Where Are We Now? Survey on Rates of Faculty Self-Deposit in Institutional Repositories

Ruth Kitchin Tillman

Where Are We Now? Survey on Rates of Faculty Self-Deposit in Institutional Repositories

Ruth Kitchin Tillman

Linked Data Strategist, Penn State University Libraries, Penn State University

INTRODUCTION The literature of institutional repositories generally indicates that faculty do not self-deposit, but there is a gap in the research of reported self-deposit numbers that might indicate how widespread and common this is. **METHODS** This study was conducted using a survey instrument that requested information about whether a repository allowed self-deposit and what its rates of self-deposit were, if known. The instrument contained additional questions intended to gather a broader context of repositories to be examined for any correlations with higher rates of self-deposit. It also included questions about the kinds of labor required to populate an IR as well as satisfaction with the rates of self-deposit. **RESULTS** Of 82 respondents, 80 were deemed to fall within the study's parameters. Of these, 55 respondents' institutions allowed self-deposit, and 10 reported rates of self-deposit of more than 20 items per month. More than half the total respondents reported using at least three methods other than relying on self-deposit to add content to their repository. Respondents are generally unsatisfied with their deposit profiles, including one at a school reporting the highest rate of self-deposit. **DISCUSSION** From the responses, no profile could be formed of respondents reporting high rates of self-deposit that did not entirely overlap with many others reporting little or no self-deposit. However, the survey identifies factors without which high rates are unlikely. **CONCLUSION** The results of this survey may be most useful as a factor in administrative prioritizations and expectations regarding institutional repositories as sites of scholarly self-deposit.

IMPLICATIONS FOR PRACTICE

1. This study supports the general understanding in the community of people who maintain IRs that, despite outreach, few faculty self-deposit anywhere.

2. When attempting future projects that require broad faculty buy-in, librarians and administrators should ask themselves whether they are repeating assumptions made about IR buy-in.

3. The levels of respondent satisfaction with repository self-deposit and number of additional methods being used to fill repositories present opportunities for research about job satisfaction, emotional well-being, and retention in librarians tasked with filling repositories.

INTRODUCTION

Fourteen years since Clifford Lynch's (2003) assertion that institutional repositories (IRs) were essential infrastructure for 21st-century scholarship, where are we? Has the institutional repository transformed how faculty produce scholarship? Are faculty eagerly depositing their work, or are they at least interested once the repository has been "seeded" with enough articles to attract their attention? Will the next pilot project or the next outreach session turn things around? Can we catch them as graduate students and hold their attention?

We certainly have a broader software landscape for these repositories, from early standards like EPrints and DSpace to Hyrax (initially known as Hydra-in-a-Box). They may be hosted by a big name like OCLC, spun up by a single systems administrator in an afternoon, or run by an entire team of developers. Each repository fulfills basic functions of associating metadata and articles and each has known flaws, which one may learn by listening to the conversations of metadata librarians gathered at conferences or on Slack channels. None has proved a panacea.

The literature of institutional repositories contains such sanguine titles as "Recruiting content for institutional repositories: The barriers exceed the benefits" (Covey, 2011), "It takes more than a mandate" (Zhang, Boock, & Wirth, 2015), and, most memorably, Dorothea Salo's "Innkeeper at the roach motel" (2008). Salo's assessment of the flaws in repository strategies as then practiced—high expectations, understaffing, bad user interfaces, perceived lack of value in intended user communities—and her recommendations for next steps—partnership with campus IT, streamlining design to erect fewer barriers, focus on working with user communities who do find value—remain relevant nine years later. Echoes of her words can be seen in both Covey and Zhang et al. Richard Poynder (2016) referenced it in the preliminaries to his 2016 Q&A with Clifford Lynch about rethinking the institutional repository.

Indeed, Salo's article was one of the motivators for the research conducted through this survey instrument and on which this article is based. Perhaps best-known among those library workers most intimately connected to the work of repositories (and perhaps sought out as proof that they were neither failures nor alone), the article has not received the recognition among administrators that it deserves. Understandably, few administrators would want to classify the investment of time, money, and labor made on their watch as a "roach motel."

Lynch (2003) asserted that faculty were wasting time running their own websites, and that in a successful IR, one could easily deposit a work and be done with it. In 2007, Xia and Sun set out more explicit assessment criteria for evaluating the success of an institutional repository: a) are full-text copies of works (as compared with just citations) being deposited? and b) are the depositors also the authors? When speaking of a successful repository or deposit profile, this article focuses primarily on the second criterion—are the faculty themselves depositing the works?

The primary goal of this research and article, then, was to gather data about the landscape of faculty interactions with repositories, to test the hypothesis that almost no repository has what could be considered a healthy or "sustainable" self-deposit profile (Carr & Brody, 2007), and to examine the shared features of those few exceptions to the norm. It was to expose what the workers closest to the work already know about this landscape and give them numbers to share with administrators. And it was conducted in the hope of changing expectations and prompting reevaluations of the purposes and uses of institutional repositories for those who have not already begun the work. Many institutions have already begun this process. Indeed, the author hopes that her work may serve mostly serve as a retrospective evaluation of a generally failed proposition.

LITERATURE REVIEW

A common theme runs through the literature surrounding institutional repositories: faculty do not deposit their works in them. Much of the literature focuses on a particular institution's efforts to understand or change these behaviors. Only occasionally do such articles take space beyond the literature review to survey repositories other than their own.

In 2007, Carr and Brody conducted a survey of deposit behavior in repositories worldwide that contained 10,000 or more objects. They used Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH) extracts to determine the dates and frequency of deposit to determine whether objects were deposited in large batches (suggesting work by the IR's staff) or whether they were deposited in small numbers continually. Only three of the

twenty repositories they presented in their results demonstrated what they, referencing Xia and Sun (2007), considered patterns for sustainable success. However, those with successful deposit patterns were not contacted for any kind of follow-up to determine what local practices or factors might have contributed to these deposits.

Concerned by the lack of deposit in Cornell's IR, Davis & Connolly (2007) used log files to evaluate local deposit patterns as well as those at seven other DSpace repositories. The surveys showed that regular patterns of deposit developed only where enforced policies demanded deposit of assets such as electronic theses and dissertations, or collections coming from the library itself.

Xia conducted a 2007 assessment of four disciplines in seven UK and Australian IRs and found that much of the deposit was performed by engaged liaisons (including departmental staff), even when a culture of self-deposit in disciplinary repositories existed. Xia's 2008 follow-up focused on a single repository and found that many authors who participated heavily in disciplinary repositories did not deposit their own papers in their own depository, although liaisons might do so on their behalf.

Vincent-Lamarre, Boivin, Gargouri, Larivière, and Harnad (2015) found "a small but significant positive correlation" between percentage of Web of Science (WoS)-indexed material deposited by faculty and open access (OA) policies that included the following three points: (a) immediate deposit required; (b) deposit required for performance evaluation; and (c) unconditional opt-out allowed for the OA requirement but no opt-out for deposit requirement (p. 3). Because the 67 institutions they studied were crawled, not surveyed, researchers were unable to gather information about the extent to which these deposits were facilitated by those maintaining the repository.

The Vincent-Lamarre et al. article introduces a new and hopeful factor in the history of institutional repositories, that of campus-wide OA mandates that either require deposit or at least maintain the right of the university to deposit their faculty's work in an IR. In a 2004–2005 Australian survey, Sale (2006) had found a significant correlation of deposit with mandates *and author support*. Yet when Zhang et al. (2015) evaluated the success of Oregon State University's 2013 OA mandate, they found that in practice only 43% of articles indexed in WoS after the mandate could be found in the IR (p. 13). This number is almost identical to the percentage of university output that Sale (2006) reports as captured by the most supportive and successful IR. His graph bar shows percentages in the low 40s for 2004 and 40% for 2005. Sale does not indicate who did the depositing, although Xia's 2008 study suggests it may have been staff/liaisons. However, at OSU, this required active work from OSU library and press staff, harvesting WoS data and sending individual

requests to faculty for authors' accepted manuscripts. As their title indicates, "it takes more than a mandate."

Similarly, Neugebauer and Murray (2013) evaluated four years of deposit at Concordia University, before and after the passing of an OA mandate. They conclude that "good intentions, altruism and a sympathetic stance with regards to open access and its ideals have not yet generated enough voluntary deposits compared to the total number of research articles published per year" (p. 100). Even with mandates and library support, only those who control promotion and financial support, that is, the universities and funders, have the power to enforce such policies.

From these articles, one may conclude that the one of the most important aspects of an open access mandate is the clarification for institutional repository managers on whether or not they have the right to deposit a work. As the literature surveying faculty indicates, "the barriers [often still] outweigh the benefits" when attempting to convince faculty to deposit their work (Covey, 2011). Faculty may not understand how to deposit, see no utility in depositing their article in a local repository, and even consider it a risk (Covey, 2011; Casey, 2012; Zheng & Li, 2015). Even in fields whose competitive practices lead to a culture of preprint deposit, this takes place in centralized disciplinary repositories rather than local IRs (Fry, Spezi, Probets, & Creaser, 2015).

Authors who have previously conducted broad surveys collected their research data through surveys without any kind of human mediation, for instance by crawling publicly available repository data. What the literature lacked was a survey conducted on a broad scale, addressing both quantitative and qualitative factors. While those analyzing OAI-PMH fields could, through additional research, obtain data about the institution and the age of the IR, their research relies on deductions made about self-deposit numbers, based on whether or not records had been created from apparent batch loads on a handful of days or regularly over time (Carr & Brody, 2007). This inference relies on a potentially flawed assumption of workflow, namely that the data was being batch loaded instead of hand-keyed by those responsible for the IR. Aspects such as workflow and perceptions of the work are most commonly addressed in case studies, such as Neugebauer & Murray (2013) and Zheng & Li (2015), and have never before been paired with broader survey data.

METHODS

Design of the Survey Instrument

The survey instrument was designed to fill a particular gap in the research to date. This instru-

ment sought to capture data about (a) the institution itself; (b) the environment of the IR, such as its age, software type, types of materials supported, support of self-deposit, and types of approved users; (c) actual self-deposit numbers and refining details if known; (d) other methods of adding materials to the repository, types of outreach and support conducted by those working on the IR, and satisfaction levels with the current situation; and (e) attempts to capture the scope of labor required in maintaining an IR, particularly when self-deposits are low and it must be filled through other means.

Unlike research data collected through crawling, research data created through this instrument relied on human knowledge of the repository and the understanding of the respondent (or respondents if the institution's response was collaborative) of activities taking place in support of the IR. In particular, it relied on the respondent's interpretation of questions on behaviors such as faculty self-deposit. The respondent may not have had access to information such as whether a faculty member deputized a graduate student to deposit their materials or a subject librarian sat down with the faculty member to arrange deposits without the respondent's knowledge.

As with all voluntary surveys, it allowed respondents to self-select. Only questions about institution type and size, whether or not the IR allowed for self-deposit, and the number of monthly faculty self-deposits (or a statement that this number was unknown or could not be determined) were required. Thus, while most individuals completed every question, some did not. Whether they did not have the data necessary to complete certain questions—a person responsible for a repository's metrics may not have a good idea of the kinds of outreach conducted—or simply chose not to is unknown. However, because so little broad data on both the environment of institutional repositories and on library response has been collected, these factors were considered an acceptable risk in pursuit of broader data. Association of Research Libraries (ARL) SPEC kits are subject to similar inherent risks, as noted by the Library Loon (2017).

Because of the differences in higher education in the United States and other countries, the study limits itself to the United States. Those interested in the European practices may find Creaser et al. (2010); Spezi, Fry, Creaser, Probetz, & White (2013); and Más-Bleda, Thelwall, Kousha, & Aguillo (2014) of interest. Additionally, because the focus of the study was on faculty, the study limits itself to institutes of higher education.

Distribution

The instrument invited responses from persons self-identifying as responsible for bringing content into an institutional repository or responsible for the metrics of an institutional repository. It was distributed to listservs for scholarly communication, metadata librarians, and the code4lib community. The first distribution occurred on the day the survey opened.

A reminder email was sent in the last week before the survey closed. The survey remained open for six weeks in June and July of 2016.

Known Issues

A flaw in this instrument was the lack of a question regarding whether the institution had an OA and whether said mandate required deposit in the institutional repository. The author had miscategorized OA mandates as being not entirely connected to library effort, although often championed by librarians, and thus not a good fit for the qualitative section.

However, although a campus-wide OA mandate is not necessarily connected to the library's work, as addressed in the qualitative section, it could be considered as much a feature of the institutional profile as its faculty size. Having that data alongside other data could have proved useful, particularly for future explorations of the subject. It is the author's hope that others will take up the subject and conduct targeted research on OA mandates and deposit, perhaps expanding on the work of Vincent-Lamarre et al. (2015) with surveys that address impact of these mandates on the labor of library workers responsible for the IR.

A question intended to separate 2-year, primarily community, colleges from colleges with at least a BA program and public/private schools caused some confusion. The wording it used to make this distinction was not necessarily in line with qualities the respondent considered fundamental characteristics of the institution. An R1 institution may also be a "public 4-year college and university," but it may think of itself as the former. Either including more questions or allowing a space for respondents to share what they considered important might have provided particular insights into the outliers with comparatively successful rates of self-deposit. A few textual responses in "Other" provide a glimpse of what might have been.

Additionally, a typographical error was made on certain questions determining ranges, where such ranges were indicated by such statements as 1–20 and 20–50. As these were broad ranges, respondents can be expected to have replied within the spirit of the range, and yet it is important to acknowledge that the error was made.

Research Questions

The breadth of data collected allows for the examination of a wide variety of correlations, and other researchers may find the published dataset useful for answering their own questions. The research questions this survey sought to address were the following:

- R1. What percentage of respondents receive 0, under 20, over 20, over 50, and over 100 faculty self-deposits per month? i.e. what is the general landscape of faculty self-deposit?
- R2. Is there any correlation between faculty self-deposit rate and faculty size, institution type, or the combination thereof?
- R3. Is there any correlation between faculty self-deposit rate and the age of the institutional repository?
- R4. Is there any correlation between faculty self-deposit rate and the software on which the repository is built?
- R5. (a) What other methods are most commonly used to add materials to a repository and (b) do any have high correlations with repositories whose rates of faculty self-deposit are over 20 articles per month?
- R6. What are satisfaction levels with the rate of faculty self-deposit? Does the actual rate of self-deposit make any impact on the level of satisfaction expressed?

The instrument included two questions about the institution and URL for the IR, scrubbed from all released data, which served as an attempt to eliminate multiple submissions from a single IR. The proposed method determined for analyzing data in such situations had been to extract and analyze that data, note any differences in the quantitative section, and more closely examine qualitative differences, as individual perceptions of the work around repositories and their success often vary at the same institution. However, upon completion of the survey, 70 of the 80 respondents completed the deduplication fields, and no repository was represented twice in the data.

RESULTS

The survey received 82 responses from institutes of higher education within the United States. Using institution name, submitted for deduplication, the author removed two responses whose institutions did not meet this qualification, leaving 80. One was in Europe, not the United States. The other was a government agency. Of the respondents, 55 answered "yes" to the required question "Does your institution allow faculty to self-deposit materials?" The results are drawn from these 55 responses. R1 and R5 include respondents who allow self-deposit but cannot determine or do not collect information about it.

Data

http://doi.org/doi:10.7274/R08K771B

R1. What percentage of respondents receive 0, under 20, over 20, over 50, and over 100 faculty self-deposits per month?

This question allowed two nonnumeric answers for repositories with self-deposit enabled: "We collect information about self-deposit rates but cannot determine how many come from faculty" and "We do not collect information about self-deposit rates."
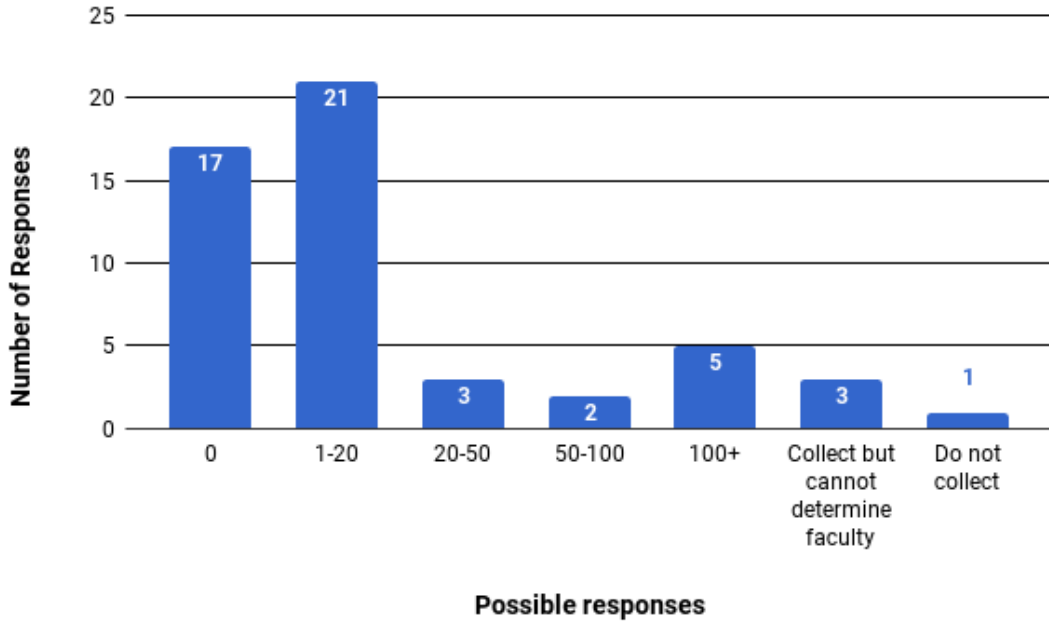

Figure 1. Monthly Rates of Self-Deposit and Number of Responses (52 Respondents)

A total of 10 respondents reported that they could determine that at least 20 self-deposits monthly came from faculty, with only seven of those over 50 and five over 100. The majority reported either no faculty self-deposits or under 20 faculty self-deposits.

R2. Is there any correlation between faculty self-deposit rate and faculty size, institution type, or the combination thereof?

The following table shows the monthly rate of self-deposit for each combination of faculty size and institution type that had at least one response.

	0	1–20	20–50	50–100	100+
Private 4-Year, < 500	4				
Private 4-Year, 500–1,500	2				
Private 4-year, > 1,500	1				1
Public 4-Year, < 500	3				
Public 4-Year, 500–1,500	3	3			
Public 4-year, > 1,500	2	7	3	1	4
Public 2-Year, 500–1,500		1			
Other, < 500	1	1			
Other, 500–1,500	1				
Other, > 1,500		2		1	

Table 1. Institutional Profile and Rates of Faculty Self-Deposit

The row that stands out in this table is the public four-year institution with more than 1,500 faculty members. Not only does this row represent four of the five institutions that report faculty self-deposit rates of over 100 items a month, it also represents all three that report 20–50, and one of the two that report 50–100. The other two rows of note are private four-year institutions and "other," both with over 1,500 faculty members, which contain the other responses of 100+ self-deposits and 50–100 self-deposits, respectively.

However, one must note that for each of these three rows, as many or more responses fall into the categories on the lower end. Ultimately, this represents correlation, not definitive conclusion. This is when additional data about the type of institution—research, liberal arts, etc.—would have been useful for clarification.

One of the five repositories reporting 100+ self-deposits declined to answer nonmandatory questions, leading to only four respondents in that column for subsequent research questions.

R3. Is there any correlation between faculty self-deposit rate and the age of the institutional repository?

This research question seeks to address the cumulative effect of existence and outreach on an institutional repository. The question on age made it clear that these were to be considered years after the repository was publicly available and promoted, not its developmental stages. It did presume that if self-deposit was available, it was upon time of public release and promotion. Is it reasonable for managers of institutional repositories (and the administrations) to expect successes to occur early on? The answer here is a resounding "no."

	0	1–20	20–50	50–100	100+
< 2 years	6	3	0	0	0
2-5 years	5	6	0	1	0
> 5 years	6	10	3	1	4

Table 2. Age of Institutional Repository and Rate of Faculty Self-Deposit

R4. Is there any correlation between faculty self-deposit rate and the software on which the repository is built?

The software on which an institutional repository can be built comes in a variety of flavors, from hosted and "turnkey," which can be set up in a matter of days, to highly customizable software, which requires extensive development. Any correlations between type of repository software used and rates of self-deposit may occur for a number of reasons.

First, the ease with which a faculty member may deposit items will have an effect on that faculty member's desire to deposit items in the future. Second, the features or lack thereof that the faculty member experiences as a website user may influence their decision to deposit. Third, the amount of time and money spent on repository software may reflect the administration's overall investment in the project and whether they expect the repository to simply start working (turnkey) or consider it a longer project with trial and error and investment.

	0	1–20	20–50	50–100	100+
CONTENTdm	0	1	0	0	0
Digital Commons (bepress)	8	6	0	0	0
DSpace	8	8	2	1	2
EPrints	0	0	0	0	0
Fedora (Hydra)	0	3	0	0	0
Fedora (Islandora)	0	0	0	1	0
Fedora (custom build)	1	1	1	0	0
Vivo	0	0	0	0	0
Other	0	1	0	0	2

Table 3. Institutional Repository Software and Rate of Faculty Self-Deposit

The nine institutions that report deposits of more than 20 articles per month use either DSpace, Islandora, or "Other," not including one institution that reported over 100 faculty self-deposits monthly, which declined to answer this question. In the category "Other," textual responses clarified that two were locally developed solutions and one a locally customized Fedora instance. This answer was changed in Table 3 to "Fedora (custom build)" in the table but left as-is in the raw data.

R5. (a) What other methods are most commonly used to add materials to a repository and (b) do any have high correlations with repositories whose rates of faculty self-deposit are over 20 articles per month?

Figure 2 shows the methods used to find article citations that are added to the repository, among all respondents, including those whose systems do not allow for self-deposit. Individuals performing the deposit for faculty members and working from faculty CVs and websites proved the most common. Of all respondents, 73.75% of those who answered this question used two or more additional methods. Over 50% used three or more, and 5% reported using all six.
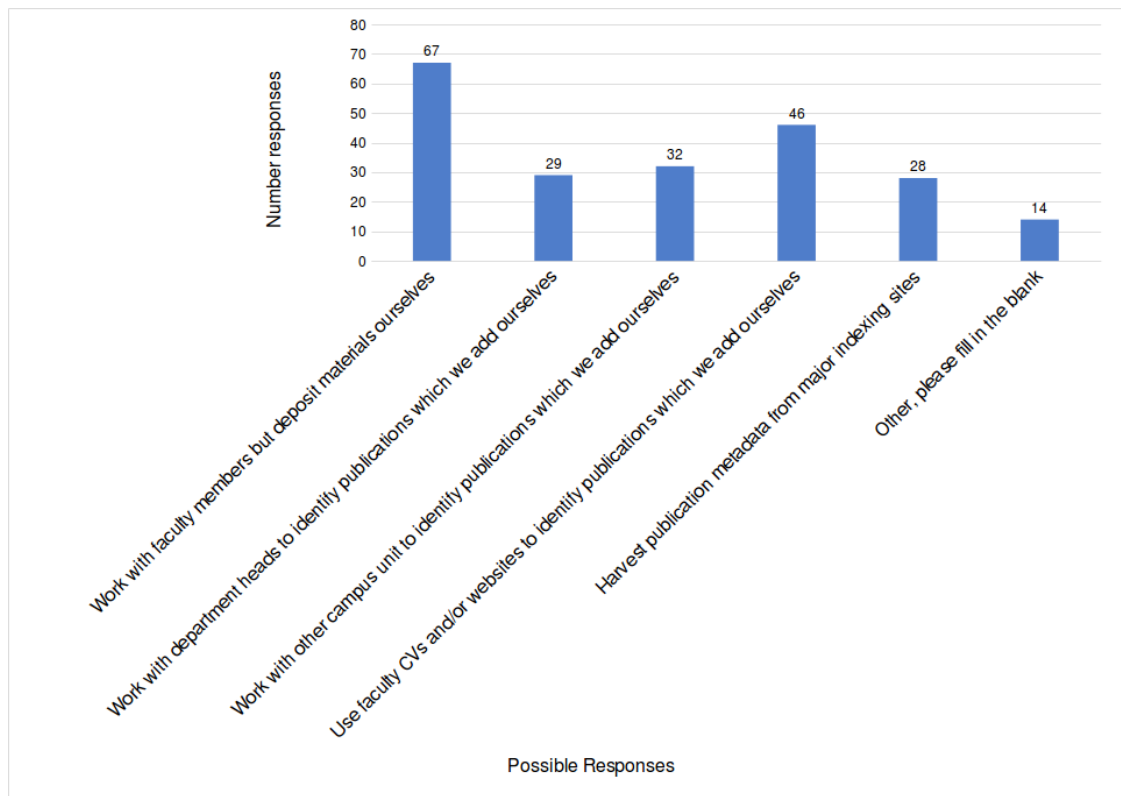

Figure 2. Alternative Sources of Deposit

Answers of "Other" included having liaison librarians or administrative professionals do the work, using alerts and active searches, harvesting OA-only, and using ORCID IDs to sync and harvest. Table 4 shows data for just repositories that report receiving over 20 faculty self-deposits per month. Perhaps more important than the numbers here is the number of additional methods for adding materials used by even those receiving robust deposits.

	20–50	50–100	100+	Total
Work with faculty members but deposit materials ourselves	1	1	3	5
Work with department heads to identify publications that we add ourselves	0	2	2	4
Work with other campus unit to identify publications that we add ourselves	0	0	2	2
Use faculty CVs and/or websites to identify publications that we add ourselves	1	1	2	4
Harvest publication metadata from major indexing sites	2	0	2	4
Other, please fill in the blank	1	0	2	3

Table 4. Other Methods Used to Deposit and Rate of Faculty Self-Deposit

R6. What are the IR staff satisfaction levels with the rate of faculty self-deposit? Does the actual rate of self-deposit make any impact on the level of satisfaction expressed?

Unlike the other research questions, which were formulated in an attempt to gather information about the landscape of self-deposit and then identify commonalities between respondents whose rates might be considered "successful," this question addressed outcomes for the programs and staff working on those repositories.

	0	1–20	20–50	50–100	100+
Completely satisfied	1	1	0	1	0
Satisfied but planning further outreach	1	4	1	0	2
Needs outreach to be satisfactory	4	8	0	1	1
Dissatisfied	5	6	1	0	1
Extremely dissatisfied	4	2	1	0	0

Table 5. Reported Rates of Satisfaction and Rate of Faculty Self-Deposit

DISCUSSION

As the results of the survey demonstrate, the short answer to the question "is our faculty depositing?" is "not really," or the even more straightforward "no." Of the 55 respondents, 45 answered that they either received fewer than 20 faculty self-deposits a month or could not tell how many they received. Should an administrator perceive their institution as lagging behind peers because of low self-deposit rates, these numbers tell a different story. Everyone is trying. Few are succeeding.

Effects of Time and Outreach

It appears to take a minimum of two years on average for repositories to have even a 50% likelihood of getting at least a single self-deposit per month, with a much greater likelihood of success after five years. This time period allows for the IR to become an established entity on campus and for responsible parties to do a variety of outreach and build new strategies after failures in the > 2-year and 2–5 year ranges. However, as 66.67% of repositories that had existed for at least five years *still* had self-deposit rates of 20 items or fewer and a full 25% had 0 average monthly self-deposits, the comparatively positive correlation of success and age should not be considered a guarantee. Age correlates even more strongly with failure.

An additional component of the question about a repository's age is the presumption of outreach conducted during those years. Of the 52 institutions that answered the question about age, only six reported not currently conducting outreach to faculty. Almost everyone is doing some kind of outreach, whether or not they see a return of self-deposit on it. Of the six who discontinued it, they all received either 0 or 1–20 deposits per month, but their ages ranged from one repository that had been public for less than two years to three that had been public for five or more years. Of those, three had conducted outreach in the past and reported it was discontinued because it was "no longer considered a priority."

Time and outreach correlate strongly with both a strong deposit profile and the lack thereof. No conclusions, therefore, can be drawn about either in general.

Software

Institutions reporting higher rates of self-deposit are more likely to report the use of either a homegrown system or one that involves high levels of developer time and engagement. This factor may indicate that the institution is investing heavily in the repository, implying that it allots more staff time and effort for other activities that promote deposit. It may also indicate that the user interface for deposit in turnkey models does not promote self-deposit. DSpace ranked better than other turnkey systems, indeed among the highest levels of deposit, but it differs from other systems in that it also has a community of engaged developers creating plugins and integrations. Whether or not these repositories made use of their own developers or others' work to enhance DSpace for self-deposit is unknown.

Of note was the row in Table 3 for Hydra, a system with a large and engaged development community. The three respondents from the Hydra community all reported receiving between one and 20 self-deposits monthly. However, as a latecomer to the repository landscape,

Hydra may have focused its development on supporting areas in which repositories have found more success, such as holding digitized cultural heritage materials.

Additional Workloads and Satisfaction

No matter how many self-deposits an IR receives, respondents still report time spent to obtain citations for deposit through additional channels. This may have an inadvertently negative effect on outreach. Xia (2007) suggested that the choice to fill repositories using nonfaculty labor may distance faculty from the idea of self-deposit, as they have no practice doing it and the citations appear anyway. Overall, IR staff satisfaction levels are low. The latter is unsurprising, as rates of self-deposit are generally low. The response that surprised the author was that of a repository reporting over 100 faculty self-deposits per month—one of only 5 respondents in 80 to make that assertion—that still reported their level of satisfaction as "Dissatisfied." No one reporting the highest level of self-deposit also reported themselves as "Completely satisfied," whereas one institution with the lowest level did so. Success, or at least achieving substantially higher levels of self-deposit than most institutions, does not necessarily lead to satisfaction within the institution.

Therefore, those planning IR programs should not expect, even if they manage to achieve enough engagement to get outlier levels of self-deposit, that they will not have to invest additional time into adding additional citations. Neither should they expect that they or their program will feel a warm surge of satisfaction when a significant number of faculty become engaged.

CONCLUSION

A small number of repositories report what could be considered successful levels of engagement. These have existed for at least five years, are large and generally public, and use repository software, which requires dedicated developer time for its creation and maintenance (or DSpace, which may require development). More repositories that fit the same profile see far lower levels or no faculty self-deposit.

Although most repositories grow in size, most are being filled by persons at the institution explicitly tasked with doing so rather than eager faculty. Those who have achieved what most peers would consider success still perform the same laborious tasks and don't feel particularly satisfied with the state of affairs. Broadly, when it comes to faculty engagement and repositories, we are neither successful nor happy. Not even, as indicated in the literature review, when we have an OA mandate which allows or requires self-archiving.

This article does not propose to address strategies by which institutions may improve the condition of their repositories, rather referring those interested to Salo (2008). In 2016, the author encountered representatives of two institutions that reported having comparative success by adopting Salo's recommendation of partnering with the broader campus cyber-infrastructure (although they did not refer to Salo). By joining academic departments in choosing and implementing new software for the annual review process, these libraries obtained the rights to harvest the records of scholarly output for each faculty member. This process still required obtaining a copy of the article which they had the right to deposit, but the libraries were aided by OA policies. This model appeared promising because it aligned with work the faculty were already doing for a goal in which the faculty saw real value—obtaining tenure or promotion. It could be compared to the work libraries already do using IRs as sites for deposit of their electronic theses and dissertations (ETDs) (Boock & Kunda, 2009; Veve, 2016). Unlike the literature around faculty self-deposit, these articles focus primarily on technical hurdles, such as system migration (Coles & Johnson, 2010) or ingesting from external sources like ProQuest (Averkamp & Lee, 2009). These partnerships with the institution's graduate school may provide some examples of how partnering with departments to capture work as part of the tenure process might look.

The case of institutional repositories may be more broadly applicable when librarians and library technologists are considering efforts for which faculty participation is critical for success. If success requires broad participation of faculty, should librarians consider themselves capable of judging what the faculty will find useful and valuable just because our community can see use and value in it? The initial levels of self-deposit in institutional repositories would suggest that we are not good at gauging levels of interest, not good at developing something people would actually use, or both.

Even if faculty do not deposit in them, institutional repositories are unlikely to vanish soon. They may contain thousands of works deposited by library employees or the complete ETD record of the institution. What's left is for library workers to decide, working with administration, what their success looks like and how what we've learned from the last decade-and-a-half will help us make better decisions in the future.

RECOMMENDATIONS FOR ADDITIONAL RESEARCH

As discussed in Methods, the survey could have been greatly enhanced by the addition of a question about whether or not the campus had an OA mandate and whether that mandate included deposit of works in an IR. There is some room to be done on a broad analysis of the practical effects of such mandates. This work might build on Vincent-Lamarre et al.'s

(2015) methodology for their survey combined with Zhang et al.'s (2015) focus on practical effects at just one institution. What do the process and workflow look like for the workers involved in managing the IR before and after the mandate? How have things improved or stayed the same?

There is also room for more research on aspects of job satisfaction and labor of those tasked with making up for lack of self-deposit by harvesting citations and loading them into an institutional repository. At HydraConnect 2016, Linda Newman described the effort of trying to find and add all possible citations to the repository as a "thankless and endless task." What do retention rates of librarians primarily tasked with this work look like? How do they compare to other rates of retention of library professionals?

ACKNOWLEDGMENTS

REFERENCES

Averkamp, S. & Lee, J. (2009, June 26). Repurposing ProQuest metadata for batch ingesting ETDs into an institutional repository. *Code4Lib Journal. 7*. Retrieved from http://journal.code4lib.org/articles/1647

Boock, M., & Kunda, S. (2009). Electronic thesis and dissertation metadata workflow at Oregon State University Libraries. *Cataloging & Classification Quarterly*, 47(3-4), 297-308. https://doi.org/10.1080/01639370902737323

Carr, L., & Brody, T. (2007). Size isn't everything: Sustainable repositories as evidenced by sustainable deposit profiles. *D-Lib Magazine*, *13*(7–8). https://doi.org/10.1045/july2007-carr

Casey, A. (2012). Does tenure matter? Factors influencing faculty contributions to institutional repositories. *Journal of Librarianship and Scholarly Communication*, *1*(1), 1–11. https://doi.org/10.7710/2162-3309.1032

Coles, B., & Johnson, K. (2010). Moving Electronic Theses from ETD - db to EPrints: The Best of Both Worlds . Baltimore: NDLTD.

Covey, D. T. (2011). Recruiting content for the institutional repository: The barriers exceed the benefits. *Journal of Digital Information*, *12*(3), 1–18.

Creaser, C., Fry, J., Greenwood, H., Oppenheim, C., Probets, S., Spezi, V., & White, S. (2010). Authors' awareness and attitudes toward open access repositories. *New Review of Academic Librarianship*, *16*(sup1), 145–161. https://doi.org/10.1080/13614533.2010.518851

Davis, P. M., & Connolly, M. J. L. (2007). Institutional repositories: Evaluating the reasons for non-use of Cornell University's installation of DSpace. *D-Lib Magazine*, *13*(3–4). Retrieved from http://www.dlib.org/dlib/march07/davis/03davis.html

Fry, J., Spezi, V., Probets, S., & Creaser, C. (2015). Towards an understanding of the relationship between disciplinary research cultures and open access repository behaviors. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.23621

Library Loon. (2017, January 4). ARL: Please eliminate silencing potential from SPEC surveys. Retrieved from https://gavialib.com/2017/01/arl-please-eliminate-silencing-potential-from-spec-surveys/

Lynch, C. A. (2003, February). Institutional repositories: Essential infrastructure for scholarship in the digital age. *ARL Bimonthly Report 226*, 1–7. https://doi.org/10.1353/pla.2003.0039

Más-Bleda, A., Thelwall, M., Kousha, K., & F. Aguillo, I. (2014). Successful researchers publicizing research online. *Journal of Documentation*, *70*(1), 148–172. https://doi.org/10.1108/JD-12-2012-0156

Neugebauer, T., & Murray, A. (2013). The critical role of institutional services in open access advocacy. *International Journal of Digital Curation*, *8*(1), 84–106. https://doi.org/10.2218/ijdc.v8i1.238

Poynder, R. (2016, September 22). Q&A with CNI's Clifford Lynch: Time to re-think the institutional repository? Retrieved from http://www.richardpoynder.co.uk/Clifford_Lynch.pdf

Sale, A. (2006). Comparison of content policies for institutional repositories in Australia. *First Monday*, *11*(4). Retrieved from https://doi.org/10.5210/fm.v11i4.1324

Salo, D. (2008). Innkeeper at the roach motel. *Library Trends*, *57*(2), 98–123. Retrieved from https://doi.org/10.1353/lib.0.0031

Spezi, V., Fry, J., Creaser, C., Probets, S., & White, S. (2013). Researchers' green open access practice: a cross-disciplinary analysis. *Journal of Documentation*, *69*(3), 334–359. Retrieved from https://doi.org/10.1108/JD-01-2012-0008

Veve, M. (2016) Harvesting ETD Metadata from Institutional Repositories to OCLC: Approaches and Barriers to Implementation. *Journal of Library Metadata*, 16, 69-79. http://doi.org/10.1080/1051712X.2016.1215730

Vincent-Lamarre, P., Boivin, J., Gargouri, Y., Larivière, V., & Harnad, S. (2015). Estimating open access mandate effectiveness: The MELIBEA score. *Journal of the Association for Information Science and Technology*, *67*(11), 2815-2828. https://doi.org/10.1002/asi.23601

Xia, J. (2007). Assessment of self-archiving in institutional repositories: Across disciplines. *The Journal of Academic Librarianship, 33*(6), 647–654. https://doi.org/10.1016/j.acalib.2007.09.020

Xia, J. and Sun, L. (2007). Assessment of self-archiving in institutional repositories: Depositorship and full-text availability. *Serials Review 33*(1), 14–21. https://doi.org/10.1080/00987913.2007.10765087

Xia, J. (2008). A comparison of subject and institutional repositories in self-archiving practices. *The Journal of Academic Librarianship*, 34(6), 489–495. https://doi.org/10.1016/j.acalib.2008.09.016

Zhang, H., Boock, M., & Wirth, A. (2015). It takes more than a mandate: Factors that contribute to increased rates of article deposit to an institutional repository. *Journal of Librarianship and Scholarly Communication*, *3*(31), 1–16. https://doi.org/10.7710/2162-3309.1208

Zheng Ye (Lan) Yang, & Yu Li. (2015). University faculty awareness and attitudes towards open access publishing and the institutional repository: A case study. *Journal of Librarianship and Scholarly Communication*, *3*(1). https://doi.org/10.7710/2162-3309.1210

APPENDIX

Survey on Faculty Self-Deposit Practices in Institutional Repositories

Default Block

CONSENT INFORMATION

Background

You are being invited to participate in a research project conducted by Ruth Kitchin Tillman, Digital Collections Librarian, University of Notre Dame. This survey gathers basic and broad information about patterns of faculty self-deposit, or the lack thereof, in institutional repositories at institutions of higher education in the United States. The survey will be available through July 8, 2016. This survey is open to anyone at an institution of higher education in the United States self-identifying as responsible in some way for bringing content into an institutional repository or otherwise involved in handling metrics of an institutional repository. The total number of participants will depend on responses to the survey.

Study Procedure

This is a virtual, multi-location study. This study will ask respondents basic information about their institution and repository, information about self-deposit (including whether it is possible in their repository), and information about satisfaction with levels of self-deposit in the repository, whether they have developed alternative ingest methods, and whether outreach is currently being conducted to faculty members. The survey should take 5-20 minutes to complete. Respondents are asked to take the survey only once.

Risks

While participating in a study carries the risk of a loss of confidentiality, the investigator has minimized risk by not collecting IP or other personally identifiable information. Questions about institutional affiliation are optional.

Benefits

While you will not directly benefit from participation, your participation may help the investigator understand trends in faculty self-deposit in institutional repositories. As the resulting data will be shared through publication, your participation may also help the field understand patterns which will allow for the development of better practices in creating and promoting institutional repositories.

Confidentiality

The results of this study may be published in professional journals. It may also be used for educational purposes or for professional presentations. However, no individual respondent will be identified. Your participation in this project is anonymous. The survey will not be collecting any personally identifying information, however, the first section contains optional identifying institutional information. These identifying fields are entirely optional and will be used for deduplication. Once deduplication has been performed, the identifying fields will be removed from the Qualtrics site. They will not be exported, saved, shared with any other researcher, or published.

Persons to Contact

If you have any questions, you may contact Ruth Kitchin Tillman at rtillman@nd.edu or 574-631-6067. If you have questions about your rights as a research participant, please contact the University of Notre Dame Institutional Review Board (IRB), Notre Dame Research Compliance, (compliance@nd.edu), phone (574-631-1461) IRB Number: FWA 00002462.

Voluntary Participation

Your participation is voluntary and you may refuse to participate or withdraw at any time. You may refuse to answer all but 3-4 core questions (the number depends in part on your responses) and still complete the survey. You may refuse to answer the core questions and withdraw from the survey without penalty.

Costs and Compensation

There are no costs and there is no compensation for participating in this survey.

Limitations on Participant Age

This survey is intended to be taken by persons 18 years of age or older. By continuing, you agree that you are 18 years of age or older.

Definitions: Questions about self-deposit specify faculty and should not be construed to include graduate students, undergraduate students, and staff. For the purposes of this survey, a faculty member is defined according to the institution's definitions and may include anyone from part-time lecturers and adjuncts to tenured full professors. If librarians in the institution hold faculty status, they may be counted as faculty.

Identifying Information

The following two fields are not mandatory but will be used for deduplication, in the

event that more than one representative of the institution answers the survey. After initial deduplication, the information in these fields will be discarded from the Qualtrics results set and will not be shared with other researchers or published.

Institution name (this information will be discarded within Qualtrics after deduplication) [fill in the blank]

Repository URL (this information will be discarded within Qualtrics after deduplication) [fill in the blank]

Demographic and Repository Information Institution Type and Size

Type of institution (required)

Private 4-year college

Public 4-year college

Private 4-year college

Public 2- year college

Other, please fill in the blank

Faculty size (required)

1-500

500-1500

1500+

Repository Information

On what software is your institutional repository built?

Digital Commons

ContentDM

DSpace

Fedora (Hydra)

Fedora (Islandora)

Fedora (custom build)

Other, fill in the blank

How many years has your institutional repository been publicly-available? (not counting any initial development time and beta testing)

0-2 years

2-5 years

5+ years

What types of materials does your repository accept? (please select all that apply)

Articles and article-type objects (journal articles, preprints, conference papers, white papers)

Books

Research data (including datasets)

Institutional records

Software

Other, fill in the blank

From whom does your repository accept materials? (please select all that apply)

Faculty and researchers

Staff

Graduate students

Undergraduate students

Alumni

Non-academic departments

Questions About Self-Deposit

Does your institutional repository allow faculty to self-deposit materials? (required)

Yes

No

[if no, the survey directs to questions about satisfaction and alternative ingest]

How many faculty self-deposits do you receive in the average month? (required)

0

1-20

20-50

50-100

100+

We collect information about self-deposit rates but cannot determine how many come from faculty

We do not collect information about self-deposit rates

What is the primary type of material self-deposited by faculty? (please select only one)

Articles and article-type objects (journal articles, preprints, conference papers, white

papers)

Books

Research data

Institutional records (teaching materials, departmental materials, etc.)

Software

Other, fill in the blank

What school or department at your institution is most heavily represented in faculty self-deposit? (Leave blank if unknown)

Do you plan to develop or add a self-deposit system for your repository?

Yes

No

Unsure

Questions About Satisfaction, Alternative Ingest

Methods, and Outreach

[if respondents selected that they did not allow for faculty self-deposit, they were brought

to this sections]

You answered that your repository does not allow faculty self-deposit. When answering questions about satisfaction and outreach, please address them in the light of any interactions you have with faculty to encourage them to engage with the repository—working with you to deposit their materials, sharing CVs, developing a self-deposit system, etc.

What is your level of satisfaction with your current rate of faculty selfdeposit?

Completely satisfied

Satisfied but planning further outreach

Needs outreach to be satisfactory

Dissatisfied

Extremely dissatisfied

Which other methods do you use to add materials to the repository? (please select all that apply)

Work with faculty members but deposit materials ourselves

Work with department heads to identify publications which we add ourselves

Work with other campus unit to identify publications which we add ourselves

Use faculty CVs and/or websites to identify publications which we add ourselves

Harvest publication metadata from major indexing sites

Other, please fill in the blank

Does you plan to use any of the following methods to add materials to the repository in future? (please select all that apply)

Work with faculty members but deposit materials ourselves

Work with department heads to identify publications which we add ourselves

Work with other campus unit to identify publications which we add ourselves

Use faculty CVs and/or websites to identify publications which we add ourselves

Harvest publication metadata from major indexing sites

Other, please fill in the blank

When faculty work with librarians or repository staff to deposit materials, do they (please select all that apply)

Approach the library independently

Approach the library after outreach events

Agree to deposit after targeted outreach (e.g. subject librarian identifies and approaches a faculty member whose research data could be deposited in the repository; conference presenters are asked whether they would deposit a copy of their paper or presentation in the repository.)

Other, please fill in the blank

Do you conduct outreach about the institutional repository to faculty?

Yes

No

Who conducts the outreach to faculty? (please select all that apply)

Librarians and/or staff responsible for the institutional repository

Subject librarians/liaisons

Other, please fill in the blank

What kind of outreach to faculty do you conduct? (please select all that apply)

Events advertised to all faculty at the institution

Events advertised to all faculty in a particular college or school

Events advertised to all faculty in a particular department

One-on-one meetings

Other, please fill in the blank

Have you seen an increase in deposits after outreach?

Yes, self-deposit

Yes, deposit working with library or institutional repository staff

Unsure

No

Do you plan to conduct outreach to faculty in the future?

Yes

No

Was outreach conducted in the past?

Yes

No

[if no plans for the future]

Was outreach discontinued because it was . . .

Successful, we felt we could stop

Unsuccessful, we felt it would be a waste of time to continue

No longer considered a priority