Start Them Young, Raise Them Right

Undergraduates & Data Sharing

Christy Caldwell, Science & Engineering Librarian
University of California, Santa Cruz
caldwell@ucsc.edu

Abstract

The practice of teaching data literacy and research data management to university undergraduates is well documented. There has been substantial work accomplished towards pedagogy standards and integration into curriculum standards[1,2,3,4,5].

The main challenge to this effort is sustainability. The ratio of undergraduates to data specialists or subject liaisons is very large at most institutions. Solutions are often found with asynchronous teaching tools, curriculum integration and train the trainer approaches. This works if the goal is pedagogical. But some of these data undergraduates collect are valuable to the institution and/or the discipline, beyond curricular goals. Data sharing often involves more facilitation and consultation, so not all undergraduates can have this level of service. What criteria can we use to identify those populations?

Proposed Criteria

Identify what classes or populations of undergraduate students are collecting local, unique or institutionally relevant data that would be impossible or expensive to obtain again.

- Ecology or Environmental Sciences: data collection of local flora / fauna.
- Social Sciences: data collection of any local community or population
- Computer Science: code, games, or other cultural software objects.
- Film footage of communities that would be expensive or impossible to re-shoot.

Case Study

Senior capstone class Computational Media/Playable Media 80K: Foundations of Video Game Design[6].

Criteria fulfilled: Innovative unique game design approaches preserved for future game design students; preservation requested by faculty. Additional benefit of students linking to curated objects for employment opportunities.

Other outreach examples:

Environmental Studies Senior Theses
- Studying the Effects of Rock Doves on Pigeon Guillemots along West Cliff Drive, Santa Cruz, California.
- Understanding the toxicity of Scotch Broom to Douglas Fir seedlings.

Environmental Studies Senior Theses
- Examining the intersection between agricultural land management practices and bat species community composition.

Anthropology & Sociology Capstone Projects
- An Ethnography of Direct Trade Coffee in Santa Barbara, Honduras[7]
- Measuring foreclosures and loan modifications for immigrants in Watsonville, California.
- Documentary films interviewing marginalized communities about food justice.

License / Copyright

Unlike faculty and staff, students can own the data they collect under certain circumstances, and cannot be forced to upload their data[8].

General guidelines for students owning their intellectual property in the United States:

- If they used non-specialist equipment in an incidental way;
- They were not employees, did not sign any IP agreements, or enter into sponsorship agreements.

These issues can become complicated. Be prepared to address IP and copyright questions in regards to *student* work, and be sure to have students agree to the standard license for the data repository they will use.

Quality Control

There may be questions about how to assess data quality for non-expert researchers. It is important to incorporate quality assurance/quality control practices into the data collection workflow. This is best done in consultation with the research advisor, and is likely already be part of a research data management instruction.

Additionally, the Citizen Science communities are a good source of how non-experts can collect high quality data[9,10,11].

Literature Cited

1. New England Collaborative Data Management Curriculum | Lamar Soutter Library - University of Massachusetts Medical School https://library.umassmed.edu/necdmc/index (accessed Oct 13, 2017).
2. Lukyanenko, R.; Parsons, J.; Wiersma, Y.F. *Conservation Biology* **2016,** *30* (3), 447-449.
3. Piorun, M.; Kafel, D.; Leger-Hornby, T.; Najafi, S.; Martin, E.; Colombo, P.; LaPelle, N. *Journal of eScience Librarianship* **2012**, *1* (1).
4. Carlson, J.; Johnston, L.; Westra, B.; Nichols, M. *International Journal of Digital Curation* **2013**, *8* (1), 204–217.
5. Mooney, H.; Collie, W. A.; Nicholson, S.; Sosulski, M. R. *Advances in Social Work* **2014**, *15* (2), 368–389.
6. UC Santa Cruz: CMPS 80K [eScholarship] https://escholarship.org/uc/ucsc_games_cmps80k (accessed Oct 17, 2017).
7. Slocum, K. From Streetlevel to Farmlevel https://streetleveltofarmlevel.wordpress.com/ (accessed Oct 17, 2017).
8. Guide to Intellectual Property as a Student at the University of California | UCOP http://www.ucop.edu/research-policy-analysis-coordination/policies-guidance/intellectual-property-ex/intellectual-property-as-a-student-at-the-university-of-california-faq.html (accessed Oct 17, 2017).
9. Bonney, R.; Shirk, J. L.; Phillips, T. B.; Wiggins, A.; Ballard, H. L.; Miller-Rushing, A. J.; Parrish, J. K. *Science* **2014**, *343* (6178), 1436–1437.
10. Kosmala, M.; Wiggins, A.; Swanson, A.; Simmons, B. *Front Ecol Environ* **2016**, *14* (10), 551–560.
11. Barnes, L. LibGuides: Citizen Science: Evaluating Citizen Science http://guides.library.illinois.edu/citizen-science/evaluation (accessed Oct 16, 2017).