Scaling research data management services along the maturity spectrum: Three institutional perspectives.

Introduction

Research data services promise to advance many academic libraries' strategic goals of becoming partners in the research process and integrating library services with modern research workflows. Academic librarians are well positioned to make an impact in this space due to their expertise in managing, curating, and preserving digital information, and a history of engaging with scholarly communications writ large. Some academic libraries have quickly developed infrastructure and support for every activity ranging from data storage and curation to project management and collaboration, while others are just beginning to think about addressing the data needs of their researchers. Regardless of which end of the spectrum they identify with, libraries are still seeking to understand the research landscape and define their role in the process.

This article seeks to blend both a general perspective regarding these issues with actual case studies derived from three institutions (University of Cincinnati - UC, Oklahoma State University - OSU, and Florida State University - FSU) all of which are at different levels of implementation, maturity, and campus involvement. We propose herein that scholarly communication and research data services are deeply intertwined, and more so since the OSTP memorandum titled *Expanding Public Access to the Results of Federally Funded Research* (link). In this article, we offer several recommendations to build this area of data support, based on experiences at our respective institutions. We believe that a well-constructed, mature model for public access research support will focus on outreach, resources (technology, new positions, operational units/spaces), partnerships, and professional development. This article highlights the development of scholarly communications, research data services and public access policies through examples from one or more of the three institutions. At this uncertain time period related to public policies aimed at open data, there is a clear opportunity for academic libraries to deeply connect with our research colleagues throughout the research lifecycle, engaging with public/open access to data and publications.

Outreach

Outreach initiatives play a large part in an academic library's ability to be a useful resource on campus and in the community. And so, "academic libraries are catalysts for partnerships and collaborations with faculty, students, various campus departments, and external organizations. Libraries tend to have welcoming cultures and are willing to cross disciplines and services borders to create results," (Doherty, 2016). Well-constructed collaborations can also offset the challenges of a lack in funding and limited staff, obstacles that academic libraries commonly face. By having personnel dedicated to fostering outreach efforts, potential collaborators are met with a knowledgeable, dedicated person that is prepared to handle the partnership (Meyer, 2014).

Following the OSTP mandate, FSU Libraries made a concerted shift away from advocacy and the moral imperative of openness toward open scholarship as a compliance issue. Although sometimes feeling at odds with a principled librarianship, this decision was an effective outreach strategy, opening doors for further conversations about open access, open data, and altmetrics. Working rapidly to get ahead of OSTP responses from funding agencies, FSU Libraries put together a Public Access Support Team in 2015 and in one month launched an outreach support program called Federal Open Policies Operation (FOP-Op). The Public Access Support Team was created with no administrative mandate (or permission), pulling together functional specialists and subject experts for high-funded research disciplines. Annually, that team reconvenes and conducts another round of FOP-Op outreach, often including targeted emails to faculty actively working on proposals, and refreshing outreach materials like slideshows and talking points. In 2017, we released the FOP-Op materials on the Open Science Framework (link).

A large-scale reorganization effort was underway at Oklahoma State University Libraries, and in conjunction with the changes in funding mandates from the federal government, the library took the opportunity to delve into the world of research data management issues. As a first step, OSU Libraries defined its research data management services and developed a web presence with multiple contact streams. The decision was made that a one-stop website should anchor all information about data services, and provide a cohesive user approach at one URL. In addition, OSU Libraries developed a fully-realized communication plan: 1) departmental visits were coined as "RLS Roadshow" centered on a one page presentation; 2) faculty folders were compiled for individual visits, framing common talking points and serving as a leave behind resource; 3) email templates became a key tool that could be sent out to invite faculty to events or publicize a specific service or tool. This approach provided a consistent message from the libraries, and offered basic information about services, events, and important contacts. In recent iterations of the concept, the focus is on distributing core messages at key points in the semester, allowing continued communication with departments and units on an ongoing basis. In parallel with this work, the OSU libraries quickly developed a partnership with the High Performance Computing Center and the office of the Vice President for Research. This collaboration lended an outward focus to data services efforts and built on existing relationships with individual faculty and other campus offices such as Sponsored Research Officers and the Associate Deans for Research. This led to such initiatives as jointly sponsored Software Carpentry workshops and instructor certification in the Carpentries model. The most recent extension of these efforts is a campus-wide hybrid computing and data group Coalition for Advancing Digital Research and Education (CADRE), providing researchers with assistance in all aspects of computational and data-intensive research and education, including, but not limited to: high performance computing, high speed networks, data analysis and visualization, creating and implementing data management plans, applying best practices for organizing their data, as well as storing, archiving, and sharing that data once their project is completed.

Partnerships

While research data services and scholarly communication are an evolving focus for academic libraries, librarians are also tasked with forming general intra-university partnerships to better serve their patrons. One method of collating resources is to build effective partnerships with other research support units, which library administrators can leverage as campus-facing representatives. That said, there is an overall lack of published literature on research partnerships between faculty and librarians (Hollister & Schroeder, 2015). One of the few studies, from Cox & Verbaan, states that "academic libraries are changing how they support research… their involvement in research data management (RDM) implies a much deeper relationship with researchers throughout the research lifecycle. Perhaps we are witnessing a shift from support to partnership" (2016).

Across the board, partnerships with the Office of Research, campus IT units, and high performance computing groups prove essential. Naturally, many of the data support activities that libraries are investing in are connected to these kinds of campus units. Other points of connection may include sponsored research officers, grant managers, as well as computing or cyberinfrastructure centers. It's important to think about both the types of partners that are needed in order to build an effective infrastructure for supporting open scholarship, and strategic ways to approach these offices in order to gain their support.

As a prime example of a deep collaboration, the Libraries at Cincinnati partnered with central IT to support research computing needs and data, which led to combined efforts in these areas. Relationships between IT@UC and UC Libraries grew as two IT staff became part of the UC Researcher Services group, and the co-leaders for the UC Researchers Services group also serve on the IT Research and Develop committee. Information Technology hired an NSF-funded Cyberinfrastructure Educator and Engineer (CIEE) to develop high performance computing and data support educational efforts, and to partner with the informationist team on data related workshop development. Additionally a faculty librarian was appointed Assistant Vice President for Integrated Research in the Office of Research in June 2016. The Digital Humanities and Digital Scholarship Center co-director is a joint hire between Arts & Science department of English and UC Libraries. As part of the relationship building efforts of the UC Researcher Services group, UC Libraries also offered support to the Department of Geography's annual National Geographic Information Systems (GIS) Day Celebration for the past three years..

FSU's most valuable partnerships are the Offices of Proposal Development and the Office of Research Compliance. Following early work on self-training, developing instructional and outreach materials, data management plan consultation workflows, and DMPTool implementation, identifying opportunities for cross-pollination with the Office of Proposal Development (OPD) was essential to establishing data expertise. OPD strongly endorsed the libraries plans to develop data management services, and became a key outreach partner,

assisting with marketing established services through direct referrals and use of campus-wide communication channels. As a result of this partnership, the library received over 40 requests for data management planning consultations from 2014 - 2016, mostly through direct referrals and word of mouth. Deepening that partnership, the library collaborated with FSU's Office of Research Compliance on a suite of university policies on Research Data Management Policy, Public Access to Research Publications, and Authorship and Research Integrity. The Research Data Management Policy was drafted almost entirely by a team from FSU Libraries, addressing data collection, description, retention, security, publication, and access. Similarly, the Public Access to Research Publications Policy serves to reinforce the terms of FSU's Faculty Senate Open Access Policy, and establishes guidelines and procedures for complying with federal public access mandates. In both cases, these policies gave the Libraries' Public Access Support team a formal mandate to assist in grants compliance, in addition to providing PIs with step-by-step guidelines for ensuring that their sponsored research outputs are made publicly accessible in accordance with the requirements of specific agencies.

Resources

Staffing and Organizational Structures

As noted by the Association of Research Libraries (ARL 2013), "new roles in research services, digital humanities, teaching and learning, digital scholarship, user experience, and copyright and scholarly communication are being developed at research libraries across the country," to account for the continued interest in open research practices. Developing an effective and sustainable public access support program is impossible without dedicating teams and people. At all of the authors' institutions, a key aspect of program development involved identifying and reprioritizing positions. Once staff are identified and tasked with public access support, there is a need to "brand" related services and define how and where they fit within the broader organizational structure of the library. Although the nature of this branding and structuring will necessarily vary from institution to institution, the importance of these steps to public access program development should not be underestimated. Similar conversations are ongoing around digital humanities in libraries (Erway and Schaffner, 2014), and as the field shifts toward capturing these scholarly communication and data curation activities under terms like "digital scholarship", we expect to see a variety of models for how these units are administratively coordinated and organized. For example, the Association of Research Libraries (ARL) has published SPEC Kits on topics such as: the organization of scholarly communication services, research data management services, and supporting digital scholarship ([SPEC Kit 332](#) [SPEC Kit 334](#) [SPEC Kit 350](#)).

UC Libraries administration planned how to transition the library workforce to meet the needs of research and teaching faculty in data management planning. The library hired three informationists, two based at the Health Sciences Library and one based in the STEM libraries. The deliberate choice to use the term informationist signaled to researchers that the library had

staff with first hand experience in research and a strong knowledge base in one of the discipline areas the informationists would support. For example, the research informationist has a PhD in Pathobiology and Molecular Medicine. Additional new staffing support include a Scholarly Communications/ Digital Publishing Strategist and Digital Humanities/Digital Scholarship Center co-director. As previously described, the move from strategic to operational status opened this kind of financial and human resource investment. Coinciding with these positions, new spaces were launched to give physical homes to this work, including the Data & Geographic Information Systems (GIS) lab, Digital Humanities and Digital Scholarship Center, and the Health Science Libraries Informatics labs. In the Fall 2017, a data visualization wall will be installed in close proximity to the Data & GIS lab.

One of the main challenges faced at Oklahoma State libraries was that the libraries were unable to hire new positions into these areas. Librarians in the RLS unit therefore had to collaborate with both internal library partners such as the division of Digital Resources and Discovery Services as well as external partners such as the High Performance Computing Center in order to increase the capacity and extend the skillset beyond what is currently available in RLS. Research data services is seen as a component of everyone's duties, and primary assignments were re-written to include a focus on integrating these activities into the work of the entire division, keeping in mind variations in need among different disciplines and departmental priorities. The unit has since re-purposed one existing position to include these elements on a more formal level that will be hired in the near future as a Data Management coordinator. Because this is the unit's first functional specialist role, the group is examining how this position will function alongside liaisons to differentiate types of interactions and interventions with faculty and researchers. For example, a liaison might be expected to make recommendations and a referral regarding data storage options as opposed to knowing exactly what the data deposit workflow might entail which is something the coordinator will be able to do.

Tools and Technology

Information technology resources are another critical consideration in public access program development. Public access compliance is highly dependent on sophisticated technical infrastructure, from DMP planning tools to institutional, subject, or funder-specific repositories for publications and data. Finding funding and development resources to implement such technical infrastructure can be challenging for many academic libraries, particularly in cases where research and learning service divisions have primary responsibility for providing RDM support.

To meet the data needs of UC researchers, UC Libraries sought out community standards and open source resources to fill gaps in the data support provided to researchers. UC Libraries set up an instance of the DMPTool (developed and maintained by the California Digital Library) with shibboleth login to assist researchers with constructing a data management plan. In partnership with IT@UC,UC Libraries also purchased an ORCID institutional membership. This membership enabled integration with UC's institutional repository, Scholar@UC, and an in-house researcher profile product, Research Directory, through the enhanced ORCID API. For project

management support, the Research Data Services team facilitated the setup of a UC instance of the Open Science Framework (OSF - osf.uc.edu) also with shibboleth login. Integrations were developed to streamline metadata transfer between the repository, ORCID and Research Directory Profile and content transfer between the OSF and Scholar@UC. In November 2016, the Data & GIS Collab (derived from collaboration) was established. UC Libraries envisioned this workspace would be used as a collaboration space by researcher to support data intensive analysis in applications such as GIS and data visualization. The library purchased high-end workstations with dual large screens and a 3D mouse and installed software such as ArcGIS and Matlab. Also in November 2017, a visualization Wall will be installed very near to the Data & GIS collab.

FSU and OSU worked on implementing similar tools and technology to support data management needs, with the notable exception that OSU also implemented Altmetric for Institutions.

Professional Development

Providing public access support requires a wide range of skills in scholarly communication and research data management. This requires individuals to spend significant time researching and training in order to develop a baseline level of knowledge to build and implement these new types of programs and services. Professional and informal communities like #datalibs on Twitter, the Datacure community, and the attendees at the Research Data Access and Preservation (RDAP) Summit have become invaluable to skill- and knowledge-building. As with most forms of learning, however, significant skill development in public access support is most likely to occur through hands-on professional practice, rather than through prior training.

At UC, the Director of the Science and Engineering Libraries, the Chief Information Officer for the university and the Dean of Libraries attended the ARL-DuraSpace E-Science Institute in Dec 2012 (http://duraspace.org/). The focus of the institute was to help research libraries assess needs and develop strategic plans for an e-science research support system. Funding for UC Libraries staff involved in these services to engage in professional development came from the budgets supporting library strategic initiatives. Library Administration encouraged staff to seek out non-traditional events such as RDAP and AMIA (American Medical Informatics Association) , National Center for Biotechnology Information (NCBI) training courses, and the Gordon Research Conference on Data Visualization. Informationists attended a Science Boot camp which focused on current research in biology and engineering as well as data management best practices.

One of the first projects initiated by the UC Researcher Services team was an introductory, two-day workshop for UC liaison librarians on data management using the New England Collaborative Data Management Curriculum (NECDMC Jan 2014). The Assistant Director of Research and Informatics and the Director of the Science and Engineering Libraries led 25

members of the UC library staff through five of the seven modules from the NECDMC, gathered feedback on the workshops, and planned how to extend the workshops to research and teaching faculty at the university. This initial workshop became the basis of several modified versions of the workshop offered to the greater UC research community. The Research Data Services team has developed new workshops focused on areas such as data management, bioinformatics, and tools and software such as RedCAP, ArcGIS, Spreadsheet Best Practices, Open Science Framework, and DMPTool.

At OSU, the libraries' internal training program encompasses a set of lesson plans and materials that are not meant to be prescriptive, but rather give liaisons a jumping-off point so that they can use as much or as little as they wish. The materials are accompanied by in-person workshops to go over the handouts and answer any questions liaisons might have. Due to the abundance of information about DMP's and other resources to draw upon, OSU librarians initially conducted this training on their own, drawing from what was already available: http://info.library.okstate.edu/c.php?g=401548. As a final effort, the libraries are also connecting with national movements and initiatives and have been able to bring such programs as the ACRL Building Your Research Data Management Toolkit Roadshow, as a way to help build capacity in these rapidly changing areas. As a result of this training, the RLS unit undertook a mapping, affectionately termed "fun with stickers", activity which allowed them to utilize the data skills matrix presented during the session. The matrix included not only the topic in question, but the degree of depth and complexity to which the topic could extend ranging from individual activities to developing campus-wide policies and practices. This process enabled the unit to determine liaisons' comfort levels with the various areas listed in the matrix and develop a customized training program for the following semester addressing the areas they identified as needing more attention.

Recommendations for (Best) Practice

Action

Throughout this article, there has been a consistent theme of erring on the side of action. Each organization made a strategic decision to identify areas of low hanging fruit (easy wins) and started from there. Depending on institutional culture and resources available, each university chose a different starting point. It is important to note that OSU, UC, and FSU began providing these types of services by identifying low barriers for entry.

The four areas of focus identified in this article are: identifying outreach opportunities, building strategic partnerships, engaging in professional development activities, as well as the allocation of resources toward providing research data services. The first three areas do not require significant financial resources beyond the allocation of staff time. Most library organizations lack a surplus of resources, making the allocation of limited resources a difficult task. One strategy for reprioritizing resources to advance new initiatives such as research data services is to

identify current staff with relevant skillsets and task them with forming an exploratory team of librarians and functional specialists. This team can then identify low barriers of entry, conduct an environmental scan to identify opportunities (and forecast demand) for new services, and make recommendations to leadership on next steps.

Scale

It is also important to plan for scaling these efforts so that they become sustainable and perhaps even renewable approaches to dealing with institution-specific challenges. In this regard, the first step of the process is to determine when there is enough information to make a decision about the future direction of the effort in question, including what additional resources are necessary and what changes need to be made based on the feedback received during the initial phase to transition the initiative from an exploratory project to an established program. In many cases, minor calibrations might be needed in order to tweak a particular aspect that is not working as well as anticipated. In other cases, an organization may conclude that a project will not work on a larger scale or in another context, and it is always better to back out during these early stages than to wait until significant time and effort has been invested into a program or project with larger, and more visible, ramifications. The question of scalability is also different than that of maturity, as a fully developed program covering all necessary elements may still fail to achieve widespread adoption into researcher workflows, with the result that the library is left out of all but a small portion of the equation. Engaging in the strategic efforts described above should assist not only in making this distinction for your institution, but also in examining what level of support and action is needed for each part of your program.

Assessment

A final consideration is that of impact. In spite of the most well-meaning activities, if students and faculty are not taking advantage of these offerings, they are of little value to the institution. First, a determination needs to be made regarding what it is that you would like to measure. Combining both quantitative as well as qualitative data will yield the richest results and allow you to determine how your program is impacting campus efforts for compliance with OSTP requirements. This can range from capturing the number of attendees at workshops and consultations to the number of DMP's reviewed, and can also include data from satisfaction surveys that measure the value provided to participating researchers. It is even more difficult, however, to get a sense of how the institution as a whole is doing towards being in compliance, and that perhaps might be measured by the number of deposits of data sets into the institutional repository and via more in-depth activities such as focus groups, interviews, or even ethnographic studies which directly observe how researchers behave in these contexts. A broad idea that might be taken into consideration in this instance might be something along the lines of "We know our institution as a whole is generally compliant because we have evidence in these three of four areas that activity is occurring, there are grants being funded, and researchers are finding these efforts useful". As Anne Kenney discussed in an article on measuring the impact of liaison work, scaling labor-intensive efforts, quantifying goals and measuring success towards progress, building iteratively, layering successively deeper levels of engagement and impact is a

way to address the nuances which emerge by the very nature of trying to measure abstract notions such as relationships and success of compliance with guidelines (Kenney, 2014).

Commitment

It goes without saying that we are facing a time of uncertainty regarding the future of the OSTP and federal mandates for open data. Libraries may feel a certain amount of unease putting resources and personnel into programs that may have little to no value if the government were to decide to roll-back requirements for making research data open and accessible. Fears of OSTP disappearing or not being in force, and of research infrastructure being acquired by commercial entities such as Elsevier. But it can also be argued that we have entered a new era where even if that were the case, we cannot afford to look back. In light of these developments, it has never been more important for Scholarly Communication and Data professionals to advocate for public access policies at the local level and invest time and resources in the development of locally hosted, nationally interoperable infrastructure for the public dissemination and long-term preservation of publications and data.

Works Cited

"Principles and Strategies for the Reform of Scholarly Communication 1", American Library
      Association, September 1, 2006.
http://www.ala.org/acrl/publications/whitepapers/principlesstrategies
Document ID: e34e8161-fa32-5cd4-19d7-8614fd62e9c3

Boock, M., Zhang, H., & Clark, E. (2016). Article deposit services in support of federal agency
      public access requirements. *Journal of Librarianship and Scholarly Communication*, *4*,
      eP2147. http://doi.org/10.7710/2162-3309.2147

Cox, A. M., & Verbaan, E. (2016). How academic librarians, IT staff, and research
      administrators perceive and relate to research. *Library & Information Science Research*,
      38(4), 319–326. https://doi.org/10.1016/j.lisr.2016.11.004

Doherty, B. b. (2016). Optimizing library services—collaborations and partnerships for the
      modern academic library. *Against The Grain*, 28(6), 53-55.

Erway, R. & Schaffner, J. (2014). Does every research library need a digital humanities center?
      OCLC Research.
      http://www.oclc.org/content/dam/research/publications/library/2014/oclcresearch-digital-h
      umanities-center-2014.pdf.

Finnie, E. (2016, March 23). What organic food shopping can tell us about transforming the
      scholarly communications system. Retrieved from
      http://intheopen.net/2016/03/what-organic-food-shopping-can-tell-us-about-transforming-
      the-scholarly-communications-system/

Hollister, C. V., & Schroeder, R. (2015). The impact of library support on education faculty
      research productivity: An exploratory study. *Behavioral & Social Sciences Librarian*,
      34(3), 97-115. doi:10.1080/01639269.2015.1062584

Johnston, Lisa R., Jake R. Carlson, Patricia Hswe, Cynthia Hudson-Vitale, Heidi Imker, Wendy
      Kozlowski, Robert K. Olendorf, and Claire Stewart. 2017. "Data Curation Network: How
      Do We Compare? A Snapshot of Six Academic Library Institutions' Data Repository and
      Curation Services." *Journal of eScience Librarianship* 6(1): e1102.
      https://doi.org/10.7191/jeslib.2017.1102

Kenney, A. (2014). *Leveraging the Liaison Model: From Defining 21st Century Research
      Libraries to Implementing 21st Century Research Universities*. Retrieved from
      http://www.sr.ithaka.org/wpcontent/mig/files/SR_BriefingPaper_Kenney_20140322.pdf

Meyer, E. (2014). Low-Hanging fruit: Leveraging short-term partnerships to advance academic library outreach goals. *Collaborative Librarianship*, 6(3), 112-120.

Pooley, J. (2017, August 15). Scholarly communications shouldn't just be open, but non-profit too. Retrieved from http://blogs.lse.ac.uk/impactofsocialsciences/2017/08/15/scholarly-communications-shouldnt-just-be-open-but-non-profit-too/