Classifying Data Deposited by Scientists into a Library's Data Repository

Andrew T. Creamer  
*Brown University, andrew_creamer@brown.edu*

Ann C. Caldwell  
*Brown University, e_caldwell@brown.edu*

Joseph Rhoads  
*Brown University, Joseph_Rhoads@brown.edu*

*See next page for additional authors*

Follow this and additional works at: [http://escholarship.umassmed.edu/escience_symposium](http://escholarship.umassmed.edu/escience_symposium)

Part of the [Life Sciences Commons](http://escholarship.umassmed.edu/lifesciences), [Medicine and Health Sciences Commons](http://escholarship.umassmed.edu/medicinehealthsciences), [Physical Sciences and Mathematics Commons](http://escholarship.umassmed.edu/physicalsciencesandmathematics), and the [Scholarly Communication Commons](http://escholarship.umassmed.edu/scholarlycommunication)

[http://escholarship.umassmed.edu/escience_symposium/2016/posters/19](http://escholarship.umassmed.edu/escience_symposium/2016/posters/19)

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in University of Massachusetts and New England Area Librarian e-Science Symposium by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
Presenter Information
Andrew T. Creamer, Ann C. Caldwell, Joseph Rhoads, and Ben Cail

Keywords
data repository, data management, digital scholarship, digital research products

Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 4.0 License.
Classifying Data Deposited by Scientists into a Library’s Data Repository

AT Creamer, EA Caldwell, J Rhoads, and B Cail
Brown University Library, Providence, Rhode Island USA

Introduction

Libraries that are developing their own data repositories benefit from data on the nature of objects in established data repositories so that they can prepare their future platforms, services, and policies accordingly. While repository reports usually focus on the number, formats, and sizes of files deposited, we have instead endeavored to describe other useful attributes, such as whether the deposited data are associated with a publication or whether data collections are related to a sponsored project.

This poster reports the results of our review and analysis of the individual data deposits (n=20) as well as collection-level data libraries (n=11) deposited by scientists into Brown Library’s data repository, the Brown Digital Repository (BDR), from April 2014 to April 2016 using a rubric, developed by the authors, and information provided by depositors.

About the Brown Digital Repository

By depositing their data in open, sharing, and preservation-friendly formats, along with appropriate documentation, in the Brown Digital Repository (BDR), scientists can maximize the potential for their data’s discovery, access, use, and reuse.

BDR platform relies on Fedora Commons software and offers an open API for programatically retrieving both structured metadata and digital content in public collections. The BDR is searchable using an Apache Solr index and via a public search API.

Items in the BDR are stored redundantly in off-site storage. Audit trails are maintained for each object to document changes and deletions, and older versions of files are retrievable in the event of unintentional modification. Objects also receive a checksum value to allow for periodic auditing of data integrity.

Objects receive a permanent link and a digital object identifier (DOI) to maximize their potential for discovery, access, and citation.

Rubric Criteria for Classifying Deposits

U: Data underlying a publication
Data results reported in a publication, necessary for the peer review and reproduction of the paper’s reported results.

S: Supplementary files accompanying a publication
Data such as images, tables, or visualizations that were not able to be included in the published paper or were cited as supplementary in the paper.

DS: Digital scholarship
Digital objects that are equivalent to a publication, of a scholarly nature, and intended for dissemination (e.g., video, animation, database, etc.)

DC: Data libraries (data collections)
A collection of scientific data, not connected with a single publication or a grant, and is available for other researchers or the public to access and reuse. This category includes legacy data.

GF: Grant-funded project (data collections)
A collection of data that is associated with a specific funded-project; the purpose is to make these available for other researchers and the public to access and reuse.

Results

n=20 individual project deposits

- Digital Scholarship (equivalent to a publication): 35%
- Related-to-Publication: 65%

n=13 publication-related project deposits

- Supplementary Data: 31%
- Data Underlying Reported Results: 69%

n=13 publication-related deposits

- Monograph: 9%
- Journal Article: 95%

Future Directions

- The BDR is a means to support reproducibility
- The BDR helps faculty to meet grant-funders’ data sharing mandates and publishers’ requirements for retaining, citing, and sharing research data
- The BDR offers researchers the means to archive and publish their digital scholarship

- Begin archiving more digital objects related to undergraduate honors theses, master’s-level theses, and doctoral-level electronic dissertations
- Linking and harvesting BDR assets into faculty profile reporting systems, VIVO, and ORCID scholarly profiles
- Enable Google Scholar indexing of BDR assets
- Provide users with analytics, metrics, and evaluation
- Provide users with suggested citations
- Faculty use BDR instead of personal websites and cite BDR assets on their CVs, personal websites, and ResearchGate profiles
- Create digital preservation-related policies

Print your poster

To order your poster, go online to PosterPresentations.com and click on the “Order Your Poster” button. Choose the poster type that best suits your needs and submit the form. Select a size of 48” x 48” or 36” x 36” and submit your order. Once your order is placed and paid for before noon, Pacific, Monday through Friday, your order will ship out that same day. Next day, Second day, and Same Day options are available.

After your order is printed, go to VIEW > SLIDE MASTER. After you finish working on the master be sure to go to VIEW > NORMAL to continue working on your poster.

How to change the template color theme

1. Click on the HOME menu.
2. Under the CONTENT section, click on the STYLE button.
3. Click on the second style in the second row.

How to change the column configuration

1. Right-click on the poster background and select LAYOUT to see the column options available for this template. The poster columns can also be customized on the MASTER. VIEW > MASTER.

How to add Text

1. Click on the HOME menu.
2. Under the TEXT section, click on the INSERT TEXT button.
3. Enter the text into the text box.

How to remove the info bars

1. Click on the HOME menu.
2. Under the FORMATTING section, click on the INFO BAR button.
3. Click on the “OFF” button.

How to add Graphics / Charts

1. Click on the HOME menu.
2. Under the INSERT section, click on the IMAGE button.
3. Select the image you want to add.

How to add Video

1. Click on the HOME menu.
2. Under the INSERT section, click on the VIDEO button.
3. Select the video you want to add.

How to add a table from scratch

1. Click on the HOME menu.
2. Under the INSERT section, click on the TABLE button.
3. Enter the row and column options.

How to add Copied and Pasted Text

1. Click on the HOME menu.
2. Under the EDIT section, click on the COPY button.
3. Select the text you want to copy.
4. Go to the location where you want to paste the text.
5. Right-click and select PASTE.