Supporting Research Data: A Collaborative Library Approach

ANNA CRAFT
METADATA CATALOGER
THE UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

LYNDA KELLAM
DATA SERVICES LIBRARIAN

Questions:

• How many of you work in institutions that are currently providing research data management (RDM) support services for users?

• Is RDM support a cross-departmental effort in your institution, or is it centralized?
Overview

- UNCG and the University Libraries
- UNCG’s research data support service model
- Outreach efforts
- Institutional repository/Dataverse integration
- Challenges, value, future directions

https://www.flickr.com/photos/canon_duo/4950617867  (CC BY-ND 2.0)
UNC Greensboro

• Public, four-year research university

• One of seventeen campuses in the UNC System

• FTE is 19,000 students
UNCG University Libraries

Collections:
- 2.8 million items (books, govdocs, and microforms)
- 42,666 electronic journals
- 250+ electronic databases

Collection strengths:
- Local and regional history
- Music/performing arts
- University history
- Women’s history

Staff:
- Approximately 90 FTE
- 30 tenured or tenure-track faculty
WT Cox Subscriptions Acquires Journal Finder From UNCG

WT Cox Subscriptions (www.wtcox.com) has entered into an agreement to purchase Journal Finder from The University of North Carolina Greensboro (UNCG; http://journalfinder.uncg.edu/uncg). Journal Finder, the first OpenURL Link Resolver to go into production in the U.S., gives library patrons easy access to the electronic resources to which a library subscribes. Journal Finder was created and designed and has been maintained by the Electronic Resources and Information Technology staff of UNCG.
Why research data management? And why the libraries?

• The Libraries are already in the business of preserving, providing access, and supporting faculty research.

• Pre-existing data discovery support.

• Culture of support for open access, desire to extend those offerings.

• Funder requirements (NSF, NIH, and other granting agencies).

• Demonstrated need on campus.
  ◦ Survey results
  ◦ Requests from faculty
Secondary data discovery

• New Data Services & Government Information Librarian position created in 2007.

• Focus on supporting data discovery primarily for government data, data archives, and subscription sources.
From UNCG Faculty Research Data Support Needs Survey report, October 2013
We’re part of a bigger picture

Other partners in campus research data management services:

• Office of Sponsored Programs
• Information Technology Services
• Departmental programs
  ◦ School of Health and Human Sciences
  ◦ School of Nursing
What do we do?

• Educate about data and data management best practices.
• Assist in the creation of data management plans.
• Provide an option for sharing and archiving data (NC DOCKS/ Odum Institute partnership).
• Connect researchers to other (non-library) services.

http://uncg.libguides.com/RDM
Library roles

- Research, Outreach, and Instruction
- Technical Services
- Electronic Resources and Information Technology
- Research Data Support
Research, Outreach, and Instruction (ROI)

- Serves as connection to the faculty and research offices regarding data needs.
- Provides training to graduate students and faculty on RDM best practices.
- Provides assistance with data management plans or archiving options as requested.
- Works within the liaison team structure to support researchers in every discipline.
Technical Services

◦ Sets metadata policies and creates documentation.
◦ Manages item and metadata loading in NC DOCKS (IR) and Dataverse.
◦ Provides assistance with research data questions.
◦ Targeted outreach to faculty.
◦ Training, training, training.
Electronic Resources and Information Technology (ERIT)

- Manages NC DOCKS server, develops and supports NC DOCKS product.
- Consults on requests involving local technology support.
Whom do we support?

Students, faculty, and staff.
Whom do we support?

Institutional partners
Research data management task force

- **Personnel:**
  - Associate Dean for Public Services
  - Assistant Dean for Electronic Resources and Information Technology
  - Health Sciences Reference Librarian
  - Science Liaison Librarian
  - Data Services and Government Information Librarian
  - Metadata Cataloger

- Develop a plan for supporting RDM in the library
Targeted outreach

• Open Access Policy initiative
• Departmental meetings
• Outreach to research offices, Research Advisory Council, and Office of Research Integrity
• Research data management best practices workshops
• “Responsible Conduct of Research: Taking the High Road to Better Data” presentation series
Open access policy initiative

Open Access: Sharing your scholarship

What about research data?

- Data management planning resources
- Research data hosting (through a partnership with the Odum Dataverse at UNC Chapel Hill)
- There are other options - just ask to learn more!
Outreach to campus departments and research groups
RDM workshops

Research Data
@ UNCG Univ. Libraries

Lynda Kellam, Data Services Librarian, lmkellam@uncg.edu
RESPONSIBLE CONDUCT OF RESEARCH

Responsible Conduct of Research (RCR) involves shared responsibility for the responsible conduct of research among the institutional administration, the research oversight committees and the researchers. An effective RCR program includes comprehensive institutional policies and procedures for the conduct and review of research, coordination of all units (administrative offices, oversight committees, etc.) that impact the responsible conduct of research, systematic education programs in RCR and systematic quality assurance/quality improvement activities.

Responsible conduct of research (RCR) includes:

- Conflict of interest (personal, professional, financial)
- Policies regarding human subjects, live vertebrate animal subjects in research, and safe laboratory practices
- Mentor/mentee responsibilities and relationships
- Collaborative research, including collaborations with industry
- Peer review
- Data acquisition and laboratory tools; management, sharing and ownership
- Research misconduct and policies for handling misconduct
- Responsible authorship and publication
- The scientist as a responsible member of society, contemporary ethical issues in biomedical research, and the environmental and societal impacts of scientific research

The UNCG Office of Research Integrity (ORI) offers training to meet the regulatory requirements of both the National Institutes of Health and the National Science Foundation. Additionally, this program meets the institutional goal of developing a university-wide approach to meeting the ethical standards that it was founded upon.
Types of requests

• Informational – “How can you help me?”
  ◦ What services do you provide in this area?
  ◦ Will you look at my data management plan?
  ◦ What repository should I use?

• Technical support – “This isn’t working.”
  ◦ Why won’t Dataverse load my data set?
  ◦ Why isn’t this file type displaying?

• Training – “I need help doing this.”
  ◦ How do I write a data management plan?
  ◦ How do I get my data into Dataverse?
Questions we ask: DMP creation stage

• What type of data will be produced?
• What standards will be used for documentation and metadata?
• What steps will be taken to protect privacy, security, confidentiality, intellectual property, or other rights?
• If you allow others to reuse your data, how will the data be accessed and shared?
• How will the data be archived for preservation and long-term access?

A note on DMPs

- RDM LibGuide offers general DMP information and tips
- We offer access to DMPTool, but we are not a partner, so we don’t offer templates for researchers
- No formal requirement for researchers to bring DMPs to the library

Questions we ask: Researchers who already have data

- What field is the researcher working in?
- What format is the data in?
- What is the size of the dataset?
- Is there existing metadata? Is there a codebook?
- Are there privacy concerns with the data, such as PII?
UNCG’s research data partnership
NC DOCKS at a glance

• Open access
• Home-grown system
• Collaborative partnership bringing together eight UNC system schools
• Faculty, staff, and student works
• Articles, presentation slides, ETDs, audiovisual materials, and more
• Discovery via Google
http://libres.uncg.edu/ir/
Metadata for NC DOCKS

• Metadata entry is handled in the Cataloging Department
• Harvested by the WorldCat Digital Collection Gateway
• MARC XML export option
• OAI-PMH

Please select an administrative function below.
Fossil GPS

UNCG Author/Contributor (non-UNCG co-authors, if there are any, appear on document)
Robert Anemone, Professor and Department Head (Creator)

Institution
The University of North Carolina at Greensboro (UNCG)
Web Site: http://library.uncg.edu/

Abstract: The article focuses on paleontology and discusses how luck has played a big role in many of the world's great fossil discoveries. It also discusses a new technique that improves the odds of finding ancient bones. It reports on the use of modern computer models that look for hidden patterns in satellite images and generate maps of fossil sites, helping fossil hunters narrow their search.

Additional Information

Publication
Scientific American, 310(5), 48-51
Language: English
Date: 2014

Keywords
Paleontology, Computer simulation, Remote-sensing images, Fossils, Paleoarcheology

University of North Carolina at Greensboro University Libraries LibQual+ results

UNCG Author/Contributor (non-UNCG co-authors, if there are any, appear on document)  Kathryn Crowe, Associate Dean for Public Services (Creator)

Institution  The University of North Carolina at Greensboro (UNCG)

Web Site: http://library.uncg.edu/

Abstract: This LibQual+ library survey data was collected by the Association of Research Libraries (ARL) for the University Libraries at the University of North Carolina at Greensboro. Survey respondents were students, faculty, and staff at the University of North Carolina at Greensboro. This data was collected from October 15, 2012 to November 10, 2012. This data set and its associated documentation are available for download. To view and download the file, follow the link below and choose the "Data & Analysis" tab.

Link to data and documentation  http://hdl.handle.net/1902.29/11520

Additional Information

Publication  Language: English
Date: 2013

Keywords  Library surveys, Library assessment, gap analysis surveys, Library service quality, Academic library assessment

This item requires:
University of North Carolina at Greensboro Dataverse

UNIVERSITY OF NORTH CAROLINA AT GREENSBORO UNIVERSITY LIBRARIES
LIBQUAL+ RESULTS
hdl:1902.29/11520
Version: 1 – Released: Mon Aug 12 09:30:42 GMT-05:00 2013

If you use these data, please add the following citation to your scholarly references. Why cite?


Citation Format

Title
Study Global ID

Data Citation Details

http://arc.irss.unc.edu/dvn/do/UNCG/faces/study/StudyPage.xhtml?globalId=hdl:1902.29/11520
## Metadata in two systems

<table>
<thead>
<tr>
<th><strong>NC DOCKS</strong></th>
<th><strong>DATAVERSE</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Holds basic metadata about the data set, as well as a link to Dataverse.</td>
<td>• Holds metadata and the actual data set(s), with option for much more detailed metadata.</td>
</tr>
<tr>
<td>• Data entry by Libraries’ personnel.</td>
<td>• Data entry by faculty member or other user (with support from Libraries, if needed).</td>
</tr>
<tr>
<td>• Metadata fields controlled and defined by Libraries’ personnel.</td>
<td>• Metadata fields controlled and defined by Dataverse.</td>
</tr>
<tr>
<td>• Fields are general, to accommodate many types of scholarly works.</td>
<td>• Fields are specific to research data and are DDI compliant</td>
</tr>
</tbody>
</table>
## Documentation

<table>
<thead>
<tr>
<th>Field</th>
<th>Field Part</th>
<th>Dataverse: Required / Recommended / Other</th>
<th>NC DOCKS: Required / Recommended / Other</th>
<th>Dataverse description</th>
<th>DC map</th>
<th>Data field format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>NA</td>
<td>Required</td>
<td>Required</td>
<td>Full authoritative title for the work at the appropriate level: marked-up document; marked-up document source; study; other material(s) related to study description; other material(s) related to study. The study title will in most cases be identical to the title for the marked-up document. A full title should indicate the geographic scope of the data collection as well as the time period covered. Title of data collection (2.1.1.1) maps to Dublin Core Title element. This element is required in the Study Description citation.</td>
<td>none</td>
<td>text box - free</td>
</tr>
<tr>
<td>Subtitle</td>
<td>NA</td>
<td>Other</td>
<td>Other</td>
<td>A secondary title used to amplify or state certain limitations on the main title</td>
<td>none</td>
<td>text box - free</td>
</tr>
<tr>
<td>Publication</td>
<td>Citation</td>
<td>Recommended</td>
<td>Recommended</td>
<td>Bibliographic reference that includes author, title, year, and publisher of the publication associated with this study and data. Use any standard bibliographic format.</td>
<td>none</td>
<td>text box - free</td>
</tr>
<tr>
<td>Publication</td>
<td>Replication data for Original Publication</td>
<td>Recommended</td>
<td>Recommended</td>
<td>This will apply &quot;Replication Data for:&quot; to your title</td>
<td>none</td>
<td>check box</td>
</tr>
</tbody>
</table>
What we don’t do

• Mandate library review of all DMPs
• Accept legal responsibility for data deposited in NC DOCKS/Dataverse
• Know the researcher’s data/research area as well as they do
• Store private/confidential/sensitive data
  ◦ Personal Identify Information
  ◦ Electronic protected health information (ePHI/HIPAA data)
• Maintain restricted data access contracts
The User Submissions do not contain information that could, directly or indirectly, identify subject except where the release of such identifying information has no potential for constituting an unwarranted invasion of privacy and/or breach of confidentiality. And for all personally-identifiable information in the data at least one of following conditions applies:

- this information has been previously released to the public;
- this information describes public figures, where the data relates to their public roles or other non-sensitive subjects;
- a sufficient length of time has passed since the collection of the information so that the data can be considered "historic";
- all identified subjects have given explicit informed consent allowing the public release of the information in the dataset;
- all information was collected with an explicit statement concerning the public nature of the data, such as information collected for governmental regulatory purposes;
- or, for federal records (data created by a U.S. federal government agency or under a federal contract) only, all identified subjects are deceased and no federal statute explicitly restricts the release of the data.

By clicking the "I agree" checkbox below, I confirm that I have read and understood each and every term set forth in the terms and conditions for the use of data and other materials found above, and I agree to be bound by all of such terms and conditions.

If I do not understand or agree to all of the terms and conditions, I must not use or download any data or other materials.

☐ I agree and accept these terms of use.

Continue  Cancel
NC DOCKS/Dataverse process

• Researcher contacts liaison librarian.

• Liaison librarian explains our services or contacts DS Librarian.

• Liaison librarian contacts Metadata Librarian.

• Further communications with researcher, via email or in-person.

• Metadata Librarian and/or liaison librarian assist researcher in loading materials into Dataverse and completing metadata entry.

• Metadata librarian completes associated metadata record in NC DOCKS.
From the UNCG Faculty Research Data Support Needs Survey report, October 2013
We aren’t the only option

- Dryad
- figshare
- ICPSR
- ...and more!
How is it going?

• Outreach sessions offered?
  • 2013-2014 – 1 session
  • 2014-2015 – 5 sessions
  • 2015-2016 – 6 sessions

• Information requests received?

• Access RDM page?

• DMPs created?

• Data sets loaded?
Challenges

• Limited resources (both budgetary and personnel)
• Communication and publicity
• “It’s one more thing to do”
• Changing expectations/requirements for data management
• The Research Data Lifecycle
The Research Data Management Lifecycle

1. Research Question
2. Data Search / Reuse
3. Data Management Plan
4. Collection
5. Description
6. Data Storage
7. Analysis
8. Re-collection
9. Publication
10. Archive
Value

• Strengthen relationships between Libraries, researchers, and scholarly communities
• Meet needs of campus constituencies
• Support new publishing/sharing models
• Demonstrate commitment to open access and the research data lifecycle
Looking forward

• Continue to follow developments in this area.
• Consider further extension of service offerings when budget allows.
• Publicize services to campus researchers and offer educational opportunities.

Photo by flickr user bengtham (CC BY-NC-ND 2.0)
Thank you!

Anna Craft
Metadata Cataloger
arcraft@uncg.edu

Lynda Kellam
Data Services Librarian
lmkellam@uncg.edu
UNCG Data Services Links

- NC DOCKS (Institutional Repository)
  - [http://libres.uncg.edu/ir/uncg/](http://libres.uncg.edu/ir/uncg/)

- Data management LibGuide
  - [http://uncg.libguides.com/RDM](http://uncg.libguides.com/RDM)

- UNCG Faculty Research Needs Survey Report

- Scholarly Communications LibGuide
  - [http://uncg.libguides.com/scholarlycomm](http://uncg.libguides.com/scholarlycomm)

- Odum Institute at UNC Chapel Hill:
  - [http://www.odum.unc.edu/odum/home2.jsp](http://www.odum.unc.edu/odum/home2.jsp)