

# THE LIBRARY'S ROLE IN BUILDING A DATA COMMUNITY

## 1. PLANS

Developing a data science education plan with the goals of developing a community of practice to improve data and software skills for early career researchers. By fostering a data community on campus, we make the library more relevant and credible in the data science education space.

- How do libraries and librarians respond to supporting data science?
- The fact that research is more and more computationally enabled -- some are claiming that computational thinking is integral to solving problems and understanding human behavior.
- Knowing how to create software and use/manage research data are foundational skills for being able to create knowledge.
- We identified patterns in the use of similar research tools and methods across many traditionally unrelated disciplines.
- We recognized there was an opportunity for the library to play an important role in fostering communities around software tools utilized across multiple disciplines.

## 2. DRY FITTING

We initiated our community building plan by offering data workshops in the library and targeted departments/schools on campus. The goal was to develop audiences, collaborators, and advocates for utilizing the library as a discipline neutral resource and space for data science training and education. We gained credibility from stakeholders by teaching Software Carpentry (SWC) workshops to demonstrate the need, impact, and value of the Carpentry workshops. This credibility led to partnerships with academic departments who sponsored additional funding for continued library data training initiatives for the campus research community.



- Initially the library paid for 1 year of Software Carpentry membership to start the data training initiative.
  - 3 members of the Library earned Software/Data Carpentry instructor certification by completing the instructor training program.
  - Certified instructors are committed to the campus community by teaching Software Carpentry workshops, improving lessons and related activities.
  - The 6 affiliate membership instructor training spots will help us grow the Software/Data Carpentry footprint on campus. These reserved spots will create a larger pool of Carpentry instructors, with discipline specific specializations, to increase workshop capacity on campus.
- School of Global Policy and Strategy and Research IT contributing to 2nd year for affiliation/professional development.

## 3. SQUARING

Positioning Data Librarians and the library as a primary data training stakeholder in the future of research communities.

- Helps researchers learn how to do research. This is a traditional function of libraries remade for a new environment: swap instruction on research methods that involved books/journals for research tools that involve code and data.
- "Library as Switzerland" - Fernando Perez (creator of Jupyter programming tool). The library is a common good neutral entity. Its neutrality lets the library partner with many different campus actors in addition to the longstanding model of liaising with departments. Supporting data science education and facilitating training for research computing skills should be a natural fit.
- The Carpentries provides tested lessons for re-use and an active community of instructors to use as a resource to learn from.
- Carpentry lessons and international community can accelerate the development of a library data science training program for the research community.
- Crowdsourced lessons benefit from many eyes on tested and proven content, but also provide ways of 'giving back' to a community with a common goal and interest.

## MATERIALS



## 4. CLEAN UP

Offering the various types of Carpentry and data workshops does have positive benefits for the library and community.

- After the workshops, we have noticed an increase in students asking for help or seeking consultations on the tools we taught.
  - For example, after teaching a 9 week data skills course, we have helped students with text analysis of twitter data using our subscription to the social media analytics platform, Crimson Hexagon, and provided data analysis support with R and Python.
- Graduate students also contact the library about their research:
  - To help them find more data (using services like Crimson Hexagon), needing help with analysis or data cleaning, web scraping data, or additional data related assistance.
  - To develop ongoing relationship from good word of mouth in the research community
- One of the primary benefits from teaching Carpentry workshops is building a community - for example, participating with SWC has connected us to their larger community of instructors and has given us ideas for more local community development.
  - As an example, we are planning to start up a UC San Diego Chapter of The Hacker Within (THW), a weekly meeting for sharing skills and best practices for scientific computation, after meeting with a researcher from UC Berkeley during a SWC instructor training event.
  - We connected with the undergraduate Data Science Student Society to sponsor their data science events held in the library.