Date of Application: 4/30/2013

Title of Proposal/Project: Professional development in data curation

Expected Length of Project: 6 days

Total Funds Requested: $2,850

Primary Applicant

Name: Amanda L. Whitmire
Academic Rank and Working Title: Assistant Professor, Data Management Specialist
Campus Mail Address: 121 The Valley Library, Corvallis, OR 97331-4503
Telephone and E-mail Address: 514-737-3133; amanda.whitmire@oregonstate.edu

Co-Applicant(s): None

Project Summary & Rationale

Goal 3 in the OSU Libraries & Press Strategic Plan 2012-2017 describes a commitment to “curate the intellectual output of the Oregon State University community … to support the entire life cycle of that data including selection, certification, maintenance, preservation, discovery, and long-term access.” In order to succeed in this aspect of the goal, OSUL&P needs to have faculty on board who have expertise in data curation. I propose to use a Lundeen Award to facilitate professional development in data curation by attending a five-day workshop, Curating and Managing Research Data for Re-Use1, at the Inter-university Consortium for Political and Social Research (ICPSR). The workshop is intended to, “assist individuals working with data to apply efficient curation practices to ensure the usability and safekeeping of data resources. Participants will learn about best practices for managing research data, how to apply them to daily operations, and the types of tools that can assist in curation efforts. This workshop is for individuals actively engaged in the curation and management of research data, particularly data librarians, data archivists, and data stewards with responsibilities for data management” (see [1] or Appendix I for full course description).

The Lundeen Award is intended to boost professional capacity in areas that are critical to advancing the Library mission and that benefit our patrons2. While I have significant practical experience in data curation, I lack formal training in this area. Attending this workshop will enhance my ability to knowledgeably curate research data in ScholarsArchive@OSU (SA@OSU) and to assist patrons with curating their own data in the event that SA@OSU is not a suitable option for them. This both addresses our Strategic Plan and benefits the faculty and students who will increasingly rely on OSUL&P to assist with the curation of the products of

their research.

Aside from explicit training in data curation, there are two important indirect benefits of attending the workshop that have value which should not be underestimated: gaining hands-on experience with social science data (my background is in the natural sciences), and networking with the instructors and other participants. The workshop is designed to be interactive and hands-on, using a social science dataset as a case study in the process of curation through the ICPSR deposit process. While it is difficult to imagine that curating social science data is significantly different from curating data produced by scientific research, the workshop will broaden my experience and increase my data literacy. This workshop will provide an opportunity to gain familiarity with a social science dataset in a scenario that doesn’t involve taking risks with a ‘real’ dataset. The ICPSR workshop is being taught by experts in data curation and a colleague from a peer institution that is leading the way in developing academic data services (see Appendix II for an instructor list). Attendees of this workshop will presumably be peers at universities who are also developing library-based data curation services. Data curation is relatively new to the scope of library support services, and a significant amount of practical information is being exchanged peer-to-peer. The value of these interactions can’t be assessed or quantified, but they are essential to faculty professional development and are impossible to achieve via email. There is no substitute for being present.

In summary, my proposal to attend the ICPSR workshop Curating and Managing Research Data for Re-Use will build professional capacity within OSUL&P that will empower us to achieve our goal of supporting the whole lifecycle of data (Goal 3 of the Strategic Plan) and provide more knowledgeable support for patrons in the curation of their data. Participation will also facilitate productive interactions with peers within a burgeoning community of practice.

Goals & Expected Outcomes

The main goal of this proposal is to facilitate professional development in the area of data curation, which will improve the capabilities of OSUL&P to act as a steward of the research outputs of the university.

The expected outcomes include:

- Gain understanding of and ability to apply the best practices for managing research data, including management of confidential data and effective documentation practices;
- Increased ability to both conduct and teach data review and preparation for deposit into a repository (SA@OSU or otherwise);
- Increased capability to create, comply with, and evaluate required data management plans;
- Understanding of data repository requirements and assessment;
- Ability to identify and implement the tools that can assist in curation efforts.

---

3 “Curating and Managing Research Data for Re-Use.”
Methodology & Timeline

The professional development aspect of the proposed work will primarily occur through attendance of the workshop. Basic methods of active learning will be exercised. Upon return to OSU, new knowledge and skills will be integrated into my daily workflow. The timeline is shown below.

July 28 – August 3, 2013: Travel to ICPSR on the University of Michigan campus, Ann Arbor, Michigan (dates inclusive of travel from/to Oregon)

August 2013: Summarize the information gathered in the workshop and share with Center for Digital Scholarship & Services (CDSS) members (and other interested parties as appropriate)

Fall 2013: Integrate data curation best practices information into ScholarsArchive@OSU data deposit guidelines, data management education and training materials, and the CDSS Research Data Services web site (see dissemination plan).

Dissemination Plan

While the information gathered in this workshop will be of the most interest and direct benefit to me, there are a few potential venues for dissemination. As stated in the timeline, a summary of the workshop will be shared with CDSS, and will likely be of interest to those department faculty who are directly involved with SA@OSU oversight and implementation (digital applications librarian, metadata librarian, and head of the department).

If there is interest, I am happy to give an LFA Seminar Series and/or an in-service presentation about what I learned, and make the presentation openly available in ScholarsArchive@OSU. I will need assistance in gauging the level of interest and the scope of an appropriate topic for the presentation.

Perhaps more importantly, information gathered during the workshop will be disseminated to our patrons via “help text” on the SA@OSU web site. I will prepare a section of the site titled, “How to Prepare Your Data for Submission to the ScholarsArchive@OSU Repository,” which will cover topics such as how to prepare metadata, addressing privacy concerns, how to organize the files for upload, how to decide on embargo and license/rights options, and ultimately, how to submit data to the repository. Information will also be shared with graduate students via a data management course and workshops. Lastly, this information will trickle down to the Research Data Services web page, where we provide data management best practices information on many subjects that overlap with the scope of the workshop content.

Does the proposal require any of the following?

- Use of Human Subjects □ yes □ no
- Release time □ yes □ no

---

Budget

Total amount requested from Lundeen Award Fund: $2,850

Salaries: $0
Supplies: $0

Travel/Lodging

Flight*: $550
Booking fee (Azumano Travel): $25
HUT Shuttle (OSU/PDX): $85
AirRide Shuttle (Detroit/Ann Arbor): $24
Lodging**: $666

* Based on a survey of airline pricing for flights to/from PDX/Detroit on the days before and after the workshop.
** Based on Business Affairs per diem lodging rates for MI ($111/night x 6 nights). If travel is approved for this trip, I will try to find less expensive options, including dorm accommodations on campus and sharing a room or house with other workshop participants.

Other Expenses

Workshop registration: $1500 (ICPSR member rate)

Signatures of all applicants: 

Other Required Signatures: N/A
(e.g. Release Time Authorization)

Attachments: CV
Appendices: ICPSR workshop description (from web page); ICPSR workshop instructor profiles

NOTE: Grant proposals are confidential until funding decisions are made.
APPENDIX I - ICPSR workshop description (from web page)

Curating and Managing Research Data for Re-Use

Instructor(s):

- Jake Carlson, Purdue University
- Jared Lyle, ICPSR
- Ron Nakao, Stanford University
- Mary Vardigan, ICPSR

This workshop is for individuals interested or actively engaged in the curation and management of research data, particularly data librarians, data archivists, and data stewards with responsibilities for data management. The course will assist individuals working with data to apply efficient curation practices to ensure the usability and safekeeping of data resources. Participants will learn about best practices for managing research data, how to apply them to daily operations, and the types of tools that can assist in curation efforts.

This five-day workshop will explore and apply the concepts and benefits of life cycle principles for data curation, from preparing data for archiving to optimizing data for reuse. An ICPSR social science dataset will serve as a case study and participants will track the dataset as it makes its way through the ICPSR data pipeline. Examples from other scientific domains will also be integrated. Participants will learn about data review and preparation, confidential data management, effective documentation practices, how to create, comply with, and evaluate required data management plans, and repository requirements and assessment. Emphasis will be placed on hands-on exercises demonstrating curation practices and on small group discussions for sharing local experiences and learning from others.

An earlier version of this course, titled ‘Applied Data Science: Managing Research Data for Re-Use’, was offered in 2012.

Fee: Members – $1500; Non-members – $3000

Tags: curation, data science
APPENDIX II – ICPSR Workshop Instructor Profiles

Jake Carlson
Associate Professor of Library Science and Data Services Specialist, Purdue University

Jared Lyle
Director of Curation Services, ICPSR

Dr. Ron Nakao
Data and Computational Social Science Librarian, Stanford University

Mary Vardigan
Assistant Director, ICPSR; Director, Collection Delivery
Director of the Data Documentation Initiative (DDI) Alliance. (DDI is an international effort to establish a standard for social science metadata)
Amanda Lea Whitmire

Education

Oregon State University, College of Earth, Ocean and Atmospheric Sciences, Oceanography, Ph.D., July 2008.

University of California at Santa Barbara, College of Letters and Science, Department of Ecology, Evolution, and Marine Biology, Aquatic Biology, B.S. with honors, December 2000.

Employment

Assistant Professor, Data Management Specialist, Oregon State University Libraries, Corvallis, OR, September 2012 – present.

Research Associate (post-doctoral), College of Earth, Ocean and Atmospheric Sciences, Oregon State University, Corvallis, OR, July 2008 – August 2012.

Graduate Research Assistant, College of Earth, Ocean and Atmospheric Sciences, Oregon State University, Corvallis, OR, September 2001 – July 2008.


Publications

Refereed Papers – Library and Information Science

Whitmire, A. L. (in prep), “Thoughts on e-Science: a Scientist’s Perspective”

Refereed Papers – Oceanography


Non-refereed Publications – Oceanography (selected)


Presentations – Library and Information Science

Posters – Library and Information Science
Whitmire, A.L., "Training the next generation: the essential role of academic libraries in educating future scientists in research data management," ACRL Annual Meeting, Indianapolis, IN, April 2013. (30% acceptance rate)

Posters – Oceanography (selected)


Grants
National Sea Grant Industrial Partner Fellowship, 2005-2007
National Defense Science and Engineering Graduate Fellowship, 2002-2005
NASA Oregon Space Grant Fellowship, 2001

Service
Committees
OSU Libraries Awards Committee
OSU Libraries Poster Gallery Implementation Task Force
ALA | ACRL | STS | Science Resources: Discovery & Access Committee
(member: July 1, 2013 – June 30, 2015)
ALA | ACRL | Digital Curation Interest Group (member)

Professional Memberships
American Libraries Association (ALA)
Association of College and Research Libraries (ACRL)
The Oceanography Society (TOS)
American Association of University Women (AAUW)