

# JLSC

ISSN 2162-3309 | JLSC is published by the Pacific University Libraries | <http://jisc-pub.org>

**Volume 8, General Issue (2020)**

## **Research Data Management Services and Strategic Planning in Libraries Today: A Longitudinal Study**

Elise Gowen & John J. Meier

Gowen, E. & Meier, J. J. (2020). Research Data Management Services and Strategic Planning in Libraries Today: A Longitudinal Study. *Journal of Librarianship and Scholarly Communication*, 8(General Issue), eP2336. <https://doi.org/10.7710/2162-3309.2336>

This article underwent fully-anonymous peer review in accordance with JLSC's peer review policy.



© 2020 Gowen & Meier. This open access article is distributed under a Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>)

# Research Data Management Services and Strategic Planning in Libraries Today: A Longitudinal Study

Elise Gowen

*Earth and Mineral Sciences Librarian, Penn State*

John J. Meier

*Head of STEM for Engagement and Outreach, Penn State*

**INTRODUCTION** Research data services have been adopted by many academic libraries. This study tracked the changes in research data management services and staffing among Association of American Universities (AAU) libraries over the past 5 years and compared them to the libraries' goals for research data management (RDM) in their strategic plan. **METHODS** This quantitative study examined libraries at the 60 U.S. AAU institutions. In order to examine longitudinal changes, portions of Briney et.al. (2015a) were used as a basis for measuring data librarian staffing and services. These trends were compared to the contemporary strategic priorities of libraries interviewed by Meier (2016), as well as against strategic plans of 2014 and 2019 available online. **RESULTS & DISCUSSION** While there have been modest increases in libraries in the sample population offering data services, most of those gains have been among the libraries that did not consider RDM a priority in 2014. Interestingly, some of the libraries that mentioned RDM as a priority in 2014 have lost data librarian positions. Over half of the libraries in this study now provide or support a data repository. Many library strategic plans that mentioned RDM as an explicit goal 5 years ago now no longer mention it. **CONCLUSION** Data librarian positions, data services, and data repositories have now become common features of large research university libraries. However, research data services are no longer as prominent in many library strategic plans at institutions where such services are more established, and libraries instead seem to be moving on to the work of rethinking the nature of the services or expanding them.

Received: 09/16/2019 Accepted: 02/17/2020

Correspondence: Elise Gowen, 105 Deike Building, University Park, PA, 16802-2710, edg16@psu.edu



© 2020 Gowen & Meier. This open access article is distributed under a Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>)

## IMPLICATIONS FOR PRACTICE

1. University libraries need to have clear guidance on research data depositing policies for both institutional repositories and data repositories.
2. Strategic plans for university libraries should acknowledge their current goals for research data curation even if baseline research data management goals have already been met.
3. The shift in faculty data needs from simple research data management to more complex data services suggests opportunities for libraries to develop new collaborations with existing information technology stakeholders and to offer new programs and outreach.

## INTRODUCTION

Since being widely identified as a potential new area for service growth in the mid-2000s, research data management (RDM) and related services have been a frequent touchstone in the strategic plans of many libraries. Such services include consultations and workshops on how to manage research data, guidance on writing a research data management plan, and support in finding repositories for the long-term preservation of data. An individual library's capacity to integrate RDM services into the research lifecycle of its patrons is based on factors such as size, budget, and mission of the institution. Some libraries have been able to develop this at a large scale, and others have not. Some have been able to operate at a larger scale by working with a consortium of universities. While the number of institutions that offered these services has grown rapidly since the mid-2000s, there are relatively few longitudinal studies of these services and whether the steps taken to grow the services are still being implemented several years on.

This study attempted to track the changes in RDM services and staffing among Association of American Universities (AAU) libraries over the past five years and compared those changes to the aspirations about research data management each library had stated five years earlier in interviews with library directors and deans along with strategic plans where available. The AAU institutions have similar characteristics with each other, while also being diverse in geographic location and missions. It should be noted that not all libraries in this group are members of the Association of Research Libraries (ARL) and those that are have widely varying levels of funding ("Spending by University," 2017). This group of libraries has been a frequent object of study, allowing for the observance of longitudinal patterns. However, the results obtained may not be applicable to other regions or to other institutional contexts.

One of the ironies this study has revealed is how few open datasets on research data management services exist in the library literature. Most often, the results of study in this area are only presented in aggregate or analyzed form. This may be due to the lack of disciplinary data repositories for library and information sciences (note the absence of anything like “library science” as a potential subject heading in the Registry of Research Data Repositories (<https://www.re3data.org/browse/by-subject/>)). One major exception of this trend is the work done by Kristin Briney, Abigail Goben, and Lisa Zilinski, who conducted research into research data management services in 2014 and published their data along with the article (2015b). While the focus of the Briney paper was on data policies, many of the elements of their dataset provide information on library services and staffing as it pertains to RDM. The study collected data from the data services policies posted online by 206 American universities with “Very High” or “High” research activity, which included almost all of the AAU libraries. They found that data management services and data repositories had become typical for major research institutions.

## LITERATURE REVIEW

Beginning a little over a decade ago, library literature began to discuss taking a more active role in data curation for researchers. While writers like Gold 2007 noted that “Data Librarianship,” like most forms of format-specific librarianship, was often conceptualized in terms of acquiring and curating data, Gold proposed that data librarianship might also involve taking a more active role in the data creation process, or, as the article put it, the “upstream” parts of the data research cycle, as opposed to the traditional “downstream” acquisition and collection of published datasets (Gold, 2007).

Coinciding with the rise of institutional repositories (IR) in the libraries, libraries attempted to address researchers’ growing need to manage and preserve their research data. IRs have played an important role in growing interest in research data as an aspect of library services, but it has not always been clear if research data should be a key component of content preserved in institutional repositories (Shreeves & Cragin, 2008). Nonetheless, these repositories were some of the earliest examples of libraries providing research data management in the libraries (Witt, 2008). Early researchers into the new, library-housed data curation repositories noted that the libraries had “laid the groundwork for future, higher-level work to formalize data curation services for the institution” (Witt, 2008). In anticipation of an increasing need for data services, libraries began hiring traditional liaison librarian positions with data services skills as well as creating new data services librarian positions (Delserone, 2008).

Research data management was given a major boost of importance in the sciences with the announcement that the National Science Foundation would begin requiring the inclusion of a data management plan for funding. A majority of responding libraries starting RDM services in 2011 said the NSF requirement was the main reason they introduced the services (Fearon, Jr., Gunia, Pralle, Lake, & Sallans, 2013). The ARL SPEC Kit provides a useful and broad-ranging overview of the state of RDM services in 2013, the year before Briney et al. conducted their research. The Kit includes surveys of the kinds of RDM services on offer, snapshots of the strategic plans and data repositories of the surveyed institutions, and titles of data librarians.

Even though data services have expanded greatly in the past ten years, recent review of RDM studies in the library literature indicates the role of the library is still not clear (Perrier, Blondal, & MacDonald, 2018). While librarians are effective in offering education and training programs, they lack more technical expertise with data. Libraries that wish to expand their services beyond the archiving and preservation of data are developing staff with expertise in data analysis tools and services.

Libraries need to continually adapt to the changing needs of their users. Many fields have shown a growth in data dissemination practices that can involve the library as publisher (Walters, 2012). However, the majority of researchers do not use current best practices in documentation, dissemination, and preservation of data despite a surge in data generated (Shen & Shen, 2016). A study conducted at Virginia Tech in 2016 found that a multitude of services, education, and technology infrastructure are needed for effective data management and preservation, including strong metadata standards and automated processes. New technologies such as machine learning-based reviewing of data would help with the amount of data produced.

For libraries to create new services and positions around research data, long term decision-making by library leadership is needed. Strategic planning is often used by academic libraries to map out their priorities for future planning and respond to emerging trends (Saunders, 2015). Research data services have been a priority for some libraries in recent strategic planning cycles, though not a majority (Meier, 2016). Both Saunders (strategic plans) and Meier (interviews) independently found about 40% of libraries mentioned RDM services as a strategic priority in 2015. Subsequent studies have generally taken a qualitative approach to RDM services, like Bryant 2018, which conducted case studies of four institutions and noted that at those institutions, RDM services were developed in anticipation of researcher needs before they received any researcher demand, and have since undergone shifts in their focus in response to researcher input and use. (Bryant, Lavoie, Malpas, & OCLC Research 2018).

## METHODOLOGY

In order to evaluate the state of data services being offered in academic libraries 5 years ago, Briney's published dataset (Briney 2015b) was used as a starting point. Though primarily focused on data policy, the Briney data also recorded public online information such as whether university libraries had a data librarian, data services, a data repository, and accepted data in an institutional repository. Together, these fields create a picture of the data services offered by an institution.

The AAU libraries from this study were selected as a subset of the data to compare to Meier (2016) to determine the impact of strategic decision making. The two Canadian AAU institutions from Briney's dataset were excluded as they were not included in both studies, for a total of 60 libraries.

This study attempted to determine if strategic priorities of libraries, as determined by interviews with library deans and directors in Meier (2016), aligned with changes to research data services. The first hypothesis of this study was that libraries with deans who mentioned research data management services as a strategic priority would have greater gains in services than those that did not mention it in the 2015 interviews. The second hypothesis of this study was that there would be an overall increase in library staff, services, and data repositories in the population.

We also collected strategic planning documents as an additional way to understand strategic priorities for libraries, both in 2014 and in 2019. This was heavily inspired by the work of Saunders (2015) who drew from a different sample of universities and consulted their strategic plan documents to understand what libraries were saying their priorities were. We chose to make use of a similar process because it would easily allow us to compare the strategic plans that covered 2014 to those of 2019. Although we were not able to find a 2014-era strategic plan for every library whose dean was interviewed by Meier, we were able to find 2014-era plans for 35 of the 44 institutions interviewed by Meier, and 2019-era plans for 41 of the 44 institutions.

For consistency we adhered to the methodology of Briney (2015b) as much as possible in terms of coding for data services. Our key for coding is based on Briney (2015b), with some variations described in our rubric notes section.

We gathered institutional data from December 2018 through April 2019. We used the methodology from Briney (p. 7, 2015a) to independently evaluate a sample set of five institutions as a norming action. We divided the remainder of the institutions into two groups

and completed the rest independently. Following this we compared our results and evaluated them for inconsistencies.

### **Data Librarian**

In order to determine if the institutions had these positions, we relied on the university library's website, confirming through library staff directories and public profiles. In some cases, where website information was too inconclusive, or libraries showed evidence of previously having the position but no evidence of currently having the position filled, we reached out to personnel at several institutions to confirm whether there was still a data librarian position. We also counted situations in which there was no data librarian currently on staff, but the position was undergoing an active search, as a *Yes*. It was unclear what Briney did in a comparable situation.

### **Data Services**

We counted as data services such things as consultations, workshops, tutorials, and generalized guidance being actively provided by the library or librarians. A list of external resources by themselves were not enough to qualify as data services offered by that library. We again relied for this information on publicly available information findable through the university library's website.

### **Data in Institutional Repository**

In the case of searching to see if the library supported data in its IR, we first found out if there was an IR for the university and then searched the IR's "About" page to see if they explicitly accepted data. When that was inconclusive, the actual submissions to the IR were reviewed to see if datasets were specifically being accepted to the IR.

### **Data Repository**

In the cases of data repositories, we reviewed the "Home" pages and "About" pages of data repositories to see if the repository was characterized as being exclusively or primarily for research data uploaded by researchers within the institutions.

### **Data Policy**

We chose to focus on the institutional services aspects of the previous study and declined to research the institutional or library data policy.

## Strategic Plans

Unlike the other terms in our rubric and key, which were derived from the Briney study, strategic plans were a focus of Saunders's research (2015). We specifically sought out strategic plans for university libraries, not the whole university. Any webpage or documents that were classified as strategic plans or strategic goals or equivalent, were reviewed and searched for mentions of data management or research data services. A strategic plan was counted as the 2014-era if the scope of years spanned included 2014. Plans were counted for the 2019-era if they included that year in their date range or if the current strategic plan had no year specified. Those which referenced data management services were counted as *Yes*, those which did not were counted as *Nos*, as were libraries which did not have a strategic plan posted publicly.

## RUBRIC NOTES

We kept a key attached to the rubric for consistent definitions of each category. Both the key and the rubric were based on Briney's data (with the exception of *strategic plans*). We made our own modifications to the key for clarification when needed.

For instance, we changed the data librarian definition slightly so that in addition to the exact title "data librarian," the definition included all full-time positions supporting RDM in the libraries, regardless of whether the position included the title *librarian*, or whether it was faculty or staff (Federer, 2018). Federer's survey of data librarians and equivalent positions in North America demonstrated that there is great disparity in the job titles attached to such positions, with only two specific job titles out of 81 interviewees recurring more than once. In some regions, other emerging titles like data steward or data curator are also used. Multiple "data" positions that added up to a full-time equivalent position also counted. We excluded pure developers, librarians who are primarily subject specialists, or administrators. The key also defines the difference between a data repository and data in an institutional repository. Data repository is a dedicated data-only repository, while IR with data is an institutional repository that accepts publications and other scholarly output as well as explicitly accepting data.

## ANALYSIS/DISCUSSION

In order to analyze trends over time in the data, we assigned four classifications for changes in each field. A field was marked "No" if a library did not have the service in 2014 and still did not have it in 2019, "Lost" if the library offered something in 2014 but no longer offered it in 2019, "Kept" if they offered something in both 2014 and 2019, and "Gained" if they did not have something in 2014 but had it in 2019 (Table 1).



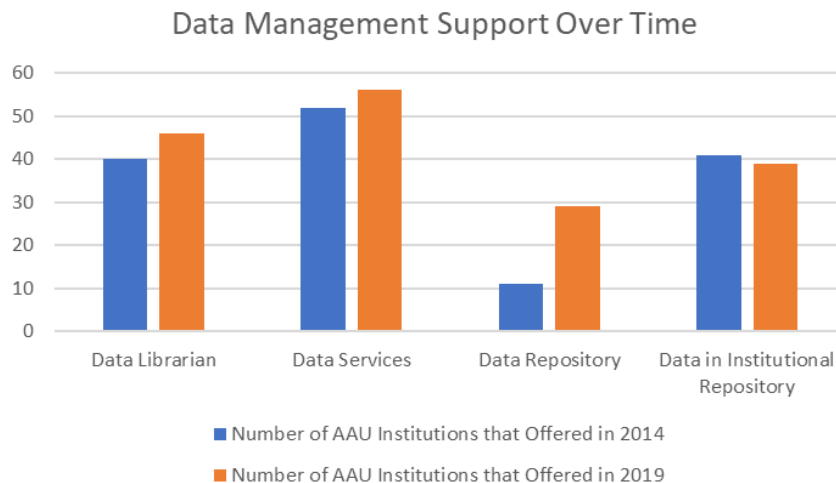
Field	Description
Gained	The library added this between 2014 and 2019.
Kept	The library had this and there was no change.
No	The library did not have this and there was no change.
Lost	The library lost this between 2014 and 2019.

**Table 1.** Coding scheme for trends over time

In situations where universities did have data services or repositories, but they were not provided by the library (usually they were provided by a central IT unit) we coded it as *Yes* or *Gained*, on the grounds that the services were still being provided to the university community and the library was not “missing out” on an opportunity to provide the service; however, we rarely encountered this. Note that Briney (2015b) was more stringent in only counting services or data repositories offered by the library.

### Overall Trends

Between 2014 and 2019, the overall gains for libraries have outpaced the losses (see Figure 1) in all categories with the exception of data in institutional repositories.



**Figure 1.** Data Staff and Service Comparison from 2014 to 2019

There were no dramatic changes in the number of libraries that offer data services or that employ data librarians. The only category where there was a significant increase was in

the number of AAU libraries that offered a data repository, which more than doubled between 2014 and 2019. Some libraries also experienced losses in at least one of the categories between 2014 and 2019, although only one of these categories, data in institutional repositories, experienced a net loss. As noted, this category is where most losses from 2014 to 2019 occurred. In all other categories, some losses occurred but were outnumbered by the gains at other libraries.

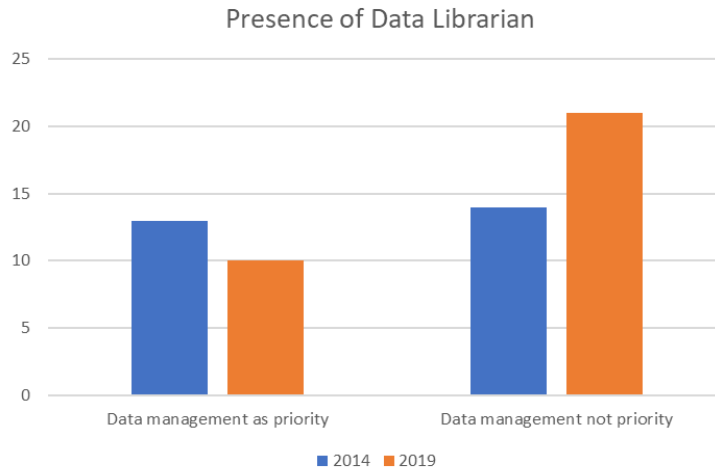
The institutions that mentioned RDM in interviews as a priority in Meier (2016) were grouped together and compared to those that did not specifically mention data as a strategic goal. Of the 60 libraries included in this study, Meier interviewed 44 of their deans or directors. Of those libraries, 15 mentioned data management as a priority in 2014, while 26 did not specifically mention any type of data services. These two sets were used in the analyses below to determine alignment of goals and action.

### **Data Librarians**

Most institutions, 34 (of 60), already had a data librarian in 2014 and retained their data librarian into 2019. Twelve institutions added a data librarian between 2014 and 2019. Six institutions lost a data librarian position between 2014 and 2019. Eight libraries did not have a data librarian in 2014 and still did not have one by 2019.

We then compared these numbers to which libraries stated in interviews that they considered RDM services a priority (see Figure 2). It is notable that 4 of the 15 libraries that mentioned RDM as a strategic priority in 2014 (Meier, 2016) now no longer have a dedicated data librarian. For example, Carnegie Mellon University has taken a few years to rethink the position after losing their research data services librarian to industry (M. Marsteller, personal communication, May 2, 2019). Drawing from this example, we speculate that the results may not indicate a lower priority for data management among this population, but rather indicate that libraries that said data management was a priority in 2014 are now changing their approach to research data support and rethinking data librarian positions. In 2014, 12 of those 15 libraries already had a data librarian, so there were few options for the number of data librarians to increase.

For those 26 libraries that did not specifically mention data services during the interviews, there were more gains to be made as 11 did not have data librarians (see Figure 2). Of these 11, a total of eight gained a data librarian position, which suggests data management became a goal over the past 5 years even though it was not mentioned in the interviews.



**Figure 2.** Changes in the Presence of Data Librarian, 2014-2019, among Institutions that Mentioned RDM as a Priority in Interviews or Did Not

Many institutions have their own language for talking about the kinds of research data management services they provide and those who provide them. As a result, it can be difficult to tell through online research who in an institution is working to support RDM, how much of their time is dedicated to it, and if their responsibilities are primarily technical or service oriented.

Due to this variability, it can also be difficult to directly compare the positions to each other except in situations where the job description is available online. As a result, there remain questions about the similarities between the positions or what else they may do as part of their job (such as liaison librarianship or other digital projects). In addition to RDM services, libraries also provide services pertaining to data such as statistical analysis, survey design and dataset acquisition that do not fall under our definition of RDM.

### Data Services

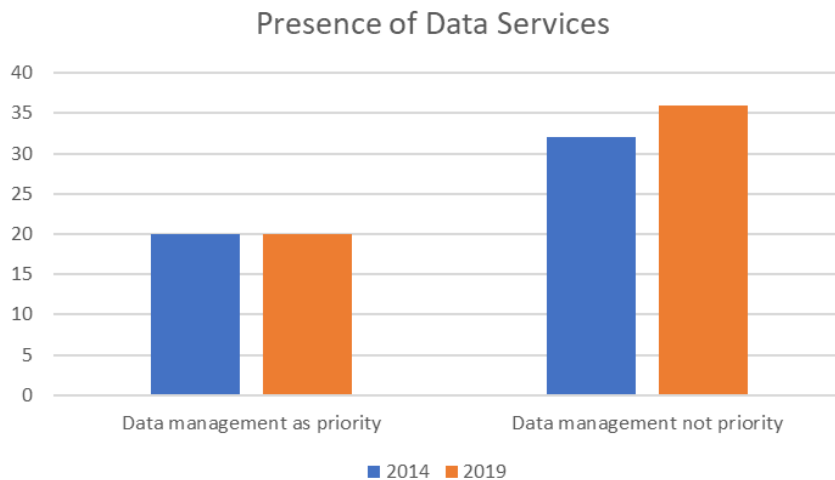
The category of *Data Services* was by far the most prevalent category of research data management services in these AAU libraries, both in 2014 (87%) and 2019 (93%) (see Figure 1). This is also growth from 2013 when 74% of respondents to the ARL SPEC Kit survey indicated they offered RDM services (Fearon, Jr. et al., 2013). While “data management” was identified as the most common name for these types of services in ARL universities and universities with high levels of research activity in 2017 (Yoon & Schultz), we found that “data services” was the most common term among our population. We also noted that some data services webpages proved surprisingly difficult to find, either due to idiosyncrasies in naming or due to being

hosted in LibGuides without local links, but most were easily navigated to from the library's main page, usually via a "Research" drop-down menu.

In our sample population, libraries were the predominant providers of research data services in a university. Even in the situations where the IT unit or other university units were the primary research data services providers rather than the library, the library usually had a presence in the unit, usually in the form of a librarian or library staff serving in the unit. University of Wisconsin is an example of a robust interdisciplinary program that includes not just library employees but Department of IT and Department of Academic Technology staff (<http://researchdata.wisc.edu/>). There is also the UC system that has shared data services across multiple institutions with individual campuses having multiple staff and librarians, depending on their budget (<https://www.ucop.edu/information-technology-services/services/it-staff-services/data-services.html>).

We found that the amount and frequency of data services programming and workshops were one way to measure activity within the library related to RDM, and often seem to be correlated with having dedicated staff or having more staff overall. Another potential area of study would be to observe the extent of participation in these programs across the country.

All 15 of the libraries that mentioned RDM as a priority in the 2015 interviews (Meier, 2016) also provided data services in 2014 (Briney, 2016) and kept them into 2019 in our results (see Figure 3). For the 26 libraries that did not mention data management as a priority, only three did not have data services in 2014 (Briney, 2016) and two of those gained data services by 2019 (Figure 3).



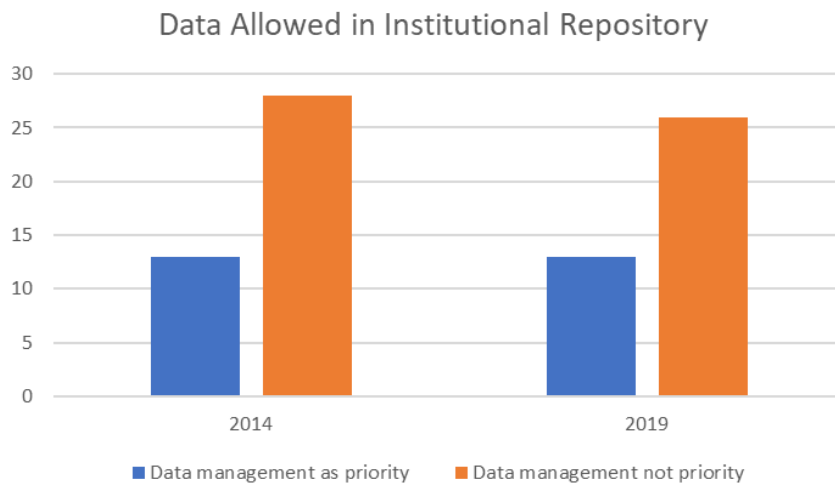
**Figure 3.** Changes in the Presence of Data Services, 2014-2019, among Institutions that Mentioned RDM as a Priority in Interviews or Did Not

### Institutional Repositories

The category of Data in Institutional Repositories was an outlier among our results, as the only category that experienced an overall loss in the five years, with the percentage of libraries that allowed data in IRs going from 68% to 64%. Most likely, this represents not a step backward in data management services in the library, but the sign of an increasing number of libraries that adopted data repositories and thus no longer needed to allow data in their IRs.

In cases where we explored a library’s IR, it could be hard to determine if an IR explicitly allowed data due to lack of documentation of data policies. Most IRs examined did not directly specify what formats are accepted. However, in many cases, we were able to locate datasets currently being stored in the IR even if datasets were not explicitly welcomed.

Those libraries that mentioned research data management as a priority in interviews showed almost no movement in this category between 2014 and 2019, with only one library that allowed data in institutional repository changing to one that did not, and only one library gaining it, for a total that did not change in the five years between studies (see Figure 4).



**Figure 4.** Changes in Allowing of Data in Institutional Repository, 2014-2019, among Institutions that Mentioned RDM as a Priority in Interviews or Did Not

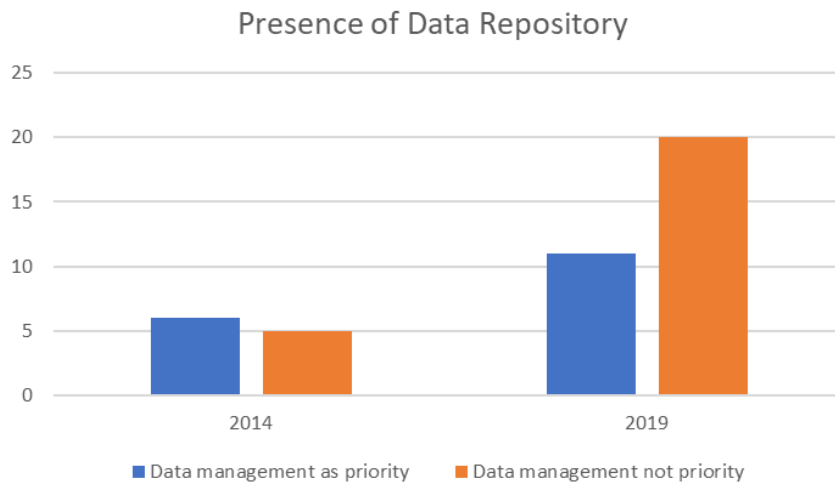
Libraries that did not mention RDM in interviews made up most of the losses that occurred, although it should be noted that these occurred overwhelmingly in libraries that gained a data repository between 2014 and 2019 (see Figure 5).

## Data Repositories

The number of data repositories grew substantially from 18% in 2014 to 35% in 2019 (see figure 5). An increase can also be seen from 2013 with 13% of ARL SPEC kit respondents with data repositories (Fearon, Jr. et al., 2013).

As noted, some of the losses in the category of data in institutional repositories is likely due to the addition of a data repository and new rules explicitly making clear the difference in what is allowed in a data repository versus an institutional repository. For example, the University of Texas at Austin changed to a *yes* for data repository and a *no* for data in institutional repository, which reflects the launch of The Texas Data Repository (TDR) (<https://legacy.lib.utexas.edu/about/news/libraries-launches-texas-data-repository-support-campus-research>). Other studies have found that some libraries create a data repository in parallel to an institutional repository and other research collections (Fallaw et al., 2016).

The 26 library deans and directors that specifically mentioned data management as a priority were more likely to have a data repository at their institution already (6 of 15, or 40%, compared to 5 of 26, or 19%). Possibly as a result of this difference, the libraries that did not mention research data management service experienced much larger gains in data repositories added by 2019, both in total numbers and as a percentage of the population (See Figure 5).



**Figure 5.** Changes in Presence of Data Repository, 2014-2019, among Institutions that Mentioned RDM as a Priority in Interviews or Did Not

## Strategic Plans

While collecting data, we noticed some general trends within data services in the libraries we observed.

We noticed a downward shift in the percentage of strategic plans among AAU libraries that mention RDM as a priority between 2014 and 2019. While searching online, we were able to locate a 2014-era strategic plan for 34 of the 60 AAU libraries and a 2019 strategic plan for 56 of the 60 AAU libraries. In 2019, a smaller percentage of the strategic plans mentioned research data management specifically. In the strategic plans for 2014, 22 out of the 34 plans mentioned data management services (65%), as opposed to 26 out of the 50 strategic plans or goals for 2019 (43%). However, 43% is still higher than Saunders (2015) or Meier (2016) found in their respective research (roughly 40%). Also, it should be noted that our methodology of seeking out 2014-era strategic plans in 2019 may have biased the sampling toward PDFs of strategic plans, which stay online longer and tend to have a longer page length and cover more topics overall, which may have given us a higher percentage of plans that mention RDM than we would have found if we had conducted the research in 2014.

One interesting trend we did notice is that as data services become a regular operation of the library, they are sometimes no longer identified as a strategic development area in strategic planning documents. For example, the University of Minnesota Libraries has been a leader in RDM services for many years yet does not mention them in their 2019 strategic plan. Strategic plans tend to focus on future development rather than maintaining past initiatives, so a nascent program such as Tulane University with few data services and no data librarian still mentions RDM in their 2019 plan. Roughly half of the 2014-era strategic plans that mention RDM no longer mention it in 2019-era strategic plans.

The libraries that have already achieved their 2014 goals for research data service and research data management planning have shifted their focus away from mentioning RDM in strategic plans, possibly because it's considered a goal that has been achieved. Some have shifted towards stated goals relating to data science, data visualization, and data curation, indicating a focus on more ambitious and sophisticated data service options.

This focus on data science was not observed in the earlier strategic plans and was only observed in a few of the 2019 plans. The possibility of focus shifting or expanding to data science should be observed as a potential future trend. We may be moving into a "post-RDM" future where libraries like the University of Arizona don't mention RDM services, but do say in their strategic plans that they intend to "establish strategic alliances with cam-

pus partners... to develop campus-level support for data science.” University of California at Berkeley Library’s strategic plan says, “As Berkeley pioneers break new ground in areas such as data science and digital humanities, the Library must facilitate collaborations and provide scholarly resources, tools and spaces.” University of Illinois Urbana Champaign Library frames their goal for data services as one of growing what already exists into new areas: “expand research data services and accompanying education initiatives, focusing on data curation, use, and dissemination.”

These services often require higher technical and specialty skills to support their activities, from more specialized staff. We assume that as initial goals are met, some libraries are pushing forward with higher level data goals, while others are maintaining services as is. Many of the emerging collaborative efforts to make data findable and reproducible, such as the FAIR data principles or Open Science Framework, were not referenced in any plans.

## CONCLUSIONS

### Broad Trends

Between 2014 and 2019, research data management support grew at university libraries in the AAU. However, there have been individual institutions that lost services in all categories, most notably in terms of losing a data librarian position where they once had one. It is not clear if this work has been absorbed into other roles. New staffing models involving campus IT, consortial efforts, or entire departments have emerged to support data services and repositories. Stated goals in library strategic plans or in interviews with library deans and directors in 2015 were not a good predictor of increases in data librarian positions or data services by 2019, and the most significant growth across all categories occurred among the libraries that had not identified research data management as a priority in 2015.

While there have been only modest overall increases in the variety and amount of research data services within the population, many institutions have shifted focus toward data repositories in addition to using institutional repositories for data. We found that over half of the libraries in this study now provide or partner with a data repository. While the library may not be the host of institutional repository (IR) or data repository platforms in every instance, librarians and library staff are frequently key partners in these programs.

Data services are almost ubiquitous in AAU libraries with 93% of the population offering services for research data management, including consultations, workshops, and tutorials. These libraries may also offer services like survey design, statistical analysis, and dataset acquisition that touch many other stages of the research lifecycle.



Overall, the way our population of libraries frame data services as a key strategic goal has shifted, and there's been an overall decrease in the percentage of strategic statements that make mention of RDM. This may be an indication that research data services have become so core as to no longer need mention in library strategic plans. It also possibly reflects a change in the overall level of depth and detail in library strategic planning documents.

### **Limitations of the research**

Our research was based on a binary “yes/no” approach to data librarianship and data services that did not track, for instance, whether there were multiple data librarians at an institution or just one, or the extent or variety of data services offered. These nuances are worth studying in more depth, perhaps through more interview-based exploration. As data services have become almost a given for libraries in major research institutions, future library science research will need to shift beyond whether libraries have these services, and instead explore the quantity and nature of services and staffing. There may be particular value in assessing the number of patrons who make use of these services and the programming supporting them. One potential model already in use in higher education for self-assessment of capability of research data services is the RISE model, developed by the Data Curation Centre, which identifies 21 areas of research data support and asks libraries to assess them at three different levels of capabilities, from basic compliance to leadership (Maxwell, Norton, & Wu, 2018).

We also did not conduct a follow-up to the interviews conducted in 2014 with library deans and department heads. This might have given greater insight into how the administration prioritizes RDM services in their own words and would have given greater insight into how administration sees research data management fitting into the current library services landscape, in the absence of more detailed strategic plans.

### **Recommendations for libraries**

If libraries consider RDM services to be a core activity, they likely will benefit from ensuring that their Data Services page can be easily navigated to from the library's homepage, and that the page be easily discoverable through search engines. While this was true of many of the pages visited in the process of collecting data, a notable minority of pages were difficult to find, suggesting that user experience testing for users seeking RDM services information in the library might be worthwhile. It is also recommended for libraries to provide contact info for personnel users can reach out to for data services. Due to uncertainty about whether institutional repositories accept data, and norms varying between institutions, it should

be noted on a library's institutional repository if datasets are accepted or not, and if not, where the data should be deposited. Librarians and other library researchers should lead by example in following good data management practices and ensuring that their research data has been made freely available in institutional repositories or other sources.

Data repositories are increasingly replacing institutional repositories as the home for research data. Yet library workflows still need to integrate the repository into research practice. We noticed that this is more common in institutions when interdepartmental collaboration within a university produces the data repository.

Data librarian positions have become widespread within academic libraries, and so have non-librarian specialists who perform many of the same functions. The lack of librarian status for these specialists should not be a barrier to their professional development in the position or their role in the library. Professional support networks and organizations should be strengthened based on this.

AAU libraries in the 21<sup>st</sup> century have made significant strides in the widely stated goal to offer support in the research data life cycle. As RDM services move from an up-and-coming service being initiated in anticipation of researcher needs to a set of widely implemented services and staffing, library's attention seems to be shifting, as reflected in the changing focus of strategic plans. Libraries should put long-term support behind programs and positions in order ensure sustainable growth. Now is the time to assess whether the library is serving the desired role in the research data life cycle that led to the creation of these services and positions. There are also indications that libraries are looking ahead to the future of what lies beyond RDM services to supporting a fuller suite of data science services, and libraries may want to consider such an approach for their institution.

## REFERENCES

Briney, K. (2016). Strategic planning for research data services. *Bulletin of the Association for Information Science and Technology*, 42(4), 39–41. <https://doi.org/10.1002/BUL2.2016.1720420411>

Briney, K.; Goben, A.; Zilinski, L. (2015a). Data from: Do you have an institutional data policy? A review of the current landscape of library data services and institutional data policies. *Harvard Dataverse*. <https://doi.org/10.7710/2162-3309.1232>

Briney, K., Goben, A., & Zilinski, L. (2015b). Do you have an institutional data policy? A review of the current landscape of library data services and institutional data policies. *Journal of Librarianship and Scholarly Communication*, 3(2), eP1232. <https://doi.org/10.7710/2162-3309.1232>

- Bryant, R., Lavoie, B. F., Malpas, C., & OCLC Research. (2018). The realities of research data management. Part three, Incentives for building university RDM services. In *OCLC Online Computer Library Center, Inc.* Retrieved from <https://eric.ed.gov/?id=ED589137>
- Delserone, L. M. (2008). *At the Watershed: Preparing for Research Data Management and Stewardship at the University of Minnesota Libraries*. Retrieved from <http://digitalcommons.unl.edu/libraryscience/245>
- Fallow, C., et al. (2016). *Overly honest data repository development*. Retrieved from <https://www.ideals.illinois.edu/handle/2142/91684>
- Fearon, Jr., D., Gunia, B., Pralle, B., Lake, S., & Sallans, A. (2013). *SPEC Kit 334: Research Data Management Services (July 2013)*. <https://doi.org/10.29242/spec.334>
- Federer, L. (2018). Defining data librarianship: a survey of competencies, skills, and training. *Journal of the Medical Library Association : JMLA*, 106(3), 294–303. <https://doi.org/10.5195/JMLA.2018.306>
- Gold, A. (2007). Cyberinfrastructure, data, and libraries, part 2. *D-Lib Magazine*, 13(9/10). <https://doi.org/10.1045/september2007-gold-pt2>
- Maxwell, D., Norton, H., & Wu, J. (2018). The data science opportunity: Crafting a holistic strategy. *Journal of Library Administration*, 58(2), 111–127. <https://doi.org/10.1080/01930826.2017.1412704>
- Meier, J. J. (2016). The future of academic libraries: Conversations with today's leaders about tomorrow. *Portal: Libraries and the Academy*, 16(2), 263–288. <https://doi.org/10.1353/pla.2016.0015>
- Perrier, L., Blondal, E., & MacDonald, H. (2018). Exploring the experiences of academic libraries with research data management: A meta-ethnographic analysis of qualitative studies. *Library & Information Science Research*, 40(3–4), 173–183. <https://doi.org/10.1016/j.lisr.2018.08.002>
- Saunders, L. (2015). Academic libraries' strategic plans: top trends and under-recognized areas. *The Journal of Academic Librarianship*, 41(3), 285–291. <https://doi.org/10.1016/j.acalib.2015.03.011>
- Shen, Y., & Shen, Y. (2016). Strategic planning for a data-driven, shared-access research enterprise: virginia tech research data assessment and landscape study. *College & Research Libraries*, 77(4), 500–519. <https://doi.org/10.5860/crl.77.4.500>
- Shreeves, S. L., & Cragin, M. H. (2008). Introduction: institutional repositories: current state and future. *Library Trends*, 57(2), 89–97. <https://doi.org/10.1353/lib.0.0037>
- Spending by university research libraries, 2015-2016. (2017, August 13). Retrieved from <https://www.chronicle.com/article/Spending-by-University/240829>
- Walters, T. (2012). *The Future Role of Publishing Services in University Libraries*. 12(4), 425–454. <https://doi.org/10.1353/pla.2012.0041>
- Witt, M. (2008). Institutional repositories and research data curation in a distributed environment. *Library Trends*, 57(2), 191–201. <https://doi.org/10.1353/lib.0.0029>

Yoon, A., & Schultz, T. (2017). Research data management services in academic libraries in the u.s.: A content analysis of libraries' websites. *College & Research Libraries*, 78(7). <https://doi.org/10.5860/crl.78.7.920>

