



# Conducting a Science Data Services Needs Assessment

Brian Westra | Lorry I. Lokey Science Data Services Librarian | [bwestra@uoregon.edu](mailto:bwestra@uoregon.edu)

## Project Description

The needs assessment was an initiative by the University of Oregon Libraries to lay the foundation for a data services program. The intent of the project was to create relationships with partners outside the library, develop a team within the library for data support, and gather information on the data assets and needs of our researchers.

More specifically, the goals of the project were to:

- define the information services needs of science research faculty;
- inform the Libraries and other stakeholders of gaps in the current service structures;
- identify research groups or faculty whose datasets would be good subjects for a pilot data curation project

## Methods

1. Review of other similar projects, and question-set development.

The Data Audit Framework (DAF) methodology (University of Glasgow/DCC/JISC) was judged to be the most comprehensive set of questions with a field-tested methodology. The set of questions used in our assessment was based largely on the DAF, with input from Campus Information Services, and the Electronic Records Archivist who had recently worked with social scientists on a related project. The question set was further refined based on the interview questions employed in the Data Curation Profiles project by Purdue University Libraries and the University of Illinois, Urbana-Champaign.

2. Outreach

One of the recommendations of the DAF was to create a business case to formalize the “justification for undertaking the data audit, based on the expected costs, risks, benefits and savings” (DAF Methodology). Familiarity with these topics in the local context was helpful for outreach to deans, chairs, and others. Establishing common definitions for the topics and issues at hand improved the entire process of outreach and the interviews themselves.

Presentations to new research faculty, and at a meeting of the chairs of research centers and institutes facilitated participation. Library subject specialists’ knowledge of their respective science departments also aided outreach and the selection of interested faculty who were good candidates for the assessment.

3. Interviews

A web-based survey tool was briefly considered for collecting information from faculty. However, initial interviews established that more complete and accurate information about data management

approaches and related issues would be best gathered through an interview process. All information was recorded via a set of forms created with Drupal and accessed via login for each participant. This allowed faculty to log in and review and add/refine records about their data assets. The data was exported from Drupal into a .csv file for analysis and reporting via pivot tables and graphs in an Excel spreadsheet.

## Participating departments and institutes/centers:

- Biology
- Chemistry
- Computer & Information Science
- Geological Sciences
- Human Physiology
- Materials science/nanotechnology
- Museum of Natural and Cultural History
- Physics

## Data sets

As expected, a wide range of data sets, file formats, and domain-specific issues were represented in the assessment.

Examples:

- Biomechanics data
- C. elegans embryonic development videos
- Diel activity patterns
- DNA sequence data
- EMG, EEG, Motion Capture
- Environmental/organismal data
- Filemaker Pro databases
- Fossil collection database
- Geophysical and seismological data
- Glioma histology
- LiDAR and related GIS data
- Life history data
- Motion analysis data
- NMR Spectra
- Satellite-derived solar radiation data
- Scanning electron microscope data
- Stickleback crossing database
- Ultrasound avi files
- Windaq heart rate data

## Results

Significant benefits have been realized from the assessment:

- increased awareness of data management issues among research faculty and stakeholders in the campus organization
- requests for training for graduate students and lab personnel
- increased communication between research faculty and the Libraries about data management issues and data curation

Gaps/top issues are graphed below, and can be summarized as:

1. Infrastructure (storage, backup, network speed)
2. Need for data management tools and services in the research setting (annotation/metadata, [re-]discovery, version control). Faculty were most interested in solutions to address their needs for data management at the beginning of the data life cycle.
3. Faculty expressed some interest/needs related to publishing data sets (i.e., repositories), but this was far down the list of expressed needs in the interviews.

Improved communication about these issues, and the prospect of increased scrutiny and data management plan requirements from granting agencies, have raised the level of awareness and interest in these topics among faculty.

