Communication with Researchers

Or, Bridge Building for Librarians and Researchers


Patricia Hswe, Digital Collections Curator, Penn State University Libraries | phswe@psu.edu
Librarians and library services of the past

http://www.flickr.com/photos/nypl/3110116862/

http://www.flickr.com/photos/nypl/3110117236/
Librarians and library services now/future

Patricia Hswe | phswe@psu.edu

5/17/2011
One increasingly important set of users:

RESEARCHERS!

Patricia Hswe | phswe@psu.edu
What do we mean by “researcher”?

- Faculty
- Post-doctoral fellow
- Graduate student
- Independent researcher
- Advanced undergraduate (juniors, seniors)
- Others?
A tripartite approach

Be where they are (connect)
- Embedded librarians
- Liaison librarians

Find/Learn a common language (collaborate)
- Data, data, data!
- Data interview
- Research data lifecycle

Know their research practices (cohere)
- Survey for needs
- Respond with services
- Assess, adapt, evolve
brings about
connection, collaboration, cohesion
for libraries and librarians
(i.e., bridge building)
Library cuts threaten research
As journal cancellations sweep across the US, scientists worry about how they will affect research

Federal budget cuts threaten UNC research
While President Barack Obama is promising to limit federal cuts to scientific research, a new proposal from the Republican leadership would do just the opposite.
A proposed budget in the U.S. House of Representatives calls for a 9 percent cut in non-defense federal spending, which would lead to cuts in funding for institutions that provide funds to universities for research.
The cuts would affect organizations like the National Science Foundation, NASA and the National Institutes of Health.
U.S. Rep. David Price, D-N.C., said the cuts, which are meant to help alleviate the federal budget deficit, are a "disaster."
Including discontinuation of archives, repositories, data sets

Sequence Read Archive (SRA) and Trace Archive repositories have been discontinued

Due to budget constraints, NCBI will be discontinuing its Sequence Read Archive (SRA) and Trace Archive repositories for high-throughput sequence data. Closure of the databases will occur in phases. SRA and Trace will stop accepting some types of submissions in the coming weeks, and all submissions within the next 12 months. Over the next several months, NCBI will be working with staff from NIH Institutes that fund large-scale sequencing efforts to develop an approach for future access to and storage of the existing data. NCBI will continue to support and develop information resources for biological data derived from next-generation sequencing such as genotypes, common variations, rare variations, sequence assemblies and gene expression data. We therefore encourage the research community to continue submissions of these data to the applicable databases, including:

1. RNA-Seq and epigenomic data to GEO
2. Variants, genotypes, phased haplotypes, and polymorphisms to dbVar, dbGaP and dbSNP
3. Genomic assemblies to GenBank/WGS
4. Transcript assemblies to GenBank/TSA
5. 16S ribosomal RNA and other targeted locus survey assemblies to GenBank

NCBI expects new applications will continue to emerge for next generation technology. We are excited to work with the community to develop strategies for archiving other summary experimental measures that are informative, efficient, and valuable to the biomedical research community.

For further information about submissions, contact NCBI's Help Desk.
All the more reason . . .

for librarians to know how to communicate with researchers
Communication with researchers

means being where they are, to connect and collaborate with them
We’ve been connecting with researchers where they are in the form of Embedded Librarians and Liaison Librarians (both a type of Subject Librarian)
Embedded Librarians
June 9, 2010

BALTIMORE — Nancy Roderer is one for bold predictions. As a library consultant in the 1980s, Roderer predicted that all academic journals would be electronic by the mid-1990s.

A decade into the 21st century, Roderer’s opinion might now be considered prescient, if a bit off on the timing. It may have taken a little longer than she predicted, but every relevant academic journal now publishes an electronic version, and many journals only publish in the digital format.

Now, as director of the Welch Medical Library at Johns Hopkins University, Roderer has taken the opportunity not only to forecast sea changes in library science, but to pioneer them. By now, most librarians agree that the role of the library is changing, and that e-journals and e-books are poised to turn the library building into study space and librarians into e-sherpas, and many academic libraries have begun moving in that direction.
‘Embedded Librarian’ on Twitter Served as Information Concierge for Class

February 25, 2011, 2:08 pm
By Jeff Young

What if a reference librarian was assigned to a college course, to be on hand to suggest books, online links, or other resources based on class discussion? A media-studies course at Baylor University tried the idea last semester, with an “embedded librarian” following the class discussion via Twitter.

At the start of each class session, the professor, Gardner Campbell, asked the 11 students to open their laptops, fire up Twitter, and say hello to their librarian, who was following the discussion from her office. During the hourlong class, the librarian, Ellen Hampton Filgo, would do what she refers to as “library jazz,” looking at the questions and comments posed by students, responding with suggestions of links or books, and anticipating what else might be helpful that students might not have known to ask.

Also: the embedded researcher


A Bimonthly Report from ARL, CNI, and SPARC

A Special Issue on Liaison Librarian Roles

Introduction: Positioning Liaison Librarians for the 21st Century...1
Karla Hahn, Assistant Executive Director, Research, Teaching, and Learning, ARL

A Framework for Articulating New Library Roles ...3
Karen Williams, Associate University Librarian for Academic Programs, University of Minnesota Libraries

Amplifying the Educational Role of Librarians ...9
Elizabeth A. Dupuis, Associate University Librarian for Educational Initiatives and Director of the Doe/Moffitt Libraries, University of California, Berkeley

The Last Mile: The Liaison Role in Curating Science and Engineering Research Data ...15
Tracy Gabridge, Co-Head, Engineering and Science Libraries, Massachusetts Institute of Technology

Scholarly Communications: Planning for the Integration of Liaison Librarian Roles ...22
Joy Kirchner, Librarian for Collections, Licensing and Digital Scholarship, University of British Columbia Library

New Roles of Liaison Librarians: A Liaison’s Perspective ...29
Kara M. Whatley, Life Sciences Librarian and Head of the Coles Science Center, New York University

University of Saskatchewan - Music and Education Liaison Librarian

Posted: Monday, April 18, 2011
Music and Education Liaison Librarian (Librarian I or II)
Applications are invited for a tenure-track appointment with the University Library at the University of Saskatchewan. We are seeking a dynamic, innovative and flexible individual with the appropriate qualifications and background to work collaboratively as part of the team based at the Education and Music Library, one of our seven branch libraries. The appointee will have primary responsibility for the music collection and liaison.

Faculty Liaison Librarian – Science and Technology

Employer: Anglia Ruskin University
Location: Cambridge
Published Date: 20 Apr 2011
Status: Permanent
Ref: 000018-7

Job Description
Faculty Liaison Librarian – Science and Technology

University Library

Based in Cambridge


Our vibrant, modern University has ambitious plans for its future. Our two main campuses in Cambridge and Chelmsford are being transformed with major capital investments. With an annual income of £145m, 31,000 students and 2,100 staff, we are a major force for higher education in the East of England.

We are looking for an innovative and resourceful individual who can demonstrate expertise in building effective partnerships with academic staff, collection management and advanced information literacy principles and practice. Leading a team of assistant librarians across sites, you will be responsible for the delivery of high quality academic library services to the Faculty of Science and Technology.
means collaborating: finding common ground, possibly learning a new “language”
## Data, data, data!

<table>
<thead>
<tr>
<th>Reality</th>
<th>Methods (&quot;language&quot; training)</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Common language for some librarians, for others a new language to learn</td>
<td>- U. Virginia Library – “How to Retool Librarians for Data Curation”</td>
</tr>
<tr>
<td>- Need internal education/outreach programs for librarian colleagues</td>
<td>- Penn State – “Research Data Management Services Team”</td>
</tr>
</tbody>
</table>
INTRODUCTION
Librarians at Purdue University are beginning to identify the scientific datasets that are being generated by our faculty and researchers as information assets to be collected, preserved, and made accessible as a function of the library’s collection development. These librarians are subjectarea specialists, and many have advanced degrees in their respective disciplines in addition to a degree in library science. They have all been trained in collection management; however, much of this training was related to traditional formats such as monographs and serials and not data sets. In our experience, one of the most effective tactics for eliciting datasets for the collection is a simple librarian-researcher interview. In this poster, we share a set of questions that a librarian can use as a starting point for such a “data interview.” It is not a comprehensive strategy but instead a practical tool to draw out information that needs to be considered in order to evaluate the suitability of a dataset for the collection and the requirements for the infrastructure and services that will be needed for data curation.

#2 What form and format are the data in? What computing environments (e.g., software) are required to use the data? If the data are in proprietary structures, you may consider reformating them into a format that can be made available for review. If there are any associated metadata, either external to the data or description that could be extracted from it? Ideally the data could be described to be discoverable by researchers from another discipline.

#3 What is the expected lifespan of the dataset? In many cases, there are distinctions in the utility of a dataset as it begins in a raw state and then is analyzed and processed into new forms and versions as a result of different steps in the research workflow. Different entities may have custody of the data and use it for different purposes at different times, affecting its provenance. Funding agencies may require that data be archived for a prescribed period of time or you may forecast its future value and the amount of time it should be retained. The data may be described and archived for effective preservation to ensure its accessibility and integrity over time.

#4 How could the data be used, reused, and repurposed? This is a primary selection criteria that also impacts how the data are accessed and what policies may be needed to govern its use. As data are archived and shared, new and unintended uses for the data may increase its value. For example, a research dataset may be repurposed as a learning object.

#5 How large is the dataset, and what is its rate of growth? It is important to quantify the size of the data for storage and network provisioning if you intend to ingest it into your repository. What is its physical [lite] and logical [records] scale? Is the dataset static or dynamic? Ask for a sample of the data to examine.

#6 Who are the potential audiences for the data? Information regarding potential users of the data and the needs is paramount. Along with potential uses for the data, this is another primary selection criteria. In some cases, the data may need to be embargoed or restricted to a limited group of users who are granted permission to access it.

#7 Who owns the data? Establishing and maintaining the intellectual property represented by the data should be discussed at the earliest opportunity, and any conflicts should be resolved up front. Many organizations have a submission policy that asks the contributor to verify that they own the data and have the right to submit it.

#8 Does the dataset include any sensitive information? All data should be reviewed for information that violates confidentiality, such as identification information on human subjects. Data curation activities should be informed by institutional review board requirements.

#9 What publications or discoveries have resulted from the data? The researchers may have a list regarding the importance of their data. The purpose of this question is to establish an objective metric for determining the value of the data for the collection. Different metrics may be more appropriate in determining the selection criteria for different kinds of data and data collections.

#10 How should the data be made accessible? There is value in making data accessible using a conventional web-based user interface, but machine-readable interfaces should also be evaluated. These methods of access will be informed by the answers to the previous questions and this question can be asked in an open-ended manner to fill in any gaps remaining at the conclusion of the interview.

SUMMARY
Although building robust collections of datasets presents several complexities and challenges to resolve, the process of looking at scientific datasets as information assets and exploring what is needed to develop and manage data collections is similar to the traditional collection development practices that have been successfully employed by librarians for decades. We offer these ten “data interview” questions as a springboard for librarians to explore data curation in greater depth and specialization.

Michael Witt (witt@purdue.edu) Assistant Professor of Library Science
Jake Carlson (jcarlso@purdue.edu) Data Research Scientist
Purdue University Libraries
Distributed Data Curation Center
http://c2r2.is.lib.purdue.edu

"Conducting a Data Interview"
Michael Witt & Jake Carlson, Purdue University Libraries, West Lafayette, Indiana, USA
Creating data

- design research
- plan data management (formats, storage etc)
- plan consent for sharing
- locate existing data
- collect data (experiment, observe, measure, simulate)
- capture and create metadata
Communication with researchers

means cohering programs and services in libraries with what we know about research practices
Survey for needs: Data Asset Framework

1. PLANNING THE AUDIT
   A business case is developed, meetings scheduled and preliminary research conducted. The online tool has a calendar and address book to schedule meetings and track progress.

2. IDENTIFYING AND CLASSIFYING ASSETS
   An inventory of assets is created in this stage. The online tool can generate and circulate a questionnaire to help populate the inventory.

3. ASSESSING MANAGEMENT OF DATA ASSETS
   Vital assets are assessed in detail to investigate data management approaches. The online tool pulls through information already collected.

4. REPORTING AND RECOMMENDATIONS
   A report with recommendations is created. The tool analyses data collected to produce statistics and generate an appendix.
Welcome to the Data Curation Profiles community!

Welcome to the community for Data Curation Profiles Toolkit!
This website is an environment where academic librarians of all kinds, special librarians at research facilities, archivists involved in the preservation of digital data, and those who support digital repositories can find help, support and camaraderie in exploring avenues to learn more about working with research data and the use of the Data Curation Profiles Tool.

A Data Curation Profile is essentially an outline of the “story” of a data set or collection, describing its origin and lifecycle within a research project. The Profile and its associated Toolkit grew out of an inquiry into the changing environment of scholarly communication, especially the possibility of researchers providing access to data much further upstream than previously imagined. If researchers are interested in sharing or forced to provide access to data sets or collections, what does that mean for the data, for researchers, and for librarians?

Data Curation Profiles can:
- provide a guide for discussing data with researchers
- give insight into areas of attention in data management
- help assess information needs related to data collections
- give insight into differences between data in various disciplines
- help identify possible data services
- create a starting point for curating a data set for archiving and preservation
Survey to Assess Data Services Needs at Penn State

The University Libraries and its Research Data Management Services Team are conducting this survey for current NSF grant awardees. The objectives are to assess data management needs at Penn State, particularly in terms of services for storage, back-up, sharing, data documentation and description of data, and long-term archiving of data.

The answers to this survey will help the libraries develop sustainable data services that meet the needs of Penn State’s research community. For more information about what the libraries are doing about data management planning see: http://www.libraries.psu.edu/pui/scholar/datamanagement/libraries_help.html.

The survey will close on May 12.

Note: A red asterisk (*) beside a question denotes that an answer is required.

1. What is your department and/or research institute affiliation? Your position at Penn State? *

2. Are you aware of the National Science Foundation's data management plan requirement? *
   - Yes
   - No

3. Has your project budgeted for management of research data? If so, then how? *

4. Who currently manages, or who will manage, data for your project? *
   - Research assistant / Graduate assistant
   - Post-doctoral researcher
Respond with new services (informed by survey results) - 1

Workshop and Training

The library offers data management workshops periodically throughout the year as free one-hour sessions. See the library's full list of workshops to register for the next available session. Or view the recorded versions.

Next Available Session

Workshop: Creating a Data Management Plan

In January 2011, the National Science Foundation (NSF) began requiring a data management plan with all new grant proposals. This workshop, geared toward graduate students, faculty and researchers, will provide an overview of the questions to consider when creating a data management plan with a focus on data preservation issues and tools for sharing your data, like subject-based repositories. This workshop is discussion-based and we work through each of the five data management questions via group and paired activities.

Register for this Course

Fri, 05/20/2011 - 2:00pm - 3:30pm
This course qualifies for RCR credit.
Room 314 Walter Library
Lisa Johnston

Recorded Workshops and Slides

- Introduction to Data Management for Scientists and Engineers
- Creating a Data Management Plan
- Data 101: Best practices for managing your digital research data (coming Spring 2011)

Training & tools

Training - internal and external

- University of Oxford - Data Management training
- OUCS IT Learning Programme Courses
- Bodleian Libraries courses
- Research Skills Toolkit
- Oxford Supercomputing Centre (OSC)
- External training

Tools & guidance

- Tools & guidance
Respond with new services (informed by survey results) - 3
Assess, adapt, evolve

- Assessing services helps keep libraries accountable

- Adapting in response to changing researchers’ needs reinforces trust, loyalty, relevance

- Evolving means continually learning, improving, and delivering

But what does the above have to do with communication with researchers?
Assessing, adapting, and evolving

all result from listening to researchers, and listening is communicating
“Liaisons have a long track record of building bridges with others in similar roles at other institutions in order to bring best practices into local practice.”

Images

- “Text a Librarian Booth at ALA Midwinter,” [http://www.flickr.com/photos/41993437@N00/3687133143/](http://www.flickr.com/photos/41993437@N00/3687133143/), (CC BY 2.0)