

4-20-2018

What to Do with Research Data

Julia A. Lovett
University of Rhode Island, jalovett@uri.edu

Harrison Dekker
University of Rhode Island, hdekker@uri.edu

Follow this and additional works at: https://digitalcommons.uri.edu/lib_ts_presentations

Recommended Citation

Lovett, Julia A. and Dekker, Harrison, "What to Do with Research Data" (2018). *Technical Services Faculty Presentations*. Paper 59.
https://digitalcommons.uri.edu/lib_ts_presentations/59

This Article is brought to you by the University of Rhode Island. It has been accepted for inclusion in Technical Services Faculty Presentations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu. For permission to reuse copyrighted content, contact the author directly.

What to Do with Research Data



Julia Lovett
Digital Initiatives Librarian

Harrison Dekker
Data Services Librarian

Search Savvy Seminar
University of Rhode Island Libraries
April 2018

The background is a solid orange color. In the top-left corner, there are three vertical bars of varying heights, each composed of several overlapping semi-transparent orange circles. In the bottom-right corner, there are four vertical bars of varying heights, also composed of overlapping semi-transparent orange circles.

Data in DigitalCommons@URI



DigitalCommons@URI, URI's institutional repository

- <http://digitalcommons.uri.edu>, URI's platform for sharing scholarly works and other digital output
- Contains faculty publications, theses and dissertations, Senior Honors Projects, departmental publications, scholarly journals with faculty editors, and more!
- Hosted by Bepress and managed by URI Libraries



Parameters for depositing data in DigitalCommons@URI

- **File size:** Recommended up to 2GB but can go up to 10-20GB
- **Multiple files:** Zipped up, or one main file plus supplemental files
- Any **file type**, but check out [recommended practices](#) for long-term preservation and accessibility and/or consult with Harrison
- Required **metadata:** Title, Author(s), Date, Abstract/description, link to article; **provide a legend to define variables**
- **Access options:**
 - Fully open with Creative Commons CC0 license
 - Hidden and accessible only through a direct link
 - Delayed access, under embargo until specified date (metadata is visible)
- **Privacy:** You must certify that you own the data or have permission to share it; we will review for privacy and/or rights issues



Registering a DOI

- **Citing data:** Digital Object Identifiers are permanent identifiers/links that persist even when URL's change; for example, if we moved from Digital Commons to another system, the DOI would stay the same
 - Example: <https://doi.org/10.23860/dataset-mclean-2017>
- We can also register a Digital Object Identifier (DOI) for your data set, whether or not you deposit it to Digital Commons -- but, the data set has to be deposited somewhere
- The Libraries create the ID and register it through the CrossRef system



Benefits of DigitalCommons@URI

- Convenience of deposit
- Quickly comply with publisher and funder requirements
- Customer service and easy consultations
- Usage information for your data set
- Long-term preservation, URI Libraries take responsibility



How to deposit?

- Contact Julia or Harrison directly
 - jalovett@uri.edu
 - hdekker@uri.edu
- Include file as email attachment if possible along with descriptive information
- From there, we will consult and deposit it on your behalf



Data management





Getting help

Library Data Services

Training/Consulting/Grant Collaboration

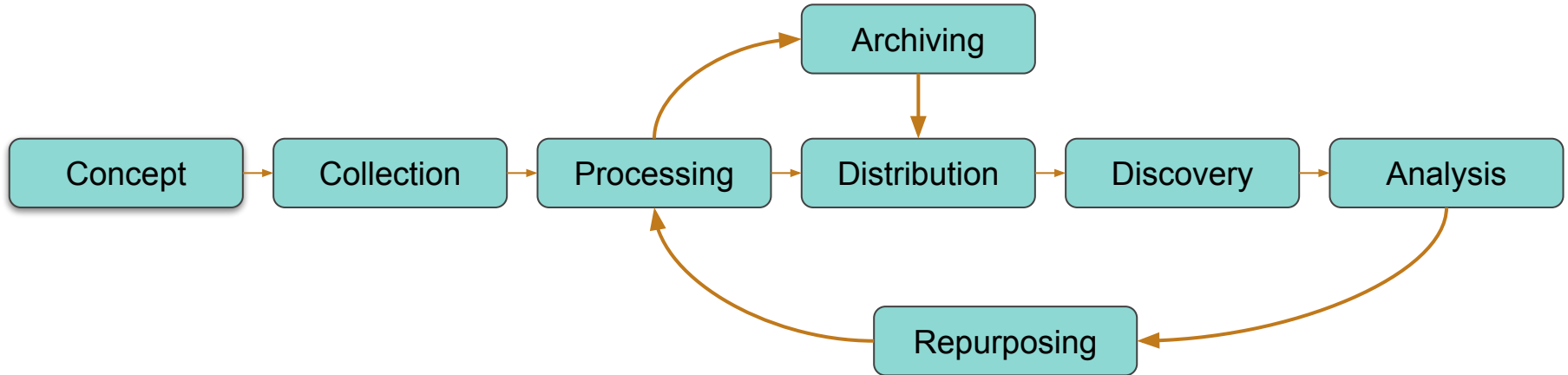
Technical: R, Python, Stata, SQL, Git, File management, Format conversion

Research: Finding data, Reproducible research, Data cleaning, Data sharing

Contact: Harrison Dekker - hdekker@uri.edu



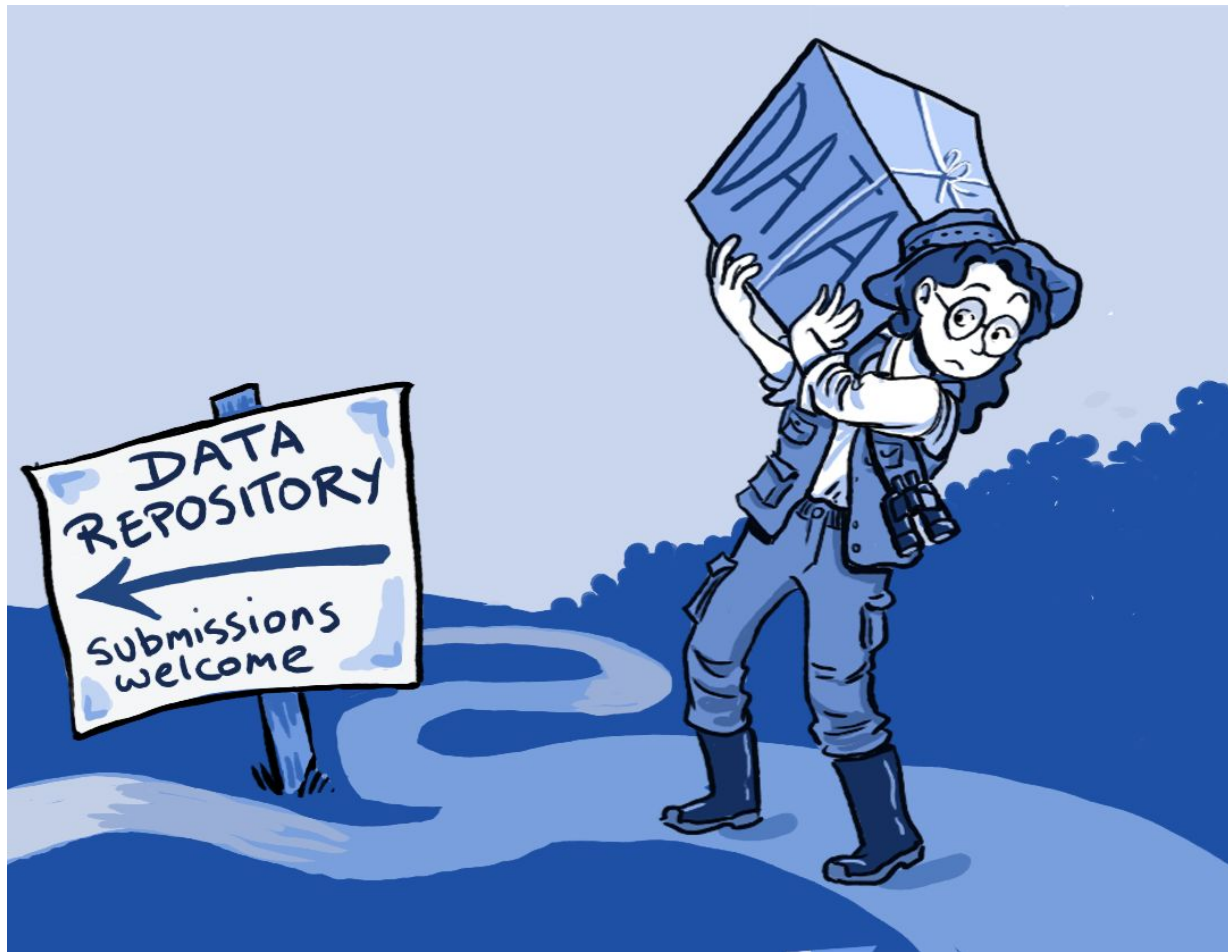
Data Lifecycle





Why share data





Roche DG, Lanfear R, Binning SA, Haff TM, Schwanz LE, et al. (2014) Troubleshooting Public Data Archiving: Suggestions to Increase Participation. *PLoS Biol* 12(1): e1001779. doi:10.1371/journal.pbio.1001779



How to share data



DIGITALCOMMONS@URI

The
Dataverse
Project 

zenodo


DRYAD

ICPSR

re3data.org

REGISTRY OF RESEARCH DATA REPOSITORIES

Search...

 Search



**2,000 Data Repositories
and Science Europe's
Framework for
Discipline-specific
Research Data
Management**

**Three new DOI Fabrica
features to simplify
account management**

Last month we launched DOI Fabrica, the modernized version of the DataCite Metadata Store (MDS) web frontend. It is the one place for DataCite

**One step closer towards
instant DOI search
results**

Art Art? You might be wondering, what this pink and green picture illustrates? A few months ago we couldn't show you this picture, the data that we used to created it,

The image features a solid orange background. In the top-left corner, there are three vertical bars of varying heights, each composed of several overlapping semi-transparent circles. A similar set of four vertical bars is located in the bottom-right corner, also made of overlapping semi-transparent circles.

Data preparation for sharing



FAIR data principles

FINDABLE

Data and supplementary materials have sufficiently rich metadata and a unique and persistent identifier.

ACCESIBLE

Metadata and data are understandable to humans and machines. Data is deposited in a trusted repository.

INTEROPERABLE

Metadata use a formal, accessible, shared, and broadly applicable language for knowledge representation.

REUSABLE

Data and collections have clear usage licenses and provide accurate information on provenance.

Documentation





Important documentation elements

- Motivation for dataset creation
- Dataset composition
- Data collection process
- Data preprocessing
- Data distribution
- Dataset maintenance

Formats





Common file format problems

- Proprietary formats that can only be read with commercial software
- Orphaned proprietary formats
- Non-proprietary formats with insufficient information (e.g. fixed width formats, poorly named variables)
- Multiple spreadsheet-related issues!

Confidentiality





Issues with data related to people

- IRB rules
- Subject consent
- De-identification problems
- Bottom line is to seek guidance before sharing any human subject data. (That said, many types of useful “people-related” data is shared in ways that are ethical and legal.)

Copyright/Licensing





Common issues

- Do you have permission to share the data, e.g. was any of portion of it acquired under specific licensing terms.
- Did any of the procedures for acquiring data violate actual laws or implied (or easily overlooked) terms of use.



Thank you!

Questions?

Julia Lovett, jalovett@uri.edu

Harrison Dekker, hdekker@uri.edu