

**Institutional Repository Use of Vendor-Based Solutions Relating to
Technical Knowledge and Digital Curation**

Mary Fender O'Brien

Johns Hopkins University: Museum Studies Digital Curation Certificate Program

AS.460.674.81. SP23: Digital Curation Research Paper

Professor Jocelyn Boigenzahn

May 2023

Abstract

Institutional repositories (IRs) are digital collections that curate and disseminate the intellectual output of an institution. They play a significant role in the open access movement and in providing access to research and scholarly outputs. Vendor-based systems (VBSs) are a popular option for managing IRs. VBSs offer a number of advantages, including scalability, security, and support. However, they are expensive and require a high degree of technical knowledge for staff to extend the IR collections into digital preservation workflows. This paper examines the impact of VBSs on digital curation and preservation in IRs. The paper begins by providing an overview of IRs and VBSs. It then discusses the benefits and drawbacks of VBSs relating to the executable connectivity for digital curation and preservation by digital librarians. Finally, the paper presents the results of a survey that gauged current technical knowledge and other aspects of those working in the librarian roles of digital curation/preservation, asset management, and institutional repository management of scholarly digital assets. The analysis of the survey suggests several factors that inhibit IRs from utilizing VBSs to their fullest for digital curation and preservation. Overall, a VBS is a valuable tool for digital curation and preservation in IRs if gaps in technical knowledge, increased resources, and stakeholder support are improved.

Keywords: Institutional Repositories, Scholarly Communications, Vendor-Based Systems, Technical Knowledge, Digital Preservation, University Libraries

Table of Contents

Abstract. 2

Introduction and Problem Statement. 4

Literature Review. 6

Research Design and Methodology. 13

Analysis. 15

Recommendations. 42

Conclusion. 44

References. 46

Appendix A: Survey. 49

Introduction and Problem Statement

Institutional repository librarians, specifically dealing in scholarly communications, are charged with many tasks, and at the top of the list are collecting, disseminating, and preserving the intellectual output of the faculty, students, and journals published by the institution. To accomplish this task, Vendor-Based Systems (VBSs) are often used to manage digital assets. These systems, which require an initial and ongoing financial commitment, are myriad in the marketplace. Institutions perform a lengthy selection process resulting in systems chosen based on the following needs of the institutional repository (IR); the “turnkey” nature of implementation (including migration of existing assets and ease of later additions), the vendor’s history and vision of longevity, the vendor’s frequency of adding/enhancing features, and the available financial resources of the institution for the system. After migrations are complete, the idea of changing vendors for these services is overwhelming, resulting in vendor lock-in. Furthermore, “services with a principally academic market seem to be especially susceptible to monopoly or oligopoly dynamics among commercial providers” (Schonfeld, 2018).

An unaccounted cost of these VBSs are the resources needed for staff to fully implement these systems’ capabilities. An article by Ruth Tillman, “Repository Ouroboros,” explains the selection and implementation process well, including the frustration by many staff with final installations - finishing in time to start the platform selection process all over again, because the next new shiny product or vendor has come out with the next best thing for the institution. The sentiment is often, “You feel so discouraged. You are sure everyone else is ahead of you. You do not yet see that you are just one more person riding the Repository Ouroboros” (Tillman, 2019). Institutions contract with VBSs while not fully understanding that these systems are not the “end all” for their existing and future IR needs. Without additional staff knowledge and time to

operate and manage these systems, workflows for digital preservation are not realized. “System training” and user support groups are offered up by the vendors and the IRs assume this will be enough. But this is not the norm. The lack of time and knowledge required quickly add up and the systems do not get the attention needed. This often leads to the resignation of many IR managers to “set it and forget it,” only scratching the surface of what these systems offer. With this, data may be curated but the next steps toward preservation actions are not taken.

This paper will discuss how this non-practice affects digital curation and preservation of the repository assets at these institutions and detail the barriers prohibiting staff from benefiting more from these VBSs. In addition, the requirements necessary for IRs to perform true digital preservation (instead of warehousing), and other actions that will benefit the institutions more fully will be shared. To this end, existing and additional author survey research conducted for the topic will be used and analyzed, detailing the disparity between the current capabilities and those required for digital librarians to reduce the gaps of knowledge needed to maximize VBSs, and more productively use these systems to benefit the institutions.

Literature Review

This literature review is based on the identified issues: non-practice effects on digital curation and preservation of the repository assets at institutions, the barriers prohibiting staff from utilizing the features more completely from these VBSs, the requirements necessary for IRs to perform true digital preservation through advanced connectivity to digital preservation services, and other actions that will benefit the institutions. The literary resources collected are as timely as possible since the nature of the issues discussed are recent developments in this field. Initial research and interviews conducted by the author revealed dissatisfaction in the management of VBSs and their relation to incorporating digital preservation.

VBSs are utilized in many areas of data collection management in libraries, museums, and archives, and the features and functions of these are specialized for these areas. Many university libraries that maintain institutional repositories (IRs) are using the VBSs to publish their content to the internet, and as they are able, are utilizing the cloud-based storage features. Lynch defines an institutional repository (IR) as “a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution” (Lynch). More than a set of services, IRs, are policies, staff, and the digital material that is capable of being extracted from one system and placed into another system. As Trevor Owens states in his first of sixteen Guiding Digital Preservation Axioms, “A repository is not a piece of software” (Owens, 2018). Additionally, to perform ongoing preservation actions on the digital assets within the IR, workflows need to be customized and extended beyond a VBS.

Mark Ware concurs and adds specifically regarding scholarly material, “to be a “web-based database (repository) of scholarly material which is institutionally defined (as opposed to a subject based repository); cumulative and perpetual (a collection of record); open and interoperable (e.g. using OAIS- compliant software); and thus collects, stores and disseminates (is part of the process of scholarly communication). In addition, most would include long-term preservation of digital materials as a key functions of IRs” (Ware, 2004). VBSs advertise digital preservation as a service except the offerings are not true preservation but a form of collection management.

Of the variety of systems in the scholarly communications marketplace “There is general consensus that DSpace and Digital Commons are the two most frequently used platforms at American institutions. In studies where researchers reported software platform usage, DSpace installations ranged from 43 to 58 percent with Digital Commons implementations ranging from 21 to 27.8 percent of all platforms identified” (Henry & Neville, 2017). This consensus was held in 2010 (Xia & Opperman, 2010) and ongoing as shown in the survey findings to follow.

However, even with such wide usage of these two platforms, due to a number of factors including digital preservation strategic planning, gaps in knowledge of cloud computing, and staffing, these efforts do not consistently produce true digital preservation actions, and this “non-practice” effects to digital curation and preservation puts the repository assets at risk. The definition from Giaretta explains, “if we are to preserve digitally encoded information, we must ensure it remains understandable and usable. In other words, preservation is the *sine qua non* of curation. For example, it is possible to manage and publish encoded information without regard to future use; on the other hand, if one wishes to ensure future as well as current use, one must understand the requirements for preservation” (Giaretta, 2011)

DSpace is an open-source software platform hosted by LYRASILS, which also supports DuraSpace, DuraCloud, VIVO and Fedora, created by staff of MIT Libraries and Hewlett-Packard (HP) Labs in 2002. It has since evolved into a system governed, managed, and updated by community users. As an open-source platform, it is free but paid membership and registration are encouraged, and hardware and specialized peripherals are needed. It advertises itself as “an out-of-the-box” open-source software package for creating repositories focused on delivering digital content to end users and providing a full set of tools for managing and preserving content within the application” (*Technical specifications* 2022). But looking deeper into how DSpace accomplishes these actions, they share that “DSpace itself does not guarantee the preservation of your digital materials. However, DSpace software is suited to play a central role in your overall digital preservation strategy. Keep in mind that your local digital preservation strategy should likely include a backup/restore plan, along with virus checking, etc.” (User FAQ - How does DSpace preserve digital material?).

The other top VBS, bepress’ Digital Commons, owned by Elsevier, a subsidiary of RELX, is a hosted, open access platform with a search engine friendly discovery layer, designed to highlight an institutions’ scholarly research, ETDs, archive collections, as well as conference content and publishing of journals (*Digital Commons* 2022). It came to be in 1999 with the goal to “help academic communities maximize the impact of their research and demonstrate their value. With Digital Commons and the growing Expert Gallery Suite, universities can collect, preserve, and showcase the full range of their intellectual output and expertise” (*Bepress launches new service* 2017). Preservation is one of three main efforts of the platform but to achieve this, much more is necessary.

On digital preservation, from bepress,

“Long-Term Preservation and OAIS Compliance via LOCKSS - In the face of the challenges posed by providing access to digital assets over the long term, the LOCKSS organization (<http://www.lockss.org>) has developed a very effective tool. The Open Archival Information System-compliant LOCKSS software is designed to harvest and preserve subscription-controlled journal articles, and is well-suited to preserve open access repository content. Bepress has worked with the LOCKSS organization to ensure that Digital Commons repositories are LOCKSS-compliant. We encourage Digital Commons subscribers to join together to create a Private LOCKSS Network, or PLN, to back up each other’s content for long-term preservation” (*Safeguarding your content with Digital Commons*).

In a round-about way Digital Commons points out that the platform is not Open Archival Information System (OAIS) compliant and encourages extensions to preservation platforms. In 2016, bepress launched its bepress Archive feature. This functions as a “mirrored archive” of the Digital Commons held repository and sends it to a connected Amazon S3 (Simple Storage Service) bucket in an account under the institution and allows unlimited access to a backup of this data as well as connectivity to outside platforms performing preservation actions. This seemed revolutionary but not without issues. The platform creates new file names for each digital object added to Digital Commons, however they lack unique identifiers and titles. The system also requires additional add-on features to add preservation metadata, and other additional schema to crosswalk and/or allow harvesting to outside preservation platforms. Another issue is that changes made to a file in the online Digital Commons repository will not overwrite the original file previously sent to the bepress Archive but instead, create a new file.

To correct this, the old file must be manually removed from the S3 archive. The author made these determinations, while in an internship at a university that uses bepress Digital Commons, in the scholarly communications department, during their studies of digital curation at Johns Hopkins in 2022. Because of these additional requirements of technical expertise and time, Digital Commons, even with the new Archive feature, is still not OAIS compliant.

Bepress - Digital Commons and DSpace, and other vendors, share these details on digital preservation, but in practice, the design of digital preservation strategies required to realize these actions falls on the institutions themselves, and this realization often comes later - after onboarding of these systems. While sharing knowledge of digital preservation between organizations, its importance, and efforts surrounding these actions, are ongoing and growing, according to the 2018 State of Digital Preservation (Rieger, 2018), there remain significant barriers prohibiting staff from utilizing the features from these VBSs more completely.

Most significant of these stem from lack of stakeholder support in the creation and implementation of a digital preservation strategy, and lack of training, resources, and time. In leadership, as Schonfeld states, “Fragmented decision-making cannot address issues of collective strategy” (Schonfeld, 2018), and without a strategy, limited resources are spread thin, reducing effectiveness of existing efforts. Lack of stakeholders’ comprehension of the complexities required for successful implementation and ongoing management of VBSs and digital preservation is a common theme, and logically leads to a lack of support for strategic planning, additional resources and staffing.

Another barrier is gaps in technical knowledge relating to the connection of additional platforms that would produce a digital preservation workflow. VBSs incorporate the ability to

connect to systems such as CLOCKSS, LOCKSS, Portico, the Internet Archive, Preservica, cloud, local and hybrid storage, and more, but given the vast differences in institutional sizes, staff technical knowledge, and time resources, every instance is unique and requires expertise to fulfill and manage these workflows. Community support groups for the VBSs are collaborative and helpful, but knowing what to ask and how to identify needs and find solutions in these forums is time consuming and difficult to project and plan for. It is these extensions and connections that would inevitably fulfill the goal to support the OAIS preservation information model (Lavoie, 2014).

Since every IR is unique, the requirements necessary for them to perform true digital preservation through advanced connectivity to digital preservation services and other actions vary because of size, staff resources, and the goals of each institution. However, the common requirement of IRs and at the core, is the publishing and preserving of high risk born-digital assets - electronic theses and dissertations (ETDs), and the top reason the repositories exist. In addition, journals, research data, open educational resources (OER), websites, and audio/visual digitized assets are objects managed in IRs but with different requirements of scope and accessibility. These additional assets are at risk of loss and damage due to inaccessibility and obsolescence. Given these complexities, including metadata, embargoed and persistent URL requirements, high levels of technical knowledge are required for planning, and for training existing staff or outsourcing IT needs to university departments of additional outside IT vendors, stakeholder support and budget forecasting is imperative.

As the number of IRs has increased since the early 2000s, technical knowledge of digital librarianship has increased but not always at the same rate needed to “keep up” with advances in software and systems. In initial interviews with scholarly communication librarians, frustration

was a common theme. In addition, these unrealized digital preservation strategies and lack of resources result in resignations and may appear to be a prevalent problem contributing to issues with mental health in those working in digital preservation. A recent survey conducted by the Digital Preservation Coalition, that closed in April 2023, and has yet to publish its finding, but according to the set of published queries, brings up issues regarding; clarity of roles, job description, ability to “keep up” with expectations within regular contracted hours, as well as feeling “overwhelmed by the ongoing advocacy burden relating to digital preservation”(Digital Preservation Coalition Survey questions 2023).

Research Design and Methodology

The research data collected in this study was gathered through both qualitative and quantitative methods. In the qualitative method employed, interviews with scholarly communication library professionals, managers, and administrators of their respective IRs. These interviews were open-ended but with prepared questions that led to new points of discussion. These interviews contributed to a greater understanding for the author regarding the roles of staff in IR management, IT technical requirements, stakeholders, and desire for digital preservation. These interviews also informed the development of the questions in a quantitative survey delivered to library professionals.

The survey prepared for this study was disseminated for university scholarly communications to library professionals engaged in using VBSs for institutional repository management. The main question which focused this survey was, what factors keep IRs from utilizing the features in vended solutions services more fully to perform digital preservation. The set of 50 questions in the survey (see appendix A) were tailored to gauge the current technical knowledge of those working in the librarian roles of digital curation/preservation, asset management, and the institutional repository management of scholarly digital assets, including, but not limited to, born digital objects, electronic theses, dissertations, journals, research data, open educational resources, websites, and audio/visual digitized assets. The survey consisted of ten sections, with two sections with in-depth questions about bepress Digital Commons and DSpace. The other eight sections focused on participant and institutional general information, community and collaborative attitudes and involvement, publishing and preservation platforms, related workflows, digital preservation strategy, technical skills, stakeholders, and other factors. This survey aims to gain information about the barriers that may prevent full implementation of

VBSs to preserve these assets for the long-term utilizing cloud technology, and collaborative initiatives needed for digital preservation. Additional data was collected through a meta-analysis of the valuable published scholarly research, posts on Digital Library Federation, and the DLF Announce Listserv pointing to previous studies on IRs and digital curation and scholarly communications over the past ten years, as well as posts about training and digital preservation found within Google community groups for bepress Digital Commons and DSpace. Analysis of the scholarship and other resources in combination with the interviews and survey results provides a clear state of the field picture of the use of VBS in IRs and some insights into what the field needs to continue to improve to the benefit of the faculty and future researchers who will make use of the digital objects being preserved.

Analysis

Bringing together the survey results, a summary can attempt to answer the main question: What factors keep IRs (Institutional Repositories) from utilizing the features in vended solutions services more fully to perform digital preservation. This survey was sent to seventeen potential participants via direct email and broadcast on to 3 community boards; DLF Announce Listserv, and DSpace and Digital Commons Community Google Groups. At the survey's close, 20 respondents were recorded: 4 from direct email and 16 from community boards. One respondent's answer field was blank equating to nineteen respondents with information to be evaluated.

Section 1: Basic Information

The beginning of the survey gathered basic information. Respondents self-identified with a variety of job titles all centered around digital librarianship in a scholarly communications or repository management with two outliers: Technical Services and Senior Information Assistant.

These titles included;

- Scholarly Communication Librarian
- Head of Digital Services
- Scholarly Communications Outreach Coordinator (Associate Librarian)
- Digital Collections Specialist
- Head - Metadata Services
- Senior Information Assistant
- Technical Services
- Digital Access & Metadata Librarian

- OER and Scholarly Communications Librarian
- Scholarly Publications and Metadata Librarian
- Faculty Scholarship Outreach Librarian
- Digital Repository Administrator
- Digital Records Archivist and Digital Preservation Librarian
- Student Services Librarian
- Digital & Electronic Resources Librarian (soon to be Digital Collections & Archives Librarian)
- Electronic Resources and Scholarly Communications Librarian
- Reference & Digital Repository Librarian
- Interim Head of Special Collections/Archives and Digital Resources Librarian

The length of time in these current roles varied from 0-2 years to 15+ years, with respondents reporting;

0-2 years 26.3%

2-4 years 15.8%

4-7 years 21.1%

7-11 years 15.8%

11-15 years 5.2%

15 + years 15.8%

There have been differences found relating to Carnegie Classification and so determining which group gave which answers is important. In this survey, five R1: Doctoral Universities–

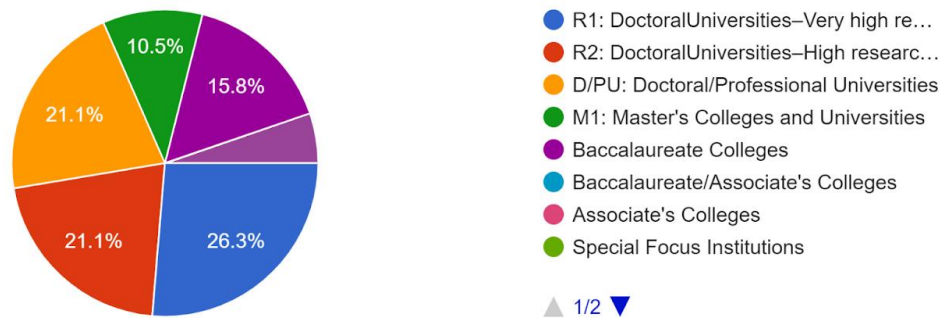
Very high research activity, four R2: Doctoral Universities–High research activity, four D/PU: Doctoral/Professional Universities, two M1: Master's Colleges and Universities, three Baccalaureate Colleges, and one Medical Library (Figure 1).

Figure 1

Carnegie Classification Graph

What is your institution's Carnegie Classification?

19 responses



The next questions relate to working groups, associations, and training. The results show overwhelming support by respondents for participation, 89.5%, in these activities except for those from 2 of the 3 Baccalaureate Colleges, 10.5% (Figure 2). However, the support from stakeholders and time provided for these skewed more neutral to negative (Figure 3).

The list of respondents' organization involvement.

- Library Publishing Coalition
- VIVA
- NAHSL

- Network of Alabama Academic Libraries for Alabama-specific training, Open Education Network is great for OER training.
- Lyrisis
- Law Repositories Caucus of the American Association of Law Libraries
- Consortiums
- For Digital Curation: NDSA, BitCurator Consortium
- Society of American Archivists and Digital Preservation & Outreach Network
- Minitex
- ORCID consortia with Lyrisis, they seem to have lots of training opportunities.
- Local/regional groups and associations are really helpful!

Figure 2

Support and Training Organization Involvement

Are you (and/or your institution) a member of working groups for associations, or consortiums, organizations where you have access to education and training opportunities?

19 responses

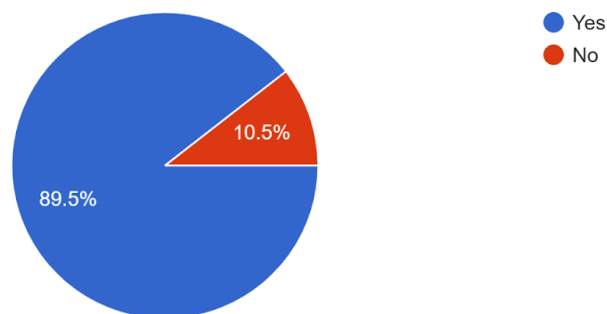
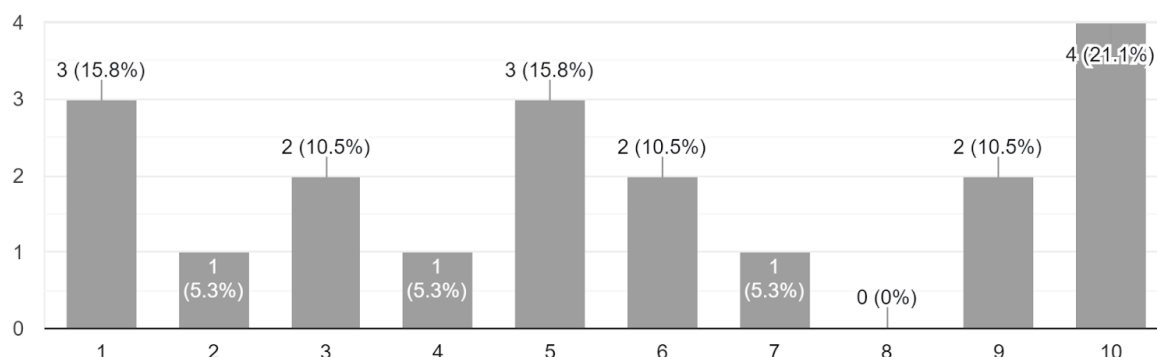


Figure 3*Stakeholder Support Graph*

In your work environment, is service to non-profit associations (through board positions, working groups, task force, etc.) that perpetuates the care ...through digital curation encouraged? To what level?

19 responses

**Section 2: Publishing Platforms**

Prior research was confirmed as this survey also shows the two leaders among VBSs (Vendor Based Systems) are Digital Commons and DSpace. Others with high instances of use were CONTENTdm, Islandora and Omeka (Figure 4). This question allowed for multiple answers and the differences in the main platform, between Digital Commons and DSpace, as a stand-alone or combined with additional VBSs, is worth noting. Specifically, in class R1, the two DSpace users also reported using between three and seven additional systems. Of the four Digital Commons users, only one listed a single additional system - Omeka. Of the eleven other Digital Commons users in other classes, three of them listed two VBSs and the other eight listed Digital Commons solely. In the management of these systems, 73.5% respondents identified as Administrators, with 15.75% as Managers and 10.5% as Authorized Users.

Figure 4

Publishing Platform Percentages

Do you use a Publishing Platform? Select all that apply.

19 responses

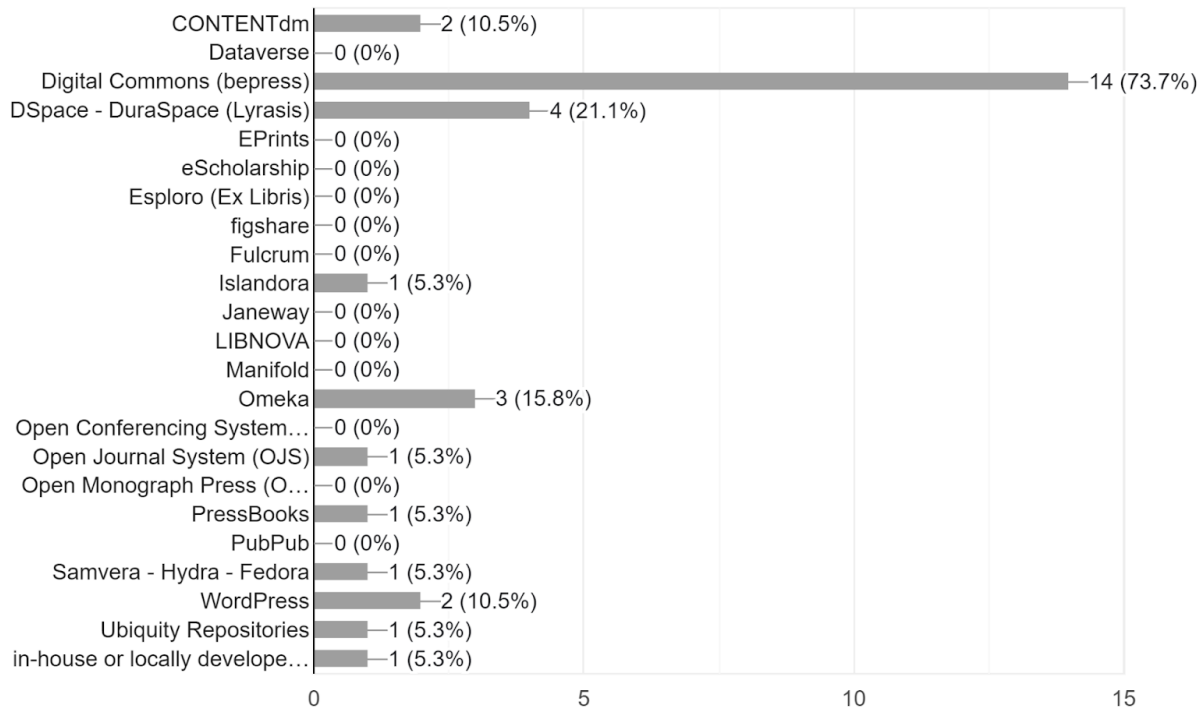
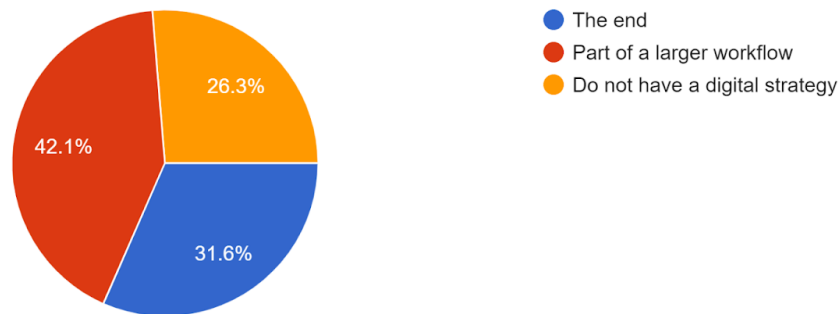


Figure 5*Publishing Platform Strategy*

Considering your institution's digital strategy, is the IR publishing platform the end of the line for the digital assets therein? Or part of a larger workflow?

19 responses



Findings show four out of five R1 institutions use these publishing platforms as parts of larger workflows. In other classes, four out of fourteen reported this with the remaining evenly split between the workflow ending with the publishing platform and having no digital strategy - which are indeed comparable results (Figure 5).

Section 3: Digital Preservation Platforms

This section focused on preservation standards, policies, and platforms. For standards and platforms, multiple responses from one participant were allowed. In the standards responses, eight participants selected “Other” - a clear favorite (Figure 6). This was a surprising return, and in retrospect, a place for a space for additional answers would have produced more data. Four participants skipped this question, while one R1 participant selected both OAIS Reference Model (ISO 14721) and NDSA Levels of Preservation. The next finding shows only four respondents have a preservation policy, two of these are in the R1 class, with one each in R2 and D/PU

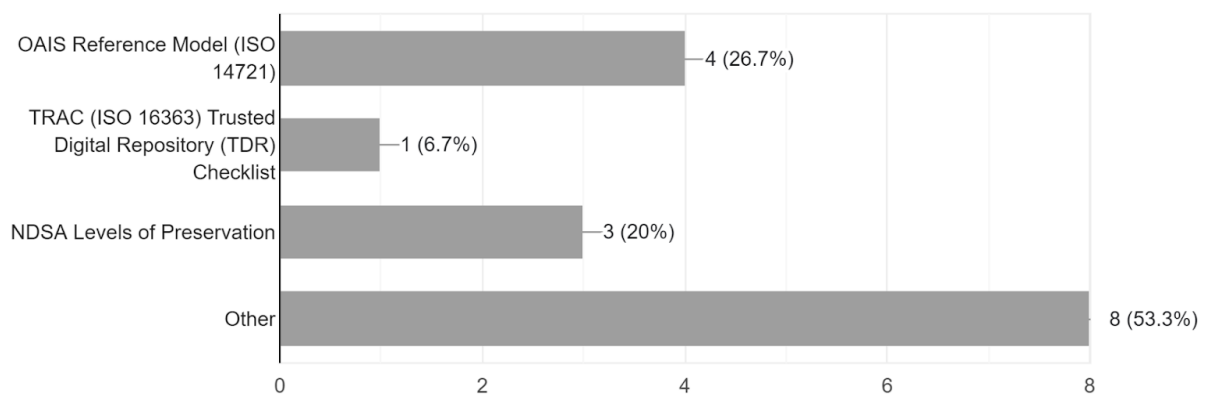
classes. Of the remaining fifteen respondents, nine have no policy and 6 stated they are creating one.

Figure 6

Preservation Standards

What digital preservation standards does your repository follow?

15 responses



For the findings for preservation platform usage, again, this question allowed for multiple answers and again, specifically, in class R1, the two DSpace users also reported using between two and eight systems. The remaining results did not fall in line with the findings for publishing platforms in that here there is a greater variety of platforms enlisted. Results showed the highest use was that of the Internet Archive at six of the eighteen responses.

A good example of difficulties of connectivity and extensions of IRs publishing platforms beyond Digital Commons and DSpace to preservation platforms becomes known here. For surveyed Digital Commons users, only one, an R1 IR, is connected to LOCKSS, though beyond is open about LOCKSS and CLOCKSS connectivity as a possibility (*Safeguarding your content*

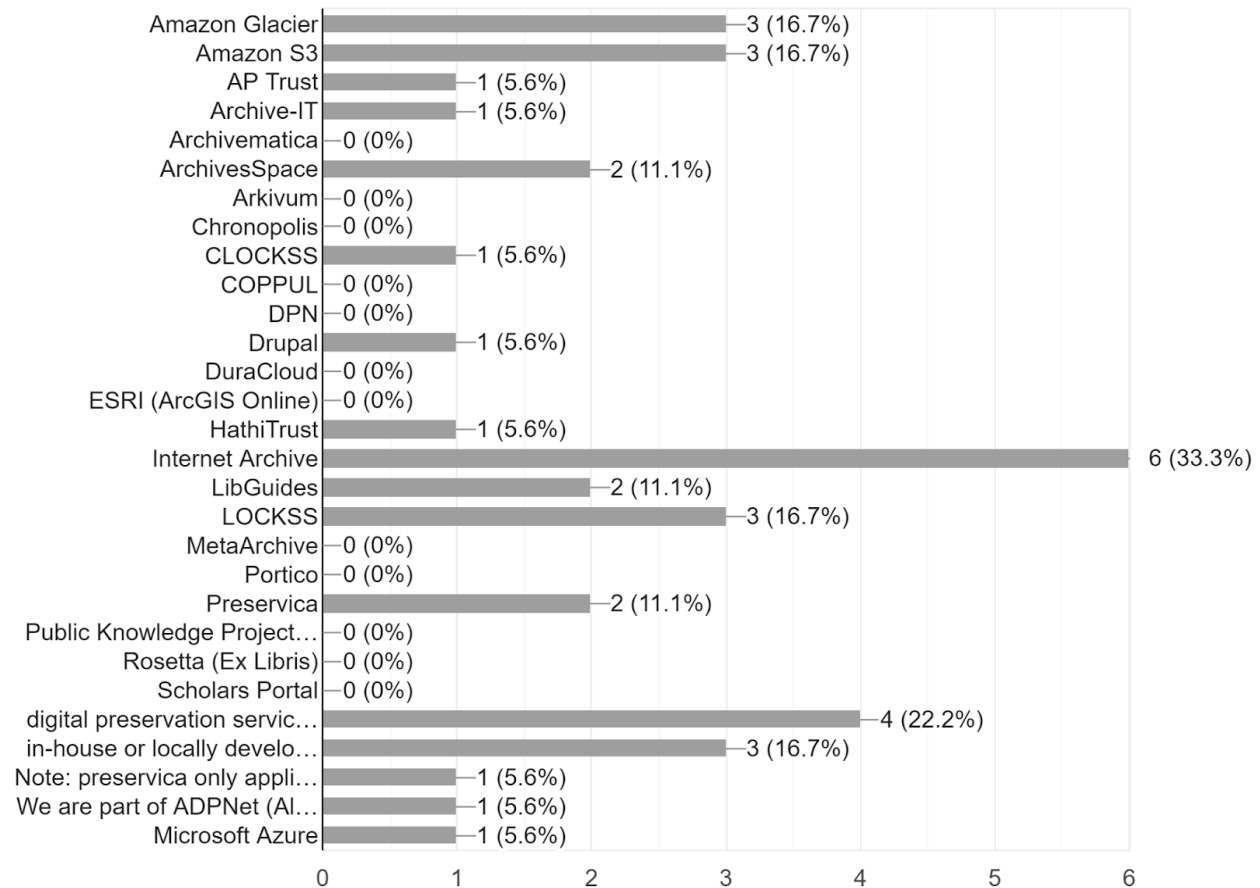
with Digital Commons). Conversely, of DSpace users surveyed, two extend to LOCKSS and one to CLOCKSS. In Cloud storage, six Digital Commons users are Amazon Web Services Glacier and S3. In DSpace users, one is using Microsoft Azure (Figure 7). One may think the launch of the bepress Archive in 2016 may be the reason for six of the fourteen Digital Commons users to use AWS (Amazon Web Services) cloud. But upon further questions in the survey, only four said they are using the bepress Archive feature.

Figure 7

Preservation Platforms

What platform(s) is used for digital preservation? Select all that apply.

18 responses



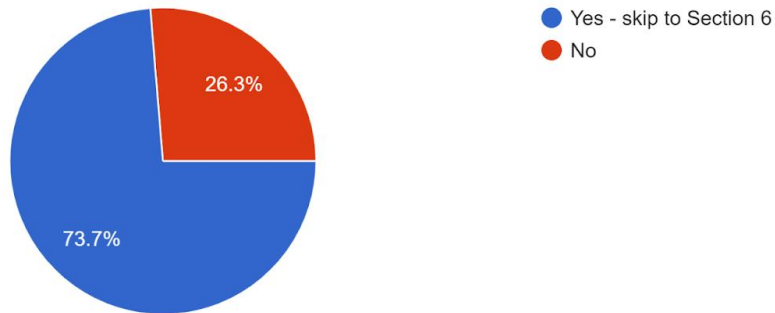
Section 4: Digital Commons (bepress)

Figure 8

Digital Commons User Graph

Is your IR using Digital Commons?

19 responses



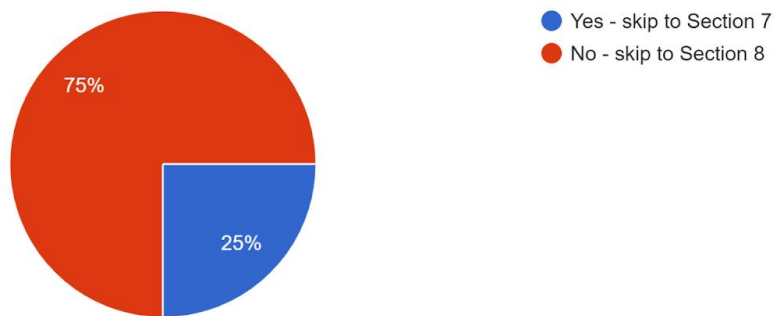
Section 5: DSpace

Figure 9

DSpace User Graph

Is your IR using DSpace?

16 responses



Section 6: Digital Commons (bepress) Questions

Sections 4 and 5 were used to separate Digital Commons and DSpace users. Section 6 contains questions specific to Digital Commons users. Findings showed that the largest group of these IRs, 35.7%, had been using Digital Commons for between 5 and 6 years and of those all have a current contract with bepress. In fact, all the Digital Commons IRs have a current contract with bepress except one. Other time frame groups were 11 years or more - 28.6%, 9 to 10 years - 14.3%, and 7 to 8, 3 to 4, and 1 to 2 years - 7.1%. All but one respondent answered that updates to the software were automatic instead of manual. The question on training for the platform was an open response and most of the answers included onboard training, self-training, and training as needed from a representative. One answer, from an R1 IR, stated “bepress manager certification course, conference programs, and self-trained.” Another unique response from a R2 IR said “Introductory, in-person and self-driven with online documentation.” These findings express that bepress has made training for their platform accessible and available during an introductory period and ongoing self-led and need for it self-determined.

In 2016 bepress announced its new feature, the bepress Archive. This a backup into a S3 bucket owned in the IR Amazon Web Services (AWS) cloud service. Response to question 21 showed that 25% were knowingly using the feature and another 25% did not know if their IR was or not. This lack of information is concerning and leads to confirmation that actions outside of the Digital Commons platform - for preservation or otherwise - are outside of staff knowledge. Eight of the 16 IRs use the publishing feature and seven stated their use was for journal publishing. Two also use the Expert Gallery Suite/"Selected Works" for individual faculty profiles.

About digital curation/preservation workflows developed around Digital Commons, response rate was 68.75% with 18.75% stating “none.’ One of these respondents put it all on bepress, with “We pay them to take care of stuff for us.” Most of affirmative responses ranged from vague to very detailed;

- Selected content from DC Archive routinely deposited into AP Trust
- In theory, bepress Archive automatically pushes a backup of all IR content and metadata to Amazon S3, where a series of custom scripts package the content for long-term preservation in Amazon Glacier. In practice, this process is no longer working properly (on our end) without expertise and dedicated time to manage the preservation scripts and processes. Instead, our workflow is simply assuming that "if some crisis takes out all of the Digital Commons servers and backups, we probably have bigger technology problems than restoring repository files". We do also try to save a copy of all submitted IR content to an external hard drive, but this is not a systematic or comprehensive workflow.
- We collect as many faculty publications as we can, and regular law school publications including our alumni magazine and law reviews
- We have scanning settings and file processing/publishing processes in place.
- I am currently developing a digital strategy to include preservation. Digital Commons will house and make available access copies only.
- files saved locally - dark archived
- We use separate systems for digital preservation, but we do use additional preservation workflows when the work being submitted to Digital Commons is a

website. In this case, we archive the website using Archive-It, and then link to the archived version from the Digital Commons record.

- Some items have been uploaded to the Internet Archive.

Respondents' use of phrases like “we try to,” “probably” and “in theory” show the lack of confidence in policies and procedures surrounding Digital Commons workflows. While other respondents have sound routines that involve outside platforms such as AP Trust, Archive -IT, the Internet Archive, and local “dark archive”. A further question on systems integration produced even fewer responses. One respondent added integration with ALMA and WorldCat.

Concepts and implementation of Open Access (OA) and Open Educational Resources (OER) are important aspects in IRs. While this subject is loosely connected to the survey's main question, the author included the topic, and how it relates to VBSs and preservation, because the movement against “pay-walled” content is in discussion at many institutions. In response, most participants mentioned how they are using OER and OA in their IRs and the pros and cons of Digital Commons, stating “DC supports OA journals well. It is limited on hosting research datasets (20GB+) OERs we have yet to use it.” and “We use Digital Commons to publish OA and OER materials created by members of our university community. We also try to only publish materials that are open.” One respondent from a D/PU class university wrapped it up plainly saying, “Digital Commons continues to be the leading option, and possibly only real option, for OA repository infrastructure at small to medium institutions which lack dedicated IT staff to maintain an open-source repository. However, the Digital Commons journal publishing tools lag behind the features and functionality of Open Journal Systems (OJS), and the product lacks any specific features to support OER creation. Pressbooks, Manifold, and other open book publishing

tools are far in advance of Digital Commons in this regard - DC offers little more than PDF hosting for OER.” This leads to the direct question of challenges and or limitations specific to Digital Commons that users have experienced. Most of the responses related to limitations.

These included;

- “Limited flexibility as a platform,” specifically; in ease of customization of the user interface, inability to initiate new collections, and other operations, without consultant assistance - and “Journal publishing tools and administration interface that are both falling behind Digital Commons alternatives.”
- The upload limit, 20GB+ for datasets and lack of integration with ORCID
- Lack of features and development of features for OER creation
- Inability to export data/recs for Integrated Library Systems (ILS)
- “It's an access/discovery only tool (would prefer a more integrated system for digital collections that includes preservation or more collections management).”
- “It feels like it's built to be so simple that it's sometimes harder for someone who understands what repositories can actually do to use it.”

There were also challenges identified;

- A persistent focus on tools for harvesting metadata-only records from commercial publishers and using the IR as a discovery portal rather than providing access to actual scholarly content.
- The bundled nature of Digital Commons products (you cannot get the repository without the journal tools, etc.) makes library budgeting difficult.
- “While Digital Commons can host many file types, many users opt to embed audiovisual content, like videos, from third party platforms rather than upload

individual files directly to Digital Commons. This can be a challenge because if the user later deletes the video from the third-party platform, it will also disappear from Digital Commons.”

Even with these issues, satisfaction with bepress Digital Commons was rated moderate to high (Figures 10 and 11).

Figure 10

Digital Commons Community Support

How satisfied are you with Digital Commons community support? Is it used widely to be active and helpful?

14 responses

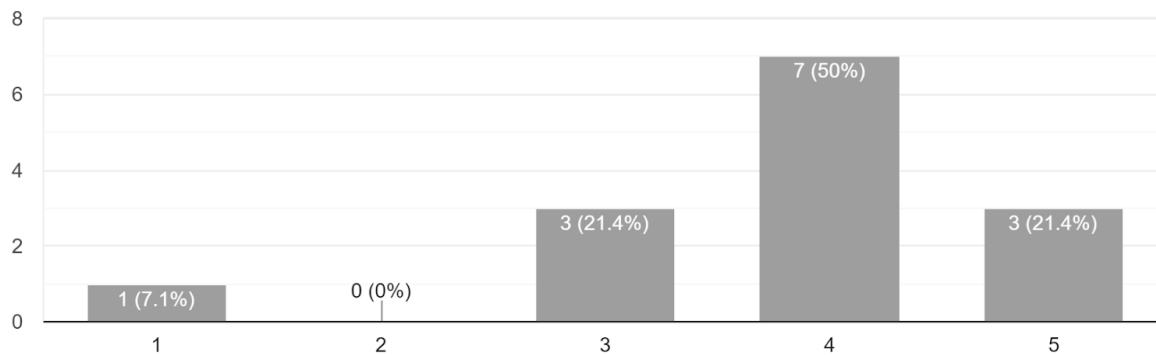
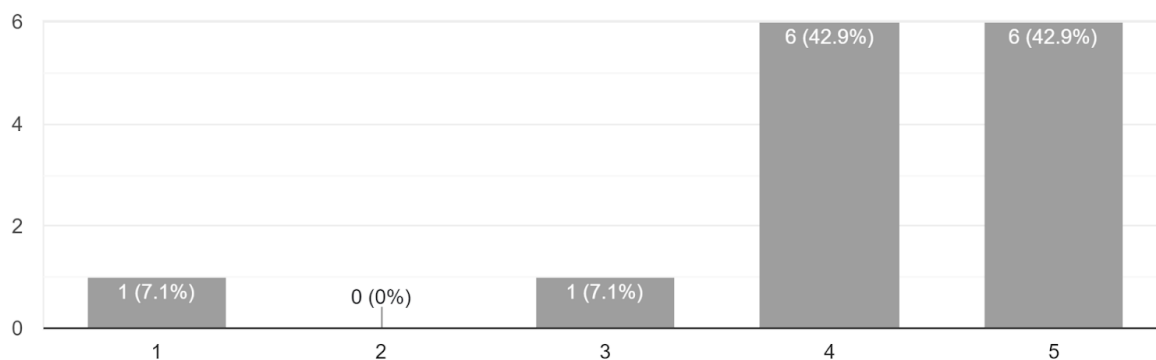


Figure 11*Digital Commons Direct Support*

How satisfied are you with Digital Commons account manager/team support?

14 responses



At the end of this section respondents were asked if they see Digital Commons as a long-term solution. The response from the R1 class leaned toward the negative, stating that the system is too closed to be useful and its parent company, Elsevier, and its actions, are concerned regarding its perceived mission of pushing clients into more expensive products in the future. For the R2 class, respondents mentioned reevaluation of continuing use of Digital Commons, “it’s an access/discovery platform, so that misses a lot of very important steps and workflows for maintaining digital collections” and acknowledged that VBSs should not be seen as long-term solutions. Interestingly, institutions in the D/PU class answered in the positive - that they would continue this VBS in the future. And as the institution’s size grew smaller, the M1 and Baccalaureate classes, respondents stated cost as a factor in discontinuation.

To wrap up both sides this response from a Baccalaureate IR is fitting,

“Yes and no. Yes, because it is a hosted solution that is easy for a small institution to use and implement. The support team is responsive and helpful, and the vendor manages all system

maintenance and technical issues. No, because in the end, it is a vendor-based system that we have little control with - for example, it could be sold, closed, or just increase in cost to a level we could not afford. While it is currently a good solution for my small institution, it is difficult to be certain of its long-term success.”

The last question for Digital Commons users was on exit strategy. Of the fifteen respondents, only one, 6.6%, had a plan, two, 13.3%, did not know and the remaining twelve, 80%, did not have one.

Section 7: DSpace Questions

Section 7 contains questions specific to DSpace users. Findings showed that 50% of these IRs had been using DSpace for longer than 11 years and the other 50%, for 7 to 8 years. Of these, 75% stated they do not have a contract. On system updates, it was again split 50-50 between automatic and manual. The question on training for the platform was an open response and all submitted that they were self-trained. These findings make sense as DSpace is an open-source software system hosted by Lyris, and on their site, readily available training resources, wikis and user guides are found.

In 2016, DSpace released an Enhanced File Storage layer feature, S3 BitStore. “In DSpace 6, the file (bitstream) storage plugins have received a major refactor, including support for Amazon S3 file storage” (Donohue, 2016). In writing the survey, the author strived to ask related questions of each user group. This feature was released the same year as bepress Archive and had similar attributes. The DSpace respondents answered with half, “no”, and the other half, “I don’t know.” This could have a few meanings; 1) the plugin did not fulfill a need, 2) the plugin did not work well, 3) it was too difficult to install and run, 4) or something else. But regardless of the author’s guesswork, this leads to confirmation, like the finding from Digital

Commons users, that actions outside of the DSpace platform - for preservation or otherwise - are outside of staff knowledge.

Regarding digital curation/preservation workflows developed around DSpace, 3 of the 4 respondents answered, “ETD self-deposit; ETD deposit from ProQuest; Open Access Scholarly publication deposit with Symplectic Elements”, “Institutional structured”, and “Item versioning”. The fourth expects workflow development in the future. There were two responses for other systems/platforms that have been integrated with DSpace, one is working on integrating ORCID and the other stated the repository API, Elements.

DSpace users surveyed responded optimistically to the question on the concepts and implementation of Open Access (OA) and Open Educational Resources (OER). One to their answer added, at least on the local level. Another shared that 99% of their IR content is OA or OER.

This leads to the direct question of challenges and or limitations specific to DSpace that users have experienced.

The limitations included;

- “Limited out-of-the-box integrations with other systems.”
- “Unable to manage the software alone. We have a contract with Atmire to host the software, and they take care of keeping the software working.”

There were also challenges identified;

- The learning curve is steep
- Limited staff to support, upgrade the software.

- “A colleague and I administer the DSpace instance. We are still on version 5, but we should go to version 7 later this year. We are not happy with the search functions, and the quality of the image thumbnails are less than great.”

Satisfaction with DSpace was rated moderate to high (Figures 12 and 13).

Figure 12

DSpace Community Support Satisfaction

How satisfied are you with DSpace community support? Is it used widely as to to be active and helpful?

4 responses

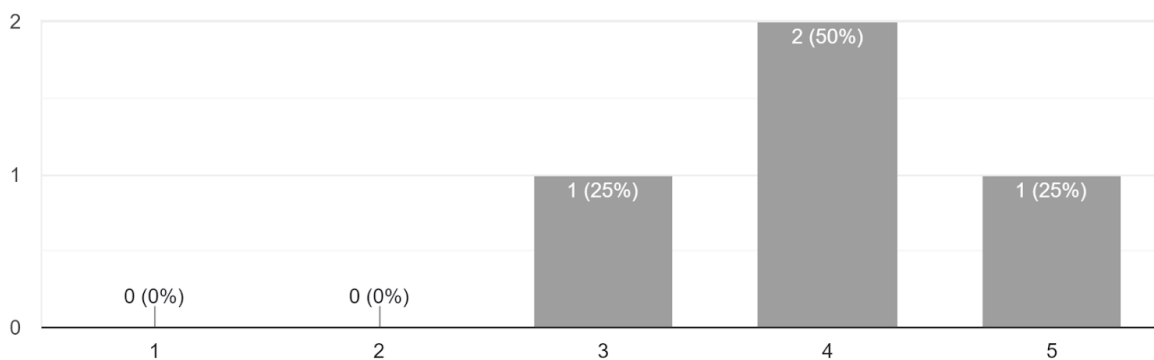
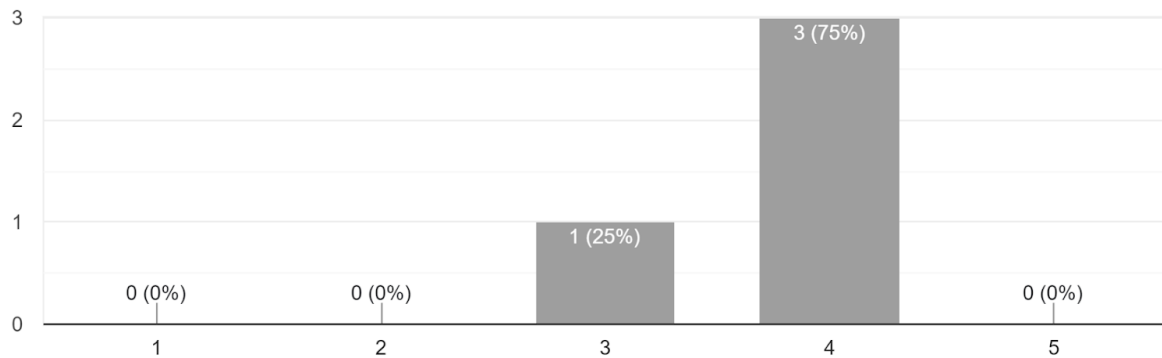


Figure 13*DSpace Team Support Satisfaction*

How satisfied are you with DSpace account manager/team support?

4 responses



Concluding this section respondents were asked if they see DSpace as a long-term solution. The one response from the R1 class resigned stating, “It is working good enough for us right now.” The M1 user was positive, “User friendly and good command of support system” while the Baccalaureate class respondent was definite about staying with DSpace because “it promotes long-term accessibility of our institutional publications”.

The last question for DSpace users was on exit strategy. Of the four respondents, two did not know and the other two did not have one.

Section 8: Technical Skills

This section of questions, and in the two sections to follow, was put to all survey participants. Also, with the findings from these sections, all the previous supporting questions complete the snapshot of issues relating to the survey’s main question.

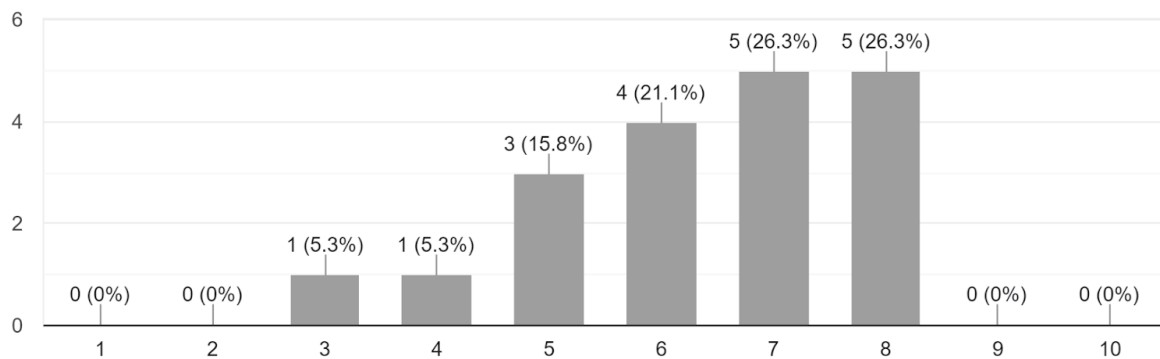
The first question queried self-assessment of technical skills relating to VBS management and digital preservation skills. The findings are on a linear scale 1-10, with 1 - Low - limited functionality and 10 - High - Complete understanding. More than half of respondents assessed high-moderate (7-8) level of skill (Figure 14).

Figure 14

Technological Skills Self-Assessment

Self-assessment: please rate your technological skills in relation to VBS systems management and digital preservation.

19 responses

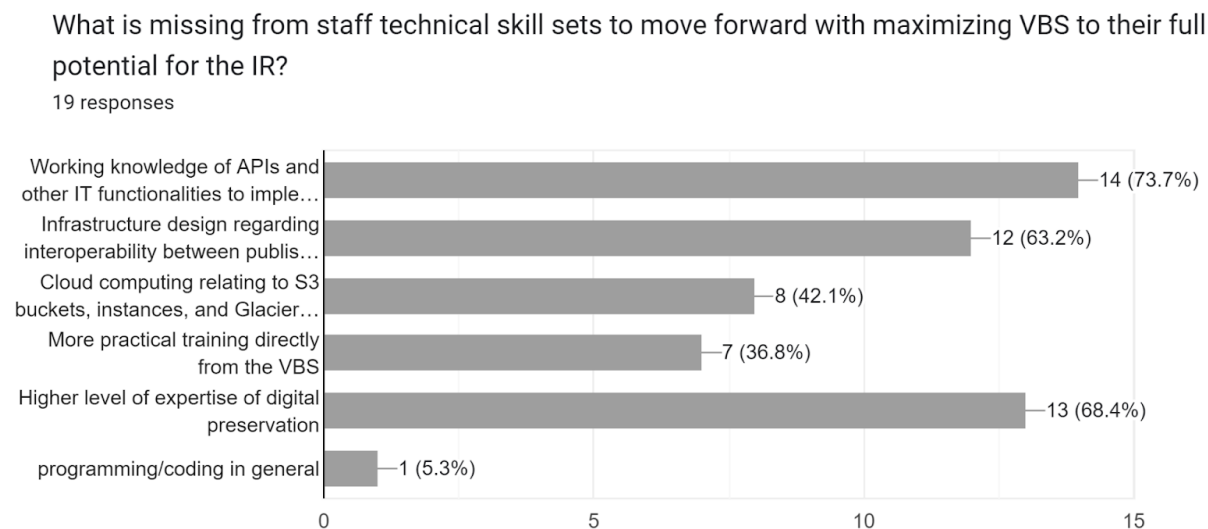


In the following question, five multiple choice options and one fill-in response were offered for the question about what the participants think is missing from staff technical skill sets to move forward with maximizing VBS to their full potential for the IR. Respondents were able to select all answers that applied and offer one of their own if desired. In the results, all but one respondent selected multiple answers and 15.75% selected all the skill set answer options. Of the R1 class, all but one selected “Working knowledge of APIs and other IT functionalities to implement workflows between publishing and preservation”. Additionally, 60% of the R1s

selected the remaining options with one selecting all options. Of the R2 class, all selected the option, “Infrastructure design regarding interoperability between publishing and preservation platforms” and an average of 50% chose the other options. And like R1, with one selecting all options. For the D/PU class, all selected “Working knowledge of APIs...” and “Higher level of expertise of digital preservation,” and zero selected “More practical training directly from the VBS.” One D/PU added another response, “programming/coding in general.” The other classes had similar responses, especially the Baccalaureate class, with one respondent selecting all the options (Figure 15).

Figure 15

Missing Technical Skills for Advancing VBS



Section 9: Other Factors

This section's answer options were based on interviews, other research from readings, as well as the author's internship. This resulted in six multiple choice options and one fill-in response for the question about what other factors prevent maximizing features of VBS for

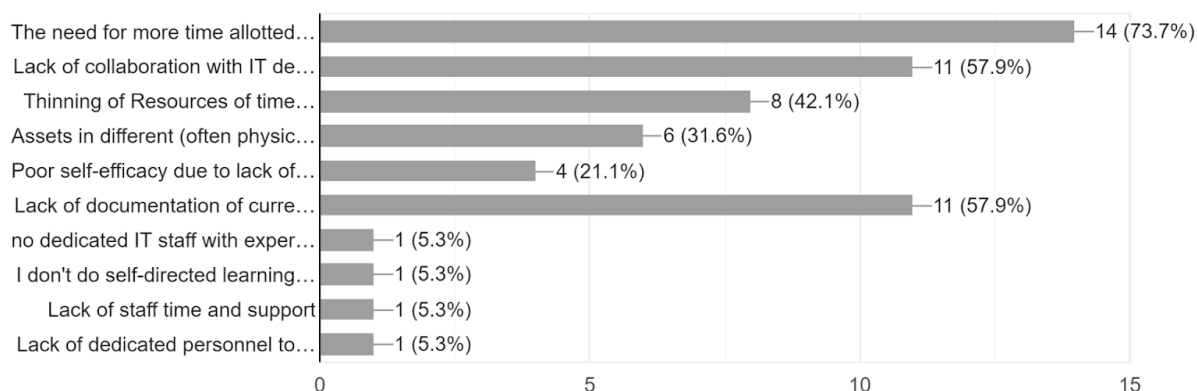
curation and preservation. Respondents were able to select all answers that applied and offer one or more of their own if desired. In the results, participants added four other answers, three of these referenced staff, time and support. The response rate (19) was like the previous question, but with an increase of similar responses across Carnegie classes. Where this differed was with the additional answers added (Figure 16).

Figure 16

Other Factors preventing VBS Maximization for Preservation

What other factors prevent maximizing features of VBS for curation and preservation? Select all that apply.

19 responses



Section 10: Stakeholders

Stakeholders’ understanding of the needs of the IR and resources required for stewardship of the digital assets held therein, is important for daily management of these assets and essential for long term preservation. In selecting which VBS to use, consideration for budgetary, technological, staff time, storage and IT requirements is paramount for stakeholders. Without this knowledge, strategies, and policies are not informed and will not succeed. The following three questions gauge stakeholders’ understanding and support levels from the

respondents' point of view. The question responses are on a linear scale 1-5, with one reflecting No and five reflecting Yes.

Results for the question “After a platform was chosen, were the requirements of ongoing management, implementation, and integration of the VBS understood by the stakeholders” were mixed ranging from no understanding to some understanding as the average response, across all Carnegie classes (Figure 17).

Figure 17

Stakeholder VBS Requirements Comprehension

After a platform was chosen, were the requirements of ongoing management, implementation, and integration of the VBS understood by the stakeholders?

19 responses

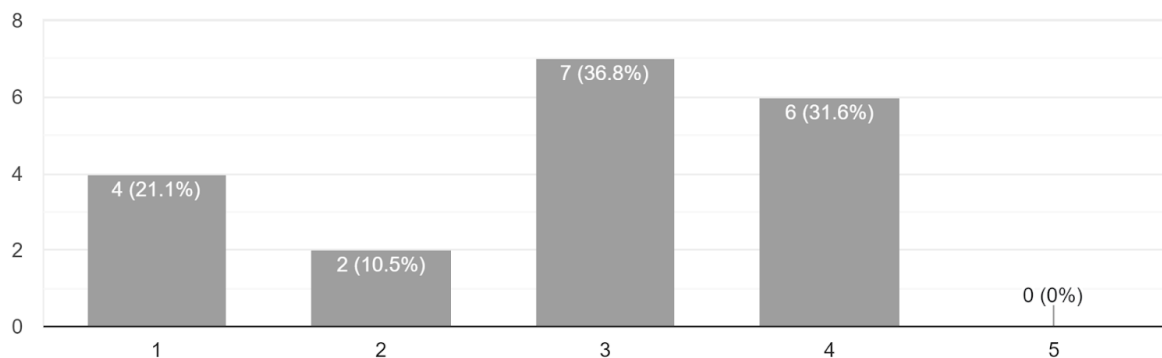
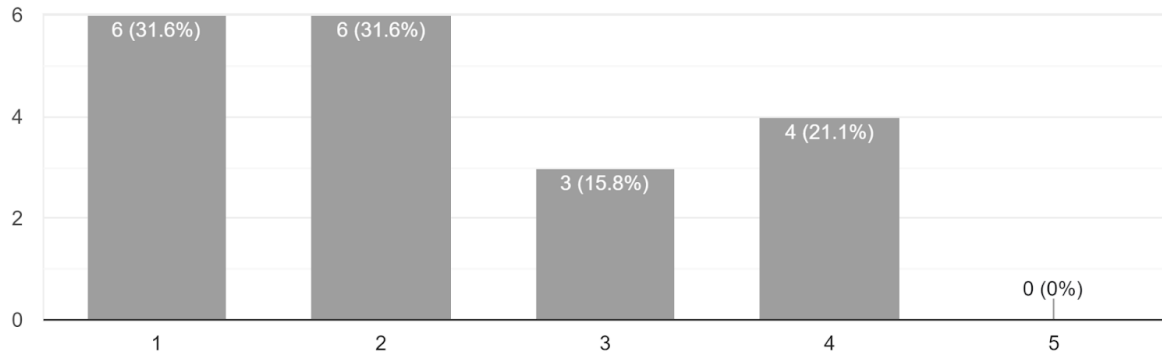


Figure 18*Stakeholder Comprehension of Digital Preservation*

Do stakeholders understand the complexities and requirements of digital preservation?

19 responses



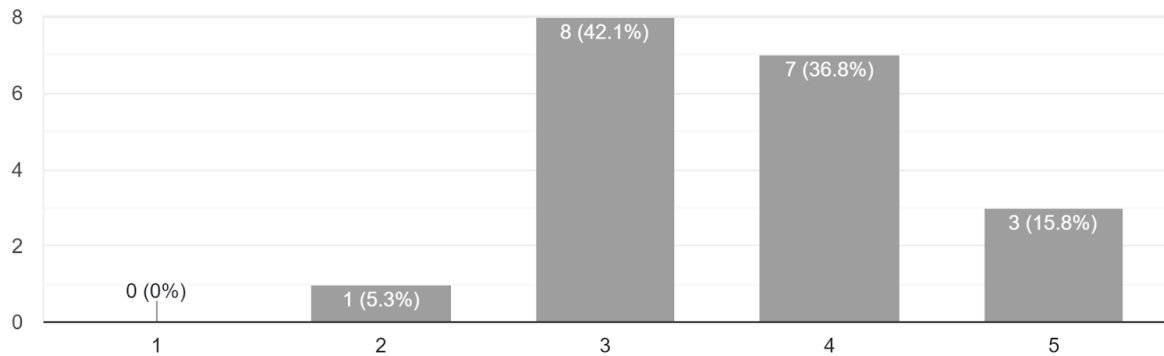
Regarding the stakeholders understanding the complexities and requirements of digital preservation, results leaned heavily in the negative with more understanding noted from R1 and M1 classes (Figure 18). Support for digital preservation was evident, as these responses had the highest positive numbers of all the stakeholder questions (Figure 19).

Figure 19

Stakeholder Support of Digital Preservation

Do stakeholders support digital preservation?

19 responses



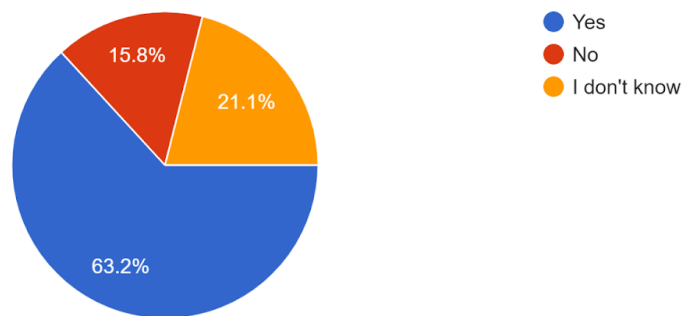
The last question of the survey dealt with a perceived need for organizational change in the infrastructure of the institution that would incorporate ongoing, not project based, digital preservation. The responses leaned to the affirmative with the highest average “yes” from R1, D/PU, and Baccalaureate Carnegie classes (Figure 20).

Figure 20

Need for Organizational Change Graph

Is there a need for organizational change to incorporate ongoing digital preservation (not project based but in architecture /infrastructure)?

19 responses



Recommendations

Start at the beginning. During the platform selection process use this questionnaire to help pilot the selection committee and serve as a starting place. The NASIGuide: Talking Points and Questions to ask Publishers about Digital Preservation (The Digital Preservation Committee, 2020).

Create a “Digital Preservation Working Group” at the IR to keep abreast of advancements and updates to VBSs used by your institution (Uzwysyn, 2021).

Or take a note from Virginia Tech: “Virginia Tech Libraries has developed a cloud-native, microservices-based digital libraries platform to consolidate diverse access and preservation infrastructure into a set of flexible, independent microservices in Amazon Web Services” (Tuttle et al., 2020).

Reconsider collaboration but on a larger scale. Even with interdepartmental collaboration and more time devoted to continued education there is a limit to resources. How long can this be sustained without major change? In a 2018 paper, Arlitsch, K., & Grant, C. discuss the concept of the “network effect” through “some kind of unification of IRs”. This would reduce costs, increase security, free up staff resources at the individual IR level and increase dissemination, sustainability as well as preservation actions. “The proliferation of individual repositories means that standards are often implemented a little differently from one repository to another, requiring crosswalks that lead to loss of granularity when attempts are made to aggregate content. Aggregators, such as Google Scholar (GS) for IR, or the Digital Public Library of America (DPLA) for cultural heritage repositories, often struggle to harvest and normalize metadata from disparate repositories where “standards” were applied inconsistently” (Arlitsch and Grant, 2018).

Small scale versions of this are happening with HathiTrust and AP Trust but “how can individual universities develop visions that are sufficiently shared that they can realize the economic and technical advantages in shared infrastructure that scales across institutional boundaries” (Schonfeld, 2018).

Final recommendation is to acknowledge that more time and resources are required to fulfill and maintain digital asset management and preservation workflows. By acknowledging, job descriptions can be reevaluated, and organizational change can be made to incorporate ongoing digital preservation.

Conclusion

Institutional repository librarians play a significant role in the management, dissemination, curation, and preservation of digital scholarly material using a wide range of platforms and systems. A valuable tool for these tasks is VBSs. However, as this paper demonstrates, there are major barriers prohibiting the extension of the IRs' digital assets they tend into digital preservation workflows.

The findings from surveying scholarly communication library professionals from six different Carnegie Classes have more similarities than differences. These issues cross over all the Carnegie Classes. There are some variances due to the size of the institutions and availability of resources, but many problems are universal. Users of two leaders in VBS publishing platforms, bepress Digital Commons and DSpace, answered questions on training, use, limitations, and challenges, as well as on workflows extended from the VBSs to preservation platforms. Overall, both had similar findings though DSpace was cited as having a steep learning curve getting started compared to responses of Digital Commons consultant onboarding training.

While these professionals are energetically involved with many digital curation organizations for training and continuing education, support is mixed from stakeholders, often citing a time issue. Again, the lack of time came up as a reason for the not activating use of extensions into preservation platforms, such as LOCKSS and CLOCKSS. These connections are readily available, according to bepress Digital Commons. DSpace users, while not having such a clear road map to these platforms, had a similar number of uses reported in the survey. Both platforms had issues reported connecting workflows to ORCID. Connectivity to cloud storage was considered by many survey participants to be a version of preservation and utilized by 39%.

Unfortunately, this is only a part of a digital preservation strategy. Most respondents reported having no digital preservation policy with only 21% having one. Few reported using the OAIS Reference Model and fewer NDSA levels of Preservation and only one using TRAC Checklist.

While these findings do not sound encouraging, many issues could be corrected by taking time to focus on preservation of the IR. Instituting a Digital Taskforce that analyzes the needs and goals of the individual institution with a goal of creating preservation policies and procedures, with documentation, would be a helpful move forward. Acknowledging the knowledge gap on digital preservation itself as well as requirements for third party IT collaboration (could be a vendor, another institution, or another university IT department) are foundational steps. “While a few inter-institutional collaborations already exist, there are many more connections that can be developed to benefit repository managers and the ecosystem of scholarly work housed within repositories” (Sterman, 2014).

But the largest barrier shown in this study is not just technological education, but also the need for more time allotted for professional development and more openness to collaboration with university IT departments to work through challenges with an interdisciplinary scope of players. Universities have technological expertise in different departments but connecting them to other departments and their specific needs is daunting. With a preservation policy these needs could be expressed and acted upon.

Again, it comes down to stakeholders and their understanding of the complexities of digital preservation and how their institution fits into the evolving IR landscape. If the university can utilize its resources toward this common goal, it would benefit through shared knowledge of its faculty and student scholarship reaching those in need of it.

References

- Arlitsch, K., & Grant, C. (2018). Why so many repositories? examining the limitations and possibilities of the Institutional Repositories Landscape. *Journal of Library Administration*, 58(3), 264–281. <https://doi.org/10.1080/01930826.2018.1436778>
- Bepress launches new service to address archiving and preservation needs.* bepress. (2017, January 26). <https://bepress.com/news/bepress-launches-new-service-address-archiving-preservation-needs/>
- bepress. (2022, October 26). *Digital Commons*. Digital Commons - bepress. <https://bepress.com/products/digital-commons/>
- Basic classification.* Carnegie Classification of Institutions of Higher Education. (2023, March 30). <https://carnegieclassifications.acenet.edu/carnegie-classification/classification-methodology/basic-classification/>
- Digital Preservation Coalition. (2023, March). Digital Preservation Community Mental Health and Wellbeing Survey questions. <https://www.dpconline.org/docs/working-groups-and-task-forces/wellbeing-task-force/2811-digital-preservation-mental-health-and-wellbeing-survey-2023/file>
- Digital Commons Consulting Services. (2023, January 17). *Safeguarding your content with Digital Commons.* bepress. Retrieved May 2, 2023, from https://bepress.com/reference_guide_dc/safeguarding-content-digital-commons
- Donohue, T. (2016, October 25). *Now available: Dspace 6.0–API enhancements and new features.* DSpace. <https://dspace.lyrasis.org/now-available-dspace-6-0-api-enhancements-and-new-features/>
- Giaretta, D. (2011). *Advanced Digital Preservation.* Springer- Verlag Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-16809-3>
- Hamill, L. (2019, June 11). *DSpace as an institutional repository.* CORE Scholar. https://corescholar.libraries.wright.edu/irday/spring_2019/june11/2
- Henry, D. B., & Neville, T. M. (2017). Repositories at master’s institutions: A census and analysis. *Library Resources & Technical Services*, 61(3), 124. <https://doi.org/10.5860/lrts.61n3.124>
- Lavoie, B. (2014, October). *Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition).* DPC Technology Watch Report 14-02. <https://www.dpconline.org/docs/technology-watch-reports/1359-dpctw14-02/file>

Lynch, C. A. (2003). "Institutional repositories: Essential infrastructure for scholarship in the digital age", *Portal Libraries Acad.*, vol. 3, no. 2, pp. 327-336.

Lyrasis. (n.d.-a). *User FAQ*. User FAQ - DSpace - LYRISIS Wiki.
<https://wiki.lyrasis.org/display/DSPACE/User+FAQ#UserFAQ-AboutDSpace>

Lyrasis. (n.d.-b). *User FAQ*. User FAQ - DSpace - LYRISIS Wiki.
<https://wiki.lyrasis.org/display/DSPACE/User+FAQ#UserFAQ-HowdoesDSpacepreservedigitalmaterial?>

Lyrasis. (2022a, October 27). *Technical specifications*. DSpace.
<https://dspace.lyrasis.org/technical-specifications/>

Lyrasis. (2022b, November 16). *Join the community*. DSpace. <https://dspace.lyrasis.org/join-the-community/>

Owens, T. (2018). Sixteen Guiding Digital Preservation Axioms. In *The theory and craft of digital preservation* (pp. 4). Johns Hopkins University Press.

Rieger, O. Y. (2018, October 31). *The state of Digital Preservation in 2018*. Ithaka S+R. Retrieved May 2, 2023, from <https://sr.ithaka.org/publications/the-state-of-digital-preservation-in-2018/>

Schonfeld, R. C. (2018, November 6). *Big deal: Should universities outsource more core research infrastructure?* Ithaka S+R. Retrieved May 2, 2023, from <https://sr.ithaka.org/publications/big-deal-research-infrastructure/>

Sterman, L. (2014). Institutional repositories: An analysis of trends and a proposed collaborative future. *College & Undergraduate Libraries*, 21(3-4), 360–376.
<https://doi.org/10.1080/10691316.2014.943919>

The Digital Preservation Committee. (2020, January). *NASIGuide: Talking points and questions to ask publishers about digital preservation*. NASIG.org. <https://nasig.org/Talking-Points-and-Questions-about-Digital-Preservation>

Tuttle, J., Chen, Y., Jiang, T., Hunter, L., Waldren, A., Ghosh, S., & Ingram, W. A. (2020). Multi-tenancy Cloud Access and preservation. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020*, 557–558.
<https://doi.org/10.1145/3383583.3398624>

- Uzwyshyn, R. (2021). *Creating infrastructures for long term digital preservation for libraries, museums and memory institutions*. Presented at the Internet Librarian 2021 Connect Conference. <https://digital.library.txstate.edu/handle/10877/14739>
- Ware, M. (2004). Institutional Repositories and scholarly publishing. *Learned Publishing*, 17(2), 115–124. <https://doi.org/10.1087/095315104322958490>
- Xia, J., & Opperman, D. B. (2010). Current trends in institutional repositories for institutions offering master's and baccalaureate degrees. *Serials Review*, 36(1), 10–18. <https://doi.org/10.1080/00987913.2010.10765272>

Appendix A

Survey

Section 1: Data Survey Scope

Survey Title:

The Institutional Repository use of Vendor-Based Solutions relating to technical knowledge and Digital Curation

The Main Question: What factors keep IRs (Institutional Repositories) from utilizing the features in vended solutions services more fully to perform digital preservation?

Questions:

Section 1

1. What is your job title?
2. How long have you been in your current role?
 - 0-2 years
 - 2-4 years
 - 4-7 years
 - 7-11 years
 - 11-15 years
 - 15+ years
3. What is your institution's Carnegie Classification?
 - R1: Doctoral Universities–Very high research activity
 - R2: Doctoral Universities–High research activity
 - D/PU: Doctoral/Professional Universities
 - M1: Master's Colleges and Universities
 - Baccalaureate Colleges
 - Baccalaureate/Associate's Colleges
 - Associate's Colleges
 - Special Focus Institutions
 - Tribal Colleges
 - Not Applicable
 - Other
4. Please share your organization's name to remove duplication from the results of this survey.

5. Are you (and/or your institution) a member of working groups for associations, or consortiums, organizations where you have access to education and training opportunities?
6. If yes, which is most beneficial?
7. In your work environment, is service to non-profit associations (through board positions, working groups, task force, etc.) that perpetuates the care and preservation of cultural heritage through digital curation encouraged? To what level?
 - Linear scale 1-10
 - 1 - Not encouraged - no time for this
 - 10- Encouraged and considered part of your job description

Section 2: Publishing Platforms

8. Do you use a Publishing Platform? Select all that apply.
 - CONTENTdm
 - Dataverse
 - Digital Commons (bepress)
 - DSpace - DuraSpace (Lyris)
 - EPrints
 - eScholarship
 - Esploro (Ex Libris)
 - figshare
 - Fulcrum
 - Islandora
 - Janeway
 - LIBNOVA
 - Manifold
 - Omeka
 - Open Conferencing System (OCS)
 - Open Journal System (OJS)
 - Open Monograph Press (OMP)
 - PressBooks
 - PubPub
 - Samvera - Hydra - Fedora
 - WordPress
 - Ubiquity Repositories
 - in-house or locally developed software
 - Other
9. Are you the system administrator or manager?
 - Administrator
 - Manager

- Editor
 - Authorized user
 - Other
10. Considering your institution's digital strategy, is the IR publishing platform the end of the line for the digital assets therein? Or part of a larger workflow?
- The end
 - Part of a larger workflow
 - Do not have a digital strategy

Section 3: Digital Preservation Platforms

11. What digital preservation standards does your repository follow?
- OAIS Reference Model (ISO 14721)
 - TRAC (ISO 16363) Trusted Digital Repository (TDR) Checklist
 - NDSA Levels of Preservation
 - Other
12. Is there a digital preservation policy at your IR?
- Yes
 - No in process of being created
13. What platform(s) is used for digital preservation? Select all that apply.
- Amazon Glacier
 - Amazon S3
 - Ap Trust
 - Archive-IT
 - Archivemata
 - ArchivesSpace
 - Arkivum
 - Chronopolis
 - CLOCKSS
 - COPPUL
 - DPN
 - Drupal
 - DuraCloud
 - ESRI (ArcGIS Online)
 - HathiTrust
 - Internet Archive
 - LibGuides
 - LOCKSS
 - MetaArchive
 - Portico
 - Preservica

- Public Knowledge Project Preservation Network (PKP-PN)
- Rosetta (Ex Libris)
- Scholars Portal
- digital preservation services under discussion
- in-house or locally developed software
- Other

14. Are you the system administrator or manager?

- Administrator
- Manager
- Editor
- Authorized user
- Other

Section 4: Digital Commons (bepress)

15. Is your IR using Digital Commons?

- Yes - skip to Section 6
- No

Section 5: DSpace

16. Is your IR using DSpace?

- Yes - skip to Section 7
- No - skip to Section 8

Section 6: Digital Commons (bepress) Questions

17. How long has your IR been using Digital Commons?

- 1-2 years
- 3-4 years
- 5-6 years
- 7-8 years
- 9-10 years
- Longer than 11 years

18. Do you have a contract?

- Yes
- No

19. Are system updates/upgrades automatic or performed manually?

- Automatic
- Manual

20. What training did you (are you) receive(ing)? Self-trained?

21. Does your IR use the bepress Archive feature?

- Yes
 - No
 - I don't know
22. Are you using any Digital Commons Publishing Services? If yes - which ones?
23. What digital curation/preservation workflows have been developed around Digital Commons?
24. What other systems/platforms have been integrated with Digital Commons?
25. How do you see Digital Commons integrating with the concepts and implementation of Open Access (OA) and Open Educational Resources (OER)?
26. What challenges/limitations specific to Digital Commons have you experienced?
27. How satisfied are you with Digital Commons community support? Is it used widely to be active and helpful?
- Linear scale 1-5
 - 1 - Not Satisfied
 - 5 - Very Satisfied
28. How satisfied are you with Digital Commons account manager/team support?
- Linear scale 1-5
 - 1 - Not Satisfied
 - 5 - Very Satisfied
29. Do you see Digital Commons as a long-term solution? Why? Why not?
30. Do you have an exit strategy for Digital Commons?
- Yes
 - No
 - I don't know

Section 7: DSpace Questions

31. How long has your IR been using DSpace?
- 1-2 years
 - 3-4 years
 - 5-6 years
 - 7-8 years
 - 9-10 years
 - Longer than 11 years
32. Are system updates/upgrades automatic or performed manually?
- Automatic
 - Manual
33. Do you have a contract?
- Yes
 - No
34. What training did you (are you) receive(ing)? Self-trained?

35. Does your IR use the S3BitStore cloud storage option?
- Yes
 - No
 - I don't know
36. What digital curation/preservation workflows have been developed around DSpace?
37. What other systems/platforms have been integrated with DSpace?
38. How do you see DSpace integrating with the concepts and implementation of Open Access (OA) and Open Educational Resources (OER)?
39. What challenges/limitations specific to DSpace have you experienced?
40. How satisfied are you with DSpace community support? Is it used widely to be active and helpful?
- Linear scale 1-5
 - 1 - Not Satisfied
 - 5 - Very Satisfied
41. How satisfied are you with DSpace account manager/team support?
- Linear scale 1-5
 - 1 - Not Satisfied
 - 5 - Very Satisfied
42. Do you see DSpace as a long-term solution? Why? Why not?
43. Do you have an exit strategy for DSpace?
- Yes
 - No
 - I don't know

Section 8: Technical Skills

44. Self-assessment: please rate your technological skills in relation to VBS systems management and digital preservation.
- Linear scale 1-10
 - 1 - Low - limited functionality
 - 10 - High - Complete understanding
45. What is missing from staff technical skill sets to move forward with maximizing VBS to their full potential for the IR?
- Working knowledge of APIs and other IT functionalities to implement workflows between publishing and preservation
 - Infrastructure design regarding interoperability between publishing and preservation platforms
 - Cloud computing relating to S3 buckets, instances, and Glacier storage
 - More practical training directly from the VBS
 - Higher level of expertise of digital preservation
 - Other

Section 9: Other Factors

46. What other factors prevent maximizing features of VBS for curation and preservation?

Select all that apply.

- The need for more time allotted for professional development - instead of learning as needs arise.
- Lack of collaboration with IT departments with wider knowledge to work with to accomplish IR long term goals
- Thinning of Resources of time and focus - too many organizational memberships not delivering stronger benefits
- Assets in different (often physical) locations/platforms that overwhelm curation/management planning
- Poor self-efficacy due to lack of hands-on experience in digital preservation
- Lack of documentation of current policies and procedures within the organization
- Other

Section 10: Stakeholders

47. After a platform was chosen, were the requirements of ongoing management, implementation, and integration of the VBS understood by the stakeholders?

- Linear Scale 1-5
- 1 - No
- 5 - Yes

48. Do stakeholders understand the complexities and requirements of digital preservation?

- Linear Scale 1-5
- 1 - No
- 5 - Yes

49. Do stakeholders support digital preservation?

- Linear Scale 1-5
- 1 - No
- 5 - Yes

50. Is there a need for organizational change to incorporate ongoing digital preservation (not project based but in architecture /infrastructure)?

- Yes
- No
- I don't know

Section 11: Follow Up Opportunity

51. Are you open to answering additional interview questions about your IR and VBSs? Or do you have additional thoughts on this research topic? Please share your preferred contact email or phone information below so I may reach out to you directly.

52. Thank you for your efforts in completing this survey. Time is a valuable resource and I am grateful for your investment in this process.