USER SURVEY

Executive Summary

As part of the Hydra-in-a-Box discovery process, we conducted a web-based survey of the digital library, archives, and museum community. The goal was to gather information about which, if any, digital repository software institutions are currently using, the types and sizes of content they are managing with that software, their likes and dislikes with current software, and what features they'd like to see in future repository software.

Method

The web-based survey was created using the Qualtrics online data collection and analysis product and consisted of between 12 and 65 (depending on how many repositories the respondent chose to describe) closed- and open-ended questions. On July 15, 2015 we distributed an email announcement of the survey, briefly introducing the Hydra-in-a-Box project and requesting that any interested institution that manages digital collections, whether in a repository or not, take the survey. This announcement was distributed to mailing lists in the digital library, Hydra, archives, scholarly communication, digital preservation, and museum communities. We sent one follow-up reminder email to these lists. We concluded the survey on July 31, 2015.

We exported the completed survey data from Qualtrics and used Excel and Tableau to view and analyze the results.

Participants

There were 248 completed surveys. We considered a survey complete only when the respondent completed the last page of the survey (it was not possible to get to the last page without seeing all other pages). The majority of complete surveys included answers to most of the survey questions.

Over half (54%) of respondents represented either a public or private college or university. Most respondents were located in the U.S. but responses also came from Europe, Africa, New Zealand, and several other international locations. 83% of respondents were currently using repository software. Only a very small number of respondents represented large institutions (measured by annual budget or number of FTE staff).

Summary of Findings

Based on the survey results, what are people expecting from Hydra-in-a-Box?

- A low(er)-barrier entry to Fedora
- No need for developer support
- Metadata: easy in, easy out
- Support for multiple content types
- Reduced need for managing multiple repositories
- Asset management and preservation
- Easier installation and customization than existing options
- Modular but integrated
- Scalability and migration paths

Detailed Findings

The sections below include each question provided in the survey with a brief analysis of the results to that question. Survey respondents did not necessarily answer each question, so although the number of completed surveys was 248, the number that answered each question varied and is indicated in the analysis.

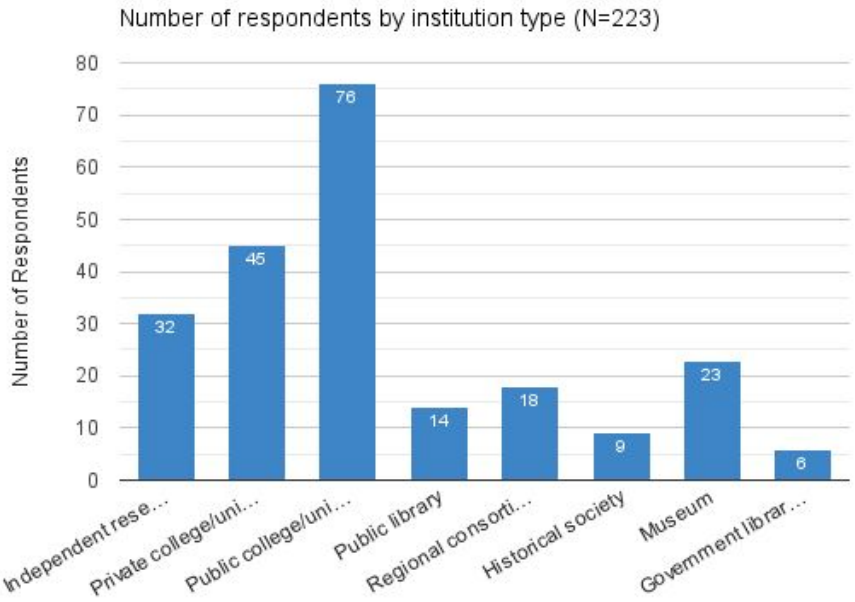
Institutional Characteristics

Who is interested in Hydra-in-a-Box?

Institution Types

> ***Q1:*** *To begin, we have a few questions to get a sense for your institution size and type. What best characterizes your institution's type?*

One-third of respondents classified their institution as a public college or university; together, **private and public colleges and universities composed just over half** of those who answered this question. There were at least six respondents for each of the institution types that were available for a response.



Institution Staff, Budget, and Collection Holdings Sizes

> ***Q2:*** *Please characterize the size of your institution and its collection holdings. If your organization is part of a much larger institution, such as a research library within a larger academic institution, please enter information describing your organization, and not the whole institution.*

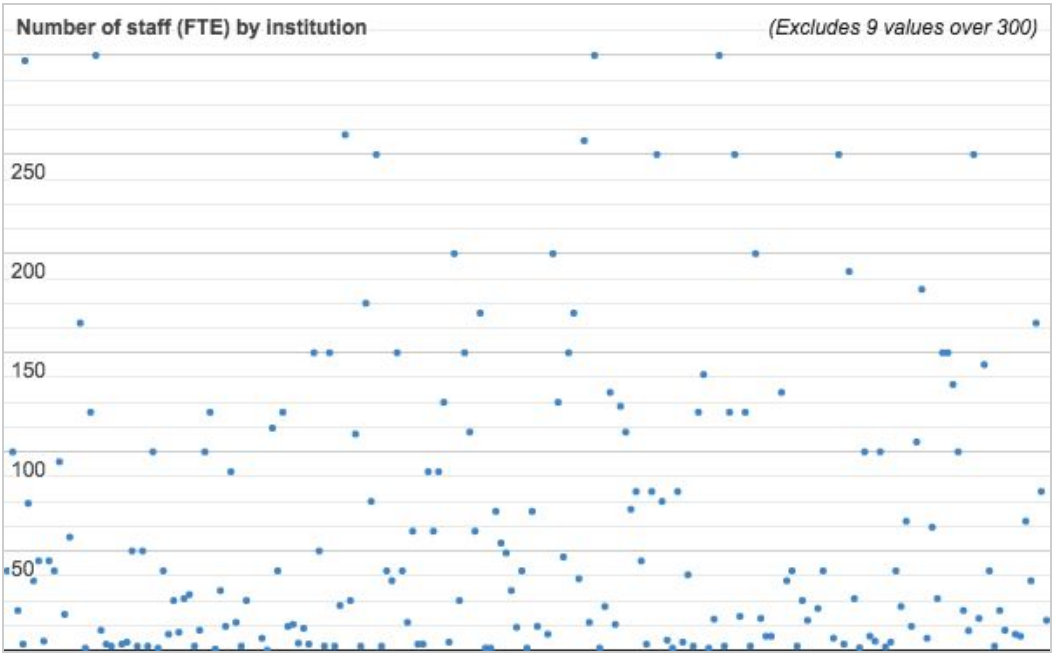
This question was presented as a series of six sub-questions related to understanding size characteristics of institutions. The answers were provided as open-ended responses so there was some inconsistency in the format of the responses, but the response set is large enough that we expect the results to be an accurate reflection of the respondents.

A. Number of staff (FTE) by institution

As shown in the table below, the mean number of staff reported was 97. However, the median was just 40, reflecting the fact that there were a very small number of institutions that reported staff sizes much larger than the average. The table below also shows that half of all reporting institutions had 40 or fewer FTE staff, and the mean value for the smallest three-quarters of reporting institutions was only 28.

Reporting institutions	Mean	Median
All institutions (N=201)	97	40
Institutions with 300 or fewer (N=192)	62	-
Institutions with 100 or fewer (N=146)	28	-

The graph below illustrates the distribution of institutions that provided staff sizes, showing that many have fewer than 20. (In this and other scatterplots included in this document, each dot represents an institution that provided a response to the question. This graph excludes the following outliers: 1886, 1600, 1321, 600, 524, 438, 400, 400, 350.)

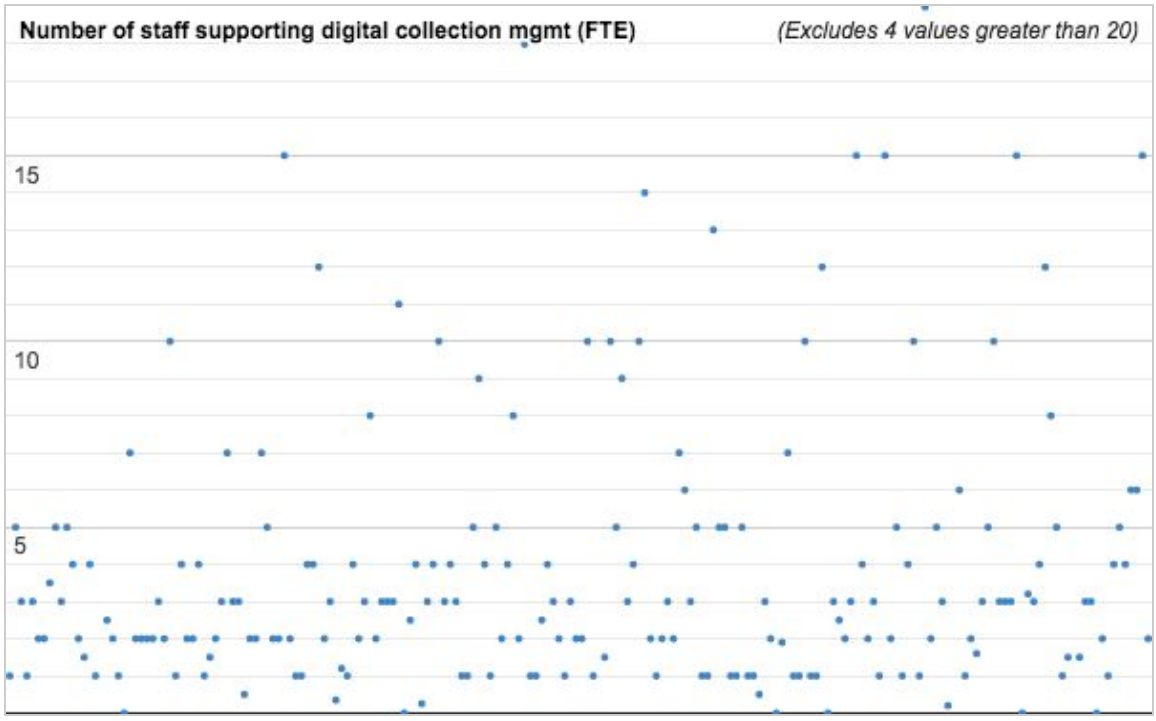


B. Number of staff supporting digital collection management (FTE)

The responses to the number of staff supporting digital collection management again show that while there are a few institutions with relatively large staff sizes, three-quarters of those responding had less than five staff members supporting digital collection management.

Reporting institutions	Mean	Median
All institutions (N=200)	4.35	3
Institutions with fewer than 20 (N=196)	3.85	-
Institutions with fewer than 5 (N=145)	2.12	-

The graph below illustrates that large number of institutions with between 0 and 5 staff members supporting digital collection management. (This graph excludes these outliers: 50, 25, 20, 20.)

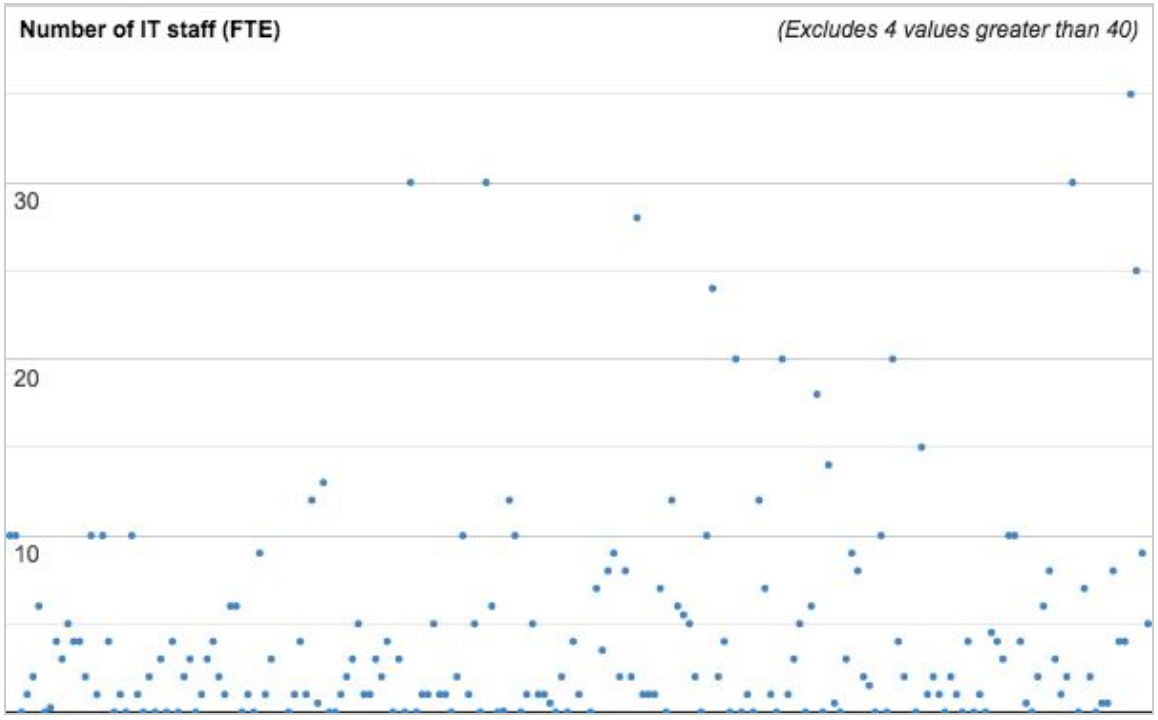


C. Number of IT staff (FTE)

Although several institutions reported having a very large IT staff, the large majority (83%) reported having an IT staff of less than 10. **Of this majority, the mean IT staff size was just a little over two**.

Reporting institutions	Mean	Median
All institutions (N=197)	6.41	2
Institutions less than 40(N=193)	4.44	-
Institutions with less than 10 (N=164)	2.30	-

The graph below illustrates how the responses were heavily skewed towards very small staff sizes. (This graph exclude these outliers: 200, 80, 75, 50.)

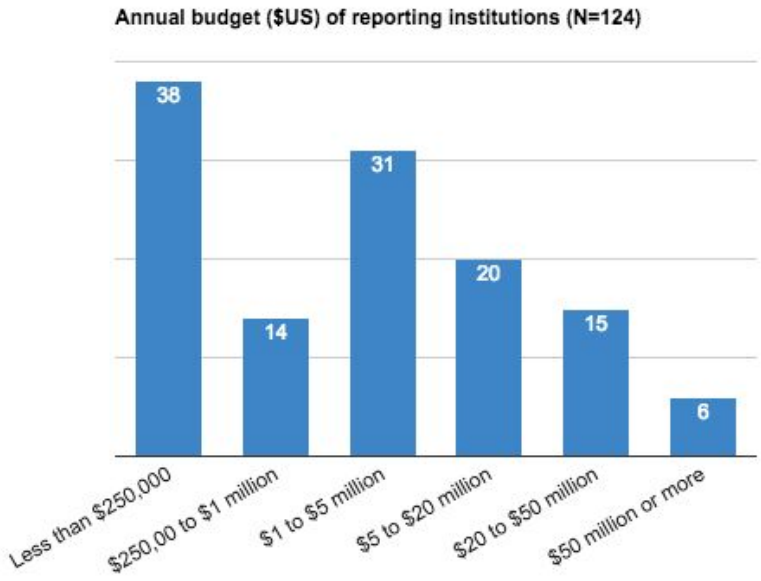


D. Annual budget ($US)

Reported **annual budgets varied widely**, from a low of $0 (five responses) to a high of $483,779,984. A substantial number of institutions reported annual budgets of less than $1 million, and **nearly one-third of respondents reported an annual budget of less than $250,000**.

Reporting institutions	Mean	Median
All institutions (N=124)	$13,958,080	$1,500,000
Institutions with less than $1 million (N=57)	$243,558	-
Institutions with less than $250,000 (N=38)	$63,895	-

Because annual budgets varied so widely, we grouped responses into six ranges to illustrate the variance. The largest range was the smallest, under $250,000, but there were a substantial number in the $1 to $5 million range, and even in the $5 to $20 million range.

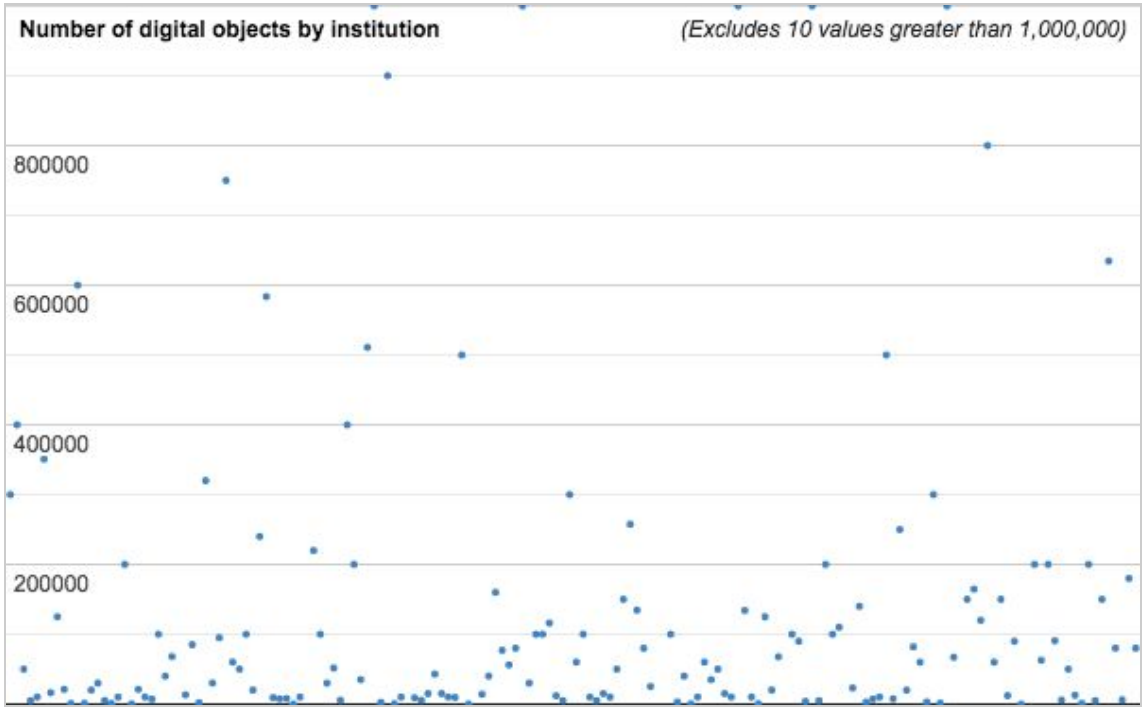


E. Number of digital objects

The final two questions of the section of the survey were intended to learn more about the amount digital objects institutions managed, and how much storage those objects required. The responses followed a pattern similar to the previous question about annual budget: there are some very large outliers (**8 million was the most**, with nine others reporting more than 1 million objects), a decent spread of values in the middle, and a substantial group with relatively small numbers. Although even the median number of digital objects reported was 60,000, **when we consider the smallest 28%, the average number of objects drops to just 5,452**.

Reporting institutions	Mean	Median
All institutions (N=168)	301,512	60,000
Institutions with less than 1 million (N=158)	133,569	-
Institutions with less than 10,000 (N=48)	5,452	-

The graph below shows the distribution of digital objects reported by respondents with fewer than 1 million objects. (This graph excludes the following outliers: 8,350,000; 4,000,000; 3,500,000; 3,000,000; 2,500,000; 2,000,000; 2,000,000; 1,800,00; 1,200,000; 1,200,000.)

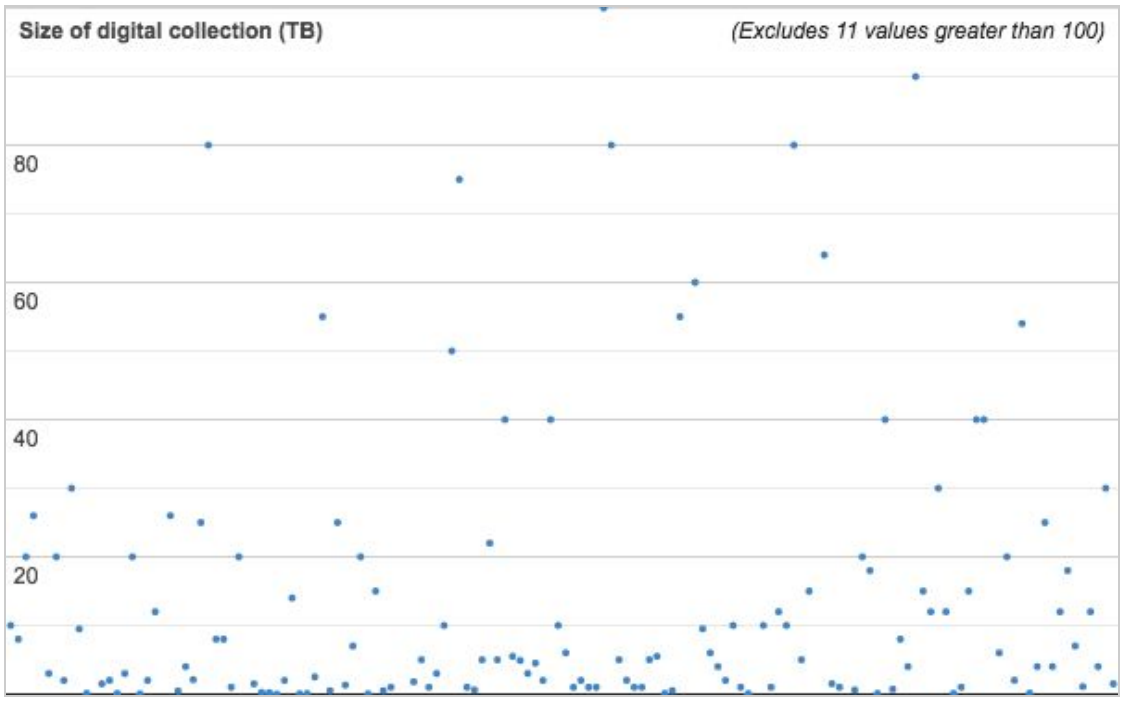


F. Size of digital collection (TB)

Not surprisingly, the pattern for the size of digital collections is very similar to that of number of digital objects. While larger institutions reported collection sizes as large as 170,000TB and 15,000TB, nearly all other institutions reported sizes below 100TB, with the median value for all reporting institutions just 6TB.

Reporting institutions	Mean	Median
All institutions (N=146)	1395	6
Institutions with 100TB or less (N=135)	14	-

The graph below shows the distribution of collection sizes for those institutions reporting sizes below 100TB. This graph excludes these outliers: 170000, 15000, 9000, 5000, 1400, 500, 363, 159, 150, 137, 120.)



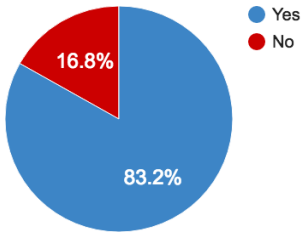
Digital Repository System Use

Who uses digital repositories? If not used, why not?

> ***Q3**: Now we are going to ask a series of questions about your use of repository systems and services in general. Reminder: For purposes of this survey, a repository is defined as a system or service that is intentionally used to manage digital resources (files and metadata) for discovery, access, and/or preservation. A repository is not the same as a file system or a back-up of a file system. A repository may be open source or proprietary. A repository may be a locally installed or hosted by a third party service provider. Do you currently use a digital repository system or service for managing digital content?*

Although **83% of the respondents said that their institution was already using at least one digital repository system**, 37 (17%) respondents reported that their institution was not currently using a repository system.

Percentage of respondents currently using a digital repository system or service (N=220)



Reasons for Not Using a Digital Repository System

> ***Q4:** [If answer to **Q3** was "No"] Why do you NOT currently use a repository system or service for managing your digital collection(s)? Check all that apply.*

The 37 respondents who answered "No" to the previous question were asked this additional question intended to better understand why their institution was not currently using a repository system or service. As shown in the table below, the reason was not because their institution didn't have a need for a repository system, but most often because they lacked the necessary staff, they weren't sure which system to use, and/or they were still in the planning stages.

Reason	Count	Percentage
Lack of funding	19	15%
Lack of administrative support	18	15%
Planning underway; no implementation to date	17	14%
Lack of staff to manage (select, describe, ingest, etc.) the content	16	13%
Lack of staff to provide technical support	16	13%
Not sure which system or service to use	15	12%
Other	8	7%
Have not found a system or service that meets needs of our content type(s)	7	6%
Lack of staff to provide researcher (end-user) support	6	5%
No need	1	1%

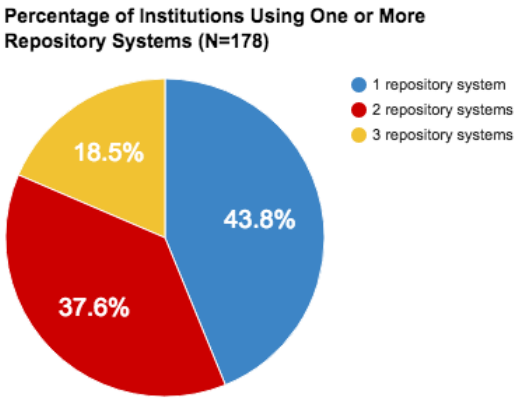
Current Use of Digital Repository Systems

Which repository systems are used and how are they managed and supported?

Repository Systems In Use

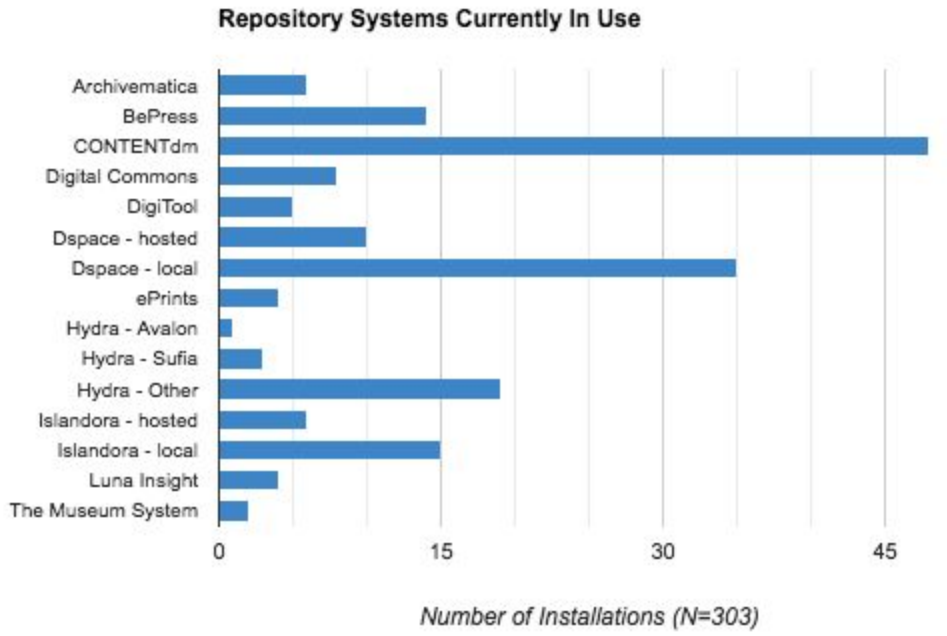
> ***Q5:*** *[If answer to **Q3** was "Yes"] For each digital repository system or service you use to manage digital content, enter a label for its name or function below. If you use more than three systems or services, describe the three that are the most critical to your organization. For each distinct repository system you indicate, you will be asked a set of more detailed questions about that repository.*

When asked to list up to three repository systems or services their institution is currently using, 178 respondents listed at least one system in use. The graph at right shows the percentage of institutions using one, two, or three systems.



> ***Q6:*** *Your Repository Contents. This set of questions refer to details of your repository. Because your answers to these questions might differ depending on how you are using a given repository, you will be shown this set of questions for each repository you indicated previously that you are using. Regarding your entry, what is the repository system or service are you currently using?*

As to the specific repository systems and services institutions are currently using, respondents were given a list of 15 systems (including separate choices for the local and hosted versions of both DSpace and Islandora) and a choice of other, for which a free-text name could be provided. As shown in the graph below, CONTENTdm and DSpace were the systems most frequently mentioned. Hydra and BePress were each mentioned by more than a dozen respondents.

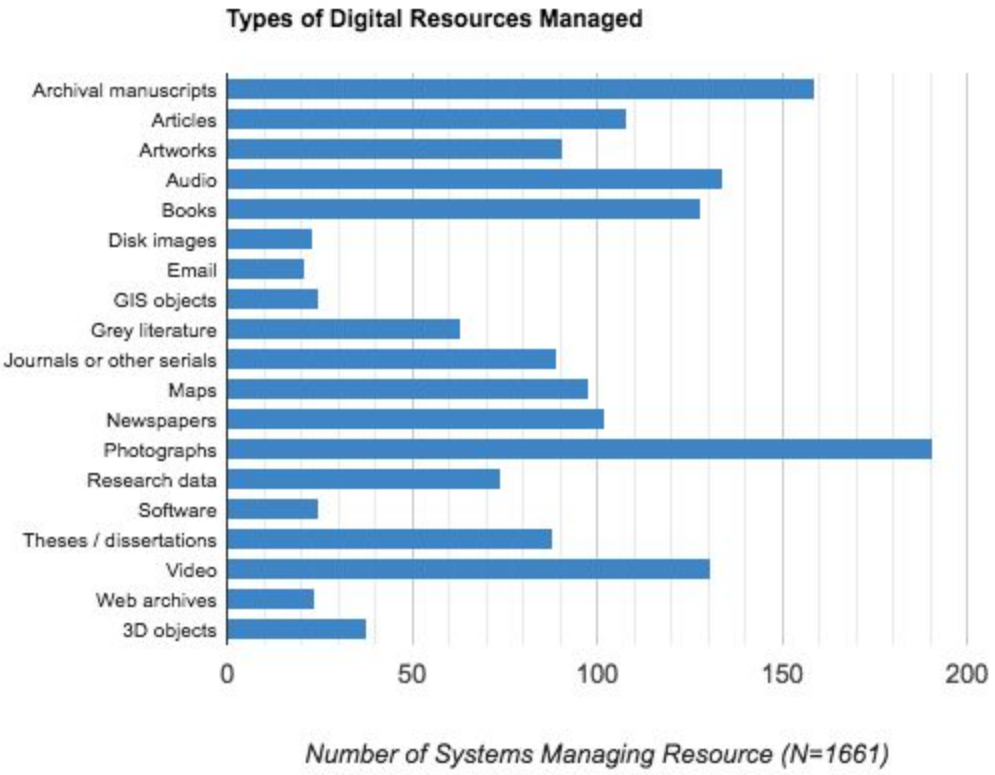


There were 123 responses of “Other,” indicating current use of a repository system not listed in the answer choices. Many of these referenced use of a “custom” or “homegrown” system, sometimes using Hydra or Drupal. Other responses mentioned more than a few times included Archive-It, ArchivesSpace, CollectiveAccess, Fedora, Omeka, PastPerfect, and Preservica.

Types of Digital Resources Managed

Q7: *Regarding your entry [name of repository system in use], what type(s) of digital resources are managed in the repository system or service? Check all that apply.*

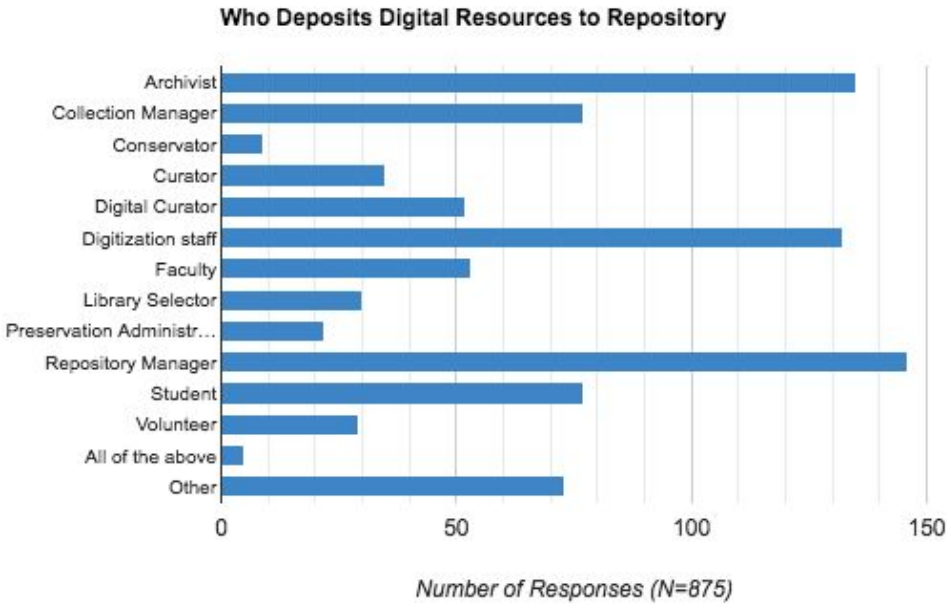
For each repository system respondents listed as currently in use at their institution, they next indicated what types of digital resources were managed by those systems. As shown in the graph below, photographs were most frequently cited (191 responses), followed by archival manuscripts (159), audio (134), video (131), and books (128). Some of the free text responses that were given more than once for the answer choice of “other” included finding aids, learning objects, and metadata-only records.



Who Deposits Content?

Q8: *Regarding your [name of repository system in use], who deposits digital resources to the repository system or service? Check all that apply.*

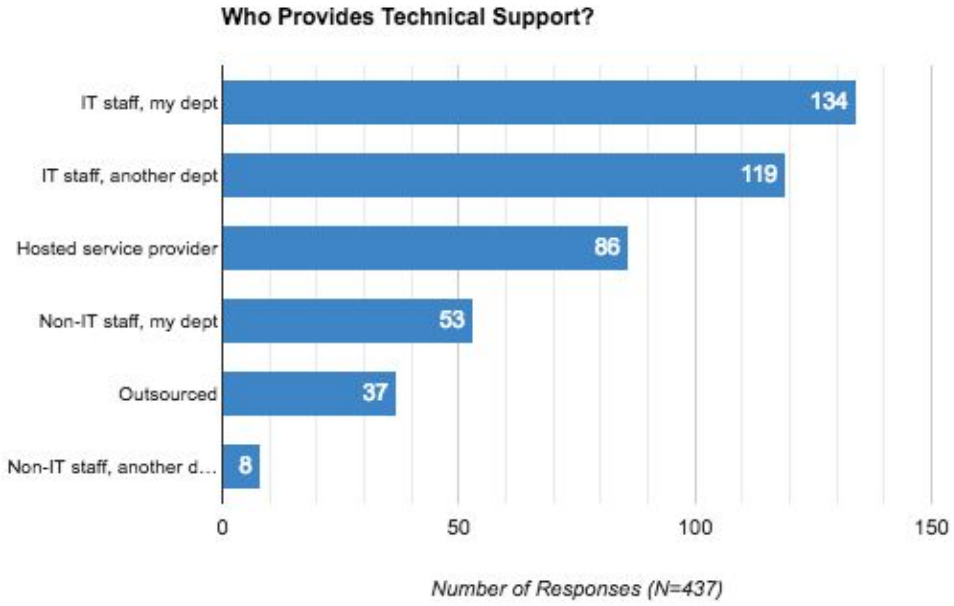
Of 875 responses to this question (respondents could select more than one answer), roles most often mentioned as being responsible for depositing digital resources to the respondent's repository system were the **repository manager (146), archivist (135), and digitization staff (132)**. As shown in the graph below, collection manager (77), student (77), and "other" (73) were also mentioned well over 50 times, with 12 of the "other" answers specifically listing "metadata librarian."



Who Provides Technical Support?

Q9: *Regarding your [name of repository system in use], who provides technical support for the repository system or service?*

Based on the 437 responses to this question, technical support for the repository systems at the institutions of respondents is most often provided by IT staff in either the respondent's department (31%) or another department at their institution (27%). About 20% of the responses indicated that technical support for a repository system was provided by a hosted service provider.



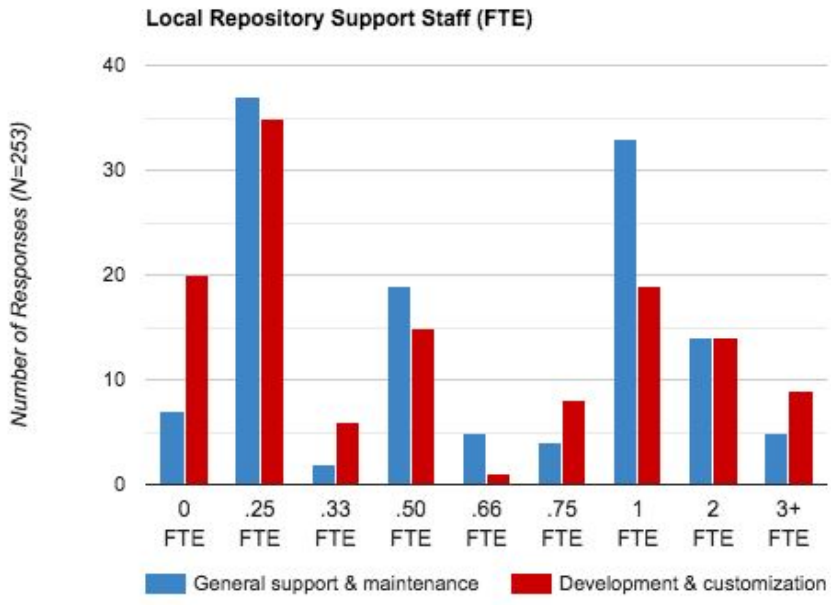
Local Staffing Resources

Q10: *[If the answer to the previous question was one of the "local staff" answers.] Regarding your [name of repository system in use], for general technical support and maintenance, approximately what number of FTE of local staff time (any department, IT staff or otherwise) is devoted to the system or service per year?*

Q11: *[If the answer to the previous question was one of the "local staff" answers.] Regarding your [name of repository system in use], for software development, including local customizations, approximately what number of FTE of local staff time (any department, IT staff or otherwise) is devoted to the system or service per year?*

The next two questions were very similar. Both asked respondents who indicated they had local support staff to estimate the number of full-time-equivalent (FTE) staff they had, to provide general support and maintenance and to provide development and customization support. The answer choices to these questions were identical and the results are combined in the graph below.

As shown in the graph, very few institutions have more than a few FTE staff providing support for their repository systems. Out of 127 responses, only five systems had three or more staff providing general support and maintenance, and only nine systems had three or more staff to do development and customization. Most frequently, repository systems had .25 FTE staff providing each form of support.

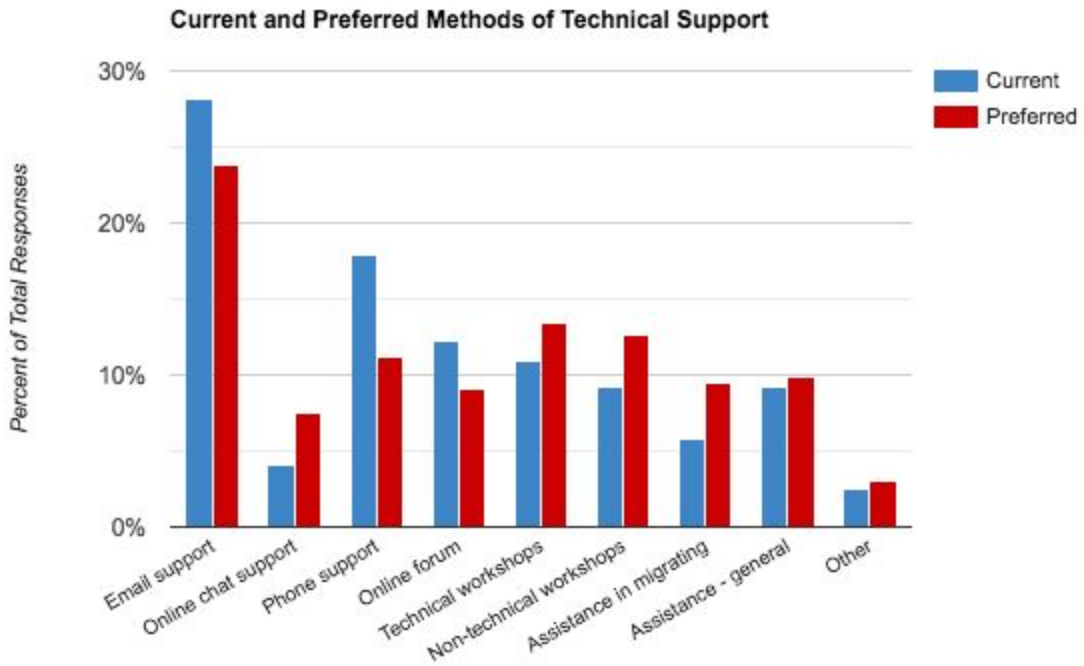


Current and Preferred Forms of Technical Support

> ***Q12:*** *Regarding your [name of repository system in use], what forms and methods of technical support do you currently rely on to keep the system or service in production? Which do you prefer to use? (Your preference may be the same or different than a method that you currently use.) Check all that apply.*

Respondents provided a total of 878 answers to the question of which methods of technical support they currently rely on, and 571 responses to the question of which methods they would prefer to use. The graph below combines the results of these two questions (the answer choices were identical) and presents them as percentages so they can be more fairly compared. Respondents most often use email support, phone support, and online forums as their current support methods. When answering which their preferred methods of support are, email support is still mentioned most often, but in this case technical and non-technical workshops and training were mentioned more often than phone support. Online chat support and assistance in migrating were two other responses that increased when comparing preferred to current methods of support.

For the “other” responses to the question of preferred methods of technical support, the one free-text response mentioned multiple times was “ticket-tracking system,” such as JIRA or GitHub.

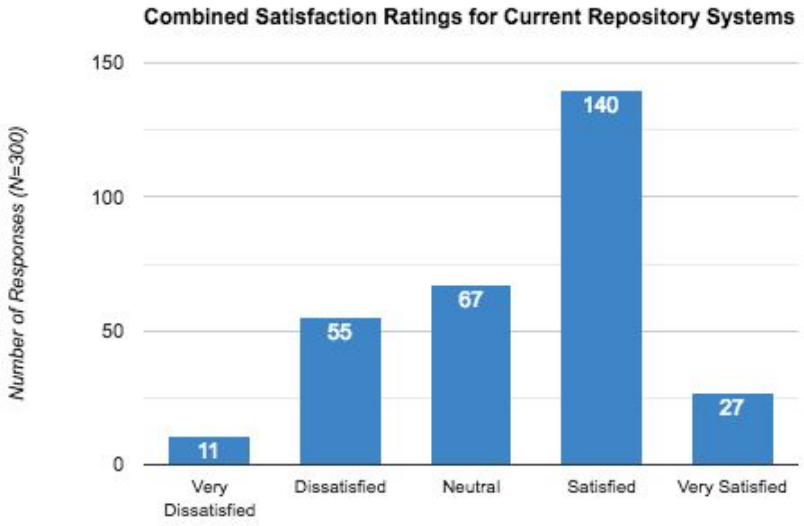


Satisfaction with Current Repository Systems

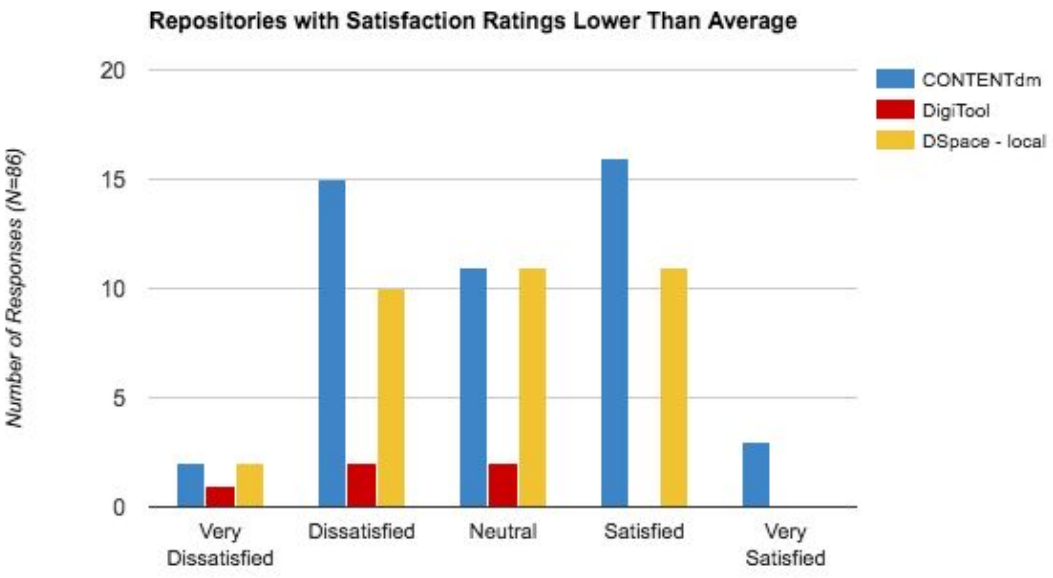
How do respondents feel about their current systems?

> ***Q13:*** *Regarding your [name of repository system in use], how satisfied are you with the repository system or service overall?*

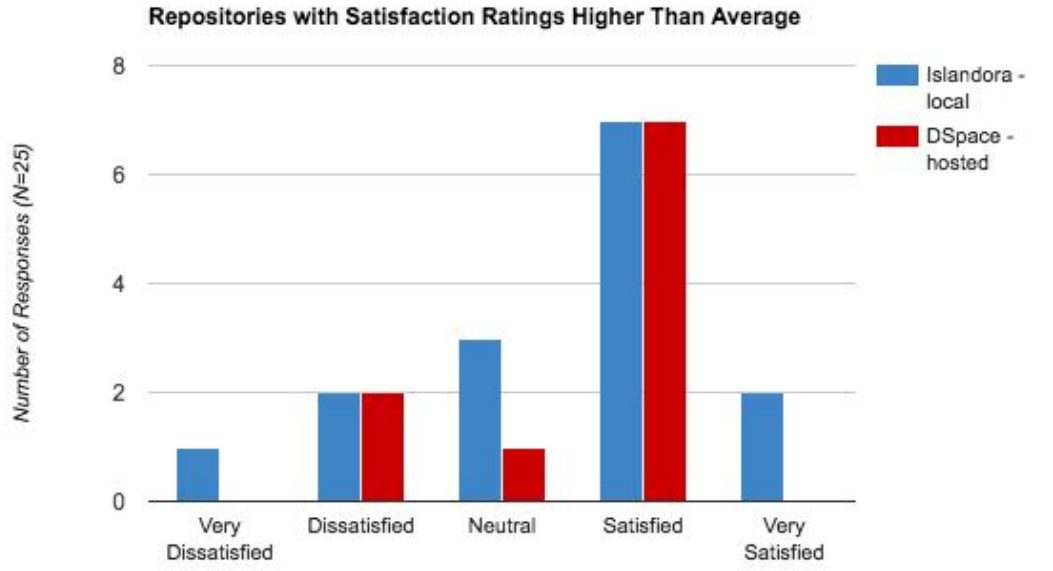
Overall, survey respondents were satisfied with their current repository systems. As shown in the graph at right, out of a total of 300 responses, many respondents (55%) said they were either “satisfied” or “very satisfied” with their repository system. In contrast, 22% said they were “dissatisfied” or “very dissatisfied.”



Not all repository systems received satisfaction ratings consistent with the overall ratings, however. For systems that had at least five responses, CONTENTdm, DigiTool, and DSpace - local received a higher percentage of “dissatisfied” or “very dissatisfied” ratings than all systems considered together.



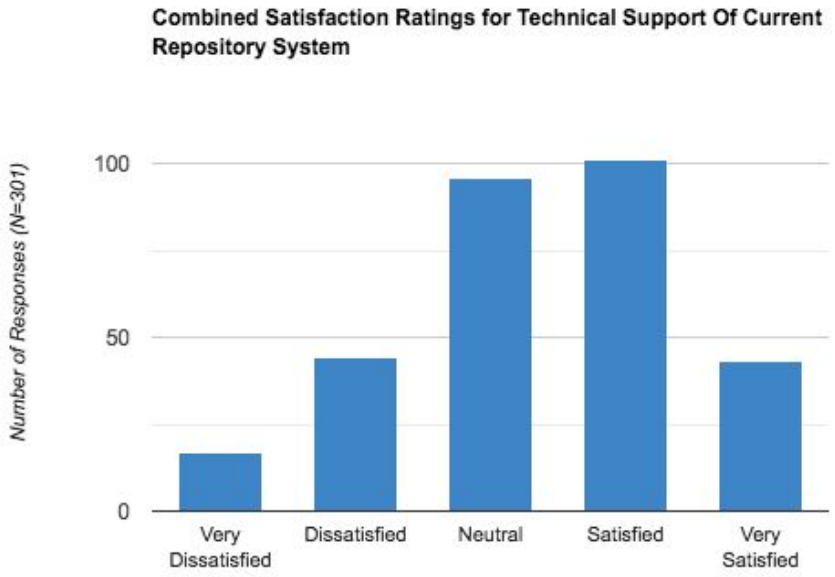
On the other hand, two systems — Islandora - local and DSpace -hosted — had a slightly higher percentage of positive ratings as compared to all systems combined.



Satisfaction with Current System Technical Support

Q14: *Regarding your entry [name of repository system in use], how satisfied are you with the repository system or service's technical support?*

When asked to provide satisfaction ratings for the technical support of their current repository system, just under 50% of 301 total responses indicated that they were “satisfied” or “very satisfied” with their technical support. About 20% said they were “dissatisfied” or “very dissatisfied.”



Among repository systems that received five or more ratings, CONTENTdm (31%), DigiTool (80%), and Islandora - local (34%) showed higher percentage of “dissatisfied” or “very dissatisfied” ratings compared to all systems combined.

Likes and Dislikes of Current Repository Systems

Q15: *What do you like about your [name of repository system in use] repository system or service? What features can you not live without?*

There were 240 free text comments in response to this question. The most common responses are summarize below by repository system.

CONTENTdm

Respondents using CONTENTdm felt that it is:

- Strong on metadata and batch operations
- Strong on images

Their comments were more mixed on CONTENTdm's:

- Customer/technical support
- Multi-page content

DSpace

Positive comments from respondents about DSpace included:

- Works out of the box
- Strong user and developer community
- Strong for IR workflows
- Strong SEO

Islandora

Respondents using Islandora were most positive about the fact that it:

- Comes with Fedora
- Supports many content types

Islandora users had mixed reviews about its staff/admin UI.

BePress / Digital Commons

Positive comments about BePress/Digital Commons included:

- SEO
- Workflows for IR content
- Customer service

Q16: *What do you NOT like about your [name of repository system in use] repository system or service? Is anything missing or in need of improvement?*

There were 242 free text comments in response to this question. The most common responses are summarize below by repository system.

CONTENTdm

Respondents using CONTENTdm mentioned a wide range of dislikes about the system:

- Poor search performance, relevancy
- Poor accessibility
- Poor SEO
- Poor navigation
- Outdated look, non-responsive
- Expensive
- Limited customizability; hard to maintain
- Lack of integration with institutional authorization

DSpace

Respondents using DSpace mentioned the following dislikes about the system:

- Limited by community/collection model
- Lack of hierarchical support
- Poor support for AV content
- Poor reporting
- Outdated look, non-responsive
- Customization is time-consuming, hard to maintain

Islandora

Dislikes mentioned by Islandora users included:

- Underdeveloped
- Difficult to install
- Customization is hard without Drupal expertise
- “Not as DIY as we had planned.”

BePress / Digital Commons

Respondents using BePress/Digital Commons mentioned a wide range of dislikes about the system:

- Metadata blackbox
- Rigid hierarchy and object relationships
- Limited batch operations
- Lack of customizability
- Expensive
- Dependency on service providers for basic changes
- Not engaged with broader repository community

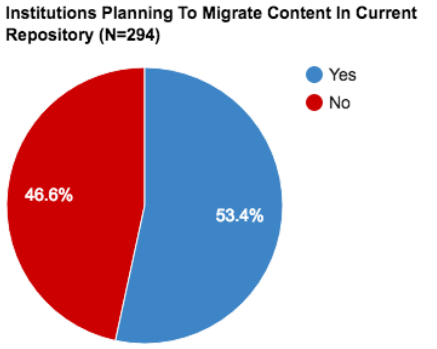
Repository System Migration Plans

Do respondents plan to migrate content to new digital repository systems?

Plans To Migrate

Q17: *Do you want to or plan to migrate the content in your [name of repository system in use] to a different repository system or service?*

Roughly half (53%) of the 294 responses to this question indicated a plan to migrate content in a current repository system or service to another one.



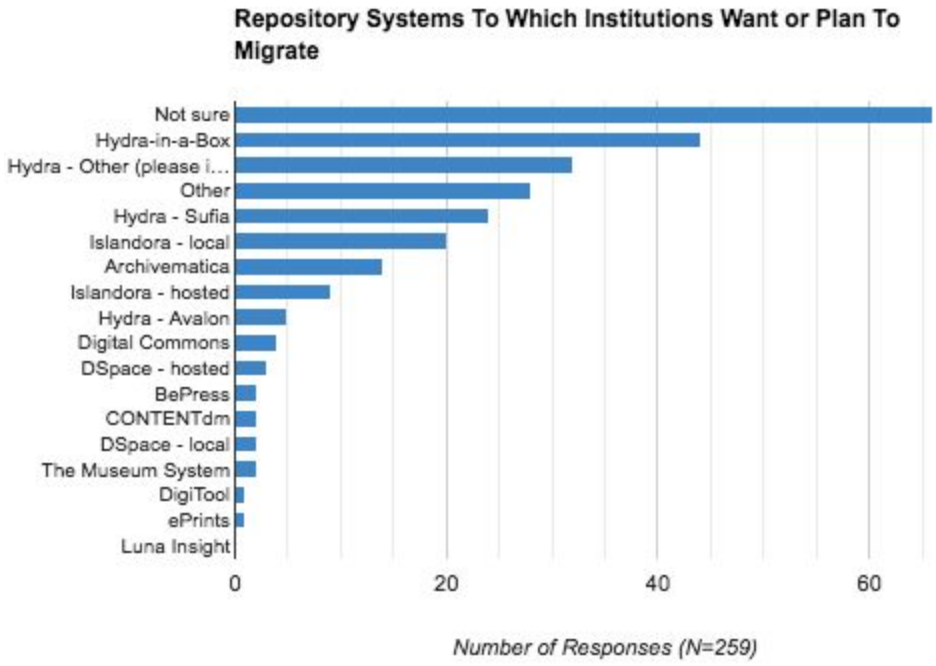
To Which Systems?

***Q18**: What different repository system or service do you want to or plan to migrate to?*

*[Next two questions presented if for **Q17.** Do you want to or plan to migrate the content to a different repository system or service? Yes was selected]*

Although a large number of survey respondents want or plan to migrate to a new repository system, they are clearly uncertain about the system to which to migrate. There were 259 responses indicating which repository system they planned to migrate to, with the most common choice being “Not sure” (25% of all responses). Hydra-in-a-Box was mentioned by 17% of respondents, “Hydra - Other” by 12%, and “Other” by 11%.

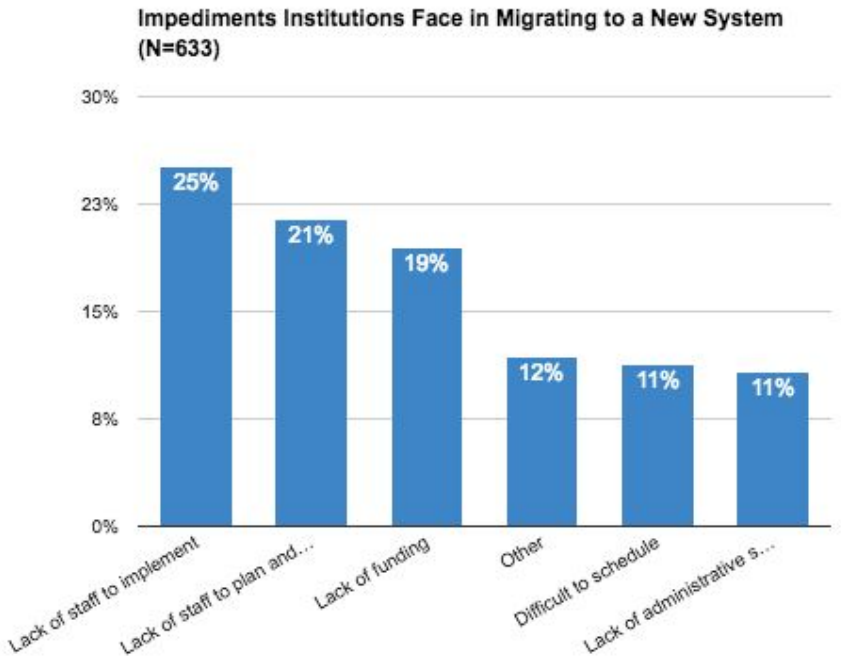
The “Other” and “Hydra - Other” choices both included several responses indicating a plan to build a custom local system, or using Fedora 4 with a custom front-end.



What Impediments to Migration Do They Face?

***Q19:** Do you face any impediments to migrating from one repository system or service to another? Check all that apply.*

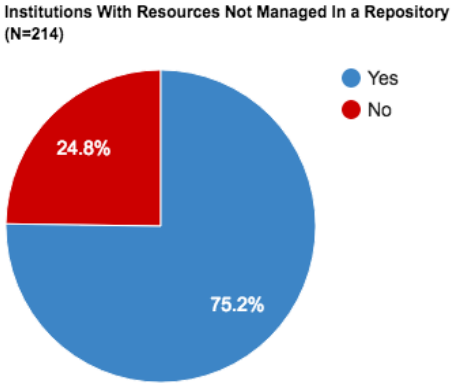
When asked what impediments to migration they faced, lack of staff to implement, lack of staff to plan and manage, and lack of funding were the most frequent reasons given, although all reasons available for selection were picked by at least 10% of the 633 total responses.



Managed Digital Resources Not In Current Repository Systems

Q20: *Are there digital resources currently in your collection that are NOT managed in a repository system or service? Reminder: For purposes of this survey, a repository is defined as a system or service that is intentionally used to manage digital resources (files and metadata) for discovery, access, and/or preservation. A repository is not the same as a file system or a back-up of a file system. A repository may be open source or proprietary. A repository may be a locally installed or hosted by a third party service provider.*

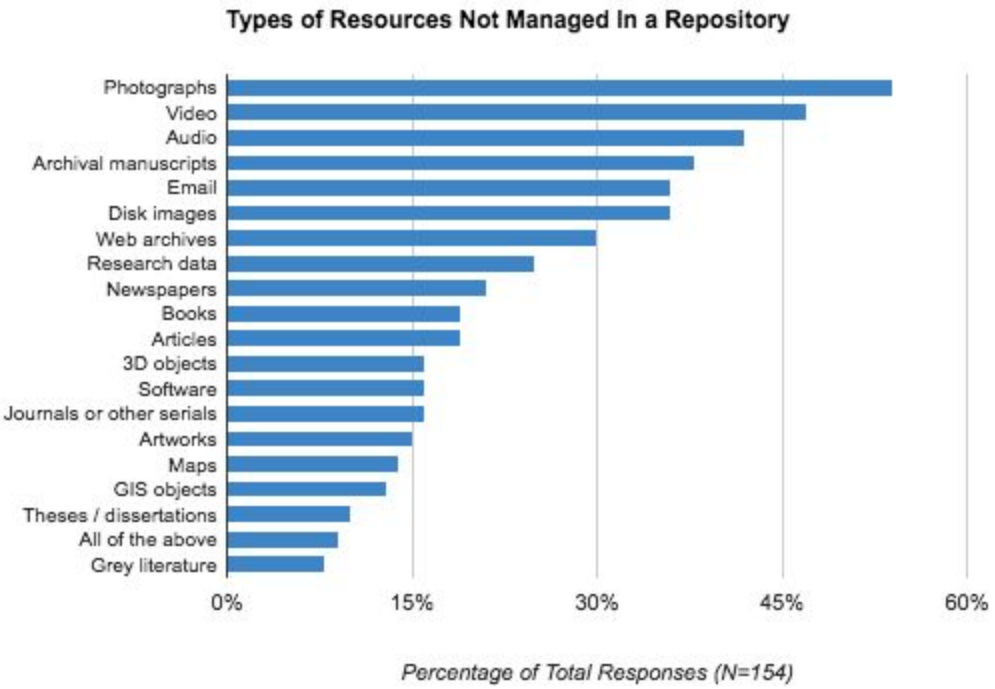
Of the 214 respondents who answered this question, 75% said they have digital resources that are not managed in a repository system or service.



Managed Resources Not in Repository

Q21: *What type(s) of digital resources are currently in your collection but are NOT managed in a repository system or service? Check all that apply.*
Asked if "Yes" was selected for ***Q20****: Are there digital resources currently in your collection that are NOT managed in a repository system or service?*

Respondents indicating that their collections had digital resources not managed in a repository system or service were next asked to indicate which types of resources those were. The graph below shows the percentage of the 154 total responses were selected for each resource type. For example, of the 154 respondents that indicated one or more resource types not managed by a repository system, 54% of them had photographs that weren't being managed in a repository system. Video, audio, archival manuscripts, email, and disk images were also mentioned by more than 30% of those responding to this question.



Reasons for Not Managed Resources in a Repository

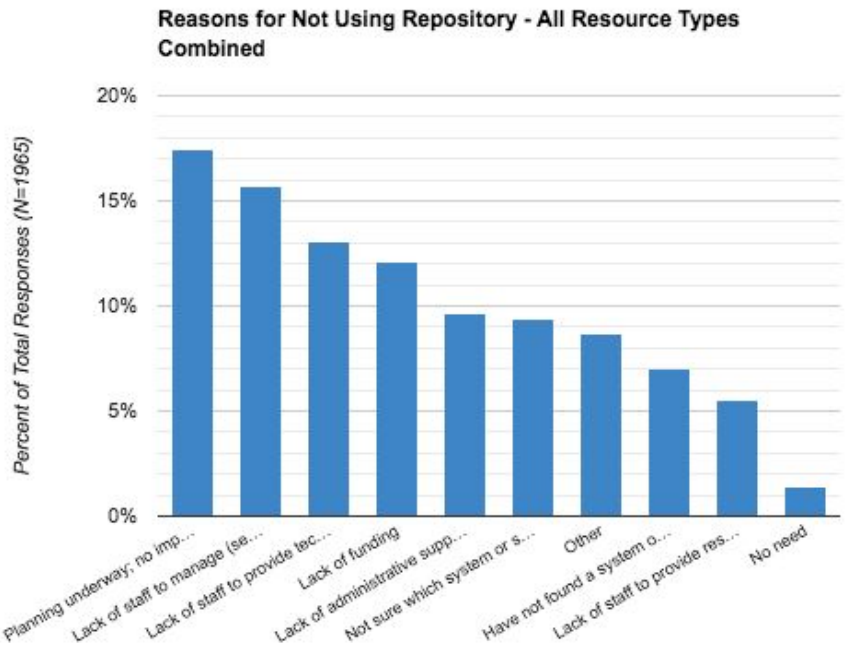
Q22: *Why do you currently NOT use a repository system or service for managing [resource type]?*
Asked if "Yes" was selected for ***Q20:*** *Are there digital resources currently in your collection that are NOT managed in a repository system or service?*

This question was actually presented as a series of questions, one for each type of resource the respondent indicated their institution is not using repository system to manage (the question text and

answer choices were identical for all resource types). There were a total of 1965 responses for all resource types combined.

Considered across all resource types, most of the available answers were selected with similar frequency.

Reason	Percent of Total
Planning underway; no implementation to date	17%
Lack of staff to manage (select, describe, ingest, etc.) the content	16%
Lack of staff to provide technical support	13%
Lack of funding	12%
Lack of administrative support	10%
Not sure which system or service to use	9%
Other	9%
Have not found a system or service that meets the needs of this content type	7%
Lack of staff to provide researcher (end-user) support	5%
No need	1%



Critical Features

What do respondents think are the most important for a digital repository system?

To understand what survey respondents thought were the most critical features for a digital repository system, we asked them to categorize a list of potential features into three categories: must have, should have, and nice to have. We split this up into four distinct areas — core features and workflows, metadata and access control, integrations, and administration — each with its own list of relevant features to be categorized.

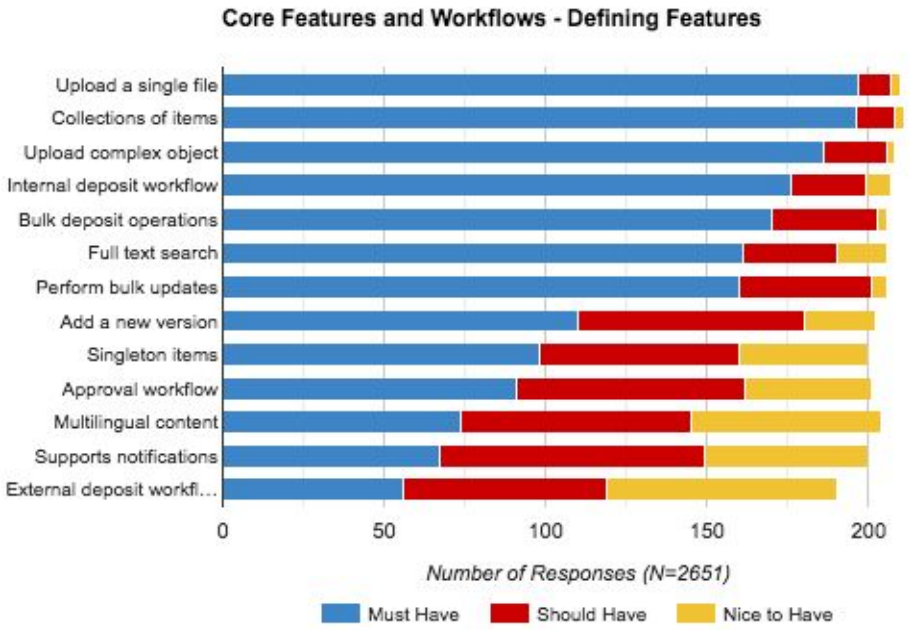
Core Features and Workflows

> ***Q23**: Your Ideal Repository - Core Features and Workflows. What are the defining features and workflows supported by the ideal repository system or service as you envision it?*

There were 2651 selections made for the core features and workflows area. There were a group of features that were **clearly judged as must haves**, based on the proportion of must-haves to should- and nice-to-haves. These were: upload a single file, supports collections of items, upload a complex (multi-file) object, supports an internal deposit workflow (by staff in my organization), perform bulk deposit operations, supports full text search, and perform bulk updates.

There were other potential features that received more must-have votes than should- and nice-to-haves, but they were not as overwhelmingly seen as must-haves as the group of features above.

Features/Workflows	Must Have	Should Have	Nice to Have
Upload a single file	197	10	3
Supports collections of items	196	12	3
Upload a complex (multi-file) object	186	20	2
Supports an internal deposit workflow (by staff in my organization)	176	23	8
Perform bulk deposit operations	170	33	3
Supports full text search	161	29	16
Perform bulk updates	160	41	5
Add a new version	110	70	22
Supports singleton items (without an associated collection)	98	62	40
Supports an approval workflow	91	71	39
Supports multilingual content	74	71	59
Supports notifications	67	82	51
Supports an external deposit workflow (by those outside my organization)	56	63	71

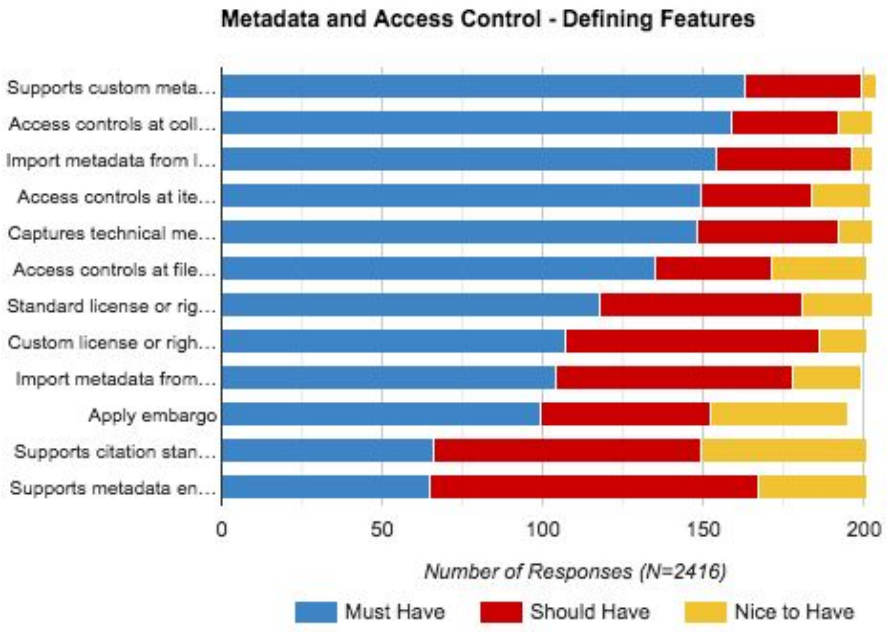


Metadata and Access Control

> ***Q24:*** *Your Ideal Repository - Metadata and Access Control. What are the defining metadata features and access controls supported by the ideal repository system or service as you envision it?*

There seems to be a similar grouping of clear must-have features for the metadata and access control area. Of the 2416 total responses in this area, there were **six features that received a noticeably higher proportion of must-have selections than the others**: supports custom metadata, apply access controls at collection level, import metadata from local system, apply access controls at item level, captures technical metadata (checksum, file format, etc.), and apply access controls at file level.

Features/Access Controls	Must Have	Should Have	Nice to Have
Supports custom metadata	163	36	5
Apply access controls at collection level	159	33	11
Import metadata from local system	154	42	7
Apply access controls at item level	149	35	18
Captures technical metadata (checksum, file format, etc.)	148	44	11
Apply access controls at file level	135	36	30
Apply a standard license or rights statement	118	63	22
Apply a custom license or rights statement	107	79	15
Import metadata from external system	104	74	21
Apply embargo	99	53	43
Supports citation standards	66	83	52
Supports metadata enrichment through other services	65	102	34



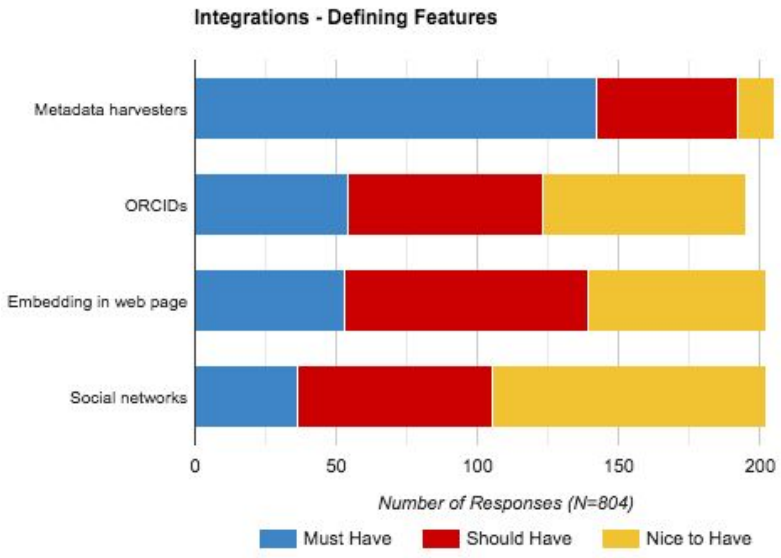
Integrations

Q25: *Your Ideal Repository - Integrations. What are the defining integration features supported by the ideal repository system or service as you envision it?*

When asked to judge how critical several integration features were to an ideal repository system, there were fewer obvious must-haves compared to the previous areas. In fact, only the potential feature "Integrates with metadata harvesters" was mentioned more often as a must-have than a should-have in the

804 total responses for this area. Of the remaining three potential features, only “Support for embedding in a web page” was seen more often as a should-have than a nice-to-have.

Integration Features	Must Have	Should Have	Nice to Have
Integrates with metadata harvesters	142	50	13
Supports ORCIDs	54	69	72
Support for embedding in a web page	53	86	63
Integrates with social networks	36	69	97



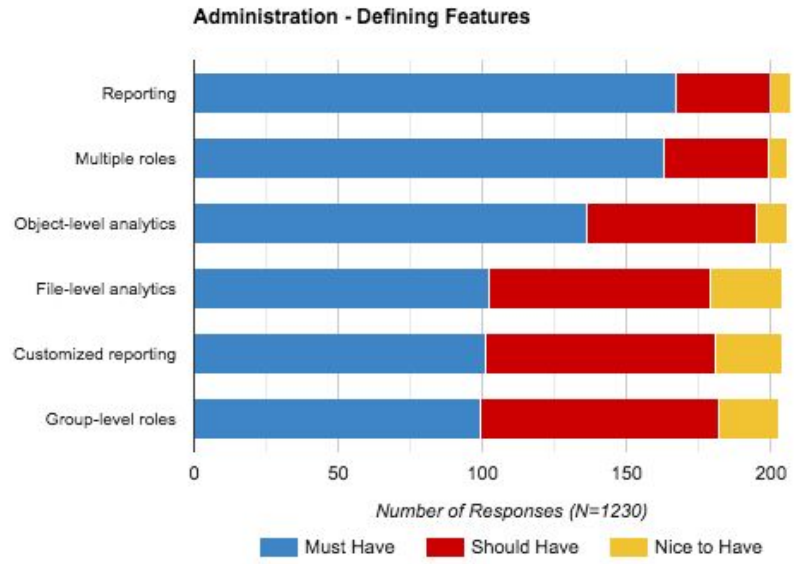
Administration

Q26: *Your Ideal Repository - Administration. What are the defining administrative features supported by the ideal repository system or service as you envision it?*

In the area of administration features, the 1230 total selections indicated that there were three potential features that clearly seem to be must-haves: supports reporting, supports multiple roles (Admin, Manager, Depositor, Reviewer, Viewer, etc.), and gathers object-level analytics.

The three remaining potential administrative features were also mentioned most often as must-haves, though these received a higher proportion of should- and nice-to-haves as compared to the features mentioned above.

Administrative Feature	Must Have	Should Have	Nice to Have
Supports reporting	167	33	7
Supports multiple roles (Admin, Manager, Depositor, Reviewer, Viewer, etc.)	163	36	7
Gathers object-level analytics	136	59	11
Gathers file-level analytics	102	77	25
Supports customized reporting	101	80	23
Supports group-level roles	99	83	21



Wrap Up

Additional comments and availability for future contact

Additional Comments

Q27: *Wrap Up. Do you have any additional comments about your current or planned use of a repository for managing digital resources at your institution?*

There were 78 respondents who provided additional comments at the end of the survey. Some of these comments detailed specific features or use cases the respondent hopes Hydra-in-a-Box will support:

> *The needs of a digital collection are very different from the needs of a digital repository for archival material. It should represent hierarchies well and should easily integrate with or allow importation from other systems that institutions might be moving out of. Embargoes and the ability to apply permissions on levels from collection to file is critical.*
>
> — Public college/university library

> *Integrated support for interoperability with GIS frameworks is essential; ability to integrate with data services APIs external to the "repository" is essential; support for IIIF (and for eventual similar approaches to other media types) highly desirable.*
>
> — Public college/university library

> *Because our planned repository will rely on staff members to contribute their permanent digital records to the repository directly, rather than relying solely on the archivist to collect and upload these materials, a simple and easy interface that staff can use for ingest is crucial.*
>
> — Independent research library/archives

> *We really want an easy way to map our images, too.*
>
> — Regional consortium

> *Our pie-in-the-sky system would allow ingest of an item, have an access copy generated and sent to an access system/portal, would connect to both on-site and off-site preservation storage for preservation copies, would allow for the preservation of metadata as well as objects, wouldn't require a specialized stack of server software or specialized IT support on-site for us, would have the functionality to migrate access files to newer format versions, would support different born digital and archival formats, would audit all activities on files, and would be affordable for a state agency with a small budget but large mandate.*
>
> — Government library/archives

Many more comments reflected a more general need from institutions for a repository software solution that is easy to use, especially for smaller institutions with little or no technical staff:

> *I like the potential power of using a repository system based on Fedora but there has been no good integration solution that is accessible to smaller institutions lacking budget or staff to support the technical requirements.*
>
> — Private college/university library

> *As a volunteer archivist supporting the institutional repository of a 40-year-old non-profit, I need to be able to recommend something that is -- above all -- easy to use.*
>
> — Historical society

> *DSpace is lacking in a number of areas, but that ease of starting and use is really key for us.*
>
> — Private college/university library

> *Additionally, making it as easy as possible for Hydra newcomers and adopters at small institutions who may have little to no technical knowledge and are looking for a (more or less) one-button install would be a huge step toward lowering the barrier for using Hydra over other solutions marketed as more “ready-to-use.”*
>
> — Public college/university library

> *We are a small office with very little in house tech support/ capabilities. Hoping hydra in a box will truly take into account organizations like us who are trying to help smaller institutions get content online and share it ie via DPLA.*
>
> — Regional consortium

> *We plan to expand the repository services, but are currently struggling with administration support and developer time.*
>
> — Private college/university library

> *I have great hopes for the Hydra-in-a-Box project! I currently manage the digital collections of more than 100 small cultural heritage organizations in Washington State, and do not have the additional technical support or funding to move these collections to an open, standards-compliant repository.*
>
> — Regional consortium

Not surprisingly, some of the institutions with smaller staff sizes expressed interest in low-cost and/or hosted solutions:

> *My research hasn't turned up an affordable software solution that will provide robust collections management, DAM, digital archiving of master files, and a highly flexible web portal/public interface with online merchant capabilities. I'm waiting to have my socks knocked off.*
>
> — Museum

> *If Hydra were also able to offer a hosted option similar to what Discovery Garden can offer for Islandora, I think many many more small to medium sized institutions would be interested. We would also be interested in hearing more about how a hosted option would work because it could potentially free-up our developers' time to write code instead of managing our AWS servers.*
>
> — Independent research library/archives

> *As a basically one-person department at a mid-size institution, I really like the model of open source software with subscription, hosted support. My library doesn't have the time or technical staff to run our own open source instances, but as far as possible I try to avoid proprietary solutions.*
>
> — Museum

Some comments conveyed respondents' frustrations with current systems:

> *There are no good options. It shouldn't be this hard to make something that doesn't look & function like it's already 20 years old.*
>
> — Independent research library/archives

> *Right now, I feel stuck with CONTENTdm - the software isn't meeting our needs, particularly on access restrictions and enhanced metadata (e.g., linked data, better support for CV, rights statements, etc) - but I don't see an ideal, affordable solution out there.*
>
> — Private college/university library

In particular, there were quite a few comments expressing frustration with having to use multiple systems to manage all of an institution's digital resources:

> *We search an access database for metadata, and have digital files elsewhere. It's insane.*
>
> — Independent research library/archives

> *We have struggled with administrators expecting one system to do all functions (preservation, access, discovery) and this has stalled the process as some in digital collections view presentation and interpretation to have highest value, others metadata management, and still others preservation. We started looking for an alternative to CONTENTdm, but due to this conflict are currently evaluating/piloting two new platforms in addition to CONTENTdm and BePress; ArchivesDirect and ArchivesSpace. These are critical systems but do not meet many in the digital collections area's priority functional areas. It seems we may be stuck cobbling together a mishmash of systems to do different things.*
>
> — Public college/university library

> *We are a relatively small organization with far too many repositories because each business unit thinks their stuff is so unique that it can't possibly all be in the same system. I am giddy at the thought of a hydra implementation that could branch out to meet each of their internal process needs, scale to support other partners as needed, and provide unified discovery options for the general public or targeted discovery options for unique audiences.*
>
> — Regional consortium

The prospect of Hydra-in-a-Box enabling institutions to reduce or eliminate the need to support multiple systems by supporting a broader range of resources was especially attractive:

> *We currently have a custom built repository to manage access to our digital collections and are getting ready to implement a preservation strategy using another system with our [regional archives]. I am excited to hear about Hydra in a Box and that it may manage both access and preservation.*
>
> — Public library

> *I think we're at a turning point with digital resources - moving away from systems that were designed for single items (digitized images, born digital PDFs) hopefully towards systems than can handle more complex collections of digital content - from research data to web archives to dynamic media, etc. I'm excited about Hydra-in-a-box and hope it can help the library world move in the right direction!*
>
> — Private college/university library

> *Would like to store all digital objects and data in one repository or at least same type of repository and be able to have public access and internal access from same technology.*
>
> — Independent research library/archives

> *Our biggest issues are lack of integration between our access and preservation repositories (as well as ArchiveSpace, our archival management system) AND lack of staff time to learn and manage a*

> *complex system, much less research, plan and do a system migration, even if a perfect integrated system came into being.*
>
> — Private college/university library

> *The goal is to find a product that manage video, audio, image, and document files for museum and archival collections in large quantities for internal and external use. Being able to integrate object metadata with EAD finding aid data is key, as is control of metadata input and display.*
>
> — Museum

> *We are currently in the very beginning stages of planning a new "digital spine" which will provide access to all of our collections, including digital objects, finding aids, library catalogue records, and museum collections.*
>
> — Independent research library/archives

In addition to being able to manage a broader range of resources from a single system, some respondents specifically mentioned a desire for a system that was flexible and modular:

> *Prefer something that is somewhat modular, so that services for streaming, image viewing, and other kinds of content can be updated as necessary.*
>
> — Public college/university library

> *One of the reasons that we chose Hydra for our repository needs is its strength as a flexible and repurposable code base. If Hydra-in-a-Box is software that provides just the most basic repository functionality with some backend administrative GUI layer for admin users (something I would consider a "must have"), and the code is structured in such a way that it becomes easy for Hydra newcomers and veterans alike to write additional features as gems to simply be included in a Gemfile (similar to how Drupal modules are written and can be included in a Drupal site to extend functionality), that would be a greatly improved workflow for developing on, using and re-using Hydra for a variety of digital repository projects, as additional features can be built out and re-used.*
>
> — Public college/university library

> *Also prefer a system that allows sharing of resources between other systems without duplicating resources (for example, an object lives in the repository but can be accessed by Omeka to feature in an exhibit, without having to duplicate in the Omeka system).*
>
> — Public college/university library

Overall, the free-text comments reinforce findings found elsewhere in the survey results: that there is no existing ideal digital repository system, especially for smaller institutions with limited staff or budgets. Existing systems lack the ease-of-use and features people are looking for, and institutions often need to use multiple systems to manage all of their resources. Survey respondents are excited about the possibility that Hydra-in-a-Box will provide the solution they're looking for:

> *We are really struggling to find our next system. This survey was fun to take because it gave me an opportunity to express my frustration and current swamp we are in with all the material that is waiting to get into the repository.*
>
> — Public college/university library

Hydra seems out of reach for us given the infrastructure here. So I am looking forward to seeing how the Hydra-in-a-Box project progresses as it seems like a promising solution.

— Private college/university library

I would really love to see Hydra-in-a-Box succeed. If I understand the goals of the project correctly, it would allow an institution like mine to take advantage of non-proprietary, more flexible repository features than proprietary repositories (such as CONTENTdm) currently allow without requiring that we have robust local IT support. (We don't have a developer or a programmer on our IT staff, for example. Our IT staff spends most of its time dealing with public and staff computer support/infrastructure. There is no one to help build out our digital collections infrastructure.)

— Public college/university library

We would very much like to migrate ContentDM content to Hydra. We've looked at Islandora and have decided we prefer Hydra's flexibility in terms of Linked Data and other features. [...] Our biggest hurdles are 1) lack of training - it's very difficult from the Hydra/Fedora documentation to figure out the 'big picture' of how Hydra really works; 2) lack of programming staff [...]. We would absolutely love 'Hydra in a box' because it would significantly lower the barriers to getting started with a project.

— Public college/university library

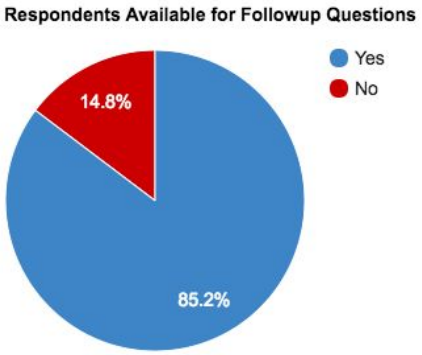
We are excited about Hydra-in-a-Box. We rolled our own Fedora / Solr based digital collections platform a couple years ago, and we've been quite happy with it. But quite precisely because things have been going well, and we started getting more uptake on it, collaborations with the archives on campus, additional features and content types, security, etc., we saw the writing on the wall and realized a small team such as our would not be able to maintain a system [...]. So we're looking to Hydra, and we're excited. Though we preferred Python, we'll glad accept the learning curve of Ruby and Rails to implement a digital collections platform that is a) extensible to the extent we've become accustom, and b) has a RAPIDLY growing community behind it. Moreover, a community we can contribute back to.

— Public college/university library

Follow-Up Questions

Q28: *May we contact you with follow-up questions if we have any?*

This survey was conducted early in the Hydra-in-a-Box design process. Because we also had plans to conduct user interviews and anticipated wanting to explore some user needs in more depth, we asked a couple of questions to obtain contact information of willing respondents. A total of 156 respondents (85% of those answering this question) indicated willingness to be contacted for follow-up questions. These respondents provided a name and an email address.



Interview Availability

Q29: *Would you be available for a follow-up interview by phone or online meeting (e.g., Skype, Google Hangout)?*

Nearly all of the respondents willing to be contacted for follow-up questions indicated that they were also available for a follow-up interview (142 respondents to this question, or 77%).