Mitigating the Risk: Identifying Strategic University Partnerships for Compliance Tracking of Research Data and Publications

Andrea Payant
andrea.payant@usu.edu
@rusros25

Betty Rozum
betty.rozum@usu.edu
@brozum

Liz Woolcott
liz.woolcott@usu.edu
@lizwoolcott

The Problem

Dog Desk  photo by buzbedo, CC-by-2.0
Researchers generate Lots of Data

Not all researchers organize and manage their data

Faculty *may* be required to deposit data to make it publicly accessible
It’s Here!

Dancing Discovery … “Explored!”  photo by Stephanie, CC-by-NC 2.0
What Keeps Our VPR Up at Night?
Can you find where your institution’s data is hiding?
How Can Libraries Help?

Outside a dog … photo by Bruno, CC-by-NC 2.0
Librarian Skills

- **Data Management Skills**
  - Familiar with Data Management Plans
  - Familiar with Data Repositories
  - Familiar with Depositing Data
  - Discipline agnostic

- **Cataloging Skills**
  - Understand how to make things discoverable
Partnership

Bone 2 photo by jb CC-by-2.0
USU Data Task Force

- 16 members with representatives from
  - Library
  - Research Office
  - Information Technology
Working Group

- Smaller group met regularly
  - Library
    - Data Services Coordinator
    - Metadata Specialist
  - Research Office
    - Associate VPR
    - Director Research Development
    - Sponsored Programs Director
    - Programmer
Key Resources

- Kuali: an electronic award management system; USU’s official record for Sponsored Programs
- DigitalCommons@USU: USU’s institutional repository
- USU’s Integrated Library System: Sierra
Kauli captures basic elements for future Primary Master Record

DSP notifies PI of requirements, requests DMP; DMP and/or Primary Metadata Document sent to Library; Library creates records in Digital Commons
DSP sends PI notice every 6 months to update PMD, sends updated PMD to Library; Library verifies data, creates records
DSP continues to notify PI, even after closeout, until all data deposited
DMP - Guiding Document

Gone 3 photo by Conor Lawless. CC-by-2.0
Records the Library Creates

- Workflow calls for the Library to create
  - Digital Commons:
    - Master Record – represents the PI’s Grant
      - Includes the DMP, if allowed
      - Includes the “Primary Metadata Document”
    - Dataset records
      - Metadata only
      - Metadata plus data files
    - Metadata records for publications
  - ILS:
    - Dataset Records
### Master Record

Represents the grant in its entirety

Includes DMP (available through the “Download” link)

Primary Metadata Document (available through “Additional Files”)

---

<table>
<thead>
<tr>
<th>Document Type</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication Date</td>
<td>5-9-2016</td>
</tr>
<tr>
<td>Grant Number</td>
<td>USDA 2006-35503-17194</td>
</tr>
<tr>
<td>Funding Agency</td>
<td>USDA NIFA</td>
</tr>
<tr>
<td>Comments</td>
<td>Data management plan from USDA NIFA Grant 2006-35503-17194</td>
</tr>
<tr>
<td>Recommended Citation</td>
<td>Broadbent, Jeff. “CHARACTERIZING STRESS RESPONSES OF INDUSTRIAL STRAINS OF BIFIDOBACTERIA AND THEIR USE FOR EXTENDING THE SURVIVAL OF BIFIDOBACTERIA IN FOODS” (2016). Federally Funded Research. Paper 1. <a href="http://demo.usu.boxress.com/federally_funded_research/1">http://demo.usu.boxress.com/federally_funded_research/1</a></td>
</tr>
</tbody>
</table>

**Additional Files**

- [PMD_Test_Jeff_Rev2.xlsx](PMD_Test_Jeff_Rev2.xlsx) (30 KB)
- Primary Metadata Document
Generated by Kuali and PI

Updates requested every 6 months

Used for
  - Setting up initial Master Record
  - Verifying and creating records for data deposits
  - Adding funding information to publication records in Digital Commons
**Title/Name assigned to grant**
Characterizing Stress Responses of Industrial Strains of Bifidobacteria and Their Use for Extending the Survival of Bifidobacteria in Foods

**Place where data originated**
Logan, UT

**Primary institution name**
Utah State University

**Project start and stop dates**
Sep 1 2006-Aug 31, 2010

**Granting Agency, grant award number**
USDA 2006-35503-17194

**Subject of research data**
food products, bacteria, quality maintenance in storing and marketing food products, bifidobacterium, probiotic, stress response

**Agency Progress and Final Report Location (URL)**

**Publications**

<table>
<thead>
<tr>
<th>Publication Citations (repeatable)</th>
<th>Details</th>
</tr>
</thead>
</table>

**Data Deposits (or Other Associated Data)**

<table>
<thead>
<tr>
<th>Title/Name assigned to data set</th>
<th>Expression data from Bifidobacterium longum strains exposed to hydrogen peroxide stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description (100 word limit)</td>
<td>Stress survival tactics in bacteria utilize the up- and down-regulation of stress response genes. In bacterial that lack classical stress response genes for oxidative stress, other cellular systems can be used for cell survival. We used custom microarrays to study the regulation of genes in Bifidobacterium longum strains to oxidative stress to elucidate novel stress response mechanisms.</td>
</tr>
<tr>
<td>Year of publication/deposit</td>
<td>2013</td>
</tr>
<tr>
<td>File type (ex. Txt,XML,PDF)</td>
<td>TXT, XML</td>
</tr>
<tr>
<td>Is a special program or software needed to access this data? If yes what is it?</td>
<td>No</td>
</tr>
<tr>
<td>Link to associated Journal Article (repeatable)</td>
<td><a href="http://doi.org/10.1016/j.jbiotec.2015.06.405">http://doi.org/10.1016/j.jbiotec.2015.06.405</a></td>
</tr>
</tbody>
</table>
## Dataset Record

Metadata records created in DigitalCommons

**Location of datasets verified**

---

### Grant Number
- **USDA 2006-35503-17194**

---

### Abstract
- Bifidobacterium longum D2957, whole genome shotgun sequencing project.
Online Catalog Record

Author: Oberg, Taylor S., researcher.
Title: Transcriptional Responses of Bifidobacterium longum Strains to Hydrogen Peroxide Stress / T. S. Oberg, J. R. Broadbent, R. E. Ward, J. L. Steele.

Connect to: Dataset deposited here: http://www.ncbi.nlm.nih.gov/nuccore/AQGL0000000
Link to associated article: http://www.sciencedirect.com/science/article/pii/S0168165615300316
Article DOI: http://doi:10.1016/j.jbiotec.2015.06.405

Description: 1 dataset.
Content: computer dataset
Carrier: online resource
This entry is the master record for a whole genome shotgun sequencing project and contains no sequence data. See additional links below for access to full sequence data.
The Bifidobacterium longum D2957 whole genome shotgun (WGS) project has the project accession AQGL0000000. This version of the project (01) has the accession number AQGL01000000, and consists of sequences AQGL0100001-AQGL0100013.
Summary: Bifidobacterium longum D2957, whole genome shotgun sequencing project.
Local Note: USU Data Deposits.
Funding: USDA Cooperative State Research, Education, and Extension Service Improving Food Quality and Value Program National Research Initiative Grant 2006-35503-17194.
Subject: Broadbent, Jeff R., researcher.
Added Author: Ward, Robert E., researcher.
Steele, James L., researcher.
Publications

Many will already have metadata records in DigitalCommons

Add funder information to existing records

Optional – add URL from agency repository to DigitalCommons record

If no record exists, create record with appropriate funder and agency repository information
Library Staff and Work Flow

- Creation of “Master Record” – including adding DMP and PMD and PMD revisions
  - Student Assistant

- Verification of data links and creation of dataset records
  - Student Assistant **

- Creations and/or editing of publication records, including verification of deposit in agency repository
  - Student Assistant **

- Creation of ILS records
  - Student assistant

  (** problems addressed by supervisor)
### Cataloging Data

<table>
<thead>
<tr>
<th>OCLC</th>
<th>NEW</th>
<th>Save File</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Computer Files</strong></td>
<td><strong>Rec stat</strong></td>
<td><strong>n</strong></td>
<td><strong>Entered</strong></td>
</tr>
<tr>
<td>Type</td>
<td>m</td>
<td>ELvl</td>
<td>K</td>
</tr>
<tr>
<td>BLvl</td>
<td>m</td>
<td>Form</td>
<td>o</td>
</tr>
<tr>
<td>Desc</td>
<td>i</td>
<td>File</td>
<td>u</td>
</tr>
</tbody>
</table>

| 007 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 040 | 1 | UUS | ib | eng | ie | rda | ic | UUS |
| 090 | 1 | ib |
| 049 | 1 | UUSA |
| 100 | 1 | [Name of Researcher], ie [relator term] |
| 245 | 1 | [Name/Title assigned to data] / ic [Researcher(s)/Authors of the data] |
| 264 | 1 | [Place where data originated]: b [Institution name], ic [year of data publication] |
| 300 | 1 | 1 dataset |
| 336 | 1 | computer dataset 2 rdacntent |
| 337 | 1 | computer 2 rdamedia |
| 338 | 1 | online resource 2 rdacntent |
| 347 | 1 | [Digital file characteristics = [file type] lb [encoding format] ic [file size] |
| 500 | 1 | [Granting Agency/Grant Award # Information] |
| 500 | 1 | [Any additional information pertinent to the data/dataset that is not otherwise related in the MARC record - use readme file, if available, as a reference] |
| 500 | 1 | [Add Citation/Reference Information for associated article] |
| 520 | 1 | [Summary/Abstract] |
| 538 | 1 | [System details note — include information about the characteristics of computer files for example: mode of access, software programming language, and computer requirements - use readme file, if available, as a reference] |
| 590 | 1 | USU Data Deposits |
| 590 | 1 | [University Department] |
| 650 | 0 | [Subject of research data] |
| 700 | 1 | [Name of Additional Researcher], ie [relator term] |
| 856 | 4 | 0 | Dataset deposited here: tu [insert url for dataset] |
| 856 | 4 | 0 | Dataset DOI: [insert doi for dataset] |
| 856 | 4 | 2 | Link to the associated journal article here: tu [insert link for associated article] |
| 856 | 4 | 2 | Article DOI: tu [insert doi for associated article - must include "http: " or the record will not validate] |

**Required**

**Recommended**
## MARC/Dublin Core Mappings

<table>
<thead>
<tr>
<th>MARC Mapping</th>
<th>DC Mapping</th>
<th>Field Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>Creator</td>
<td>1st Author/Researcher listed</td>
</tr>
<tr>
<td>245 $a</td>
<td>Title</td>
<td>Title/Name assigned to data set</td>
</tr>
<tr>
<td>245 $c</td>
<td>All authors/researchers listed</td>
<td></td>
</tr>
<tr>
<td>264 $a</td>
<td>Place where data originated</td>
<td></td>
</tr>
<tr>
<td>264 $b</td>
<td>Primary institution name</td>
<td></td>
</tr>
<tr>
<td>264 $c</td>
<td>Date</td>
<td>Year of publication/deposit</td>
</tr>
<tr>
<td>347 $a</td>
<td>Digital characteristics - file type, refer to the file extension</td>
<td></td>
</tr>
<tr>
<td>347 $b</td>
<td>Digital file characteristics - encoding format</td>
<td></td>
</tr>
<tr>
<td>347$c</td>
<td>Format.Extent</td>
<td>Digital file characteristics - file size</td>
</tr>
<tr>
<td>500</td>
<td>Description</td>
<td>Granting Agency, grant award number</td>
</tr>
<tr>
<td>500</td>
<td>Description</td>
<td>Any additional information pertinent to the data/dataset that is not otherwise reflected in the record. Refer to the readme file, if available.</td>
</tr>
<tr>
<td>500</td>
<td>Relation.IsReferencedBy</td>
<td>Citation for original publication based on this data set</td>
</tr>
<tr>
<td>538</td>
<td>Relation.Requires</td>
<td>Include information about the characteristic of the files, noting mode of access, software or computer access. Refer to the readme file, if available.</td>
</tr>
<tr>
<td>520</td>
<td>Description.Abstract</td>
<td>Include any summary information about the content of the dataset, such as an abstract.</td>
</tr>
<tr>
<td>650</td>
<td>Subject</td>
<td>Subject of research data</td>
</tr>
<tr>
<td>700</td>
<td>Creator</td>
<td>Name(s) of additional researcher</td>
</tr>
<tr>
<td>856</td>
<td>Description</td>
<td>URL for location of dataset</td>
</tr>
<tr>
<td>856</td>
<td>Identifier</td>
<td>Dataset DOI</td>
</tr>
<tr>
<td>856</td>
<td>Relation.IsReferencedBy</td>
<td>Link to associated Journal Article</td>
</tr>
<tr>
<td>856</td>
<td>Associated Journal Article DOI</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>Indicate the DCMI type (typically &quot;Dataset&quot;)</td>
<td></td>
</tr>
<tr>
<td>Format</td>
<td>Indicate the file format of the dataset, refer to the MIME types</td>
<td></td>
</tr>
<tr>
<td>Publisher</td>
<td>Indicate where the data set is housed</td>
<td></td>
</tr>
<tr>
<td>Coverage.Spatial</td>
<td>If reported, include the spatial coverage for the dataset</td>
<td></td>
</tr>
<tr>
<td>Coverage.Temporal</td>
<td>If reported, include the date coverage for the dataset</td>
<td></td>
</tr>
</tbody>
</table>
Rewards (Benefits)

“Another fun dog” photo by Joan Valencia CC-by-SA 2.0
Benefits - Partnerships

“Fight for your right to bite”  photo by Eddy Van 3000  CC by-SA-2.0
Benefits for University

- Verify compliance
- Capture the location of data while it’s (relatively) fresh in the mind of researchers
  - Opportunity to “rescue” data insecurely stored
- Create permanent records of data
- Increase discoverability of data
- Reporting functions help University to understand and analyze research data creation and lifecycles
- Sharing successful DMPs = better future DMPs
Benefits for Library

- Increases value on campus, strengthens partnership with Research Office
- Increases interactions with faculty and demonstrates value by securing data and helping faculty comply with DMP
- Gains opportunity to help faculty learn about better options for data deposit
- Cross campus synergistic relationships
Library will assess:
- Staff time (current and projected future), costs to library
- Stakeholder satisfaction with workflow & services
- Change in quality of Data Management Plans (DMPs)
Summary

Project Goal

- Create an audit system that efficiently tracks data and publication deposits resulting from federally funded research with as little impact as possible on staffing of any group involved (PI, Research Office, Library)
  - Library – most of the work will be handled by student employees
- Going live Fall 2016
- Developing benchmarks to define success or identify areas for improvement
Questions?

“Rio the black mixed breed dog with newspaper” photo by Found Animals Foundation CC by-SA-2.0
Contacts

Andrea Payant  
Data Management Metadata Specialist  
@andrea.payant@usu.edu  
@rusros25

Betty Rozum  
Data Services Coordinator and Undergraduate Research Librarian  
@brozum

Liz Woolcott  
Head of Cataloging and Metadata Services  
@lizwoolcott