UC San Diego
Books and Book Chapters

Title

Filling the Gaps: Data and GIS Services for the Undergraduate Researcher

Permalink

https://escholarship.org/uc/item/60f281r1

ISBN

978-0-8389-3933-8

Author

Sklar, Annelise

Publication Date

2023

UNDERGRADUATE RESEARCH & THE ACADEMIC LIBRARIAN:

Case Studies and Best Practices
Volume 2

co-edited by Merinda Kaye Hensley, Hailley Fargo, and Stephanie Davis-Kahl

UNDERGRADUATE RESEARCH & THE ACADEMIC LIBRARIAN:
Case Studies and Best Practices
Volume 2

co-edited by Merinda Kaye Hensley,
Hailley Fargo, and Stephanie Davis-Kahl

Table of Contents

PART V – BUILDING AND SUSTAINING PROGRAMS

CHAPTER 10

Filling the Gaps:
Data and GIS Services for the Undergraduate Researcher

Annelise Sklar

Introduction

Given that the University of California San Diego ranks as *Surfer Magazine*'s top college for surfers, the casual connoisseur of university rankings might expect that we are a party school at the beach. While we do have a prime location in La Jolla, CA, we are also the Times Higher Education's top "Golden Age" university and consistently rank highly on lists such as the Shanghai Academic Ranking of World Universities and *U.S. News & World Report* Best Colleges guidebook. In fact, like all ten University of California campuses, UC San Diego has a student body of high achievers: most students come to UC San Diego directly out of high school (the average age for undergraduates is twenty-one and only 4 percent of undergraduates are over twenty-five) and with a high school GPA of 4.08. Every freshman graduated in the top ten percent of their class and over 70 percent of students go on to graduate within four years. Transfer students are equally, if not more, motivated, with half graduating in two years and over 80 percent graduating within three years of entering the university.

At the same time, only 19 percent of the undergraduate student body is white, 38 percent identify as first-generation college students, and 72 percent of all undergraduate students receive some type of financial assistance. Perhaps not surprisingly, though 37 percent of students have majors that fall within the Division of Social Sciences, as a "STEM school," 49 percent of undergraduate majors fall into the fields of engineering, biology, or physical sciences. Our undergraduate students typically view college as the foundation of their successful futures,

153

and many seek additional opportunities to differentiate themselves from their equally high-achieving peers so as to be more appealing to employers and/or graduate school admissions officers.

Undergraduate students often seek opportunities to do original research, through capstone projects, honors programs, internships, faculty mentorships, research apprenticeships with graduate students, and academic enrichment programs, such as McNair, the California Louis Stoke Alliance for Minority Participation (CAMP), UC Scholars, Triton Research & Experiential Learning Scholars (TRELS), and a variety of other scholarships. While highly engaged with their projects, these undergraduate researchers are also typically new to the entire research process, from methodology to tools, and data collection, analysis, visualization, and management. Though they are typically paired with advisors and/or mentors, they often lack formal classroom training and prior hands-on experience in these areas, and they may not have access to the technology that their graduate and faculty counterparts do.

Student researchers, whether they are working on class assignments, their own research projects, or as part of a larger faculty-driven research project, often seek help for the following: finding data; understanding, using, converting, and even just opening the datasets they do find; choosing the right software and then learning to use it well enough to complete their tasks; troubleshooting software and code; advice on storage solutions and meeting data management requirements; and general workflow. The UC San Diego campus has an assortment of computer labs, makerspaces, storage solutions, etc., but the library is seen as disciplinarily agnostic, and only we have the mission to support all affiliates regardless of academic status, discipline, or experience level. We try to fill the gaps by providing access to collections of data, software, and higher-performance computing as well as expertise in using these tools.

Background

Traditional public services librarianship has typically focused on the early stages of the research lifecycle—discovery, topic refinement, and maybe a little bit about searching the literature for methodology or acquiring primary sources, especially for humanities research. In the past decade or so, libraries have also shifted to support the end of the research lifecycle with the dissemination of research findings and data, especially via open access initiatives. Our team of data and GIS functional specialists supports the middle of the research lifecycle, helping student researchers identify or collect data, locate storage solutions (which may or may not be those provided by the library), and analyze and visualize their data, which may be geospatial, quantitative, textual, audio/visual, or a hodgepodge of formats. Though our services have evolved over time, as open data became a movement and powerful open-source software

like R and Python became ubiquitous, the UC San Diego Library has been providing data services and GIS support for decades. The data science, GIS, and map librarian positions have undergone several organizational shifts over the years, and they are now part of the Scholarship Tools and Methods Program, which also includes format/functional specialists for the Digital Media Lab, Digital Collections, digital reformatting, digital archiving, and scholarly communication services.

The Data & GIS Lab itself is a fairly small space not far from the library's Research Assistance Desk and adjacent to the office space that houses most public services librarians. This high-traffic area of the library includes the reference, government information, microforms, CD-ROMs, current periodicals, and maps collections; general computing as well as printers and scanners; and open study space for both groups and individuals. The Data & GIS Lab has five higher-end fixed computers with an assortment of commonly used software for geospatial, statistical, and qualitative analysis, a large wall-mounted monitor, and a service desk. The door is always open: the Data & GIS Lab maintains the same self-service hours as the library building, and computers are available on a first-come, first-served basis for anyone with a campus login.

Our service desk is staffed by student assistants during daytime hours, usually mid-morning until early evening, as a supervisor must be available while they work. These students provide basic tech support and troubleshooting for geospatial and statistical analysis software and make referrals to librarians for more complicated user questions. Because we cannot compete with graduate programs that offer tuition remission to teaching and research assistants, most of our student assistants are upper-division undergraduate students or terminal master's students. They are often students who used the Data & GIS Lab heavily for a course or a project in prior quarters.

The Data & GIS Lab is primarily co-managed by the GIS librarian and data science librarian, with input from the maps librarian and the team's supervisor, the assistant program director of the Scholarship Tools and Methods Program (both of whom have served extended terms as interim GIS librarian and data services librarian, respectively). The librarians are responsible for decisions about the Data & GIS Lab space and services and regularly meet students for consultations in the Data & GIS Lab, but our jobs are fairly dynamic and, like our subject librarian colleagues, are not location dependent. In addition to providing consultations to researchers of any level by appointment, librarians offer course-integrated instruction upon request from faculty, teach and host workshops and related events, coordinate activities with communities of practice including student clubs, acquire and manage all formats of data and analytic software, maintain LibGuides and tutorials, and develop and manage a collection of guides, manuals, and other reference items.

Partnerships

Members of the team work closely with colleagues throughout the library, on campus, and across the UC system. For example, undergraduate researchers often first learn that the library provides data and GIS services from the subject librarians who work with their honors or capstone class, or who they otherwise seek out when beginning a research project. Subject and functional librarians will often work together to identify and acquire datasets requested by student researchers, splitting the costs and the collection management work, which often includes creating and monitoring accounts for services that are not IP-authenticated, such as WRDS, Linguistic Data Center, Brandwatch, and ProQuest TDM Studio. The data science and GIS librarians both co-teach and/or coordinate course-integrated instruction sessions with subject librarians, and several courses have now adopted ArcGIS StoryMaps as a platform for hosting final projects. The data science librarian also has a 0.2 FTE appointment within the library's Research Data Curation Program and has worked closely with her colleagues to curate data created by data science students for deposit into the library's Digital Asset Management System. Wearing both programmatic hats, she also jointly runs the library's Software, Data, and Library Carpentry (now collectively known as The Carpentries) computational training, which we in turn provide in partnership with the campus Research IT Services and is open to researchers of all academic levels.

We could not maintain or improve our technology offerings without working closely with the library's Technology & Digital Experience (TDX) program and campus IT Services (ITS), both of whom have been particularly crucial in making ArcGIS software ecosystem more widely available. Campus ITS provides campus-wide access to software often used in courses, such as Stata and ArcGIS, through computer labs, a virtual computing environment for courses, and downloads for individual students and researchers, but it is the library that provides expert help in using this software and fills gaps by providing access to more niche tools such as Stat/Transfer or Atlas.ti. Library TDX provides all of the Data & GIS Lab computer, software, and technical infrastructure and maintenance and controls the library's funding for equipment and software, including web-based tools that are not considered library e-resources/collections. While TDX maintains a traditional help desk ticketing system and workflow, both teams have a highly collaborative and consultative relationship when it comes to pursuing new technology or technical workflows. Collaborative ventures in the works include a virtual computing setup to provide additional and more flexible computing power for Data & GIS Lab users, the development of a GeoBlacklight instance for hosting and managing our geospatial data collections (which is also a joint venture with our colleagues at UC Santa Barbara), and the launch of a library-based ArcGIS Enterprise system to support researchers—often undergraduate honors or capstone students, without other computing access to sup-

port their more complex geospatial computing needs. The GIS librarian, TDX, and campus ITS also partnered to institute single sign-on access to ArcGIS GIS Online, which streamlined the access process and tripled the number of registered users in a little over a year.

Finally, the team also maintains relationships with other experts on campus. The GIS librarian hosts a twice-per-quarter "Pulse of GIS" meeting for practitioners of all levels to share their work, and GIS instructors from different departments often share information about campus resources. The data science librarian is also liaison to the Halıcıoğlu Data Science Institute and data science major, maintaining communication with the undergraduate advisor and faculty, contributing to the newsletter for students, and attending various faculty functions and events.

Assessment

Admittedly, some of our best evaluative data is anecdotal and observational—these students graduated with honors, those students were accepted to graduate school or found an impressive professional position, this student reworked their paper for formal journal publication, that student thanked the librarian in their acknowledgments. As the Data & GIS Lab is adjacent to the restroom nearest our offices, we notice when the Data & GIS Lab is empty (mornings when most of the student body is in class) and full (afternoons, like the rest of the library). We also occasionally peek at the very active UCSD subreddit looking for mentions of the library or courses with data and GIS components, and we seek informal feedback from our student assistants. We have used the Library Student Advisory Council as a focus group, and we also look for mentions in the library section of the campus Faculty and Staff Customer Satisfaction Survey and Student Satisfaction Survey.

More formally, the librarians record the same statistics on consultations/reference questions, instruction/workshops, and outreach that the subject librarians do, and the student assistants keep statistics like those of the library's other service desks. All these statistics are entered into our library-wide Qualtrics forms and reported out to national organizations through central channels.

We have tried to understand the Data & Lab's usage patterns in a number of different ways. We began with a sign-in sheet where students are supposed to enter their names, times in and out, and purpose of the visit. While we continue to record this information, we know that many students do not sign in or out, especially when the Data & GIS Lab is unstaffed, and reasons like "homework" are not useful for our purposes. The DML Manager recommended we install an automated people counter, and while this does not tell us why or how students are using the space, it gives us quantitative data to reinforce our more casual observations, and it provides insight into traffic patterns after the librarians have gone home for the day. Recently, TDX implemented LabStats as part of a project to create a dashboard for

students to locate open computers throughout the library buildings. This software allows us to see when and which computers are used and which software is used and how often. In turn, this informs our decisions about how many computers and specialized software licenses we need to maintain. Likewise, after implementing single sign-on access to ArcGIS Online, we also created an ArcGIS Dashboard to visualize the demographic data (department affiliation and status as faculty, staff, graduate student, or undergraduate student) of our ArcGIS Online users so that we can have better insight into where GIS technology is being used on campus.

Reflection

We consider our services ultimately successful because most of the undergraduate researchers who seek our help successfully complete their projects. Nonetheless, we know that while we are plenty busy, most undergraduate students are not aware of our services (or most library services, for that matter). While we have existing relationships with subject librarians, instructors, and advisors and promote services through library media, social media, and events, we need to improve our marketing to undergraduate students closer to their point of need. At the same time, we continue to struggle with providing services at scale when each project has a unique set of needs and timeline, and the pace of the university's ten-week quarter is exhausting for both the students and the librarians. We also grapple with presenting our services and collections on the web in a way that is not a huge LibGuide consisting of an overwhelming list of links.

　　We have gone through two planning exercises to remodel and expand our space into a larger digital scholarship center, but space in our library building is at a premium, constantly full of students studying and just hanging out at all hours of the day, so it is difficult to expand our footprint with any activity that will have lower capacity, especially if it also has open tables, chairs, and/or computers. We have also found that the term and concept of "digital scholarship" does not resonate with our student body. We continue to grapple with how to keep the Data & GIS Lab open and inviting to the students who need it while making it less appealing to study groups who see our empty group tables and whiteboards and decide to camp out, making the space too active and loud for those who need to deeply focus on their data analysis.

Impact of COVID-19 and a Move to Remote

UC San Diego shut down and switched over to remote classes the week before finals in March of our 2020 winter quarter, and classes remained remote through the end of summer 2021. Like most university employees, librarians and library

staff flipped to working from home. The library invoked a "digital first" strategy that emphasized our existing virtual reference service and online collections as well as HathiTrust Emergency Temporary Access and instruction and other face-to-face interactions via Zoom. Like our colleagues, the data and GIS team shifted all of our communication with users and among ourselves online, using Zoom, email, and our internal Slack channels. Thanks to quick thinking by the GIS librarian and flexibility on the part of TDX, we were able to offer users remote desktop access to the computers in the Data & GIS Lab, though the workflow involved setting up accounts for each new user. Because our student employees had mostly moved back to their hometowns and were immersed in Zoom classes, we shifted them away from user support and toward various maintenance projects, such as downloading GIS data from local governments, collecting our local daily COVID-19 statistics into a spreadsheet, and enhancing metadata on a large historic dataset. We instituted a short weekly team check-in meeting via Zoom to touch base with everyone and track progress or roadblocks on projects. Because the librarians had limited capacity to imagine and manage projects, we did not replace the two of our four students who graduated until the university's return to in-person classes and the library's associated reopening in fall 2021.

Early in the pandemic, we recognized that teaching and learning via Zoom are not ideal and decided that we would mostly pause our workshop offerings and events other than course-integrated instruction. However, a silver lining to the remote environment was that we could partner with our colleagues at other University of California campuses to host much larger events than we have ever been able to on our own. The Carpentries instructors joined forces to offer UC-wide Carpentries workshops via Zoom, with a higher-than-in-person ratio of helpers to students. Likewise, thanks to Zoom, UC-wide celebrations for both GIS Week and Love Data Week each grew from one or two local events with spotty attendance at each campus into a full conference-like week with a schedule that combined and opened those individual campus workshops and talks for attendees from all campuses as well as later recorded viewing.

Recommendations and Best Practices

There are a number of professional opportunities to guide library professionals new to data and GIS services. ICPSR's celebrated *Providing Social Science Data Services* summer boot camp is offered bi-annually. Many books have been written on the practice of data librarianship and GIS/geospatial librarianship (see further reading section). Organizations such as IASSIST (International Association for Social Science Information Service and Technology), the ALA Map &

Geospatial Information Round Table (MAGIRT), the ACRL Digital Scholarship Section (DSS), and the Western Association of Maps Libraries (WAML) offer community through listservs, conferences, and training. Every campus is slightly different, though, so it is important to customize services to the specific needs of local undergraduate researchers and the requirements of the programs that enroll them. Likewise, it is critical to be aware of other campus service offerings to undergraduate researchers as well as service gaps, so that the library can provide complementary, not competing services. For example, central IT may offer specialized computing labs and software distribution, or an academic department such as mathematics might offer statistical consulting, and these kinds of services may or may not be available to undergraduates. It is also useful to be aware of larger structural strengths and limitations. For example, librarians at UC San Diego are academics but not tenure-track faculty, so we are not allowed to teach for-credit courses that might lead to a formal certificate, but we can offer attendees a record of their Carpentries workshop completion through the campus Teaching + Learning Commons' Co-Curricular Record.

Staffing will always be a major driver of services. As we noted, the UC San Diego Library can currently only hire students at an hourly wage, which limits our ability to hire graduate students because we can't compete with TA/RA-ships that include tuition remission. This, in turn, limits the level of assistance we can reasonably expect our student staff to provide. We have not yet been able to make the case for a staff position to supervise our student employees, so librarians have to juggle supervision with their public services duties, and sometimes lab staffing falls through the cracks, especially at the beginning of the academic year when student hiring is bottlenecked, and onboarding takes a backseat to class visits and user consultations. We recognize that having separate data science and GIS librarian positions is a luxury many institutions do not have, and UC San Diego was only able to make the case for both positions after trial and error with other staff classifications. Librarians typically do not learn data science or GIS skills through their MLIS coursework, and, though programs like IMLS LEADING strive to remedy this, the pool of qualified applicants to fill these kinds of librarian positions is small. At UC San Diego, we have been lucky to have flexibility in interpreting "ALA-accredited masters or equivalent" to include candidates who held degrees such as a master of GIS combined with extensive public service experience in higher education, as otherwise our most recent searches would have failed.

Finally, it is worth mentioning that working with undergrads (and students in general) is often just as much emotional labor as it is technical skill support. Students are stressed about their grades, their personal finances, and their post-graduation job or graduate school prospects. They often suffer from imposter syndrome, worry that their classmates have made more progress on their projects than they have, and are intimidated by their project advisors and mentors. Many students, especially BIPOC and first-generation students, do not

consider the library to be an inviting place or institution, and even those who do may not think of the library as their go-to source for data and GIS support. As such, data services and GIS librarians should take advantage of training to support students in distress and make an effort to be as inclusive as possible in their outreach efforts. Librarians should be mindful in their marketing to undergraduates and identify the best options for reaching them at or near their point of need, whether this is through liaison librarians' connections to courses and instructors, relationships with undergraduate research program staff, blurbs in targeted campus newsletters or social media, signage in the library and around campus, etc.

Conclusion

UC San Diego students are eager for research experience and to do original research, and they need access to datasets, software, and technology so that they can find, collect, store, clean, analyze, visualize, and publish their data. As novice researchers, they often need training and personalized help with using these tools. At UC San Diego, the library provides expert assistance in utilizing data and GIS for researchers of all experience levels as well as a Data & GIS Lab where students can use software for GIS and statistical analysis. We continue to target our services to meet users' evolving data needs and fill the gaps between other campus technology services, and, whenever possible, we partner with colleagues based in the library or elsewhere on campus to develop, expand, and maintain our services. Our users are generally quite appreciative of the assistance they receive, but as our work is somewhat niche, we continue to experiment with strategies for both marketing and expanding our services. For example, the COVID-19 pandemic inspired us to offer users access to our computers via remote desktop and then further invest in virtual machines as a service. Institutions that would like to offer similar services are encouraged to identify their students' data and computing needs as well as opportunities unique to their own campus.

Bibliography

Aber, Susan Ward, and Jeremy Ward Aber. *Map Librarianship: A Guide to Geoliteracy, Map and GIS Resources and Services.* Chandos Information Professional Series. Cambridge, MA: Elsevier, 2017. https://doi.org/10.1016/C2014-0-00541-2.

Abresch, John. *Integrating Geographic Information Systems into Library Services: A Guide for Academic Libraries.* Hershey: Information Science Pub, 2008. http://dx.doi.org/10.4018/978-1-59904-726-3.

Bishop, Wade, and Tony H. Grubesic. *Geographic Information: Organization, Access, and Use.* Springer Geography. Cham, Switzerland: Springer, 2016. https://doi.org/10.1007/978-3-319-22789-4.

Carlson, Jake, and Lisa Johnston, eds. *Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers.* Purdue Information Literacy Handbooks.

West Lafayette, IN: Purdue University Press, 2015. https://www.jstor.org/stable/j.ctt6wq2vh.

Geraci, Diane, Jim Jacobs, and Chuck Humphrey. *Data Basics: An Introductory Text*, 2012. https://doi.org/10.7939/R3251FK8F.

Johnson, Eric O. *Working As a Data Librarian: A Practical Guide*. Santa Barbara, CA: Libraries Unlimited, an imprint of ABC-CLIO, LLC, 2019.

Kellam, Lynda M., and Kristi Thompson, eds. *Databrarianship: The Academic Data Librarian in Theory and Practice*. Chicago: Association of College and Research Libraries, a division of the American Library Association, 2016. ProQuest Ebook Central.

Rice, Robin, and John Southall. *The Data Librarian's Handbook*. London: Facet Publishing, 2016. ProQuest Ebook Central.