ISTL Issues in Science and Technology Librarianship

Tips from the Experts

Scaling Open Science Services in Academic Libraries Through Expertise Building, Partnerships, and Postdoc Hiring

Melanie A. Gainey
STEM Librarian; Director of Open Science Program
University Libraries
Carnegie Mellon University
Pittsburgh, Pennsylvania
mgainey@andrew.cmu.edu

Huajin Wang
STEM Librarian
University Libraries
Carnegie Mellon University
Pittsburgh, Pennsylvania
huajinw@andrew.cmu.edu

Abstract

Academic library programs and services can be challenging to sustain in the face of staffing turnover, financial constraints, and gaps in expertise. In this reflection, the authors discuss strategies for sustaining programs and how to foster growth of services over time. The success and evolution of a dedicated open science program at Carnegie Mellon University Libraries has depended upon forming partnerships, both internal and external to the university, and hiring postdocs with deep disciplinary expertise and fresh perspectives on open science. The authors also share practical advice for successfully recruiting postdocs in academic libraries.

Keywords: Open science, Hiring practices, Program sustainability, Library services

Recommended Citation:

Gainey, M. A., & Wang, H. (2025). Scaling open science services in academic libraries through expertise building, partnerships, and postdoc hiring. *Issues in Science and Technology Librarianship, 111*. https://doi.org/10.29173/istl2918

Introduction

Open science is a growing movement to transform the way in which research is conducted, with the goal of making it more accessible, transparent, reproducible, and reusable. Also, commonly referred to as open research or scholarship, open science is an umbrella term that encompasses principles and practices designed to make research outputs publicly available and licensed for reuse, such as open access publishing, open data, open educational resources, and open source code and software. Good practices in research data management are essential for making data open and reproducible. In the mid-2010s, the "reproducibility crisis" revealed that many studies in psychology and biomedicine were not reproducible, catalyzing the rise of the open science movement as a corrective mechanism led by the research community itself (Baker, 2016; Munafó et al., 2017). The movement has since expanded to promote transparency and improved research practices across disciplines, including those beyond STEM (science, technology, engineering, and math).

Against this backdrop, the Open Science & Data Collaborations Program at Carnegie Mellon University (CMU) Libraries was established in 2018 by the authors and their former colleague, Ana Van Gulick, to support the emerging need for transparent and reproducible research. Since the advent of the program, one of the main challenges has been sustainability, particularly during and in the aftermath of the changes in staffing created by the COVID-19 pandemic.

In this article, we first briefly describe how the program was established. We then focus on how we have approached sustainability of the program. Finally, we highlight a major priority since 2022, which is hiring and training open science postdoctoral associates to improve capacity and diversify our offerings. We hope our strategies are helpful for other libraries navigating the development and sustainability of open science initiatives and expanding their capacity and skill sets for open science or other topics that benefit from research experience.

Program Formation

We joined CMU Libraries as liaison librarians with life science research backgrounds in 2017. Our arrival coincided with increased public attention to the reproducibility crisis and the rise of the open science movement as a means to make research more reproducible and trustworthy (Baker, 2016; Committee on Toward an Open Science Enterprise et al., 2018). At that time, there were few comprehensive open science programs in academic institutions in the United States and, as far as we are aware, none based in academic libraries. Existing efforts in academic libraries were typically limited to supporting facets of open science such as open access publishing, open educational resources, and research data management (Scotti et al., 2025; Wang et al., 2022), despite the growing need and mandates for data sharing and reproducible methods (AAU-APLU Public Access Working Group, 2017). As new librarians who had first-hand experience in the reproducibility issues in research, we wanted to leverage the libraries' position in supporting research and develop programming and resources that help researchers navigate open and reproducible practices throughout the research lifecycle. Since CMU Libraries had already started establishing infrastructure, including the institutional repository and the Article Processing Charge (APC) fund (which partially or fully covers the cost of open access publishing), we first prioritized community building and research tools that support the process of daily research.

Our first major initiative was a collaboration with a faculty member from the Department of Biological Sciences at CMU to obtain a grant to host the inaugural Open Science Symposium, a two-day event where researchers from CMU and other academic institutions, along with stakeholders such as data professionals, publishers, and funders, came together to share success stories, showcase open source tools and projects, and discuss challenges and incentives of open science. It was our first step towards building an open science community at CMU. The symposium has since become a flagship researcher-facing event that takes place biennially. We then worked with mission-aligned digital tools and service providers to make sure that high quality open science infrastructure was available to the campus community, so that when researchers have the intention to practice open science, the necessary support and resources are already in place. These digital tools include Open Science Framework, a project management and collaboration tool; LabArchives, an electronic research notebook; and protocols.io, a platform for creating and sharing research protocols. In close collaboration with the data services team and the service provider, we helped to evaluate, improve, and encourage adoption of KiltHub, our institutional repository for open data and all other scholarly outputs produced by CMU researchers. Recognizing the importance of basic computational and coding skills in open and reproducible research, we also prioritized training workshops that support data science. As an institutional member of the Carpentries, a non-profit organization that supports instruction of foundational skills in open source coding languages and data science tools, we offered extended workshops in foundational coding in Python and R to our regional research community.

This initial array of activities formed the foundation of the programming we have today, built an initial user base, and helped us better understand the communities we work with and their needs in daily research. As the program matured and our user base grew, we added other mission-aligned activities that provide end-to-end support for open science activities across the research lifecycle. We further formalized the program's strategy and structure, which has four major pillars: digital tools, training, events, and collaborations. The program's framework and our initial approach to assess the success of the program is described in Wang et al., (2022) and Gainey et al., (2025).

A significant challenge to maintaining a library-based comprehensive open science program is sustainability. The challenge is three-fold: on-going financial support for digital tools, membership, and events cost; sufficient personnel capacity to support and manage program activities when most individuals involved in the program have other responsibilities; and the specialized expertise needed to stay current with the evolving landscape and support the program's expanding mission and offerings.

Reflections on Sustainability and Growth

Reflecting on the past several years, several factors were crucial in securing the resources needed to maintain and grow the program. Many of the factors below help to address more than one facet of sustainability.

1) Be clear with the vision and flexible with the implementation.

A core principle has been to maintain a clear vision for our program while remaining flexible in its implementation, allowing both the community voice and our own expertise and resources to guide program development. While our vision is to support all types of open research across the entire research lifecycle, we started with low-cost, high-impact initiatives, initially focusing our support on early adopters in the life sciences to align with our own expertise and campus connections. We then invested in demonstrating success and expanding our user base into broader disciplines and more variable types of research on campus. This paved the way for building partnerships and growing our team's expertise as we expand our services. In 2024, the Open Science & Data Collaborations Program became part of the newly established Open at CMU initiative, which aims to guide users through the full range of open services offered by CMU Libraries, enhance coordination among these services, and demonstrate the Libraries' ongoing commitment to open research and education.

2) Have support and buy-in from library and campus leadership.

We were fortunate to have the strong support of the Dean of University Libraries, who was instrumental in forming the program, championing our work on campus, and providing financial support for our activities. A critical factor of gaining support from the Dean and the leadership has been to align the program's vision to the strategic plans of both the Libraries and the university, while taking incremental steps and consistently demonstrating value and progress. We have also intentionally sought connections with existing library initiatives that are well-aligned with the goals of the open science program to collaboratively create value-added services, such as support for research data management. In recent years, we have also supported researchers in complying with federal policies for data sharing and public access with resources and guidance on best practices.

In the spirit of open science, one of our priorities has been to offer our resources and programming for free to our users and participants. For example, the initial funding for the Open Science Symposium came from a foundation grant, which later was supplemented by the Libraries, other colleges across campus, and sponsorship from mission-aligned vendors. Costs for digital tools and memberships are mostly provided from the Libraries' budget.

Support from the Dean and library leadership means not only direct budget approval, but also helps us build the visibility needed to form meaningful partnerships, build a campus presence, and gain support from campus leadership, who are able to influence research culture from top-down.

3) Blend top-down and bottom-up strategies for community building and outreach.

Our approach to community building and outreach has successfully combined top-down and bottom-up strategies. In the program's initial years and again during the Year of Open Science in 2023 (The White House, 2023), we conducted campus tours to present to key campus leadership, including Office of the Vice President for Research, Dean's Council, and college and department meetings. At the same time, we actively engaged individual researchers through a variety of channels, including events, workshops, newsletters, office hours, and one-on-one consultations. A key goal has been to build communities of practice that can help to amplify our efforts within their own disciplines. Examples of these initiatives include the Open Science Symposium (Wang, 2020), collaborative hackathons (Griego et al., 2024; Sabata et al., 2025), and establishing an Open Science Advisory Board (Gainey et al., 2025).

At the center of our outreach approach is a commitment to understanding researchers' processes and practices, meeting researchers where they are, and advocating for a gradient of openness to accommodate different disciplines and comfort levels with open science. For example, our ultimate goal for data sharing is to encourage sharing all research data publicly following the FAIR principles for data management and stewardship (Wilkinson et al., 2016). However, for individuals who have not been sharing data due to lack of awareness, training, privacy concerns, or disciplinary norms, we encourage them to openly share a subset of the data to support their research claims; for those who have been practicing some data sharing, we encourage them to share more, and follow the FAIR principles more rigorously in their data sharing practice to improve the reusability of the data.

4) Cultivate in-house expertise and build internal partnerships.

Having the right expertise on the team has been critical for our program's success and growth, and this was significantly enhanced through internal collaborations. The initial team's strength was in life sciences, largely due to the team members holding PhDs in those disciplines. As the program grew, we have been mindful of fostering cross-team collaboration, actively involving colleagues whose work intersects with open science, which has been instrumental in broadening our collective expertise and getting buy-in from across the Libraries. Liaison librarians have been essential as they bring disciplinary expertise and are directly connected with researchers. Similarly, functional specialists have been crucial to lead specific open services such as data services, open access, and open educational resources. These collaborations not only leverage existing strengths and foster cross-pollination, but also help to address capacity challenges by avoiding duplication of effort. Close cross-team collaboration has allowed us to integrate more data and code support into the broader open science offerings.

A newly launched initiative among members of the Open Science and Data Services teams, the Tartan Research Data Alliance, will foster a cross-disciplinary community for research data management and sharing on campus. Close collaboration between the Open Science & Data Collaborations Program and CMU Libraries' Evidence Synthesis Program has led to a focus on making evidence synthesis projects across campus more open and reproducible (Gainey and Young, 2025).

Furthermore, we have made significant efforts to bring in new open science team members with fresh perspectives and diverse expertise, including colleagues with backgrounds in STEM and social sciences research, as well as traditionally trained librarians with expertise in scholarly publishing and research data management. The program's initiative of creating designated program coordinator and postdoc positions further strengthened the team's data and code expertise. The postdoc position in particular will be discussed in more detail in the next section.

5) Leverage external partnerships and communities.

Finally, forming numerous partnerships both across campus and external to CMU has been invaluable. For example, our collaboration with Pittsburgh Supercomputing Center (PSC) not only provided high-performance computing resources for events that require advanced computing, but also directly led to NSF-funding for engaging the research and information science community in discussions about using Machine Learning and AI for data sharing and reuse (Wang et al., 2019). We also actively leveraged existing organizations and communities, such as partnering with The Carpentries and The Childhood Cancer Data Lab, to provide data science and programming training that is much needed for graduate students and researchers. In addition, we partnered with the cloud computing company DNAnexus to co-organize an annual collaborative biohackathon. These partnerships significantly extend the program's reach and impact beyond institutional and disciplinary boundaries.

Postdocs as a Key Mechanism for Sustainability

Hiring postdocs to bring in fresh perspectives and technical skill sets has been an essential mechanism for building program capacity. At CMU Libraries, the Director of the Open Science Program serves as the supervisor for the open science postdoc and is responsible for recruitment, training, and ongoing supervision. While these responsibilities require a significant time investment, the addition of a skilled postdoc — who, by definition, is an experienced researcher — improves the capacity of the supervisor and the broader team. Postdocs contribute to services, instruction, and collaborative research, making them valuable assets to the program and supporting its long-term sustainability.

Before establishing the Open Science Postdoctoral Associate position, we tried a few different approaches to solve the capacity and expertise puzzle, including hiring part-time and full-time staff. We then received funding approval to create the postdoc position and hired our first open science postdoc in 2022. Our library had successfully hired several postdocs through the now-inactive Council on Library and Information Resources Postdoctoral Fellowship Program (Maclachlan et al., 2015), and we considered hiring through that program again. However, we ultimately decided to use a recruitment process similar to that of staff because we wanted more flexibility in the timeline of the search and the desired qualifications of the candidates.

Both authors did postdocs in biology fields prior to joining CMU Libraries and have observed that the expectations for our postdoctoral associates differ significantly from those of traditional postdoctoral appointments in STEM (Table 1). As described below, it is important to discuss in the interview process how a postdoctoral position in libraries is different from a traditional postdoc to manage expectations and ensure that the position is a mutually good fit.

Table 1. Key differences between the Open Science Postdoctoral Associate position at CMU Libraries and a traditional STEM research postdoc position

Characteristic	STEM Research Postdoc	Open Science Postdoc in CMU Libraries
Recruitment	Informal and often driven by graduate students reaching out to potential postdoctoral advisors to enquire about possible positions	More traditional staff hire process driven by the Libraries distributing the ad since graduate students will largely not be aware of the opportunity
Scope / Mix of Activities	Often 100% focused on research	A mix of activities reflective of library work (i.e., teaching, research, and service)
Length	Variable depending on discipline and grant funding, and often negotiated with the postdoctoral advisor.	1-2 year fixed contract
Research Agenda	Typically driven by the postdoctoral advisor, with the postdoc having an independent project within the advisor's larger research agenda	Ability and flexibility for postdoc to create their own research agenda and pursue their interests as they evolve
Grant Funding	Often required to support the postdoc's salary for part or all of the length of the postdoc; funding secured by the PI or the postdoc	Typically, none; a constraint of the research might be what the postdoc can accomplish without access to grant funding
Preparation for Careers	Designed to prepare a researcher to be a professional researcher, postdocs increasingly continue their careers outside of academia but receive variable support and encouragement from PIs to do so	Not designed as a training position for a single career; could be a good pathway to many careers outside of research academia including open science, science policy, science communication, and academic libraries

Our goal with this hire was to expand our service offerings by bringing new and complementary skills to the open science team. We decided that our ideal candidate would have a PhD in a STEM discipline due to the demographics of our campus population, as well as a strong interest—and ideally practical experience—in open science or related fields such as science communication, science policy, or metascience. We also sought candidates with strong technical skill sets, such as coding in Python or R, to allow us to develop automated workflows and provide support for the campus community in improving the reproducibility of their research. We believe this position provides relevant training for many career paths and intentionally do not describe it strictly as training for a career in research or libraries. This broader perspective helps us attract a wide and diverse pool of candidates since STEM PhDs might not have much exposure to libraries or understand that their skill sets can translate to a career in libraries (Gainey et al., 2024).

Recruitment for STEM postdocs often occurs outside of traditional channels, with candidates often reaching out directly to potential postdoctoral advisors and advisors circulating postings informally through academic networks. While we occasionally receive inquiries from doctoral students about our open science postdoc position, we approach recruitment with a more structured process. We've had success posting to job aggregators such as Chronicle of Higher Education, Indeed, and LinkedIn, and we encourage colleagues to circulate the posting among relevant academic departments. There is variation in the degree to which STEM PhD programs are locked to the academic calendar, and it is fairly common for newly minted PhDs to finish work in their doctoral lab for a period of time after they graduate. This gives us flexibility with the timing of recruitment, and we have had success in hiring postdocs throughout the year. Each of our three open science postdoctoral associate hires has attracted increasingly competitive and sizable applicant pools.

In our job postings and during the interview process, we note that the open science postdoctoral associate will engage in a mix of research, teaching, and service activities, reflective of the range of work often done in academic libraries. Specifically, they can expect to spend approximately 50% of their time on research and 50% of their time on instruction, service, and community-building activities. A postdoctoral position in STEM research is often 100% research, and the research questions and methodologies are often driven by the research agenda of the head of the research group or lab. Therefore, it is important to explain the mix of expected activities in the interview process and to note that they have more flexibility for the direction of their research than they would in STEM research. We offer suggestions for research questions or projects, but give the postdocs a high degree of freedom to pursue their own interests within the broad area of open science. Given that candidates may have variable expectations for supervision, we've found it essential to clearly communicate the norms for supervision and assess for independence during the interview process. Ideally, a postdoc is highly self-motivated and benefits more from mentorship than close supervision.

Our first postdoctoral associate had a fixed one-year contract with an option to renew for a second year. We recognized that a one-year contract is unusual for a postdoc and might make prospective applicants wary, so we were able to change the length of the fixed term contract to two years for our second postdoctoral associate hire in 2023. We believe two years is an ideal time frame for a person to meaningfully contribute and gain exposure to the Libraries and complete a research project.

Our postdocs have had little experience with libraries as undergraduate or graduate students beyond access to books and journals, and their postdoctoral training has created an understanding and interest in academic librarianship. These positions can influence the career aspirations of STEM PhDs and simultaneously serve as a recruitment opportunity for filling STEM or functional specialist positions in libraries (Gainey et al., 2024). Our first two open science postdocs became very interested in academic librarianship during their appointments and ultimately joined CMU Libraries as STEM faculty librarians in 2023 and 2025, respectively. At the time of this writing, we are recruiting for our third Open Science Postdoctoral Associate.

One of our biggest challenges prior to our initial postdoc hire was having enough expertise to be able to sustain more technical types of open science programming. For example, we organized several foundational coding Carpentries workshops per year starting in 2019 using the global network of Carpentries instructors. Since 2023, we have been able to adapt and offer the entire basic Carpentries curriculum in-house, in large part due to the expertise of our open science postdocs. Other services that have strongly benefited from or been possible because of our open science postdocs' deep technical expertise include a consultation service for help with using open source data science tools and languages, computational hackathons, workflows for software curation, and automation of evidence synthesis methodologies.

Conclusion

Program sustainability requires continual consideration as a program matures and evolves. Unexpected events such as the COVID-19 pandemic or federal funding uncertainties can threaten a program's stability and success. To address this, we present a multi-pronged approach that includes vision alignment, getting buy-in, resource access, staffing, and expertise.

Expertise is crucial since implementation of the FAIR principles and open science often requires technical solutions. We have sustained the expertise of the program through a couple of methods: forming internal and external partnerships, leveraging the community, and hiring dedicated staff and open science postdocs.

A postdoc in libraries differs from a traditional postdoc in several ways. To ensure a successful and mutually beneficial hire, it is important to clearly communicate these differences and manage expectations for incoming postdocs. When a postdoc hire is successful, it can dramatically benefit a program by creating opportunities for new and creative services, as well as ensuring program sustainability.

References

AAU-APLU Public Access Working Group (2017). *AAU-APLU Public Access Working Group report and recommendations*. https://www.aau.edu/key-issues/aau-aplu-public-access-working-group-report-and-recommendations

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature, 533*(7604), 452–454. https://doi.org/10.1038/533452a

Committee on Toward an Open Science Enterprise, Board on Research Data and Information, Policy and Global Affairs, & National Academies of Sciences, Engineering, and Medicine (2018). *Open science by design: realizing a vision for 21st century research* (p. 25116). National Academies Press. https://doi.org/10.17226/25116

Gainey, M., McKee, L., Bongiovanni, E., & Young, S. (2025). An open science potluck: Building a program that engages your campus community. In E. Bongiovanni, M. Gainey, C. Griego, & L. McKee (Eds.), *The Open Science Cookbook* (pp. 13–18). Association of College and Research Libraries.

Gainey, M. & Sarah, Y. (2025). Building a community of practice for open science. In E. Bongiovanni, M. Gainey, C. Griego, & L. McKee (Eds.), *The Open Science Cookbook* (pp. 123–127). Association of College and Research Libraries.

Gainey, M. A., Griego, C., & Scotti, K. (2024). From the lab to librarianship: Advancing open science. *Issues in Science and Technology Librarianship, 108*. https://doi.org/10.29173/istl2846

Griego, C., Scotti, K., Terveen, E., Chan, J., Sheng, D., González-Espinoza, A., & Warren, C. (2024). A libraries reproducibility hackathon: Connecting students to University research and testing the longevity of published code. *F1000Research, 13*, 1305. https://doi.org/10.12688/f1000research.156917.1

Maclachlan, J. C., Waraksa, E. A., & Williford, C. (2015). *The process of discovery: The CLIR postdoctoral fellowship program and the future of the academy*. CLIR. https://www.clir.org/pubs/reports/pub167/

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

Sabata, S., Kubica, J., Gupta, R., Ericson, L., Atanda, H., Subramaniam, G., Moller, A., Abolade, R., Banka, A., Blechman, S., Brenner, R., Chikina, M., Chong, L., Cooley, N., Chang, D., Greer, P., Gupta, A., Jha, A., Kacar, E., … Busby, B. (2025). Addressing background genomic and environmental effects on health through accelerated computing and machine learning: Results from the 2025 hackathon at Carnegie Mellon University. *BioHackrXiv Preprints*. https://doi.org/10.37044/osf.io/3a8cn_v1

Scotti, K., Jiao, C., Gainey, M., Bongiovanni, E., & Slayton, E. (2025). Charting open science landscapes: A systematized review of US academic libraries' engagement in open research practices. *OSF Preprints*. https://doi.org/10.31219/osf.io/uwcfk_v1

Wang, H. (2020, November 30). Conference report: Artificial intelligence for Data Discovery and Reuse (AIDR) and Open Science Symposium (OSS) 2020. Carnegie Mellon University. https://doi.org/10.1184/R1/12092193.v1

Wang, H., Gainey, M., Campbell, P., Young, S., & Behrman, K. (2022). Implementation and assessment of an end-to-end Open Science & Data Collaborations program. *F1000Research, 11*, 501. https://doi.org/10.12688/f1000research.110355.2

Wang, H., Webster, K., Nystrom, N., & Buitrago, P. (2019, May 13–15). AIDR '19: Proceedings of the conference on artificial intelligence for data discovery and reuse. *AIDR '19: Artificial Intelligence for Data Discovery and Reuse*, Pittsburgh, PA, United States. https://dl.acm.org/doi/proceedings/10.1145/3359115

White House (2023, January 11). *Fact sheet: Biden-Harris administration announces new actions to advance open and equitable research.* https://bidenwhitehouse.archives.gov/ostp/news-updates/2023/01/11/fact-sheet-biden-harris-administration-announces-new-actions-to-advance-open-and-equitable-research/

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*, 160018. https://doi.org/10.1038/sdata.2016.18