

3-2020

## Connecting with undergraduate research: A pilot to tailor data literacy workshops in a library-led summer undergraduate research experience program

Chao Cai  
*Purdue University*, [caic@purdue.edu](mailto:caic@purdue.edu)

Megan Sapp Nelson  
*Purdue University*, [msn@purdue.edu](mailto:msn@purdue.edu)

Chaonan Liu  
*Purdue University*, [liu2313@purdue.edu](mailto:liu2313@purdue.edu)

Follow this and additional works at: [https://docs.lib.purdue.edu/lib\\_fspres](https://docs.lib.purdue.edu/lib_fspres)



Part of the [Library and Information Science Commons](#), and the [Scholarship of Teaching and Learning Commons](#)

---

### Recommended Citation

Cai, Chao; Sapp Nelson, Megan; and Liu, Chaonan, "Connecting with undergraduate research: A pilot to tailor data literacy workshops in a library-led summer undergraduate research experience program" (2020). *Libraries Faculty and Staff Presentations*. Paper 156.  
[https://docs.lib.purdue.edu/lib\\_fspres/156](https://docs.lib.purdue.edu/lib_fspres/156)

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact [epubs@purdue.edu](mailto:epubs@purdue.edu) for additional information.

# Connecting with undergraduate research: A pilot to tailor data literacy workshops in a library-led summer undergraduate research experience program

Chao Cai<sup>1</sup>, Megan Sapp Nelson<sup>1</sup>, Chaonan Liu<sup>2</sup>

<sup>1</sup>Purdue University Libraries and School of Information Studies; <sup>2</sup>Department of Biological Sciences, Purdue University

## Summary

Purdue University Libraries, in collaboration with a research unit, recently proposed to establish a summer undergraduate research experiences (URE) program with an emphasis on data literacy that serves students with limited access to research resources, i.e. underrepresented minorities, students from rural area and community colleges, and students with disabilities. The proposed URE professional development activities related to data literacy, specially data management and data ethics, are designed and tailored to this minority serving URE program based on the recommendations synthesized using reflections from librarians who have taught in data literacy workshops in similar URE programs. In recent years, many federal agencies provide funding opportunities to promote UREs in order to attract and retain a diverse pool of talented students and prepare these students for successful careers, and these URE programs usually emphasize the recruitment of students with limited access to research resources. Because of the increasing demand for data literacy in various disciplines, workshops on data science, specifically data management and data ethics, have been incorporated into the professional development activities in URE programs. Our efforts in tailoring data literacy workshops for minority serving URE programs are intended to connect data literacy education with the on-going and increasing needs in undergraduate research, especially with underrepresented groups of students with limited access to research resources.

## Questions

1. How the data literacy instructional setting in summer undergraduate research experiences (URE) programs differs from regular on-campus instructions?
2. How to tailor data literacy instructions in summer URE programs?

## Methods

- The study used a **case study** approach to synthesize recommendations to tailor data literacy instructions in URE programs.
- Participants were selected using the following **criteria**:
  1. Library faculty or staff who have experience **teaching data literacy**;
  2. Taught data literacy at least once in a **federal agency funded summer URE program**.
- Data collection was based on a semi-structured **interview**. Interviews were audio-recorded, and recordings were transcribed and coded using a **grounded theory** approach.

## URE Instructional Settings

### General settings:

- **Number of students:** 8 – 20 students
- **Length of instruction:** 3 – 4 hours

### Content Covered:

- **Data management:** storage & backup, metadata, file naming and organization;
- **Data ethics:** intellectual property, anonymity, data privacy

### URE disciplines:

- General Engineering
- General Agriculture
- Biology
- Chemistry

### Funding Agencies

- NIST
- NSF
- USDA

## URE Student Characteristics

- **URE students are indistinguishable from other student population:**  
“There wasn't anything that I can differentiate them [from other students] necessarily...” (Participant A)
- **URE students are more interactive, due to being in a small cohort:**  
“...It's a more personable interaction in the classroom...” (Participant B)
- **More often to have non-traditional students:**  
“...there tend to be a few older individuals in the cohort, so people who come back to school later...” (Participant B)
- **URE students tend to be more engaged:**  
“They (URE students) had projects, they knew this was for their own good, and they are very invested in this...” (Participant C)

## Resource Used for Material Development

Most participants claimed that there is no centralized resource that they use to prepare for the instruction. Resources are sparse out on the Internet. Participants recommended resources are listed below:

- **DataONE** (<https://www.dataone.org/>):  
“I have used DataONE's materials, not as they are, but for inspiration...I really like their tidy Excel exercise...” (Participant B, also mentioned by Participant D)
- **Data Curation Network** (<https://datacurationnetwork.org/>)
- **Data Carpentry** (<https://datacarpentry.org/>)
- **New England Collaborative Data Management Curriculum** (<https://library.umassmed.edu/resources/necdmc/index>)
- **Social media:**  
“...the other is that I follow people who are publishing in this area on Twitter...and that's also a way that I found those articles [related to data literacy], and that also happened on Facebook...” (Participant B)

## Recommendations to tailored instruction

- **Learning about the student population and student needs beforehand:**  
“...ask the host labs about their basic practices around data management...what their expectations for how people are going to get up to speed...” (Participant B)  
“...knowing what the goal is for that students means that you can tell the students what level of documentation/metadata...” (Participant D)
- **Using examples from real-life datasets:**  
“...I think it is important to show them that imperfect datasets are published and re-used all the time...” (Participant D)
- **Tailoring the instruction to the URE subject area:**  
“...It [the instruction] could vary depends on the discipline...you have to fit what researchers need...” (Participant A)
- **Talking about metadata using activities:**  
“...I think if you just talking to students about metadata, it seems so abstract...because it's an abstraction on top of an abstraction... but the real thing (metadata practice)...is what students can absolutely understand...” (Participant D)
- **Planning the timing of the instruction carefully:**  
“you should teach it [data literacy] only when they [students] have their [research] projects in hand...if you can hook them with what they are doing, it's more useful than doing it at the beginning, [when it's] abstract...” (Participant A)
- **Including library in the planning for an URE grant proposal:**  
“... We should be written into the grant... We should have been a part of the planning...” (Participant E)

## Future Improvement of the Study

- Participants interviewed **only represent one institution**. The study will benefit from a cross-institutional effort to include institutions with different sizes, locations, research emphases, etc.
- **Recommendations** for a tailored URE data literacy instruction are **not saturated**. Interviewing more participants will probably help discover more useful recommendations.
- There is a **lack of representation** of URE programs from **humanity/social sciences disciplines**. Including these disciplines may help with providing more comprehensive recommendations.

## Contact Information

For more information, please contact Chao Cai: caic@purdue.edu