were hiring a new digital preservation manager at your organization. [131 respondents]

When asked if organizations hired experienced digital preservation specialists or retrained existing staff, the results show that 68% retrained existing staff, 42% hired experienced digital preservation specialists, and 17% selected “other” (fig. 15).[8] Explanations provided for “other” included hiring recent MLS graduates versed in the theory of digital preservation who lacked hands on experience; hiring staff with an interest in learning digital preservation; while others indicated that staff were learning as they go or that they were not currently doing any hiring or training.

[8] Note that respondents were allowed to select more than one choice.

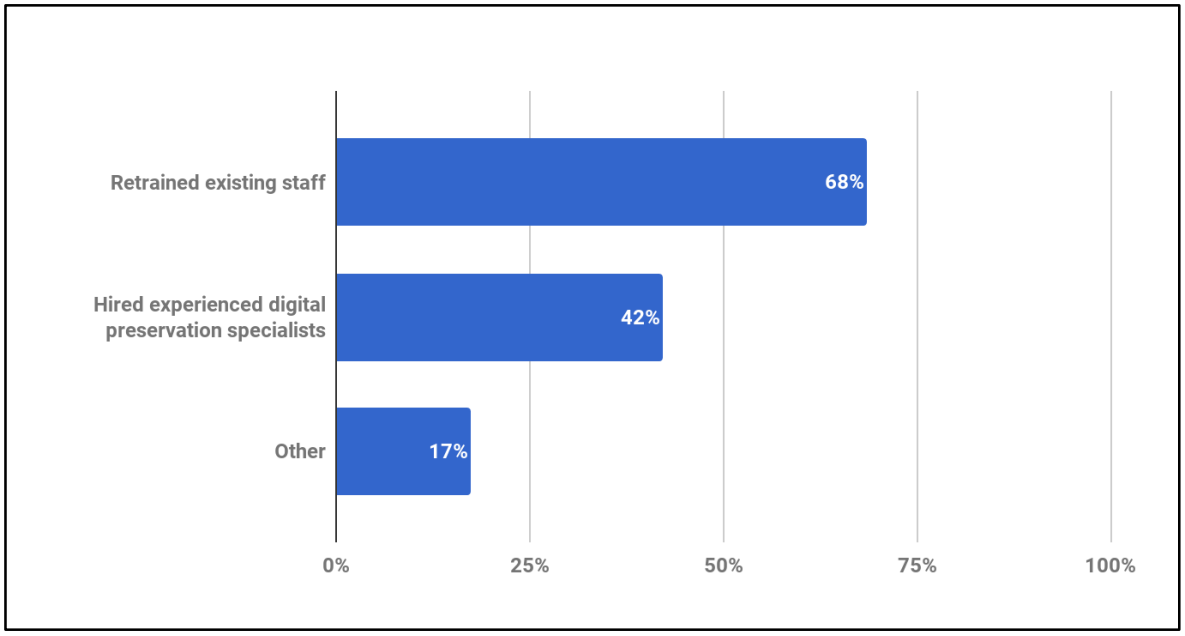

FIGURE 15: Retraining versus hiring staff

Respondents were asked to rank the relative importance of skills including, knowledge on specific topics and education qualifications if hiring a new digital preservation manager. Each skill could be ranked as extremely important, very important, somewhat important, somewhat unimportant, or not at all important (Q18). The list of skills or requirements included:

- Ability to train others
- Analytical skills
- Certificate in digital preservation or curation
- Collaboration skills
- Communication skills
- Degree in Computer Sciences
- Degree in Library and Information Sciences
- Knowledge of digital preservation standards-best practices-tools
- Leadership qualities
- Managing budgets
- Passion/motivation for digital preservation
- Professional digital preservation experience
- Project planning/management and organizational skills
- Technical abilities
- Other (please indicate)

Of the options provided, “knowledge of standards and best practices” and “communication skills” were considered among the most important qualities followed by “passion and motivation for digital preservation” and “collaboration skills.” Qualities of lesser importance to respondents related to staff acquiring or holding degrees and certificates. Figures 16 and 17 provide more details about which qualifications respondents found most and least important for a digital preservation manager.

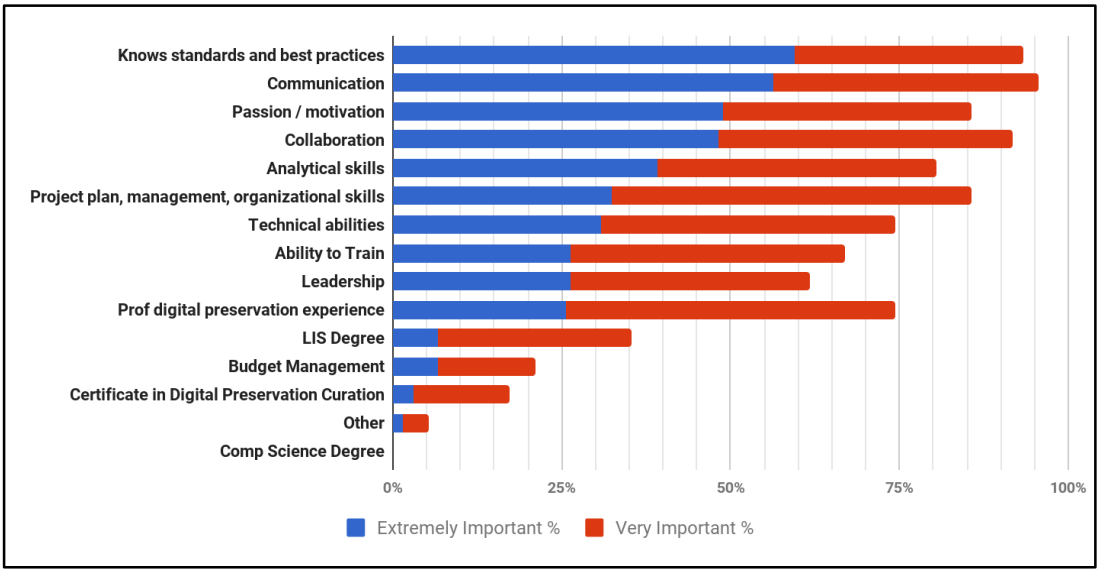

FIGURE 16: Extremely/very important qualities for digital preservation manager

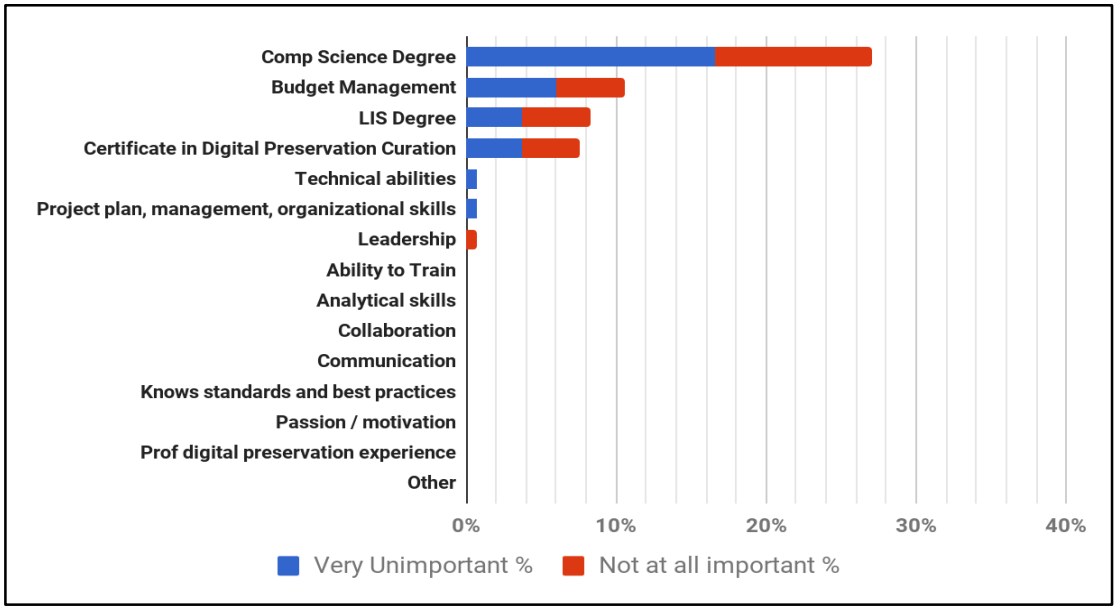

FIGURE 17: Very unimportant/not at all important qualities for digital preservation manager

Comments about Digital Preservation Program Staffing and Organization

A freeform question was asked towards the end of the survey in order to provide respondents with the ability to share their opinion, discuss their staffing organization, or ask questions of their own.

- Q19 - Is there anything else you'd like to share about the way you think an effective digital preservation program should be staffed and organized? [37 respondents]

Thirty-seven organizations responded to this question to offer insight into their programs or how they responded to the survey. We categorized their comments into the following areas:

- Comments about the survey in general (4)
- Formats (1)
- Just beginning/don't fully understand yet (3)
- Organizational - lone worker (2)
- Organizational - shared responsibilities (4)
- Organizational - preservation department (3)
- Staffing numbers (7)
- Storage and access (2)
- Support from institution (5)
- Training/skill set (6)

Together, the majority of these comments address institutional organization and support. Comments indicated that the ways in which institutions organized depended on their staffing level and abilities, but also on their organization's culture and support. Survey respondents specifically noted that they need:

- [The ability to] move from externally funded projects to scoped and well thought out internally supported programs; and
- "Buy in of constituents and administrators along with a healthy, supportive, knowledgeable infrastructure"

Within the comments, there were differing opinions as to the best organizational model for a digital preservation program. Some thought that digital preservation should be a shared activity while others thought that it should stand on its own.

Those endorsing a shared model commented:

- "Our experience has shown that shared responsibility across the organization in the areas where digital preservation is critical is important and lends itself to more effective practices and buy-in, rather than placing all ownership of digital preservation in one department or person."
- "At this time in small to medium size institutions having cross over between departments is crucial. There isn't enough funding for all the things we would like to do, so sharing resources is the only way to continually improve and provide services. Students often will be the backbone to get things done with staff guiding them."
- "While a dedicated digital preservation team is good, it can only work well if the efforts are supported throughout the entire organization and if resources from other areas (e.g. IT, collection teams) can be provided when needed."

Others who favored a dedicated department responded that:

- "Having a dedicated department where job responsibilities lie exclusively with digital preservation would help establish a more solid program and better ensure long-term preservation and access of digital items. Currently, our organization's digital preservation responsibilities are distributed among many departments; having staff members centralized in one department would enhance focus and communication and would ensure that digital preservation tasks are completed. Additionally, these services could expand to serve more functions across campus (i.e., preserve e-mail, websites, research data, etc.) if this were in place."
- "Digital preservation should be primarily housed in a preservation department."
- "Digital Preservation should fall under the Preservation dept."

Comments specifically from lone-arrangers, i.e., individuals who are doing everything within their organization themselves, state that "building an effective digital preservation program is incredibly hard"; "the real struggle is time" [to do the work that needs to be done] and larger organizations take for granted the staff they have that can do the work; and "more attention should be paid by educators and funders to the practicalities of implementing complex technical solutions with limited resources."

Other general comments of interest about staffing issues include:

- "Continuing education is a must, so that staff can stay up-to-date on current trends and the latest news in technology."
- "It should be organized keeping in mind extra skills of an individual in addition to

his/her specialist skill set. Otherwise the work can get boring or monotonous for one person working over a few years. For example, there should be a specialist for more or less all the functions mentioned at top of page [on the survey] but they must also have workable skills for 2-3 other functions. As such each function should have a scalable scope so that promotions for staff can take place as the scope of work grows."

- "Staffing is critical to success. We find that no matter the rating of digital preservation importance, the staffing increases to physical processing versus digital preservation. In addition, when hiring staff, it is just as important to ensure digital preservationists possess the required skillsets and are provided with the tools necessary to accomplish the tasks. Knowledge of tools is one thing but, if not providing access to the appropriate tools, equipment, connectivity and admin rights, it is of little use. All the pieces must be in place for a successful digital preservation effort."

ANALYSIS OF STAFFING LEVELS BY AMOUNT OF CONTENT

When reviewing the survey responses, Question 15 gave us an overview of the aggregate current staffing situation and perceived staffing needs for responding institutions. Responding institutions reported an average of 13.6 FTE working in digital preservation activities and noted that they would like to double that to an average of 27.5 FTE. This interest in doubling staffing levels was similar to the results in 2012.

Given the range of institutional missions, the amount of stored data each organization might be caring for, and likely differences of total staff sizes of the responding institutions, we felt that our overall findings deserved further analysis. We wanted to see if there was any correlation to the amount of content being managed and the amount of current or desired staffing.

To help with this analysis, we used Question 6 (the amount of online or offline storage space) to group respondents into smaller categories based on the amount of data they manage. We separated respondents into five groups: less than 1 TB (3 respondents), 1–50 TB (77), 51–100 TB (22), 101–500 TB (19), and over 500 TB (11), which were the options listed in Question 6.

With only three responses from repositories housing less than 1 TB of data, we found it hard to come to a firm interpretation of the results for this grouping. Nevertheless, the

data suggests that institutions storing a lower amount of digital content rely on existing staff to manage digital content preservation as an add-on to other responsibilities. The graph in figure 18 reveals a striking lack of overlap between the current staffing reported (in blue) which concentrated on the other IT, administrator, cataloger/metadata analyst, and archives and special collections curator positions, and the desired ideal staffing needs (in orange), which lists two completely different roles: collection needs analyst and digital preservation manager. We do not know why these organizations did not list existing positions as future positions as well. They may want to reorganize the staffing altogether or possibly misunderstood the question.

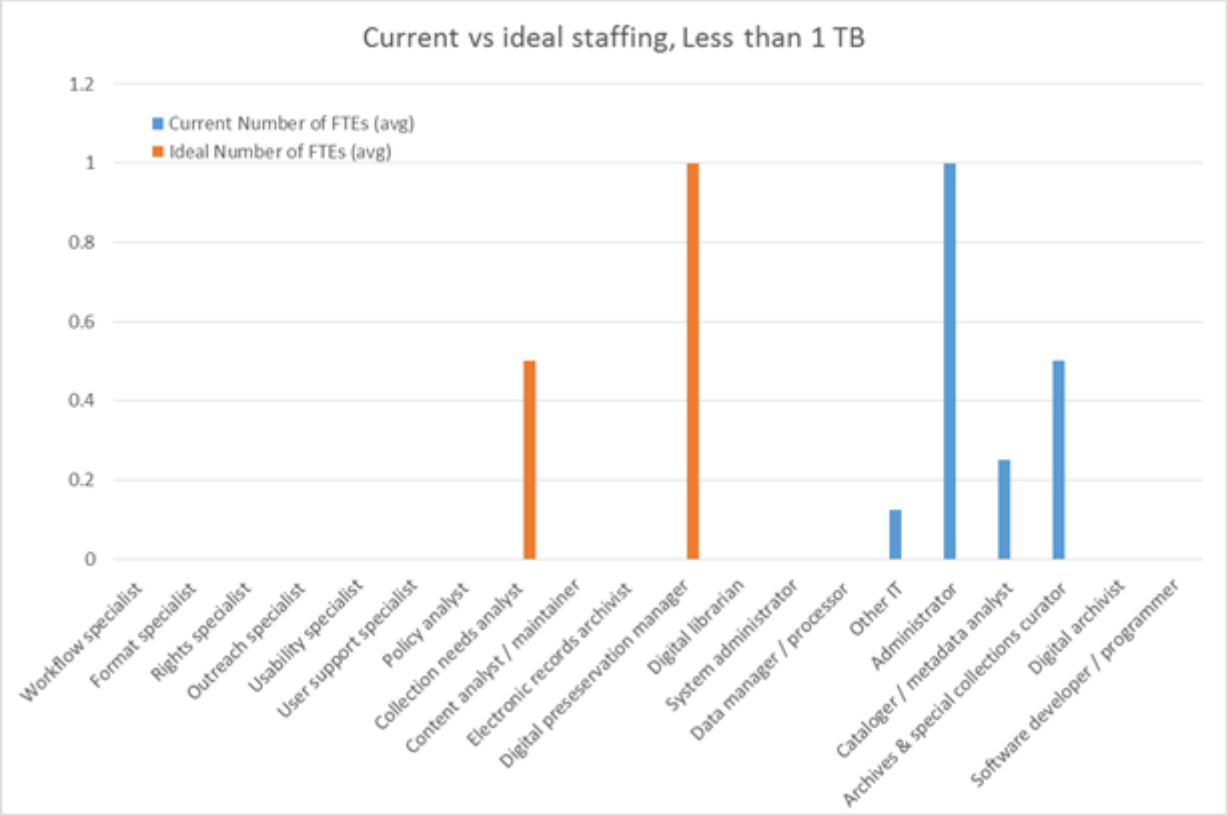

FIGURE 18: Current versus ideal level of staffing for respondents with less than 1 TB of data

Organizations preserving over 1 TB of digital content reported increased staffing to address responsibilities. Apart from the group managing less than 1 TB of data, in each aggregated group all roles listed as options in the survey reported some level of current staffing or anticipated staffing need; however, no single institution reported a need for staffing in every role listed, whether current or ideal.

For example, repositories managing between 1 and 50 TB of content demonstrated what might be termed "growing pains" as they felt the greatest need for additional staff. Although these 77 institutions had an average of 10.7 FTE staff addressing digital preservation, they reported an ideal of nearly three times as many staff, or 30.6 FTE, with a particular gap between current and ideal staffing for the roles of digital archivist, cataloger/metadata analyst, and software developer/programmer (fig. 19).

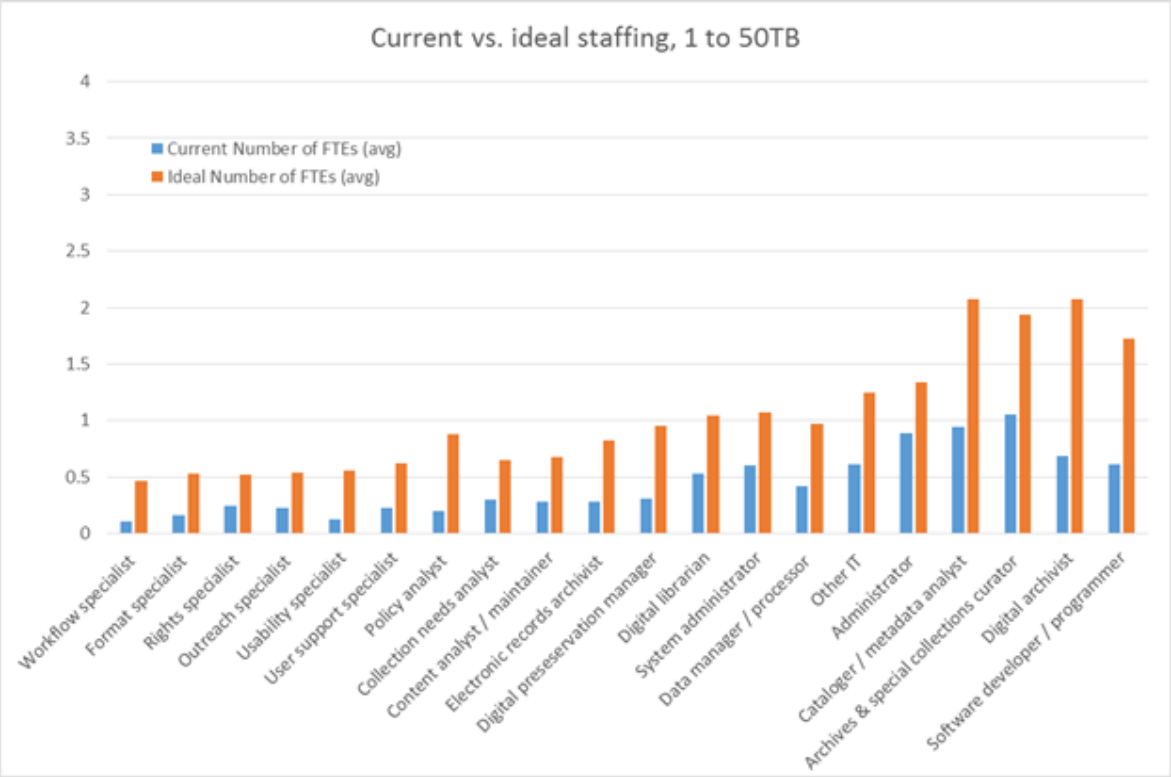

FIGURE 19: Current versus ideal level of staffing for respondents with 1 to 50 TB of data

Unsurprisingly, this need for more technical skills continued as the amount of managed content increased (figs. 20, 21, and 22). Still, we noted some differences. In our overall analysis the need for staff in the "software developer/programmer" role showed the highest current and ideal staffing levels. In reviewing the data tiers, though, we found it interesting that the "software developer/programmer" role's high ranking did not really assert itself until repositories grew larger than 100 TB. Below that threshold, the highest perceived staffing need appeared to be in hybrid roles that combine curatorial and technical processing—positions like digital archivist, cataloger/metadata analyst, data manager/processor, and digital librarian. That may suggest that the more content repositories have to manage, the more they need to customize or create software solutions to meet their needs, thereby increasing the need for software development expertise.

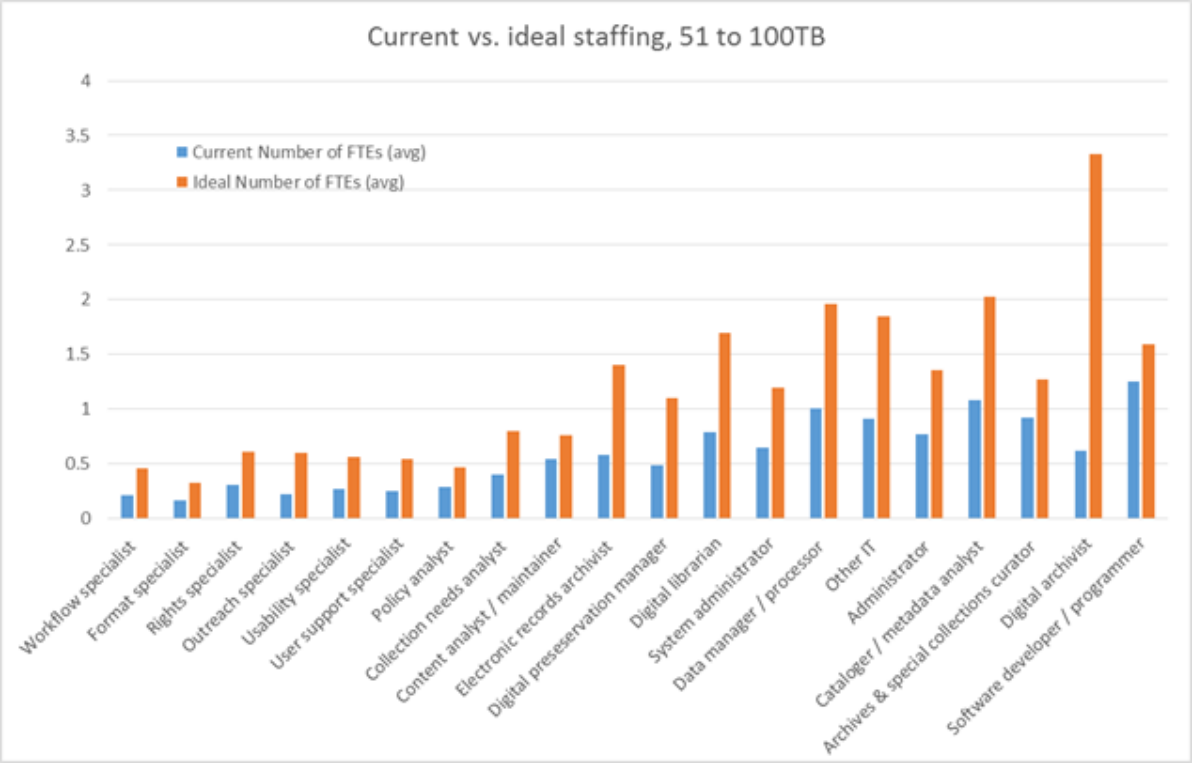

FIGURE 20: Current versus ideal level of staffing for respondents with 51 to 100 TB of data

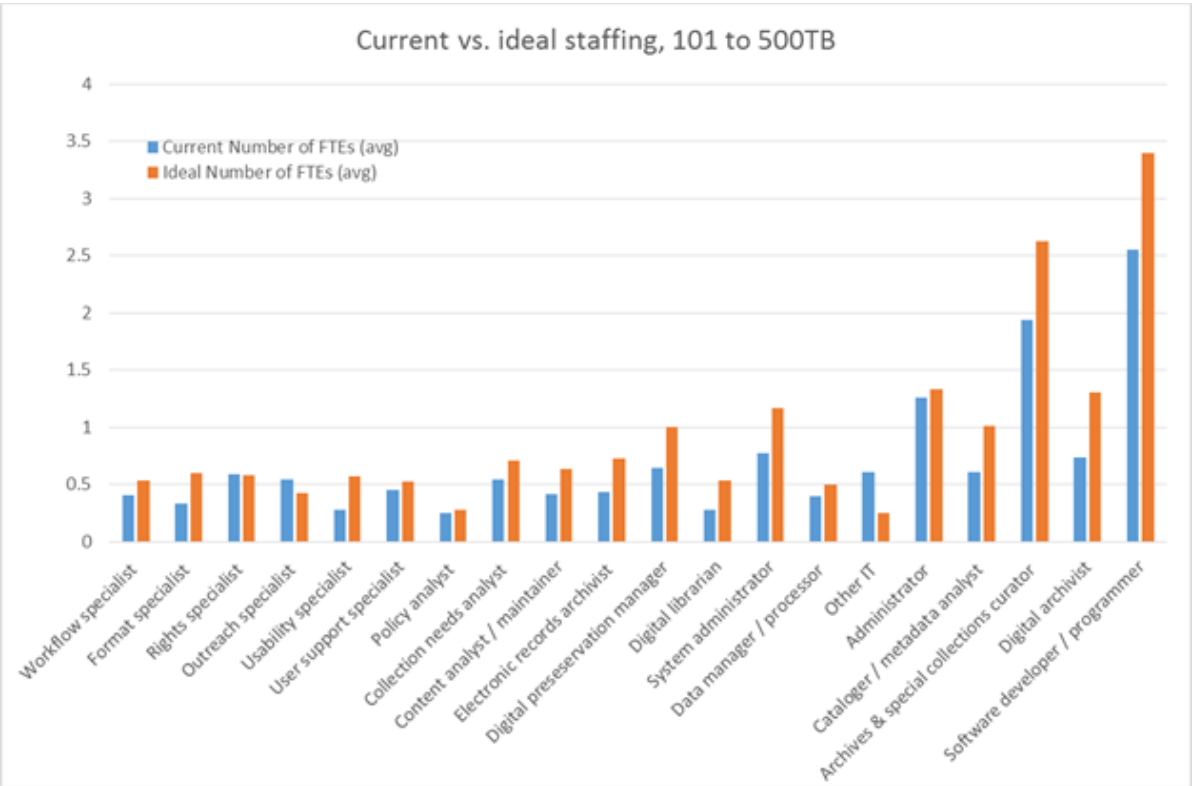

FIGURE 21: Current versus ideal level of staffing for respondents with 101 to 500 TB of data

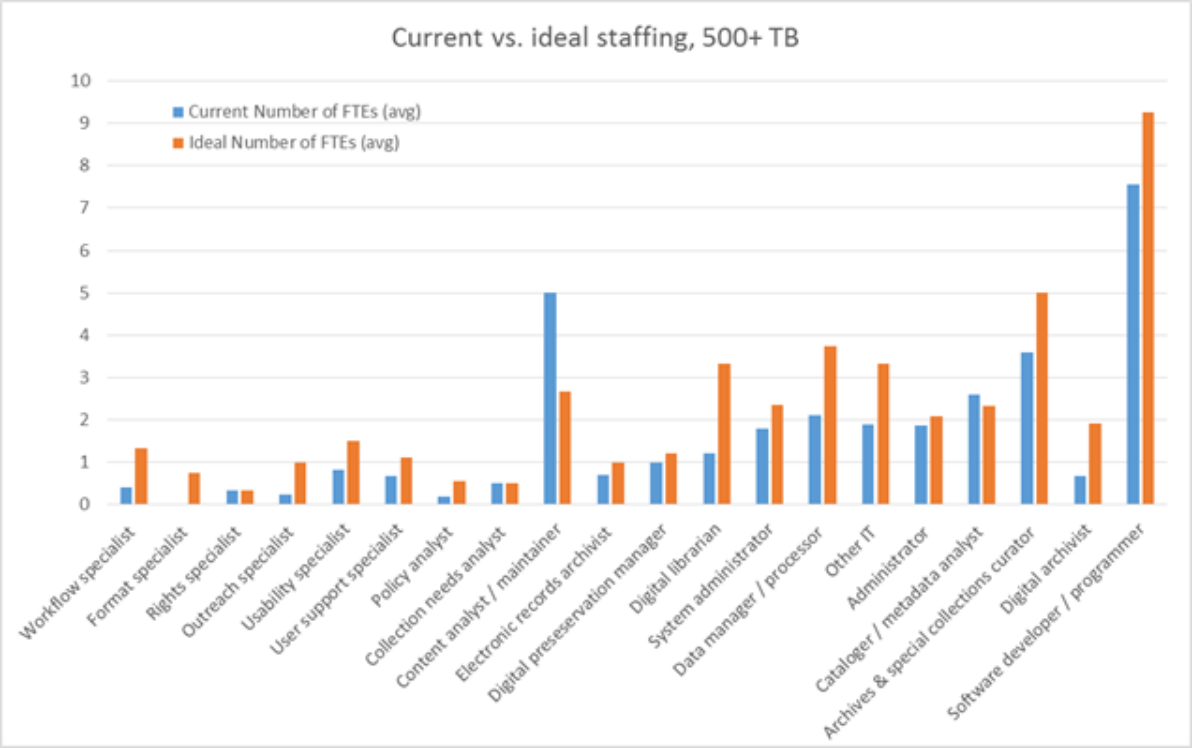

FIGURE 22: Current versus ideal level of staffing for respondents with 500+ TB of data

The charts above show that in nearly all instances, current average levels of existing staffing fall below the ideal; however, the difference between current and ideal staffing varied significantly from tier to tier (table 1). The second tier (1–50 TB) reported that ideal staffing would be nearly three times current average levels, while the next highest tier (51–100 TB) reported an ideal of two times current average staffing. The fourth tier (101–500 TB) saw a more modest 30 percent growth as ideal, and the largest storage tier (more than 500 TB) identified a 55 percent average growth for the ideal.

Amount of Managed Content	Average Current Staffing	Average Ideal Staffing	Approximate Increase (or Decrease)
Less than 1 TB	2.9 FTE	1.5 FTE	(48%)
1–50 TB	10.7 FTE	30.6 FTE	186%
51–100 TB	12.04 FTE	23.95 FTE	99%
101–500 TB	15.45 FTE	20.22 FTE	31%
More than 500 TB	34.8 FTE	54.0 FTE	55%

TABLE 1: Comparison of average current and ideal staffing by storage tier

In general, respondents from institutions in the two tiers between 1 and 100 TB appear to need the largest proportion of additional staffing. These institutions may need to staff newer digital preservation programs, or a baseline set of additional skills needed for a successful program may emerge as programs cross these storage thresholds. Programs managing more than 100 TB of data appear already to have acquired many of these sets of skills, and new staffing may be in support of anticipated growth or to provide additional services. Only institutions managing less than 1 TB of data showed a decrease in ideal staffing levels; the low number of responses (3) in this tier makes generalization about this group difficult.

COMPARISON WITH 2012 DATA

One of the reasons for using the same 2012 questions for this survey was to allow for analysis of data over time. The following section discusses some of the more interesting comparisons allowed by these two snapshots of digital preservation staffing taken five years apart.[9]

[9] The prior survey with which we are comparing data was given and completed in 2012, although the date of the published report is 2013.

Survey Response

Forty-eight more organizations completed the survey in 2017 than in 2012. Based on the responses provided for institutional name, we noted that only 8% of 2017 survey respondents had also taken the 2012 survey, which means that the survey results are not tracking change in the same organizations, just general practice in the field. This low rate of return was surprising to the team although we know that the survey was announced in different places and is assumed that it was publicized more widely in 2017.

In 2017, we had more international participation than in 2012, with respondents from thirteen countries other than the United States. Ten countries other than the United States participated in 2012. In 2017, 78% of the respondents were from the United States, whereas 86% of responses were from the US in 2012.

Organization Information

Organization Type

The types of organizations with the highest percentage of respondents for both years were academic libraries or archives and government entities. The top three for both years are listed in table 2. All other types of organizations were 6% or below for both years. Organization types were provided in Question 4.

2012	2017
Academic library or archives (45%)	Academic library or archives (46%)
Government entities (11%)	Government entities (11%)
Public Libraries (10%)	Museum (8%)

TABLE 2: Top three types of organizations in 2012 and 2017

Amount of Managed Content

The amount of content being managed by respondents stayed roughly the same (table 3). The notable change was the doubling of the percentage of organizations managing over 500 TB; from 4% to 8.3%. There was also a slight decrease in the percent of organizations caring for between 101–500 TB; from 17% to 14.3%.

Amount of Managed Content	2012	2017
0 TB	3%	2.3%
1–50 TB	59%	58.6%
51–100 TB	16%	16.5%
101–500 TB	17%	14.3%
>500 TB	4%	8.3%

TABLE 3: Amount of content being managed

Expected Volume Increase

In 2012, 20% of respondents expected the amount of content they held to increase two-fold, while 68% expected up to a 50% increase. It is impossible to determine if institutions from 2012 did increase their content holdings, especially with such limited repeat survey respondents. However the trend remains the same in that respondents expect growth in their own holdings over the next year although not as significantly (table 4).

Expected Increase in Volume	2012	2017
0–25%	*NA (grouped into 0–50%)*	*73.2%*
26–50%	*NA (grouped into 0–50%)*	*12.6%*
0–50%	68%	85.8%
51–99%	12%	1.6%
>99%	20%	12.6%

TABLE 4: Comparison of the expected one year volume increase

Type and Number of Files

The types of files being managed over the past five years has not changed much. Still/2D images, text, and video remain the most common, followed by audio recordings, websites, spreadsheets, datasets, etc.

Digital Preservation Activities

A multipart question on the survey (Q28) asked respondents to identify which of the listed digital preservation activities were considered in scope, were done in-house, were outsourced, and if they would like to outsource the activity in the future.

In reviewing these responses, there was a significant increase in the number of organizations that considered the activities listed on the survey as in scope or part of digital preservation activities between 2012 and 2017. In 2012, all of the activities on the list (with one exception) were considered in scope by at least 50% of the respondents. The one exception was emulation, considered in scope by only 22%. By 2017, all but emulation were in scope for at least 60% of respondents. In fact, all but four activities were in scope for more than 70% of the respondents. The exceptions included emulation, at 53%, research at 64%, tools development and maintenance at 65%, and descriptive cataloging at 68%.

A closer look at descriptive cataloging makes it clear that most organizations are doing this activity (80% are doing it in-house, and 8% are outsourcing), but some consider it part of another program rather than digital preservation.

In 2012, the activities *most likely* to be outsourced were digitization (37%), secure storage management (22%), and creation of access copies (20%). In 2017, the most commonly outsourced activities were digitization (51%), secure storage management (38%), and tools development/maintenance (32%, outsourced by just 17% in 2012). In aggregate, this shows more outsourcing overall but no radical changes in the types of activities outsourced.

The activities respondents are *least likely* to outsource in 2017 are preservation policies/strategy (4%), selection for preservation (4%), and preservation planning (5%). In 2012, no organizations were outsourcing preservation planning, selection for preservation, or development of guidelines for content creators. Again, the 2017 survey shows greater openness to outsourcing.

In 2012, respondents expressed the most interest in future outsourcing of digitization (65%), secure storage management (43%), format transformation/migration (34%), metadata extraction (32%), and the development and maintenance of tools (29%). In 2017, the most common responses to the question about what they might outsource in the future were emulation (35%), tools development/maintenance (32%), secure storage management (30%), and digitization (29%), which correspond to the most commonly currently outsourced activities plus emulation (table 5).

Activity to Outsource	2012	2017
Digitization	1st (65%)	4th (29%)
Secure storage management	2nd (43%)	3rd (30%)
File format migration/ transformation	3rd (34%)	6th (14%)
Metadata extraction	4th (32%)	8th (17%)
Development and maintenance of tools	5th (29%)	2nd (32%)
Emulation	12th (12%)	1st (35%)

TABLE 5: List of activities organizations would like to outsource. Rank and percent shown for 2012 and 2017.

It is interesting to note that the interest in future outsourcing fell noticeably, while the percentage of organizations actually engaged in outsourcing rose.

Department Organization and Digital Preservation Comparison

The survey team wanted to know how institutions were organizing themselves when it came to addressing digital preservation and asked if there was a dedicated digital preservation department, and if not which area of the institution took the lead for digital preservation activities.

When asked if their institutions had a dedicated preservation department, there was a slight decrease in the number of respondents answering “yes,” and in those responding “no” across the years. The uptick in “not applicable” responses showed the wide scope of strategies for addressing digital preservation. Explanations for “not applicable” ranged from a respondent noting he was a “One Man Show” to another institution stating they are “all digital, all the time” (table 6).

Dedicated Department?	2012	2017
Yes	33%	31.6%
No	59%	58.6%
NA	8%	9.8%

TABLE 6: Does your institution have a dedicated digital preservation department?

For those who do not have a dedicated digital preservation department, a library, archives, or other department that stewards the collections remains the most common response at close to 70% in both 2012 and 2017. In 2017, all of the other options fall below 20%, which is a significant change from 2012 results which stated IT departments were responsible for just over 40% of respondents. It is clear, however, that the departments most closely related to the collections themselves retain the responsibility for preserving the electronic collections (table 7).

Department with Digital Preservation Responsibilities	2012	2017
Library/Archives	73%	69.1%
IT	42%	16%
Preservation Department	20%	6.2%
Other	15%	8.6%

TABLE 7: Which department is responsible for digital preservation?

Organizations could select one or several departments in response to this question. We note that there is a significant decrease in the IT programs and preservation departments having the lead responsibility, while in both years the library or archives (or department that stewards collections) was the most common.

Current and Ideal FTE Comparison

One of the main objectives for this survey is to understand the current staffing levels for digital preservation activities as well as the desired level of staffing. Having two sets of data allows us to begin to identify any possible trends over time.

In 2012, the average responding institution had 11.4 FTE engaged in digital preservation, and felt 21.14 FTE would better serve their needs, a shortage of 9.74 FTE. By 2017, the respondents' level of current staffing had increased by 2.2 FTE to 13.6 FTE, but the ideal staffing level had increased by 16.1 FTE to 27.5 FTE. In other words, organizations reported themselves to be even more understaffed than in 2012, in spite of the increased number of staff. In 2012, organizations felt current staffing provided just over half (54%) of what they believed they needed. Despite modest staffing gains, respondents in 2017 reported that they had lost ground against ideal staffing and had only 49% of what they believed they needed.

The relative proportion of institutions in the four "size of storage" categories (Q6) remained fairly stable from the 2012 survey to this one. Consequently, the perceived need for significantly more staff in 2017 isn't driven by a higher percentage of respondents with larger collections than before. It may be that five years on, institutions have a more nuanced view of the responsibilities for which they need additional staff expertise.

More interesting are the changes in emphasis over the five years between surveys. Of the five highest needs identified in 2012, only the role of software developer/programmer remained in the top quartile for 2017. Even here, ideal staffing seems almost within reach of current staffing levels, needing only 1.11 FTE to reach the overall ideal staffing level. The greatest shortfall appears for staff serving the digital archivist's role. While the perceived need for digital archivists has grown by one FTE since 2012, the average current staffing level has grown only 0.12 FTE during the intervening five years. Meanwhile, the perceived need for people in the content analyst/maintainer role has dropped significantly, with ideal staffing coming closer to the current staffing levels. In 2017, staffing for the policy analyst responsibilities appears to have come closer to the level given as ideal in 2012.

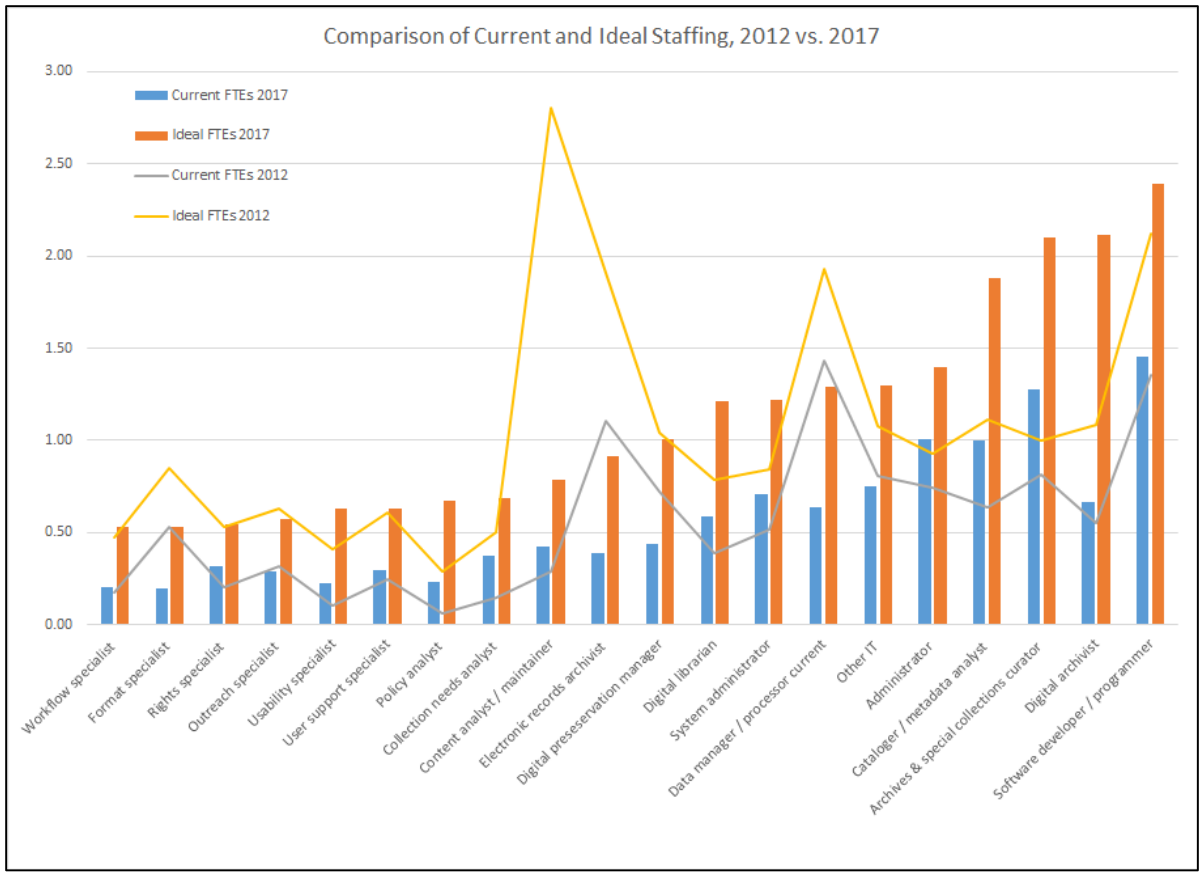

FIGURE 23: Comparison of current and ideal staffing, 2012 versus 2017

Perceived Satisfaction in Overall Organization

In 2017, 46% of respondents were not satisfied with how the function of digital preservation was organized within their institution, while only 25% were pleased, and 29% were unsure (answering “neither agree nor disagree”). This is almost opposite of what the 2012 results show in which 43% of the 2012 survey respondents agreed or strongly agreed that the digital preservation function within their institution was well organized while 34%, were not satisfied with how things were organized, and 23% were unsure (answering “neither agree nor disagree”) (table 8).

This overall lack of satisfaction is one of the most intriguing findings of the 2017 survey. Further study would be needed to identify exactly why institutions are so dissatisfied with the way the digital preservation function is organized within their institution.

Digital Preservation is Well Organized	2012	2017
Strongly Agree	6%	2%
Agree	37%	24%
Neither Agree nor Disagree	23%	29%
Disagree	28%	37%
Strongly Disagree	6%	9%

TABLE 8: Perceived satisfaction in the organization of digital preservation activities in own institution

Digital Preservation Specialist Staffing

When asked about hiring experienced digital preservation specialists and/or retrain existing staff, the 2017 responses show a decrease in retraining existing staff and an increase in hiring experienced digital preservation specialists (table 9). This may suggest that increasing demand for these skills has led to more opportunities for potential staff to have received formal training than we saw in 2012. There was also a decrease in the percent of "other" responses, decreasing from 21% to 17% although the examples provided under the "other" response were quite similar such as "hiring recent MLS graduates" or having staff "learn as they go" without any specific training.

Retrain versus Hire	2012	2017	Change
Retrain Existing Staff	75%	68%	-7%
Hire Experienced Specialists	35%	42%	+7%
Other	21%	17%	-4%

TABLE 9: Do repositories fill skill needs by hiring new staff or retraining existing staff?

When looking to hire a new digital preservation manager, the importance of qualifications selected as extremely important also have not changed drastically, with the top six only shifting positions slightly (table 10).

Importance of Qualifications	2012	2017
Knowledge of digital preservation standards/best practices	2nd	1st
Communication	3rd	2nd
Passion and motivation for digital preservation	1st	3rd
Collaboration	6th	4th
Analytical skills	4th	5th
Project planning/management	5th	6th

TABLE 10: Most important qualities for a digital preservation manager

Looking at the other end of the spectrum, the possession of specific degrees was once again rated "not very important," however in 2012, there were twice as many qualifications listed as "not at all important." In 2017, the five "not at all important" qualifications included: Degree in Computer Science, Budget management, LIS degree, Certificate in Digital Preservation Curation, and Leadership; while the qualifications of knowledge of digital preservation/best practices, communication skills, technical abilities, analytical skills, and professional digital preservation experience are no longer expressed as being "not important at all."

CONCLUSION and NEXT STEPS

The 2012 NDSA Digital Preservation Staffing Survey established a useful baseline description of how 85 organizations involved in digital preservation, mainly in the United States, addressed staffing, scoping and organizational questions. The current 2017 survey allows the community to get an updated look at these same questions, and for the first time, a sense of the direction of trends in staffing in the digital preservation field.

The survey shows that overall digital preservation programs have grown since 2012 in the amount of content they manage, as well as the number of activities they consider in scope that they perform both in-house and outsource, and have an increased number of staff available to perform digital preservation activities. Results also show that the rate of

anticipated growth in collections is lower than in 2012, although respondents still reported that they need far more staff than they have (nearly twice as many full time equivalent staff).

One of the most striking findings was the increased percentage of respondents who reported that they were not satisfied with the way the digital preservation function was organized: 46% not satisfied in 2017 (25% satisfied), up from 34% not satisfied in 2012 (42% satisfied).

Only 32% of respondents reported that their organization had a dedicated digital preservation department, but the comments made it clear that not everyone considers a dedicated department an advantage. The survey did not provide enough information to know for sure why the number dissatisfied increased. One possibility is that the growing sophistication of these programs is leading to higher ambitions for how well they should function.

As follow up activities, the data in the current survey could be analyzed further and additional questions could be asked of it. For example, the survey team did not have time to fully explore the effect of the *type* of organization on present or desired digital preservation staffing. This would be worthwhile follow-on work.

In order to further explore what changes from 2012 to 2017 are genuine trends and which are artifacts of the different pool of respondents, the survey team recommends that the same survey questions from 2017 be asked again in a new survey in several years.

In addition, the next survey could attempt to explore the more interesting results of the 2017 survey. For example, drilling down on the issue of *why* respondents are satisfied or dissatisfied with the organization of their digital preservation programs might provide actionable recommendations for strengthening programs. Possible additional questions could address the qualitative experience of why people answered as they did on the satisfaction question, or the next survey could explore the qualities that cause a digital preservation program to function well or poorly.

As we stated above, our analysis looked into staffing levels based on the amount of digital content and suggested possibly exploring staffing levels based on the type of organization. Another approach could be to include a question about the responding organization's total size. This might bring to the surface any relationship between organizations and their

staffing needs, and additional analysis would become available based on size for many areas the survey covers. This might be of broad interest, as there was such a wide variety of types of organizations responding to the survey.

More information could be gathered about the nature of digital preservation activities with additional questions, for example: Are activities centralized or decentralized in nature? What is the overall size of an organization versus the number of staff who actively participate in digital preservation activities? This might help shed some light on overall priorities—does a lone arranger or small staff choose to focus on digital preservation or on other necessary activities? More thought would need to be given to what these questions might look like or how the responses would better inform the community.

Another additional line of inquiry that was of interest to the group concerned actual job titles. The current survey asks about activities and the breakdown of FTE percentages, but it might be interesting to know what people are calling positions that work with digital preservation. Is "digital preservation" in the job title if it is the person's main focus? If digital preservation is not the position's main focus, but digital preservation activities are included under a broader position, what is that position called? Are there similarities across organizations?

A future survey of this type, repeating the core questions in the form they were given in 2012 and 2017, but adding additional questions that help provide some additional information into the organization of digital preservation programs as well as explain why some respondents are dissatisfied with their organizations and how they could be improved would be a contribution to the profession.

APPENDIX A: Survey Statistics from 2012 and 2017

General comparisons between the 2012 and 2017 survey data:

	2012	2017
Total number of respondents who completed survey	85	133
Survey started	131	168
Completion Rate	65%	79%
Percent from United States	86%	78%
List of Countries Outside the US	Australia (1), Canada (4), Denmark (1), Estonia (1), Germany (1), Ireland (1), Japan (1), Netherlands (2), Switzerland (1), and United Kingdom of Great Britain and Northern Ireland (1).	Australia (1), Canada (1), Denmark (1), Estonia (1), Germany (4), Luxembourg (1), Netherlands (2), Pakistan (1), Slovenia (1), Sweden (1), Switzerland (3), United Kingdom of Great Britain and Northern Ireland (8), and Uruguay (1).
Percent from Academic Libraries	45%	46%
Percent from Government entities	11%	11%
Percent participating in Consortia	54%	52%

APPENDIX B: Survey Questions

Note: The Qualtrics software numbers the questions in the order created, rather than in the order in which they appeared to respondents. The order in which the questions are listed below is the order in which the respondents saw them.

Q1. Initiated within the Standards and Practices Interest Group, members of the National Digital Stewardship Alliance (NDSA) based this survey on one created in 2012 to assess current and anticipated staffing needs. Data collected from this survey will provide a more recent perspective of digital preservation staffing needs and of how perspectives have changed since the 2012 survey and report.

This survey is intended for organizations that are currently responsible for digital preservation, whether that responsibility is fulfilled in-house or outsourced to a commercial, nonprofit, or consortial provider.

It will be used to understand current and ideal practice in staffing digital preservation programs. (We encourage participation from all types of organizations and you do not need to be a member of NDSA to complete this survey).

Only one response should be submitted per organization. All questions are optional unless otherwise noted.

We will make our best effort to protect your individual survey responses so that no one will be able to connect your responses with you or your organization. Any personal information that could identify you or your organization will be removed or changed before results are made public. We will combine your responses with the responses of others and make the aggregated results public, and preserve the anonymous data long-term for research purposes. Please email any questions to ndsa@diglib.org with the subject line “Staffing Survey.”

The NDSA's institutional home is the Digital Library Federation (DLF), at the Council on Library and Information Resources (CLIR). If you would like to learn about the NDSA, including how your institution can become a member, please see the NDSA web site.

Q2. (Required) What is the name of your organization?

Q3. Can we include the name of your organization in a list of organizations that responded to this survey? Knowing specific responding organizations may be helpful to people interpreting the survey results. If you agree to this we will still make our best effort to protect your individual survey responses so that no one will be able to connect your responses with you or your organization.

Q4. (Required) Which of the following most closely describes the type or function of your organization? [Options included: Academic institution department (not a library or archives); Academic library or archives; For-profit corporation; Historical society; Institutional repository; Independent library or archives; Government entity; K-12; Museum; National, federal or legal deposit library; Non-profit organization (not one of the above types; Public library; Research data repository; Research group; University; Other (please specify)]

Q5. In which country is the responding organization located?

Q6. How much online or offline storage space are you using for your digital content, not including backup copies? [Options included: 0, 1-50TB, 51-100TB, 101-500 TB, More than 500 TB (please enter the amount as a number)]

Q7. What do you expect the percent of growth to be of your preserved digital content over the next year? Please enter a whole number representing a percentage.

Q8. Roughly how much of each are you preserving, in terms of number of files? [Options included: None, A little, Some, A lot for the following categories: Still/2D Images (e.g. TIFF, JPEG), Drawings / vector graphics (e.g. CAD/CAM), Moving images / video, Audio recordings, Websites / blogs / social media, Text / documents (e.g. Word, PDF, TXT), Geographic Information Systems (GIS) data, Spreadsheets or datasets (other than GIS data), Databases, Computer games / software, Other (please indicate)]

Q9. Do you participate in any digital preservation consortial or cooperative efforts? [Options included: Yes (please enter name of consortia or cooperative), No

Q29. [If yes was selected in Q9 was provided as a follow up] What benefits do you gain from your participation? [Options included: Networking, Training, Storage space, Programming, Access interface, Communication/marketing, Consulting, Federated search, Other]

Q28. In the matrix below, please select all that apply for each activity. (A, B, C, and D can all be checked if applicable to your situation).

- Select A for activities that you consider in scope for digital preservation function at your institution, regardless of whether or not you are currently doing the activity. (This is about how you define digital preservation, not about what you do.)
- Select B for all of the activities that your organization currently does in-house.
- Select C for all activities for which your organization currently outsources.
- Select D for all activities you wish your organization would outsource in the future. Please include activities you are currently outsourcing that you want to continue to outsource.
  [Options for activities included: Selection for preservation, Digitization, Metadata creation / extraction, Descriptive cataloging, Transformation / migration of digital formats, Creation of access copies, Normalization of files, Fixity checks, File format identification, File format validation, Emulation, Content replication, Secure storage management, Technology watch, Development and maintenance of tools, Preservation planning, Development of preservation policies and strategy, Development of guidelines for content creators, Research, Preservation education / training and outreach, Other (please indicate)]

Q13. Is there a dedicated digital preservation department within your organization? [Options included: Yes (Please indicate name of department), No, Not applicable (Please explain)]

Q14. [Displayed if the answer to Q13 was Yes.] Which department(s) take the lead for digital preservation within your organization? If this is a fairly equally distributed effort choose more than one. [Options included: Information Technology (IT), A library, archives or other department that stewards the collections, Preservation department (handling both analog and digital), Other, for example the Vault department (Please indicate)]

Q15. For each of these positions, how many FTE do you currently have supporting digital preservation (or supporting digital collections, even if the activity is not considered in scope of your DP program) and how many would be ideal? FTE stands for full-time equivalent. For example a 1.0 FTE could mean one person full-time or 2 people half-time; a 0.5 FTE could mean one person half time or two people quarter-time. Please use whole numbers or decimals as appropriate. [Positions included: Digital preservation manager, System administrator, Software developer / programmer, Other IT, User support specialist, Collection need analyst, Policy analyst, Content analyst / maintainer, Data manager / processor, Cataloger / metadata analyst, Format specialist, Workflow specialist, Electronic records archivist, Archives and special collections curator, Digital librarian, Administrator / manager (other than digital preservation manager), Outreach specialist / trainer, Rights specialist, Usability specialist, Digital archivist, Other (please indicate)]

Q16. The way our digital preservation function is currently organized (staffing levels, expertise, where they are placed within the larger organization) works well. [Options included: Strongly disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree]

Q17. For in-house staff, did you hire experienced digital preservation specialists and/or retrain existing staff? Check all that apply. [Options included: Hired experienced digital preservation specialist, Retrained existing staff, Other (please indicate)]

Q18: Please rate the importance of each of these items if you were hiring a new digital preservation manager at your organization. [Items include: Degree in Library and Information Sciences, Degree in Computer Sciences, Certificate in digital preservation or curation, Professional digital preservation experience, Knowledge of digital preservation standards-best practices-tools, Technical abilities, Leadership qualities, Communication skills, Analytical skills, Project planning-management-and organizational skills, Passion / motivation for digital preservation, Collaboration skills, Ability to train others, Managing budgets, Other (please indicate) Ranking levels included: Extremely important, Very important, Somewhat important, Somewhat unimportant, Very unimportant, Not at all important.]

Q19. Is there anything else you'd like to share about the way you think an effective digital preservation program should be staffed and organized?

Q20. Do you have organizational charts or position descriptions that you'd be willing to share? Any documents you share would provide context to your answers and would be

kept private to the NDSA Standards and Practices Working Group unless you give us explicit permission in the future to share more broadly. [Options included: Yes, No]

Q21. [If yes was selected in Q20] If the organizational charts or position descriptions are on-line, what are the URL(s)? If they are not on-line, please email this supplementary data to ndsa@diglib.org with the subject line "Staffing Survey."

Q22. (Required) Please provide your contact information.
Full name _____
Organization_____
Email address_____
Phone number_____

APPENDIX C: Partial List of Respondents

This alphabetical list of respondents is provided to show the range of institutions who participated. Only those who specifically granted permission to list their organization name as a participant in the survey are included.

Alaska and Polar Regions Collections & Archives
Archaeology Data Service
Archdiocese of Mobile Archives
Archive of the University of Zurich (UZH Archiv)
Archives de l'Etat du Valais
Archives of Michigan
ARMA International Educational Foundation
Arthur H. Aufses, Jr. MD Archives, Mount Sinai Health System
Bard Graduate Center
Berea College
Binghamton University, State University of New York
Binghamton University Libraries
California Digital Library
Caltech Archives
Cambridge University Library
Citizens Archive of Pakistan
Cornell University
CSC - IT Center for Science
Cunningham Memorial Library, Indiana State University
Dartmouth College Library
Data Archiving and Networked Services (DANS)
Davidson College
Detroit Institute of Arts, Research Library & Archives
Deutsche Kinemathek - Museum für Film und Fernsehen
Dorset History Centre
Douglas County Libraries
Dryad Digital Repository
Duke University Libraries
Fontbonne University
Frick Collection

George A. Smathers Libraries, University of Florida
Getty Research Institute
Graduate Center, City University of New York
Harry Ransom Center, The University of Texas at Austin
Harvard Library
Hawaii State Archives
Houston Public Library
Indiana University
Institute of Historical Research
Internet Archive
Kansas State University Libraries
Library and Archives Canada
Marshes of Glynn Libraries
Maryland State Law Library
Mercy Heritage Center, Sisters of Mercy of the Americas
Michigan State University Libraries
Minnesota Historical Society
MIT Libraries
Museum of Fine Arts, Boston
Museum of Fine Arts, Houston
Museum of Modern Art
National and University Library
National Archives and Records Administration
National Archives of Estonia
National Library of Medicine Library Operations
National Library of the Netherlands
National Naval Aviation Museum
Natural History Museum of Los Angeles County
New York Law School
New York State Military Museum
New York University Libraries
North Central College Archives
Northwestern University
Ohio State University Libraries
Pennsylvania State Archives
Publications Office
Purchase College, State University of New York

Rensselaer Libraries and Institute Archives
Rockefeller Archive Center
Royal Library, Denmark
Schenectady County Historical Society
Senator John Heinz History Center
Sisters of Notre Dame de Namur
South Carolina State Library
South Dakota State Archives, South Dakota State Historical Society
Special Collections & Archives, Kent Library, Southeast Missouri State University
Spencer Museum of Art, University of Kansas
Stanford University Libraries
State Archives of North Carolina
State Historical Society of North Dakota
State Library Victoria
Swiss National Library
Texas A&M University Libraries
Texas State University
Texas Tech University Libraries
TIB Leibniz Information Centre for Science and Technology
UNC Chapel Hill Libraries
University College London
University of Arizona Libraries
University of Arkansas at Little Rock
University of California, Irvine
University of Houston
University of Minnesota Libraries
University of Missouri
University of Notre Dame
University of Oregon Special Collections and University Archives
University of Pennsylvania
University of Southern Maine Libraries
University of Texas at Dallas
University of Texas Libraries
University of Virginia
University of Westminster
University of Wisconsin-Milwaukee
UT Southwestern Medical Center

Utah State Archives and Records Service
Valdosta State University Archives and Special Collections
Virginia Commonwealth University
Virginia Polytechnic Institute and State University
West Virginia University Libraries
WGBH Educational Foundation
Wiener Library for the Study of the Holocaust and Genocide
Wisconsin Department of Public Instruction
Wisconsin Historical Society
Wyoming State Archives
Xavier University
Z. Smith Reynolds Library, Wake Forest University
ZBW – Leibniz Information Centre for Economics