The Data Science and Digital Scholarship Fellowship Program (DS²F): A Library-Based Model for Addressing Curricular Gaps in Data-Intensive Training and Digital Pedagogy

Megan Senseney and Jeffrey C. Oliver

The University of Arizona Libraries has conducted a pilot implementation of a year-long Digital Scholarship and Data Science Fellowship (DS²F) to address increasing interest in digital and data-intensive scholarship among graduate students. This article provides details regarding the model for the fellowship program; a description of the pilot implementation; an assessment of the program; and recommendations for libraries interested in adopting a similar approach at other academic institutions. While the program may not be fully adaptable within all academic contexts, DS²F represents a cost-effective and transferable model for graduate student engagement in digital scholarship and data science.

Introduction

Across disciplines, increased interest in digital and data-intensive scholarship creates a largely unmet need for critical training in technical skills. To address that need, librarians from the University of Arizona Libraries developed a model that provided career-relevant professional development opportunities while also affording campus-wide training through a library-based graduate fellowship program focused on building an interdisciplinary cohort of doctoral students from across campus. This paper discusses two iterations of the year-long Digital Scholarship and Data Science Fellowship (DS²F), which were implemented first in 2020 and again in 2022. The goal of the paper is to provide details regarding the model for the fellowship program; a description of the pilot implementation and iterative adjustments made between cohorts one and two; an assessment of the program; and recommendations for libraries interested in adopting a similar approach at other academic institutions. While the program may not be fully adaptable within all academic contexts, DS²F represents a cost-effective and transferable model for engaging graduate students in extracurricular teaching and learning in support of digital and data-intensive research.

* *Megan Senseney, Department Head, Research Engagement, University Libraries, University of Arizona, msenseney@arizona.edu; Jeffrey C. Oliver, PhD, Data Science Specialist, Research Engagement, University Libraries, University of Arizona, jcoliver@arizona.edu.*

Literature Review

The rapidly evolving technological landscape creates a growing need for digital and data skills across academic disciplines. For example, text data mining (TDM) is growing in popularity in the humanities and social sciences. Several powerful tools for TDM are available in the Python programming language, yet many colleges and universities have few opportunities for students outside of traditional computer science programs to learn how to program (Feng et al., 2020; Hannay et al., 2009; Prabhu et al., 2011). Also in demand are skills to leverage "big data" resources, including cloud computing and machine learning. Given the growing interest in skills development in these areas, demands for training often outstrip available training opportunities. Many technological, computational, and digital resources are evolving at a pace too rapid for traditional college curricula to keep up, leaving many scholars to "learn on their own" (Theobold & Hancock, 2019). Furthermore, the evolution in such digital and computational applications is likely to continue outpacing strategies for formal curricular integration, given that the development of new courses and programs can take years to complete. The failure to meet these training needs leaves graduate students underprepared for future scholarly endeavors (e.g., Barone et al., 2017; Davenport et al., 2019; Federer et al., 2016).

Extra-curricular efforts to address the training gap are underway, although there are considerable opportunities for growth in this area. For example, short-format workshops as offered by The Carpentries provide an entry point for data and computational skills (Baker et al., 2016; Wilson, 2016). Briefly, The Carpentries (https://carpentries.org) is a global non-profit organization that offers skills development workshops to support efficient and reproducible research. Such workshops include discipline-specific and discipline-agnostic approaches, with an emphasis on training audiences on how to solve problems as they arise. The long-term impact of the "bootcamp" format is unclear (Feldon et al., 2017), and there remain significant opportunities for intermediate-level training following such novice-level training as Data, Software, and Library Carpentry workshops (Williams & Teal, 2017). Scholars seeking to enhance their skill set often lack the necessary support to translate and apply their introductory knowledge to real-world research projects.

This landscape, where formal graduate curricula are generally outpaced by demand for the latest skills, highlights the importance of self-directed learning. Self-directed learning has been established as an effective strategy to support life-long learning (Bergamin et al., 2019; Boyer et al., 2014; Morris, 2019). Self-directed learning is not synonymous with guidance-free learning, which is less effective than instruction that includes guidance for learners (Kirschner et al. 2006). Self-directed learning is important for the development of computational and digital skills, across disciplines (e.g., Bobkowski & Etheridge, 2023; Lawlor et al., 2022). However, as self-directed learning itself is a skill, students are likely to benefit from explicit training, including such strategies as goal setting and self-reflection (Morris-Eyton & Pretorius, 2023).

Complementary to self-directed learning as a means of skills acquisition, the act of teaching provides opportunities for reinforcement. Providing students with the opportunity to teach others, often peers, can improve learning outcomes (Duran, 2017; Pahl, 2019). These approaches are often applied in teaching conceptual knowledge. For example, Rogers found that students in an undergraduate cognitive psychology course had higher test scores when they delivered a lecture to their peers on a topic than when they wrote a paper on the topic (2021). The learning-by-teaching approach has also been shown to benefit skills development, especially in computational skills. Multiple studies in undergraduate computer programming

courses report improved learning outcomes when peer instruction is incorporated into the course (e.g., Porter et al., 2013; Ruiz De Miras, 2021). Opportunities to teach further serve to augment and enhance the learning process for graduate students, reinforcing their own knowledge and skills (Shortlidge & Eddy, 2018).

Effective teaching, itself, is a skill and there is a need for those delivering training to develop instructional skills and pedagogical best practices. Rarely do those outside the field of education receive training about the science of learning and how teaching methods can influence learning outcomes (Robert & Carlsen, 2017; Robinson & Hope, 2013). Training individuals outside of the computer sciences in the technical skills required for digital scholarship and data science requires an understanding of the novice perspective and ideally involves such practices as peer-to-peer learning and practical applications (Brown & Wilson, 2018). Active learning practices are an effective means of skills development in problem solving and critical thinking (Hepner & Carlson, 2018; Prince, 2004; Styers et al., 2018), and there are growing calls to adopt such practices in academic institutions (Bradforth et al., 2015; Stains et al., 2018). Yet transitioning from traditional, lecture-based passive learning is not trivial and requires intentional training (Niemi, 2002; White et al., 2016).

Academic libraries are well-positioned to address the needs of skills development in both digital pedagogy and computational and digital literacy for two primary reasons. First, the library is largely independent of any one domain. In the absence of the territoriality that often accompanies academic departments, the library is a natural hub for interdisciplinary work. Second, academic libraries are increasingly investing in programs supporting digital pedagogy (Lach & Pollard, 2019); digital scholarship (Hannah et al., 2020); and data science (Oliver et al., 2019). Several academic libraries are already partnering with The Carpentries to introduce data and programming skills to scholars in several disciplines. Such efforts create an opportunity for libraries to further engage with the campus community to develop and provide intermediate-level digital and data skills training (Surkis et al., 2017). Library-directed programs supporting skills development are also emerging, although there are few examples in the published literature. For example, the Sherman Centre for Digital Scholarship at the McMaster University Library runs a mentorship program for graduate students to participate in interdisciplinary professional skills training (Zeffiro et al., 2022). In another example, the LIS Education and Data Science (LEADS) fellowship program provided data science training and internship opportunities for early career information professionals (https://mrc.cci.drexel.edu/research/leads/), and the DATALIS™ initiative continues to extend professional development efforts in data science for library and information science students (https://datalis.design/). Such examples demonstrate the potential for academic libraries to drive skills development and transform campus capacity for digital and data-intensive scholarship.

Introducing the DS²F Model

DS²F is a library-based fellowship program intended to build capacity for campus-wide training in data and computational scholarship. The fellowship program is designed to address two known challenges in contemporary higher education: 1. insufficient training for data-intensive, computational research within subject disciplines; and 2. uneven attention to digital pedagogy and strategies for teaching technical concepts within graduate training programs.

Our approach to addressing these challenges emerged from a series of collaborative conversations among library faculty with expertise in data science, digital scholarship, instruc-

tional design, digital pedagogy, and open research. Prior experiences with The Carpentries' approach to curriculum development includes a focus on reverse instructional design, authentic tasks, and formative assessment. These guiding principles informed our strategy to focus on teaching fellows how to approach self-directed learning for technical skills and how to build individualized self-directed learning plans. In the first stage of the curriculum, fellows are provided instruction on how to learn. We then shift toward a set of curricular modules on learning how to teach technical skills and develop open educational resources on technical topics. This is the point at which fellows also shift from being self-directed learners to teachers in their own right. During the second semester, fellows develop and deliver a workshop open to the broader campus community that introduces topics, skills, or approaches that the fellows engaged with during their period of self-directed inquiry. This approach allows us to provide scaffolding which supports fellows who are pursuing a set of learning objectives that vary in disciplinary focus, technological application, and overarching methodology. Our approach to facilitating self-directed inquiry is also intended to serve fellows over the course of their careers as they continue to pursue lifelong learning in step with the pace of change in their respective fields.

The fellowship is designed to span two semesters. In the first semester of each round of fellowships, students engage in a scaffolded process of cross-disciplinary and inquiry-based learning. During this period, library faculty facilitate a series of monthly workshops focused on self-directed learning, project management, and research management. In the interim weeks between workshops, fellows assemble to work on individual learning and to access support from library faculty mentors and other fellows. These working sessions are supported by the team of library faculty delivering workshops and sponsoring the fellowship. Regular, sustained contact between library faculty and fellows is intended to provide connection and support, and facilitation is a shared duty among the faculty leads. Outside of weekly meetings, the UA Library supports a Slack workspace to further facilitate communication and collaboration among library personnel and fellows.

In semester two, fellows focus on creating their workshops for delivery to the campus. Throughout this period, library faculty continue to deliver monthly workshops with a focus on instructional design, digital pedagogy, and open educational resources. All fellows work with library faculty to determine dates and times for the culminating open workshop series delivered by the fellows. In the final weeks of the program, fellows package their workshop materials into an open educational resource and conduct a self-assessment of their instructional efforts while also providing feedback on the overarching fellowship program. Emphasizing student-generated teaching materials as the fellowship's primary outcome leverages potential benefits associated with peer learning, active learning, engagement with information and communication technologies, and learning through teaching (Ribosa & Duran, 2022).

The fellowship program requires coordination, planning, and evaluation on the part of participating library faculty. Preparatory activities include securing funds, recruiting applicants, offering information sessions about the program, reviewing application packages against a predetermined rubric, communicating with accepted fellows, and disbursing funding. Faculty also prepare workshop materials to deliver to the fellows and develop facilitation guides for the working meetings between workshops. During the active fellowship period, library faculty are committed to maintaining active, responsive communication, often supporting students in need of assistance or providing referrals to colleagues across campus with

relevant expertise. At the conclusion of the fellowship period, library faculty participants review feedback from fellows, discuss potential changes to the curriculum, and assess capacity for ongoing fellowship programming. While the overhead of program coordination is non-trivial, library faculty participants benefit by learning about emerging needs and trends among graduate student researchers through the application process, gaining exposure to a variety of interdisciplinary research projects, and enhancing the suite of curricular offerings available for use and reuse in different contexts.

The fellowship program was designed to complement instructional sessions and online resources offered by UA Library personnel. These include hands-on workshops in data analysis and visualization in the R programming language, consultations on data management best practices, and introductions to text mining software. Library personnel also maintained online resources for data science and digital scholarship, including tutorials for geospatial data analysis, resource navigation tools for identifying datasets, and guides for writing data management plans. The program was intentionally designed to address those training gaps not covered by programming that was already offered, rather than replace existing offerings.

Implementation

A cohort of three library faculty developed an initial proposal for piloting the DS²F model in summer 2019. We wanted to develop a relatively lightweight program that targeted graduate students from a range of disciplines across campus who had completed their coursework and were working toward their field exams or dissertation proposal. The goal was to attract potential fellows who knew there was an outstanding need to develop technical skills that they had not learned in class but which they would require to successfully complete their dissertation research. The goal was to attract intrinsically motivated students for whom we could provide additional support and incentive by 1. building a cross-disciplinary cohort; and 2. compensating each fellow with a modest stipend.

The curriculum we envisioned required engagement from library faculty based in three different departments within the library. The project leads brought these faculty together to engage in early socialization, buy-in, and iterative development over a series of meetings throughout the summer. The initial collaboration included the libraries' data science specialist, digital scholarship librarian, the head of the Office of Digital Innovation and Stewardship, two liaison librarians with functional expertise in instructional design and digital pedagogy, a content and collections librarian with expertise in open educational resources (OER), and the head of the Research and Learning department. Together, we developed a road map for launching an initial call for applications in Fall 2019 for a fellowship program that would run throughout the 2020 calendar year. We initially earmarked $7,500 to support five fellows who would each receive $1,500 disbursed in two installments, one at the end of spring semester and another at the end of fall semester. We launched the call for proposals in mid-September with an open informational session scheduled in early October and final applications due on October 15. We advertised the fellowship by directly emailing graduate student coordinators in each department on campus, and disseminating the call for proposals through local graduate student listservs and interest groups.

For the first implementation, we received 27 applications from 18 distinct disciplines across campus, including representation from the humanities, social sciences, physical sciences, and life sciences. To select fellows, we used a rubric that assessed the degree to which

the applicant:

- identified a technical skill(s) that is relevant and widely applicable;
- demonstrated knowledge of identified technical skill(s);
- demonstrated how the identified technical skill(s) applies to their research;
- expressed interest in delivering open workshops;
- expressed interest in teaching;
- demonstrated diversity in background and/or perspectives; and
- submitted an overall well-written and professional application.

Each applicant was assessed individually by six of the library faculty participating in the program, and then the top candidates were assessed as a group to ensure a diverse but complementary cohort.

We offered five fellowships to students in sociology (2), humanities (1), geoscience (1), and public health (1). All five fellows accepted and began the program in January 2020, but the student in the humanities withdrew in the early weeks of the program due in part to a misunderstanding about the program's approach to self-directed learning. For the first two months, fellows convened in the library for two hours each Friday with meetings alternating between workshops, discussions, and heads-down working sessions. Midway through the first semester, the fellowship program transitioned to virtual meetings via Zoom due to the university's closure in response to the COVID-19 pandemic. During the summer, we transitioned to monthly check-ins as fellows continued their self-directed learning program and began to outline plans for instruction in fall semester. During this period, fellows indicated appreciation for both the continuity and flexibility of the fellowship program and expressed that regular meetings counteracted some of the isolation of being a graduate student during lockdown. For fall semester, the fellows and project leads agreed to reduce weekly meetings to one hour due to sustained Zoom fatigue and ongoing remote work conditions. These sessions included several activities to assist with workshop design and preparation interspersed with heads-down working sessions and opportunities for one-on-one consultations with the project leads.

In November 2020, the team launched a fully remote DS²F workshop series, which included the following contributions from fellows:

- Sabrina Nardin – Introduction to web-scraping in Python
- Sam Scovill – Qualitative research in quarantine: The ethics and technical issues in moving offline research to an online context
- Jonathan King – An introduction to GitHub for scientists
- Mario Trejo – Editing data Visualizations using Inkscape

Sabrina Nardin's workshop was conceived from her dissertation research, which required dynamic web scraping of Italian newspapers for data collection and analysis. Sam Scovill was initially interested in delivering a workshop on qualitative data analysis software but pivoted upon witnessing many qualitative researchers struggling to continue field-based research in online contexts during the pandemic. Jonathan King's workshop was inspired by the need to develop collaborative best practices for data science research with labmates. Mario Trejo wanted to help scholars make scientific data visualizations more accessible to the public by using widely available editing tools. All four workshops were enrolled at capacity, and a discussion of reception and assessment is included in the following section. In the month following the workshops, each fellow re-packaged their workshop materials as an open edu-

cational resource and registered their offerings with OER Commons (https://oercommons.org/) or Merlot (https://merlot.org/merlot/).

Upon completion of the initial pilot, the project leads conducted an informal assessment of the overall project with plans to launch another call for applicants in fall semester 2021. While the original model front loaded all library-led workshops in the first semester, the team realized that sessions on instructional design, digital pedagogy, and open educational resources would be more useful to fellows during the active workshop development period in semester two. Though the program was designed with the intention of being relatively low touch, the project team also worked to streamline expectations and time commitments for participating library personnel without negatively impacting the fellowship experience. Individual library-led workshops were also revised based on feedback and engagement during the pilot's first year. Librarian-created lesson materials and program schedule are available at https://osf.io/68ezw/. Finally, the second round of the program was intentionally designed to accommodate shifting modalities in support of remote, on-site, and hybrid engagement options.

The second call for proposals received strong applications with a significantly diminished response rate, which we attributed to ongoing stressors associated with the pandemic. We received seven applications from seven different disciplines, and six library faculty members evaluated applicants using the same rubric from the first year. We offered six fellowships to students in public health (1), higher education (1), bioinformatics (1), English (1), earth sciences (1), and astronomy (1). Once again, one fellow withdrew during the first semester of the program, this time citing personal reasons. The second cohort of DS²F fellows met virtually once a week during the spring and fall semesters and continued the practice of monthly check-ins through the summer.

In November 2022, the team hosted the second DS²F workshop series, which included the following contributions from fellows:

- God'sgift Chukwuonye – R basics: Data cleaning and wrangling with R
- Aviva Doery – Introduction to editing 360-degree video in Adobe Premiere Pro
- Emmanuel Gonzalez – Using interactive data visualization to make sense of large datasets
- Anuj Gupta – Coding & decoding: Introduction to text mining for humanists and social scientists
- John He – Creating visualizations of the solid earth using Paraview

God'sgift Chukwuonye's workshop stemmed from the need to normalize and manage research data related to studies on the impact of heavy metals exposure on the health of individuals living in mining communities. Aviva Doery was interested in AR/VR approaches that might create virtual study abroad experiences for students who are unable to travel, and she created a hands-on workshop to introduce editing techniques for 360-degree video. Emmanuel Gonzalez wanted to explore how to use data visualization to synthesize and communicate information drawn from data-intensive research on drought resistance in plants. Anuj Gupta applied text data analysis techniques to a collection of first-year writing samples to study academic writing anxiety. John He wanted to build experience using high performance computing to design, implement, and visualize three-dimensional numerical simulations and experiments of the solid earth.

By fall of 2022, many campus activities had resumed in person, and fellows were presented with the choice of conducting remote, in-person, or hybrid workshops. The project team worked with fellows to discuss the challenges and affordances of each modality. Two

fellows elected to offer in-person workshops and three elected to conduct on-line workshops. Registration rates for the second cohort's workshops were more varied than in the first round with more participants continuing to opt for online engagement. As with the first cohort, fellows utilized the month following the workshops to create open educational resources and register their offerings with OER Commons or Merlot. Links to all student-created open educational resources are available at https://data.library.arizona.edu/data-science/digital-scholarship-data-science-fellowship.

Program Assessment

Throughout the planning and implementation of the program, we sought feedback from multiple audiences to improve the experience for all involved. We focused on soliciting information from three key audiences: participants in the fellow-led workshops, the graduate student fellows themselves, and the library personnel involved in the planning and delivery of the fellowship program. For all audiences, we used both structured and semi-structured surveys to collect information at key points of the program. These surveys informed changes to the program and provided critical feedback on fellows' workshop delivery and materials.

Following the delivery of the fellow-led workshop, we distributed a post-workshop survey to all participants. These surveys included a suite of common questions as well as questions specific to the workshop material and delivery. The questions common to each workshop included an assessment of the pace of delivery, the balance between hands-on and lecture material, and the learning experience. Workshop-specific questions asked participants to rate their experience in the workshop including the level of comfort with applying the material to their own work. Over the two cohorts, 81% of the 66 workshop participants who responded to the survey indicated an appropriate balance in time spent on hands-on activities and lecture, and 85% agreed that they could immediately apply what they learned to their own work. Workshop feedback was also shared with fellows to help identify areas of improvement in their instructional practice.

To improve the experience for the graduate student fellows, we surveyed each cohort twice: once mid-way through the program and once at the end of their fellowship. Briefly, at the midpoint of the program, we asked graduate fellows to share what was working for them and what they would change about the program. Many fellows, from both cohorts, highlighted the utility of interacting with other members of their respective cohorts. In both cohorts, we attempted to increase such opportunities during the second half of the fellowship. These opportunities included structured peer review throughout the development of instructional materials and opportunities to practice portions of their planned instruction. At the conclusion of the fellowship, we asked fellows to assess their learning progress both regarding their self-directed learning goals as well as their skills as instructors. Most graduate fellows (75%) agreed that they had enough support for computational and data skills development during the program. All graduate fellows felt they had enough support in developing skills as an instructor and 87.5% of fellows thought the balance of instructional time versus "heads-down" time for self-directed learning was appropriate. These formal feedback systems, combined with more frequent informal feedback opportunities, allowed the library team to adapt and respond to the needs of current and future graduate fellows.

Finally, several library personnel contributed to the program through planning and instruction, and we sought their perspectives to improve their experience. Following the first

cohort, we asked participating library personnel to provide an estimate of the time they spent, through synchronous (e.g., meetings with fellows, instructional sessions) and asynchronous (e.g., instructional preparation, logistics, planning) activities, on the program. Library personnel participated in 10 to 50 hours of synchronous activities over the course of the year for the first cohort and four to 40 hours in asynchronous activities. Fellowship coordinators represent the higher end of the range, while instructors tended to contribute at the lower end of the range. We also asked about interest in future participation and if there were areas that were covered in too much or too little depth. Library personnel were enthusiastic about participating in future cohorts, and almost unanimously suggested decreasing the amount of synchronous meeting time with consideration for the best use of library personnel's time. Furthermore, it was suggested that instructional sessions on pedagogical best practices were temporally too far removed from the student-led workshops (instruction on pedagogy occurred in early Spring semester, while student-led workshops took place late in the following Fall semester).

Based on the information from surveys and reflection on the program, it is important to highlight some specific benefits to graduate fellows and the campus community. All graduate students who completed the fellowship program accomplished most of the learning goals they had set for themselves; several mentioned how the structure of the fellowship program helped them accomplish learning goals they otherwise would have had difficulty completing. This structure was flexible enough to accommodate shifts in goals, allowing students to learn enough about topics to know which ones would or would not be most useful for their own thesis work. The utility of the fellows-led workshops and associated materials was evidenced by multiple fellows re-using the materials to instruct in subsequent, non-library workshop series offered at the University of Arizona.

The applications submitted by graduate students interested in the fellowship provided an informative, albeit informal, landscape scan and needs assessment for the campus. Through the students' applications, library personnel identified potential areas of engagement in a variety of campus communities in need of additional skills development support. Common themes in application materials included a desire for text analysis techniques, data visualization tools, and programming language competency. Such information is useful for highlighting emerging technologies libraries can plan to support, through library personnel or strategic partnerships.

Recommendations

The DS²F program serves as a model for other institutions seeking to support digital scholarship and data science skills development in the absence of significant additions to library personnel or professional development opportunities. For successful design and implementation of similar programs, we have four recommendations for program leaders: assemble a diverse team; rely on campus networks; recognize the value of personnel resources; and compensate fellows appropriately.

Assemble a Diverse Team

To support students in programs focused on such broad areas as digital scholarship and data science, the team should include practitioners from an array of disciplines touching on all areas of the program. This diversity of expertise will be important for ensuring that the team can guide student participants in asking the right questions, even if team members do not, themselves, have all the answers to those questions. Should expertise in the team be limited,

programs should either be restricted in scope to align with this expertise or leaders should recruit team members from outside the library to fill expertise gaps.

Rely on Campus Networks

The expertise of such a team is unlikely to include a depth of knowledge on all topics. Assembling a diverse team will also provide the benefit of creating a diverse network that program personnel can refer students to when their needs exceed the team's expertise. Campus partners can include research computing centers, information technology units, and other service centers such as bioinformatics cores. Connecting with campus partners through student fellowship programs benefits not only the student participants but also the library by strengthening ties through increased collaborations and reciprocal referrals.

Recognize the Value of Personnel Resources

Student mentoring and development programs take considerable personnel time, often more time than is anticipated. Between the first and second rounds of our implementation, we greatly reduced the amount of time asked from library personnel who were not part of the leadership team. Program leaders should be mindful of team members' time and other commitments and be prepared to fill-in if team members need to reduce their level of participation. Programs such as these require coordination, including scheduling, communications, and other logistical arrangements. Such time should be acknowledged as an explicit portion of position description(s) for personnel responsible for program coordination. Based on our program experiences, a cohort of ten students warrants approximately 25% of a full-time position for effective program coordination.

Compensate Fellows Appropriately

The level of participation should dictate how graduate students are compensated. We used an estimated weekly participation of two hours, along with contemporary hourly wages for graduate student workers to arrive at our stipend amount. The estimated time students spend on fellowship activities will most likely be an underestimate: even with guidance on time management, most fellows, if not all, spent more time on fellowship tasks outside of synchronous meetings than were expected. In addition to financial compensation, the fellowship is an opportunity to mentor students, which provides another, albeit less quantifiable, means of compensation. Such mentorship is not meant as a substitute, but rather a complement to financial compensation. Such support, through financial compensation and mentoring, can be especially important for building a diverse cohort, as these support mechanisms are especially important for retention and completion for students from underserved communities (e.g., Sowell & Okahana, 2015).

Conclusion

Programs such as DS²F are well suited to academic libraries with strong collaborative engagement within the library and active networks for coordination across a range of campus units with aligned interests. The breadth of topics in which graduate students wish to engage combined with the disciplinary diversity of our cohorts is an inherent strength of the program; it also requires that program coordinators tap into their networks to connect students with the resources they most need. Self-directed learning and consolidating learning outcomes

through teaching are vital career skills for students who are effectively at the end of their formal coursework and beginning their careers as independent researchers. By leveraging library expertise in these areas and the library's role as an interdisciplinary hub for campus engagement, graduate students are better positioned to develop new technical skills without losing momentum on their primary objective: a completed dissertation. In turn, the library is better positioned to remain at the forefront of identifying and addressing the needs of graduate students as they emerge and evolve. Ultimately, DS²F has proven to be a mutually beneficial and reciprocal model for research and learning among graduate students and librarians.

Acknowledgements

The authors would like to gratefully acknowledge the many contributions of our University of Arizona Libraries colleagues to the successful development of the DS²F program. Special thanks to Jennifer Nichols and Jennifer Church-Duran for their support during the initial program development and to Cheryl Casey, Tina Johnson, Jim Martin, Yvonne Mery, Leslie Sult for their many contributions to program instruction and coordination. The authors would also like to express gratitude to the DS²F Fellows named in the article. We learned so much from each of you. Finally, the authors would like to thank the anonymous peer reviewers for their feedback in bringing this manuscript to publication.

References

Baker, J., Moore, C., Priego, E., Alegre, R., Cope, J., Price, L., Stephens, O., van Strien, D., & Wilson, G. (2016). Library carpentry: Software skills training for library professionals. *LIBER QUARTERLY, 26*(3), 141–162. https://doi.org/10.18352/lq.10176

Barone, L., Williams, J., & Micklos, D. (2017). Unmet needs for analyzing biological big data: A survey of 704 NSF principal investigators. *PLOS Computational Biology*, *13*(10), e1005755. https://doi.org/10.1371/journal.pcbi.1005755

Bergamin, P. B., Bosch, C., du Toit, A., Goede, R., Golightly, A., Johnson, D. W., Johnson, R. T., Kruger, C., Laubscher, D., Lubbe, A., Olivier, J., van der Westhuizen, C., & van Zyl, S. (2019). *Self-Directed Learning for the 21st Century: Implications for Higher Education* (R. Bailey, J. de Beer, & E. Mentz, Eds.; Vol. 1). AOSIS. https://doi.org/10.4102/aosis.2019.BK134

Bobkowski, P. S., & Etheridge, C. E. (2023). Spreadsheets, software, storytelling, visualization, lifelong learning: Essential data skills for journalism and strategic communication Students. *Science Communication*, *45*(1), 95–116. https://doi.org/10.1177/10755470221147887

Boyer, S. L., Edmondson, D. R., Artis, A. B., & Fleming, D. (2014). Self-directed learning: A tool for lifelong learning. *Journal of Marketing Education*, *36*(1), 20–32. https://doi.org/10.1177/0273475313494010

Bradforth, S. E., Miller, E. R., Dichtel, W. R., Leibovich, A. K., Feig, A. L., Martin, J. D., Bjorkman, K. S., Schultz, Z. D., & Smith, T. L. (2015). University learning: Improve undergraduate science education. *Nature*, *523*(7560), 282–284. https://doi.org/10.1038/523282a

Brown, N. C. C., & Wilson, G. (2018). Ten quick tips for teaching programming. *PLOS Computational Biology*, *14*(4), e1006023. https://doi.org/10.1371/journal.pcbi.1006023

Davenport, J. H., Crick, T., Hayes, A., & Hourizi, R. (2019). The Institute of Coding: Addressing the UK digital skills crisis. *Proceedings of the 3rd Conference on Computing Education Practice - CEP '19*, 1–4. https://doi.org/10.1145/3294016.3298736

Duran, D. (2017). Learning-by-teaching. Evidence and implications as a pedagogical mechanism. *Innovations in Education and Teaching International*, *54*(5), 476–484. https://doi.org/10.1080/14703297.2016.1156011

Federer, L. M., Lu, Y.-L., & Joubert, D. J. (2016). Data literacy training needs of biomedical researchers. *Journal of the Medical Library Association: JMLA*, *104*(1), 52–57. https://doi.org/10.3163/1536-5050.104.1.008

Feldon, D. F., Jeong, S., Peugh, J., Roksa, J., Maahs-Fladung, C., Shenoy, A., & Oliva, M. (2017). Null effects of boot camps and short-format training for PhD students in life sciences. *Proceedings of the National Academy of Sciences*, *114*(37), 9854–9858. https://doi.org/10.1073/pnas.1705783114

Feng, X., Qiao, H., & Enquist, B. J. (2020). Doubling demands in programming skills call for ecoinformatics education. Frontiers in Ecology and the Environment, 18(3), 123–124. https://doi.org/10.1002/fee.2179

Hannah, M., Heyns, E. P., & Mulligan, R. (2020). Inclusive infrastructure: Digital scholarship centers and the academic library liaison. *Portal: Libraries and the Academy*, *20*(4), 693–714. https://doi.org/10.1353/pla.2020.0033

Hannay, J. E., MacLeod, C., Singer, J., Langtangen, H. P., Pfahl, D., & Wilson, G. (2009). How do scientists develop and use scientific software? *2009 ICSE Workshop on Software Engineering for Computational Science and Engineering*, 1–8. https://doi.org/10.1109/SECSE.2009.5069155

Hepner, S. L., & Carlson, T. E. (2018). Active learning to develop key research skills in master's level computer science coursework. *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 840–845. https://doi.org/10.1109/TALE.2018.8615196

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Lach, P. R., & Pollard, E. A. (2019). Visualizing history in the classroom: A faculty-librarian partnership in the digital age. *New Review of Academic Librarianship*, *25*(2–4), 335–356. https://doi.org/10.1080/13614533.2019.1627562

Lawlor, J., Banville, F., Forero-Muñoz, N.-R., Hébert, K., Martínez-Lanfranco, J. A., Rogy, P., & MacDonald, A. A. M. (2022). Ten simple rules for teaching yourself R. *PLOS Computational Biology*, *18*(9), e1010372. https://doi.org/10.1371/journal.pcbi.1010372

Morris, T. H. (2019). Self-directed learning: A fundamental competence in a rapidly changing world. *International Review of Education*, *65*(4), 633–653. https://doi.org/10.1007/s11159-019-09793-2

Morris-Eyton, H., & Pretorius, E. (2023). Cultivating a digital promise: Promoting reflective and reflexive activities to enhance self-directed learning for the 21st Century. *Student Success*, *14*(1), 88–99. https://doi.org/10.5204/ssj.2659

Niemi, H. (2002). Active learning—a cultural change needed in teacher education and schools. *Teaching and Teacher Education*, *18*(7), 763–780. https://doi.org/10.1016/S0742-051X(02)00042-2

Oliver, J. C., Kollen, C., Hickson, B., & Rios, F. (2019). Data science support at the academic library. *Journal of Library Administration*, *59*(3), 241–257. https://doi.org/10.1080/01930826.2019.1583015

Pahl, M.-O. (2019). Learning by teaching: Professional skills and new technologies for university education. *IEEE Communications Magazine*, *57*(11), 74–80.

Porter, L., Bailey Lee, C., & Simon, B. (2013). Halving fail rates using peer instruction: a study of four computer science courses. *Proceedings of the 44th ACM Technical Symposium on Computer Science Education*, 177–182. https://doi.org/10.1145/2445196.2445250

Prabhu, P., Zhang, Y., Ghosh, S., August, D. I., Huang, J., Beard, S., Kim, H., Oh, T., Jablin, T. B., Johnson, N. P., Zoufaly, M., Raman, A., Liu, F., & Walker, D. (2011). A survey of the practice of computational science. *State of the Practice Reports on - SC '11*, 1. https://doi.org/10.1145/2063348.2063374

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Ribosa, J., & Duran, D. (2022). Student-generated teaching materials: A scoping review mapping the research field. *Education in the Knowledge Society (EKS)*, *23*, 27443. https://doi.org/10.14201/eks.27443

Robert, J., & Carlsen, W. S. (2017). Teaching and research at a large university: Case studies of science professors. *Journal of Research in Science Teaching*, *54*(7), 937–960. https://doi.org/10.1002/tea.21392

Robinson, T. E., & Hope, W. C. (2013). Teaching in higher education: is there a need for training in pedagogy in graduate degree programs? *Research in Higher Education Journal*, *21*(131564), 1–11. https://www.aabri.com/manuscripts/131564.pdf

Rogers, M. M. (2021). Teaching-to-learn: Its effects on conceptual knowledge learning in university students. *International Journal of Innovative Teaching and Learning in Higher Education*, *2*(1), 1–14. https://doi.org/10.4018/IJITLHE.289863

Ruiz De Miras, J., Balsas-Almagro, J. R., & García-Fernández, Á. L. (2021). Using flipped classroom and peer instruction methodologies to improve introductory computer programming courses. *Computer Applications in Engineering Education*, cae.22447. https://doi.org/10.1002/cae.22447

Shortlidge, E. E., & Eddy, S. L. (2018). The trade-off between graduate student research and teaching: A myth? *PLOS ONE*, *13*(6), e0199576. https://doi.org/10.1371/journal.pone.0199576

Sowell, R., Allum, J., & Okahana, H. (2015). *Doctoral initiative on minority attrition and completion* (No. 1). Council of Graduate Schools.

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., Eagan, M. K., Esson, J. M., Knight, J. K., Laski, F. A., Levis-Fitzgerald, M., Lee, C. J., Lo, S. M., McDonnell, L. M., McKay, T. A., Michelotti, N., Musgrove, A., Palmer, M. S., Plank, K. M., … Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, *359*(6383), 1468–1470. https://doi.org/10.1126/science.aap8892

Styers, M. L., Van Zandt, P. A., & Hayden, K. L. (2018). Active Learning in Flipped Life Science Courses Promotes

Development of Critical Thinking Skills. *CBE—Life Sciences Education*, *17*(3), ar39. https://doi.org/10.1187/cbe.16-11-0332

Surkis, A., LaPolla, F. W. Z., Contaxis, N., & Read, K. B. (2017). Data day to day: Building a community of expertise to address data skills gaps in an academic medical center. *Journal of the Medical Library Association*, *105*(2). https://doi.org/10.5195/JMLA.2017.35

Theobold, A., & Hancock, S. (2019). How environmental science graduate students acquire statistical computing skills. *Statistics Education Research Journal*, *18*(2), 68–85. https://iase-web.org/documents/SERJ/SERJ18(2)_Theobold.pdf?1575083627

White, P. J., Larson, I., Styles, K., Yuriev, E., Evans, D. R., Rangachari, P. K., Short, J. L., Exintaris, B., Malone, D. T., Davie, B., Eise, N., Mc Namara, K., & Naidu, S. (2016). Adopting an active learning approach to teaching in a research-intensive higher education context transformed staff teaching attitudes and behaviours. *Higher Education Research & Development*, *35*(3), 619–633. https://doi.org/10.1080/07294360.2015.1107887

Williams, J. J., & Teal, T. K. (2017). A vision for collaborative training infrastructure for bioinformatics: Training infrastructure for bioinformatics. *Annals of the New York Academy of Sciences*, *1387*(1), 54–60. https://doi.org/10.1111/nyas.13207

Wilson, G. (2016). Software carpentry: Lessons learned. *F1000Research*, *3*, 62. https://doi.org/10.12688/f1000research.3-62.v2

Zeffiro, A., Brodeur, J. J., & Montague, A. (2022). *Social, equitable, collaborative: 10 years of teaching and learning at the Sherman Centre for Digital Scholarship*. Paul R. MacPherson Institute for Teaching and Learning. https://macsphere.mcmaster.ca/handle/11375/28154