

ISSUE BRIEF

Scholars ARE Collectors: A Proposal for Re-thinking Research Support

November 28, 2018

Danielle Cooper

Oya Y. Rieger



ITHAKA S+R



Ithaka S+R provides research and strategic guidance to help the academic and cultural communities serve the public good and navigate economic, demographic, and technological change. Ithaka S+R is part of ITHAKA, a not-for-profit organization that works to advance and preserve knowledge and to improve teaching and learning through the use of digital technologies. Artstor, JSTOR, and Portico are also part of ITHAKA.

Copyright 2018 ITHAKA. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. To view a copy of the license, please see <http://creativecommons.org/licenses/by-nc/4.0/>. ITHAKA is interested in disseminating this brief as widely as possible. Please contact us with any questions about using the report: research@ithaka.org.

After fifteen years of digging into the research practices of scholars at Ithaka S+R, it is clear that scholars are collectors. We have found that they are creating and amassing increasingly complex personal collections of information over the course of their careers. These collections vary widely depending on the discipline and take many forms, including digitized archival materials, numeric data sets from experiments, audio recordings of interviews, field notes from research sites, and visual materials.

To support scholars, we need to assist them in managing their collections in all their varieties as well as the scholarly workflows behind building these collections.

The commonality across these collecting activities? Scholars are not managing them optimally, whether for their own research needs, sharing with the community, to meet funder requirements, or the long-term preservation of the research. Ninety percent of respondents to the 2015 Ithaka S+R Faculty study reported that they organize data—a significant form of information collection—on their own computers rather than seeking help either within or outside their university for the management and preservation of their data. To support scholars, we need to assist them in managing their collections in all their varieties as well as the scholarly workflows behind building these collections.

To illustrate the complexity and variety of the types of collections scholars build and how they build them, here are some examples from Ithaka S+R studies on evolving research support needs:

- An Asian Studies scholar, may create and then store hundreds of digital images of both archival content and secondary information during their fieldwork.¹

¹ Danielle Cooper, Katherine Daniel, Jade Alburo, Deepa Banerjee, Tomoko Bialock, Hong Cheng, Su Chen et al, "Supporting the Changing Research Practices of Asian Studies Scholars," Ithaka S+R, June 2018, <https://doi.org/10.18665/sr.307642>.

- An agriculture scholar may need to contend with evolving digital formats and requirements for managing, storing, and sharing data collected in the field by multiple researchers.²
- A religious studies scholar may collect a vast personal library over the course of their career.³

While for some researchers information management increasingly falls under the purview of funding regulations, these requirements are inconsistent and there is little oversight. For instance, several funding agencies such as the National Science Foundation (NSF) and the National Endowment for the Humanities (NEH) require data management plans (DMPs) to articulate how data will be managed and made accessible, but federal science agencies are uneven in their approach and policies on supporting data management.⁴ And, research suggests that in the absence of strict enforcement, even scholars who are mandated to share their data by funders rarely comply.⁵ Although large scientific datasets around environmental change, healthcare, and migration patterns get quite a bit of attention, the long tail of research data represents a large and diverse set of content that underlies scholarship in many fields. As Jane Kamensky states, “in an era where it is fashionable to talk about the power and promise of ‘big data,’ much scholarship is still ‘tiny data,’ gathered by hand, and involves touching and noticing small details.”⁶

In this diffuse landscape how can we leverage insight about current scholarly collecting habits towards making strategic decisions at the institutional level about the future models of research support? To what extent and how should academic institutions, agencies, and other external entities be involved in supporting and leveraging the work of scholars as collectors? Can we adopt a wider strategy that accommodates the diverse

² Danielle Cooper, Sarah Bankston, Marianne S. Bracke, Beth Callahan, Hui-Fen Chang, Leslie M. Delserrone, Florian Diekmann, et al, “Supporting the Changing Research Practices of Agriculture Scholars,” Ithaka S+R, June 2017, <https://doi.org/10.18665/sr.303663>.

³ Danielle Cooper, Roger C. Schonfeld, Richard Adams, Matthew Baker, Nisa Bakkalbasi, John G. Bales, Rebekah Bedard, et al, “Supporting the Changing Research Practices of Religious Studies Scholars,” Ithaka S+R, February 2017, <https://doi.org/10.18665/sr.294119>

⁴ Adam Kriesberg, Kerry Huller, Ricardo Punzalan, Cynthia Parr, “An Analysis of Federal Policy on Public Access to Scientific Research Data,” *Data Science Journal*, Vol.16, no 27 (2017): 1–13, <https://doi.org/10.5334/dsj-2017-027>.

⁵ Jessica L. Couture, Rachael E. Blake, Gavin McDonald, Colette L. Ward, “A Funder-Imposed Data Publication Requirement Seldom Inspired Data Sharing,” *PLoS ONE*, Vol. 13, no 7 (2018), <https://doi.org/10.1371/journal.pone.0199789>.

⁶ Constance Malpas and Merrilee Proffitt, *The Transformation of Academic Library Collecting: A Synthesis of the Harvard Library's Hazen Memorial Symposium*, Dublin, OH: OCLC Research, 2017, p. 14, <https://www.oclc.org/research/publications/2017/oclcresearch-hazen-symposium.html>.

collections scholars create for a range of purposes that takes into account differing imperatives for access? What are the benefits, trade-offs, and other key questions that must be taken into account when considering how to support scholarly collecting?

In this diffuse landscape how can we leverage insight about current scholarly collecting habits towards making strategic decisions at the institutional level about the future models of research support?

Understanding the Scholar as a Collector

The Scholars' Collection Defined

In order to comprehend the strategic possibilities for supporting scholarly collecting, it is important to first understand the collecting activities scholars are currently undertaking. Scholars amass vast collections of analog and digital materials, which result from the processes of designing research methodologies, conducting research, gathering data/information, analyzing/interpreting findings, developing various outputs for communication, seeking input from colleagues, and sharing their work with different audiences. Therefore, scholars' research methods are much broader and richer than simply finding, evaluating, utilizing, and managing information sources (see Figure 1). And, the personal collections they amass as part of their research may also be used towards their teaching.⁷

While scholars across fields produce and amass collections of content over the course of their research, the policies and services relating to scholarly collecting activities typically focus on one form of content: data. The term "data" continues to be closely associated with information produced and utilized through various empirical and quantitative research methodologies that are used predominantly by social scientists and scientists. These activities may fall under the purview of data management plans, and increasingly

⁷ Diane Harley, Jonathan Henke, Sharon Lawrence, Ian Miller, Irene Perciali, I., & David Nasatir, "Use and Users of Digital Resources: A Focus on Undergraduate Education in the Humanities and Social Sciences," UC Berkeley: Center for Studies in Higher Education, 2006, <https://escholarship.org/uc/item/8c43w24h>.

data sharing requirements, as mandated by funders. However, these kinds of requirements do not provide holistic support for scholars as they collect data and other forms of information over the course of their careers, from various funding sources and institutional locales.

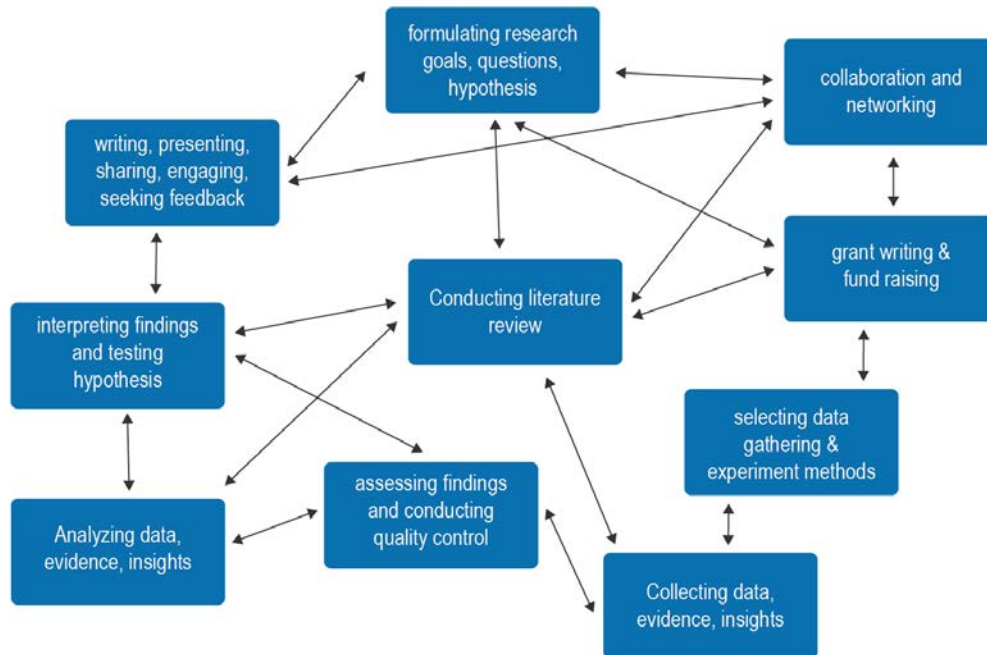
While scholars across fields produce and amass collections of content over the course of their research, the policies and services relating to scholarly collecting activities typically focus on one form of content: data.

It is also important to recognize that many scholars, particularly in the humanities, do not consider the evidence or representations of evidence they use in the work as “data” per se. However, through the very nature of their research, they are creating data, which, as Christine Borgman describes, are “representations of observations, objects, or other entities used of evidence of a phenomena for the purposes of research or scholarship.”⁸ In contrast, the NEH defines humanities data as materials generated or collected during the course of conducting research such as citations, software code, algorithms, digital tools, documentation, databases, geospatial coordinates (for example, from archaeological digs), reports, and articles.⁹ The variation in the interpretation of terminology needs to be taken into consideration when promoting and branding research data management services. Otherwise, some humanists and social scientists, who often do not perceive their work as involving “data,” may not perceive the services pertinent to their needs.

⁸ Christine L. Borgman, *Big Data, Little Data, No Data: Scholarship in the Networked World*, Boston: The MIT Press, 2015.

⁹ Data Management Plans for NEH Office of Digital Humanities Proposals and Awards, https://www.neh.gov/sites/default/files/2018-06/data_management_plans_2018.pdf.

Figure 1: The non-linear and iterative nature of scholarly research



In this age of born-digital information, many researchers continue to rely on site visits to physical archives and routinely collect information by digitizing primary materials. Some gather data during fieldwork in various formats and organize related content extracted from secondary resources owned or accessed through libraries and other sources.¹⁰ As humanists and social scientists engage in digital scholarship, they create their own personal archives and establish workflows, train student assistants in digitizing content, and make decisions on storage mediums, file formats, and metadata. Art historians and many others who work with multimodal content rely on visual image databases in identifying relevant content and use multimedia platforms such as Scalar to create online portals in support of their research projects and courses. They often merge content they

¹⁰ Danielle Cooper, Roger C. Schonfeld, Richard Adams, Matthew Baker, Nisa Bakkalbasi, John G. Bales, Rebekah Bedard, et al, "Supporting the Changing Research Practices of Religious Studies Scholars," pp. 37-38; Danielle Cooper, Katherine Daniel, Jade Albuero, Deepa Banerjee, Tomoko Bialock, Hong Cheng, Su Chen et al. "Supporting the Changing Research Practices of Asian Studies Scholars," pp. 15-17; 33-34 .

create with other pertinent material shared by peers or discovered on the web. Archaeologists and anthropologists are generating digital data in unprecedented volumes and varieties of formats as they engage with different physical sites and populations. In various forms of science, there is an increasing reliance on born-digital data created through remote sensing or at research laboratories using digital tools such as simulations, 3D modeling, and geospatial applications. Abundant online information resources coupled with more effective retrieval and information management tools are making it easy to capture and save information. For instance, citation management applications such as Zotero and Mendeley offer features to save copies, organize, annotate, and markup relevant items.¹¹

Sharing this data is not always practical—scholars may not have time to prepare the data in a format usable by others and privacy or IP-rights concerns may preclude open scholarship. For instance, public health scholars are cautious about sharing human subject data and ensuring the privacy of participants of studies. Sometimes scholars refrain from sharing due to trust issues and their hesitation about the potential misuse and de-contextualization of data. Also, it is not uncommon for scholars to want to protect their scholarship from being “scooped” by others. And, there are few professional incentives or rewards for sharing.¹² For some scholars, preserving information for the long term is not possible or necessarily desirable—they may need to purge information at some point to fulfill data gathering protocols.¹³ Even if such data are not publicly accessible, they represent a wealth of valuable knowledge that need to be efficiently and effectively managed.¹⁴

¹¹ Kornelia Tancheva, Gabriela C. Gessner, Neely Tang, Erin Eldermire, Heather Furnas, Darcy Branchini, Gail Steinhart, and Nancy F. Foster, “A Day in the Life of a (Serious) Researcher: Envisioning the Future of the Research Library,” Ithaka S+R, March 2016, <https://doi.org/10.18665/sr.277259>.

¹² Alice Meadows, “To Share or not to Share? That is the (Research Data) Question,” *Scholarly Kitchen*, 2014, <https://scholarlykitchen.sspnet.org/2014/11/11/to-share-or-not-to-share-that-is-the-research-data-question/>.

¹³ Danielle Cooper, Roger C. Schonfeld, Richard Adams, Matthew Baker, Nisa Bakkalbasi, John G. Bales, Rebekah Bedard, et al, “Supporting the Changing Research Practices of Religious Studies Scholars,” p. 29.

¹⁴ Wolfram Horstmann, Amy Nurnberger, Kathleen Shearer, and Malcolm Wolski, “Addressing the Gaps: Recommendations for Supporting the Long Tail of Research Data,” RDA Long Tail of Research Data Interest Group, 2017, https://www.rd-alliance.org/system/files/Supporting%20Long%20Tail%20of%20Research%20Data%2C%20August%202017_0.pdf; Bryan Heidorn, “Shedding Light on the Dark Data in the Long Tail of Science,” *Library Trends*, Volume 57, Number 2, Fall 2008, pp. 280-299.

Scholars' Collecting Challenges and Needs

Articulating the scope of scholarly collecting activities uncovers the challenges and needs scholars face when undertaking their research. Understanding these challenges and needs are important when thinking through the possibilities for better supporting this work strategically in the years to come. While the type of information produced, gathered, and analyzed varies, uniformly scholars feel that they are unable to manage the vast array of information they gather in the process of conducting research.¹⁵ Largely, they design and maintain their collections independently and idiosyncratically. They receive minimal training or guidance from those with expertise. Many rely on stopgap or suboptimal measures as they try to juggle multiple projects with the pressure of pending timelines and productivity requirements. Figure 2 shows the types of challenges scholars face.

Uniformly scholars feel that they are unable to manage the vast array of information they gather in the process of conducting research.

¹⁵ Danielle Cooper, Roger C. Schonfeld, Richard Adams, Matthew Baker, Nisa Bakkalbasi, John G. Bales, Rebekah Bedard, et al, "Supporting the Changing Research Practices of Religious Studies Scholars," p. 25; Danielle Cooper, Sarah Bankston, Marianne S. Bracke, Beth Callahan, Hui-Fen Chang, Leslie M. Delserone, Florian Diekmann, et al, "Supporting the Changing Research Practices of Agriculture Scholars," Ithaka S+R, June 2017, p. 29, <https://doi.org/10.18665/sr.303663>; Matthew Long and Roger Schonfeld, "Supporting the Changing Research Practices of Chemists," Ithaka S+R, February 2013, <https://doi.org/10.18665/sr.22561>.

Figure 2: Scholars' Collections: Challenges and Needs

Stage	Challenges
Storing	<ul style="list-style-type: none"> • Gathering and analyzing heterogeneous data • Working with multiple storage approaches concurrently • Lack of guidance for selecting storage and backup configurations • Navigating and fulfilling institutional requirements
Organizing, Managing, Discovering	<ul style="list-style-type: none"> • Idiosyncratic approaches • Quality control requirements • Cross-institutional collaborations • Skill gaps • Lack of infrastructure for ephemeral information • Need for internal discovery tools
Preserving	<ul style="list-style-type: none"> • Ambivalence towards long-term preservation • Opportunistic and idiosyncratic approaches • Keeping up with evolving formats • Transitioning personal collections upon retirement and departures
Sharing	<ul style="list-style-type: none"> • Suboptimal platforms for sharing • Lack of knowledge and capacity to devote to sharing • Uncertainty or reluctance about the value of sharing

Storing

- *Gathering and analyzing heterogeneous data.* Scholars work with a variety of data types and sources in a given project and struggle to obtain and manage them in a cohesive manner. Collecting, staging, and processing data in the form of text, images, AV, statistics, etc. for analysis and interpretation purposes requires adequate space and access to tools.
- *Working with multiple storage approaches concurrently.* Scholars typically employ multiple storage solutions concurrently, across digital and analog formats. It is difficult for them to keep up with distributed and unorganized storage arrangements—some struggle to remember where they stored information.¹⁶ Some scholars continue to rely on hard-copy and are having difficulty in storing and retrieving their data. This is especially a problem for those who gather evidence and data from difficult-to-access archives and research sites.
- *Lack of guidance on how to select short-, medium-, and long-term storage and backup configurations.* Scholars often choose how they store the content associated with their research based on their familiarity and comfort level with readily available options rather than a careful selection of a storage medium that fulfills their requirements.
- *Navigating and fulfilling institutional requirements.* Depending on the kind of data scholars are working with, they may have to utilize secure servers and conform to specific retention and purge policies, as mandated by their institutions or external funders.¹⁷

Organizing, Managing, and Discovering

- *Idiosyncratic approaches.* Content is typically organized idiosyncratically based on systems/devices that are easily accessible by researchers, increasing in complexity and multiplying in size over time as the researcher amasses more content and format evolve.¹⁸

¹⁶ Danielle Cooper, Katherine Daniel, Caitlin Bakker, Jaime Blanck, Chris Childs, Ann Gleason, Rosie Hanneke, et al, "Supporting the Changing Research Practices of Public Health Scholars," Ithaka S+R, December 2017, p. 21 <http://doi.org/10.18665/sr.305867>; Danielle Cooper, Roger C. Schonfeld, Richard Adams, Matthew Baker, Nisa Bakkalbasi, John G. Bales, Rebekah Bedard, et al, "Supporting the Changing Research Practices of Religious Studies Scholars," p. 28.

¹⁷ Danielle Cooper, Katherine Daniel, Caitlin Bakker, Jaime Blanck, Chris Childs, Ann Gleason, Rosie Hanneke, et al, "Supporting the Changing Research Practices of Public Health Scholars," p. 21.

¹⁸ Roger Schonfeld and Matthew Long, "Supporting the Changing Research Practices of Art Historians," Ithaka S+R, April 2014, p. 23-25, <https://doi.org/10.18665/sr.22833>http://www.sr.ithaka.org/wp-content/uploads/2015/08/SR_Support-Changing-Research-ArtHist_20140429.pdf.

- *Quality control requirements.* Managing data while conducting research (and performing other academic duties) is challenging, especially finding sufficient time for quality control to monitor, clean, and prevent mistakes in recording.¹⁹
- *Cross-institutional collaborations.* Data management tasks in collaborative research put further strain on researchers as they may need to reconcile different data gathering and quality control strategies, as well as try to adhere to different institutional data creation and sharing policies.²⁰ For instance, developing common workflows for managing sensitive data is complicated.
- *Skill gaps.* Unfamiliarity with data management techniques, as one of the public health scholars stated, “I am not a data manager ... that takes a different skill.”²¹ Some scholars rely on research assistants and post-docs for help although these individuals might not have adequate skill sets and experience.
- *Lack of infrastructure for ephemeral information.* Capturing and storing information from social media and web-based resources, some are ephemeral in nature, “easily accessible but difficult to archive.” If scholars fail to immediately capture content such as a Twitter post, they worry that they might not be able to find it again. As a result, they download videos, reports, and use snipping tools to save social media posts and comments.
- *Need for internal discovery tools.* Discovery needs are not only for external sources through search engines, information portals, or library catalogs. Scholars need to be able to discover content within their own collections as they reuse materials. They are both creators and users of their own collections.

Preserving

- *Ambivalence towards long-term preservation.* Scholars don’t have preservation in mind upstream when they are collecting or creating data—which becomes an obstacle for preservation.²² Scholars are generally unaware of preservation issues—for instance they might use Academia.edu without being concerned about the service’s business model and without an awareness of the company’s lack of long-term plans for maintaining such sites.²³ Scholars also

¹⁹ Danielle Cooper, Sarah Bankston, Marianne S. Bracke, Beth Callahan, Hui-Fen Chang, Leslie M. Delserson, Florian Diekmann, et al, “Supporting the Changing Research Practices of Agriculture Scholars,” p. 18.

²⁰ Ibid.

²¹ Danielle Cooper, Katherine Daniel, Caitlin Bakker, Jaime Blanck, Chris Childs, Ann Gleason, Rosie Hanneke, et al, “Supporting the Changing Research Practices of Public Health Scholars,” p. 21.

²² Danielle Cooper, Roger C. Schonfeld, Richard Adams, Matthew Baker, Nisa Bakkalbasi, John G. Bales, Rebekah Bedard, et al, “Supporting the Changing Research Practices of Religious Studies Scholars,” p. 29.

²³ Ibid. p. 30.

need considerable support around identifying which materials will and will not be valuable in the long-term, both for their own use and beyond.

- *Opportunistic and idiosyncratic approaches.* Some scholars hold on to information not intentionally but out of habit or uncertainty about whether or not they should be disposing it.²⁴ Some keep information with the mentality of, “you’ll never know when you’ll need it” and hesitate to weed their collections.²⁵ As information piles up, it becomes increasingly difficult to find what is needed.
- *Keeping up with evolving formats.* Scholars lack the resources and capacity to maintain personal collections over time in the face of evolving file formats and software dependencies through strategies such as refreshing (copy digital information from one storage medium to another) and migration (convert data from one technology to another).²⁶
- *Transitioning personal collections upon retirement and departure.* There is a lack of consistent policies and programs for appraising the personal information collections that scholars will no longer be maintaining and determining which information should be discarded and retained and by whom.²⁷

Sharing

- *Suboptimal platforms for sharing.* Many scholars choose to post their research on academic social networks because they are easier to use and navigate, and they also feel that these are more connected to their networks. However, many scholars are also increasingly concerned about the proprietary business models of some of the most popular networks. In contrast, scholars are more reticent to engage with institutional repositories (IRs) because they find them siloed and not user-friendly. Also, IRs often put emphasis on broadening access to institutional scholarship through open access, which might not dovetail with the professional needs of some scholars.²⁸ The proliferation of platforms is also challenging to navigate.

²⁴ Ibid. p. 29.

²⁵ Danielle Cooper, Katherine Daniel, Jade Albuero, Deepa Banerjee, Tomoko Bialock, Hong Cheng, Su Chen et al, "Supporting the Changing Research Practices of Asian Studies Scholars," p. 21; Danielle Cooper, Sarah Bankston, Marianne S. Bracke, Beth Callahan, Hui-Fen Chang, Leslie M. Delserone, Florian Diekmann, et al, "Supporting the Changing Research Practices of Agriculture Scholars," p. 18.

²⁶ Danielle Cooper, Roger C. Schonfeld, Richard Adams, Matthew Baker, Nisa Bakkalbasi, John G. Bales, Rebekah Bedard, et al, "Supporting the Changing Research Practices of Religious Studies Scholars," p. 29.

²⁷ Danielle Cooper, Sarah Bankston, Marianne S. Bracke, Beth Callahan, Hui-Fen Chang, Leslie M. Delserone, Florian Diekmann, et al, "Supporting the Changing Research Practices of Agriculture Scholars," p. 19, 29.

²⁸ Caitlin Bakker, Shanda Hunt, "Information Science Solutions to Communicating Public Health Research Findings to the Public," Medical Library Association Annual Conference, May 2018, Atlanta, GA, <https://conservancy.umn.edu/handle/11299/199761>.

- *Lack of knowledge and capacity to devote to sharing.* Some scholars want to share data with their peers however they are either not equipped to organize and archive information in a way that can be shared in a meaningful and useful way or cannot make the time to make their information sharable.²⁹
- *Uncertainty or reluctance about the value of sharing.* For many scholars sharing research openly may be inconsistent with their professional needs, or they are not yet clear on the benefits of sharing content they collect. Also, some are unsure and concerned about violating privacy, confidentiality, or copyright policies.

The Current Landscape for Supporting Scholarly Collecting

Some policies, services and tools have been developed that intersect with scholars' personal collecting activities, largely in conjunction with the worldwide movement towards open scholarship. What follows below is an outline of the four key stakeholders (funders, open data advocacy groups, external tool and service providers, and academic institutions) that support scholarly collecting and how they approach this support.

Funders

Funders have a major role in setting policy directions that impact scholarly workflows, including their collecting activities. Governments are a main branch of research funding and we include them in this category in recognition that their policies in this area have been overwhelmingly tied to their funding requirements. An emerging trend is for funding agencies to require that scholarly outputs, including data, produced through their funding, are deposited in public access repositories. The rationale for this policy position is that it maximizes the societal benefits of the research, both in terms of fostering transparency and leveraging the insights for future research. There have been mounting questions about the extent to which awardees appropriately adhere to these policies due to the lack of enforcement efforts by the funders. Funders have also been critiqued for focusing funding on innovation in digital scholarship while not funding mechanisms for maintaining and sustaining those endeavors beyond their initial development phase.

²⁹ Danielle Cooper, Katherine Daniel, Jade Alburo, Deepa Banerjee, Tomoko Bialock, Hong Cheng, Su Chen et al., "Supporting the Changing Research Practices of Asian Studies Scholars," p. 22.

Open Data Advocacy Groups

These groups perceive data as a public good and promote that wherever possible data should be open, accessible, and reusable, which has some potential to impact scholarly collecting habits in the broadest sense, depending on how data is defined. The rationale for this position is that open data facilitates interdisciplinary research and collaboration, leading to new discoveries to address society's grand challenges. They also believe that open data will enhance transparency and replicability, improving the quality of research. These positions are focused on the impacts to greater society as opposed to individual researchers, institutions, or funders. As funder policy shifts suggest, these groups have had some success in advocating for their position, particularly in the United Kingdom and Western Europe. At this point, these groups have focused on a narrower definition of data as opposed to scholarly collections in the broader sense.

External Tool and Service Providers

This group constitutes vendors and not-for-profits that are creating tools and services to support scholarly workflows, including some components of their collecting activities. The tools and services are marketed to scholars as a means of increasing their efficiency and effectiveness. These groups seek to commercialize scholarly workflows not only by directing services and tools to scholars, but also by bundling these offerings and selling to campus groups that have some oversight over their scholars' research towards establishing a competitive advantage. Some of these entities are particularly focused on how their services and tools can be further monetized by selling research content embedded in their systems and/or through analytics gathered on their usage.

Academic Institutions

Academic institutions are arguably the most reactionary of the groups associated with supporting scholars as collectors, developing policies and services largely in response to changing funder expectations around research data management and sharing. There is a lack of consensus of who, and to what extent, has responsibility and expertise towards supporting scholarly collecting as a broader activity, leading to lack of centralized coordination in what external services and tools are purchased and how they are deployed for scholars' use.

Many research institutions are also now offering their own in-house services to help scholars with the development of research data management plans and to support data management activities at various stages of the research process. Often coordinated by

academic libraries, these services are resource intensive, require various subject and functional expertise, and are difficult to sustain.³⁰ Therefore only a few libraries are able to provide the required services and expertise by themselves. Furthermore, these services are generally focused on research data management needs as they intersect with funder requirements, which is only relevant to a particular subset of some scholars' collecting activities.

Academic institutions have their own archives and special collections, through which some scholars affiliated with those institutions will sometimes donate some or all of their personal collections.

What's Next for Academic Institutions?

Scholars are engaged in creating increasingly complex collections of content in conjunction with their research, and various entities are currently working to support some portion of the work. However, this is not being undertaken in a consistently coordinated fashion nor do these various efforts currently support the holistic process of scholarly collecting. Academic institutions need to be more engaged in setting the agenda. There is a lot at stake for academic institutions if they do not take a more proactive strategic approach to supporting and leveraging scholars' collections including:

1. Loss of control or ownership of information and knowledge generated within an organization (this could even be seen at a micro-level such as within departments, doesn't always imply university-wide practices).³¹
2. Decline in research productivity due to faculty not being able to leverage their own data with effective tools, or content being locked into proprietary systems.
3. Barriers to implementing institutional information policies such as accessibility, privacy, confidentiality, etc.
4. Risks associated with not being able to connect tools with related components of the university IT infrastructure because of incompatibility issues (for instance, being able to easily move images created/stored to a learning management system for a specific course module).

³⁰ Lisa R Johnston, Jake Carlson, Cynthia Hudson-Vitale, Heidi Imker, Wendy Kozlowski, Robert Olendorf, Claire Stewart, "Data Curation Network: A Cross-Institutional Staffing Model for Curating Research Data," 2017, <https://drive.google.com/file/d/0B5Dm3XFQloc4YTh2Q1lwV0MlWFEE/view>.

³¹ Sarah Mason, "Electronic Lab Notebooks and Digital Preservation, Parts 1 & 2, Digital Preservation at Oxford and Cambridge, August 2, 2018, <http://www.dpoc.ac.uk/2018/08/02/elns-part-1/>; <http://www.dpoc.ac.uk/2018/09/12/elns-part-2/>.

5. Decreasing reliance or recognition for the relevant institutional service providers as reliance on outside vendors grow.³²
6. Duplication of work or service gaps across campus units due to uncertainty and lack of oversight of where the locus of responsibility for research support resides institutionally.

We offer the following recommendations in recognition of the importance of what is at stake for academic institutions as they develop strategies to support scholars as collectors.

Design and promote services directly targeting scholars as curators:

- Provide resources to help scholars select institutional and external storage options in support of short-, medium-, and long-term management.
- Offer easy-to-use tools such as Tropy that enable scholars to better organize information collected like photographs both to bring efficiencies to their workflows and also to make any future potential sharing arrangement easier.³³
- Connect researchers with data analysis expertise, within and outside of an academic organization.
- Create assistant positions with data management skills to participate in scholars' projects, especially by being included on grant-funded projects.
- Explore selecting and supporting platforms that also create possibilities for discovery and sharing personal collections with others.³⁴

Reframe research data management to present a more inclusive scope:

- Rebrand research data management services to cover information management more broadly (e.g. "research data and information services").
- Re-situate the locus of responsibility for these offerings within the central administration and clearly delineate how and to what extent various units are involved in these undertakings.

³² David Matthews, "Fears over potential publisher 'monopoly' in digital research tools," *Times Higher Education*, August 16, 2018, <https://www.timeshighereducation.com/fears-over-potential-publisher-monopoly-digital-research-tools>.

³³ See <https://tropy.org/>.

³⁴ Personal collections are growing in importance to scholars and pose challenges in terms of their discovery. See Roger C. Schonfeld, "Does Discovery Still Happen in the Library? Roles and Strategies for a Shifting Reality," *Ithaka S+R*, September 2014, p. 11, <https://sr.ithaka.org/blog/does-discovery-still-happen-in-the-library-roles-and-strategies-for-a-shifting-reality/>.

- Offer services customized for scholars who need to manage their own information for shorter durations or internal purposes—many scholars, especially humanists and social scientists, are not dependent on external funding for research.
- Recognize that a majority of the research and associated information gathering and management processes by researchers may not be appropriate for open sharing and does not involve public access mandates.
- Consider researchers' needs holistically beyond open science projects as the majority of scholars are involved in small-scale and time-bound initiatives.

Promote digital information fluencies for the entire academic community:

- Expand the scope of services to include all kinds of researchers as opposed to limiting to students
- Provide proactive outreach to the researchers to raise awareness of new digital research tools—digital tools are emerging far faster than researchers are adapting their practices to take advantage of them. The difficulties in developing awareness about and learning how to work with new technologies prevent scholars from adapting available tools for their information management needs.
- Offer coordinated consultancies and services to the community. Integration is an important concept as currently many universities have distributed centers with specific missions that overlap and align well--without a strong mandate to work together (beyond collaboration, build dependencies).
- Train graduate students on information management not only as future scholars but also as assistants of researchers. Many faculty members rely on their graduate students to introduce them to approaches to information management. Training and ongoing support in technologies and techniques are needed.

Seek opportunities to collaborate in the rapidly emerging research workflow support tool market:

- Work with vendors to improve the usability workflow tools in support of content management, e.g., citation management tools continue to evolve but the majority of scholars do not experience the full benefits of these improvements by virtue of their design.
- Engage the administrative and academic units in selecting, managing, and supporting research workflow support applications to systematically consider disciplinary requirements and support needs.³⁵

³⁵ Roger C. Schonfeld, "Research Infrastructure and the Strategic Decisions of Universities," Ithaka S+R, January 10, 2018, <https://sr.ithaka.org/blog/research-infrastructure-and-the-strategic-decisions-of-universities/>.

- Develop campus-wide systematic and reliable support infrastructures for “research technologies,” similar to the one that exist for “academic technologies.”

Develop university-wide policies, mechanisms and guidance for preserving information collected and produced by scholars throughout their research lifecycle:

- Assume a proactive role in these areas because who has responsibility for scholars’ information management, storage, and preservation while the information is still in the scholars’ custody is unclear.
- Take into consideration the needs of researchers at various stages of their profession, for instance early career researchers move between institutions frequently and sometimes need to carry on ongoing research projects.
- Provide adequate and user-friendly digital storage solutions to remediate scholars’ current prevalence of de-centralized storage and preservation approaches and develop programs to help manage their information.
- Focus on tools designed specifically to manage and store web-based content that also comply with data protection regulations.
- Experiment with providing faculty computer forensics and format migration services to understand needs and resource requirements. Offer consultancies and training sessions on personal archiving.
- Improve IRs for scholarly workflow and publication needs. For example, make it easier to discover through search engines and aggregators like Google Scholar, BASE, CORE, or OA finding tools like Unpaywall, OAButton, Wikipedia, SHARE, OpenAIRE etc. Accommodate different levels of access rights, more granular than internal and external (meaning being able to design access rights at project or individual researcher level).

Recognize that there is still a role for physical media in research:

- Develop guides and tools that accommodate the collection and creation of physical materials and artifacts by researchers.³⁶

³⁶ For instance, designed at the University of North Carolina at Chapel Hill, *Archivist in a Backpack*, a toolkit with basic tools for oral history and material archives, is making archive creation more accessible by offering resources that can easily launch community partners on memory projects: <https://hyperallergic.com/444554/archivist-in-a-backpack/>.

Final Thoughts and Future Directions

At the heart of our recommendations is the need to reframe “research support” in a more holistic way in order to ensure that academic institutions are strategically positioned in an evolving research support landscape that is increasingly focused on managing and leveraging scholarly workflows. We aim to expand the research community’s understanding of scholars’ daily practices in creating and amassing increasingly complex personal information collections over the course of their research. Deepening our understanding of the full research stages will be instrumental in identifying required skill sets and factoring in such needs in professional development and recruitment efforts of related service providers.

When exploring the possibilities of creating and managing personal collections with scholars, the issue of discovering others’ personal collections and sharing one’s personal collection with others was not generally presented as a challenge by those scholars. This is likely due to the fact that in the current ad hoc landscape of personal research collection management, issues pertaining to discovery and sharing of personal collections remains seemingly far off on the horizon of possibilities. However, with vendors, the open source community, and academic institutions increasing attention to scholarly workflows, combined with ongoing interest in funders, governments, and cultural institutions in making research content and collections widely open and accessible, we anticipate that this issue will become increasingly important. As noted above, evolving pedagogical methodologies that favor engaged learning through practice and engagement will also play a role in amplifying the importance of this program area.

This is why we recommend that any stakeholder involved with developing workflow support platforms or the governance of their use attend to the possibilities for building out functionality for discovery and sharing. Indeed, the potential for leveraging personal collections more widely may ultimately be one of the strongest justifications for investing external resources into scholars’ personal collection management. The value can manifest itself in multiple levels. For instance, introducing cost-efficiencies by making it easier for scholars to reuse their collections for their own purposes or with their collaborators (or within departments), or leveraging content for learning and teaching activities. How and to what extent scholars’ personal collections can be used by others, not only including other researchers, but also cultural institutions and the public at large, is grand challenge necessitating further exploration beyond the scope of this issue brief.