

Data Repositories

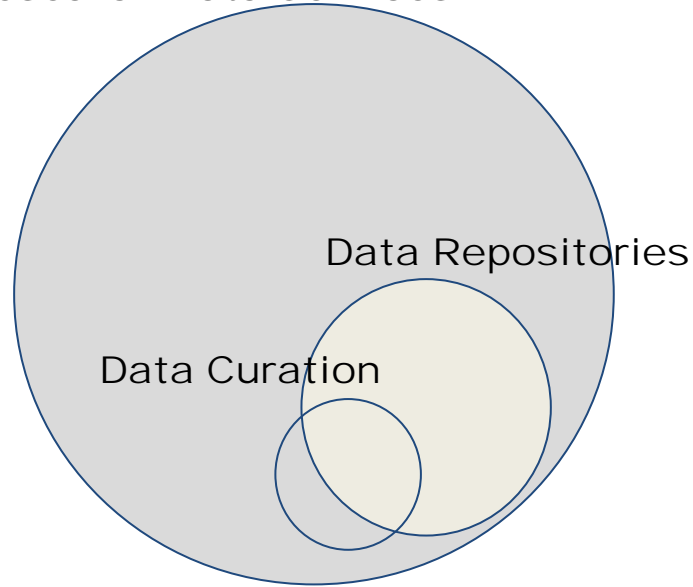
The answer that actually came with a
question

Lisa Johnston

Research Data Management/Curation Lead,
University of Minnesota

What are data repositories?

Research Data Services



“A repository (also referred to as a data repository or digital data repository) is a **searchable** and **queryable** interfacing entity that is able to store, manage, maintain and curate Data/Digital Objects. A repository is a **managed** location (destination, directory or ‘bucket’) where digital data objects are registered, permanently stored, made **accessible** and **retrievable**, and curated. Repositories preserve, manage, and provide access to many types of digital material in a variety of formats. Materials in online repositories are **curated** to enable search, discovery, and reuse. There must be sufficient control for the digital material to be authentic, reliable, accessible and usable on a continuing basis. (p3, footnote 2).”

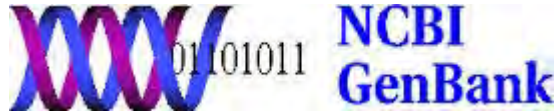
Why are data repositories useful?



- Governments have open data initiatives
 - Democratize research results
 - Release the potential of valuable/\$\$ data
- Federal/private funders increasingly require data sharing
 - Public access
 - Return on \$\$ investment ⇒ do new research
- Journals have data sharing policies
 - Increase transparency
 - Facilitate reproducibility
- Researcher/disciplinary culture shift in digital age
 - Ease of sharing ⇒ greater potential for reuse
 - Citation impact, reputation building

What kinds of data repositories exist?

Disciplinary Data Repositories



CERN *OpenData*: CMS data



And 1516 more listings at
<http://www.re3data.org/>

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

What kinds of data repositories exist?

General Data Repositories



<http://openICPSR.org>



What kinds of data repositories exist?

Institutional Data Repositories

Running...

- **Dataverse**
- Digital Commons
- DSpace
- EPrints
- Fedora
- HubZero
- Hydra
- Islandora

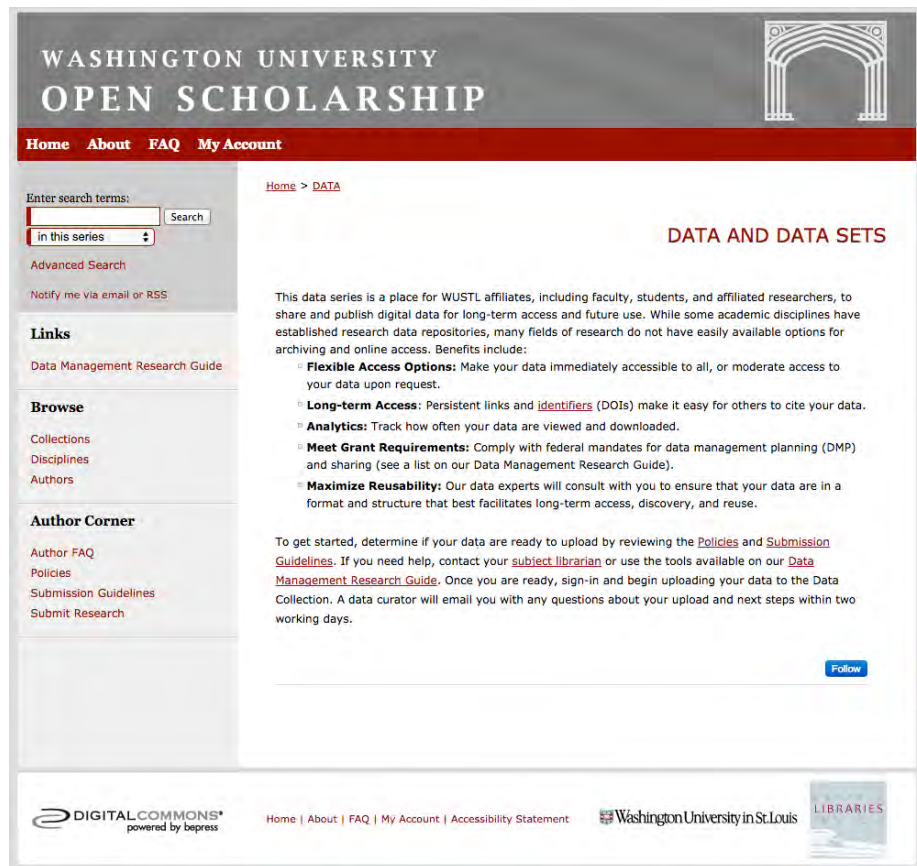
The screenshot displays the CfA Dataverse website interface. At the top, the CfA logo and name "HARVARD-SMITHSONIAN CENTER FOR ASTROPHYSICS" are visible. Below this, a navigation bar includes links for "CfA Dataverse (Harvard-Smithsonian Center for Astrophysics)" and "Home Page". The main content area features a search bar with the text "Search this dataverse..." and buttons for "Find" and "Advanced Search". Below the search bar, there are filters for "Dataverses (34)", "Datasets (111)", and "Files (1,841)". A "Dataverse Category" section lists "Researcher (5)" and "Research Project (2)". A "Publication Date" section shows counts for various years: 2011 (69), 2013 (36), 2014 (22), 2015 (9), and 2012 (7). An "Author Name" section lists authors and their counts: Thaddeus, P. (38), Dame, T. M. (25), COMPLETE team (20), Hartmann, Dap (19), and Brorfman, L. (7). The main results area shows "1 to 10 of 145 Results". The first result is "APEX CMZ SHFI-1 survey" by Ginsburg, Adam, 2015, with a link to the dataset. The second result is "Replication data for: Deep 3.8 Micron Observations of the Trapezium Cluster" by Muench, August; Alves, Joao; Lada, Charles; Lada, Elizabeth, 2015, with a link to the dataset. The third result is "Replication Data for: YSOVAR: Mid-infrared Variability in the Star-forming Region Lynds 1688" by Guenther, Hans Moritz, 2015, with a link to the dataset.

What kinds of data repositories exist?

Institutional Data Repositories

Running...

- Dataverse
- **Digital Commons**
- DSpace
- EPrints
- Fedora
- HubZero
- Hydra
- Islandora



The screenshot shows the Washington University Open Scholarship website. The header features the university's name and logo. A navigation bar includes links for Home, About, FAQ, and My Account. The main content area is titled "DATA AND DATA SETS" and contains a paragraph about the data series, followed by a list of benefits: Flexible Access Options, Long-term Access, Analytics, Meet Grant Requirements, and Maximize Reusability. A sidebar on the left includes a search bar, a list of links (Data Management Research Guide), a browse section (Collections, Disciplines, Authors), and an author corner (Author FAQ, Policies, Submission Guidelines, Submit Research). The footer includes the Digital Commons logo, navigation links, and the Washington University in St. Louis logo.

WASHINGTON UNIVERSITY
OPEN SCHOLARSHIP

Home About FAQ My Account

Enter search terms: Search

In this series

Advanced Search

Notify me via email or RSS

Links

Data Management Research Guide

Browse

Collections
Disciplines
Authors

Author Corner

Author FAQ
Policies
Submission Guidelines
Submit Research

Home > DATA

DATA AND DATA SETS

This data series is a place for WUSTL affiliates, including faculty, students, and affiliated researchers, to share and publish digital data for long-term access and future use. While some academic disciplines have established research data repositories, many fields of research do not have easily available options for archiving and online access. Benefits include:

- **Flexible Access Options:** Make your data immediately accessible to all, or moderate access to your data upon request.
- **Long-term Access:** Persistent links and [identifiers](#) (DOIs) make it easy for others to cite your data.
- **Analytics:** Track how often your data are viewed and downloaded.
- **Meet Grant Requirements:** Comply with federal mandates for data management planning (DMP) and sharing (see a list on our Data Management Research Guide).
- **Maximize Reusability:** Our data experts will consult with you to ensure that your data are in a format and structure that best facilitates long-term access, discovery, and reuse.

To get started, determine if your data are ready to upload by reviewing the [Policies](#) and [Submission Guidelines](#). If you need help, contact your [subject librarian](#) or use the tools available on our [Data Management Research Guide](#). Once you are ready, sign-in and begin uploading your data to the Data Collection. A data curator will email you with any questions about your upload and next steps within two working days.

Follow

DIGITALCOMMONS*
powered by bepress

Home | About | FAQ | My Account | Accessibility Statement

Washington University in St. Louis

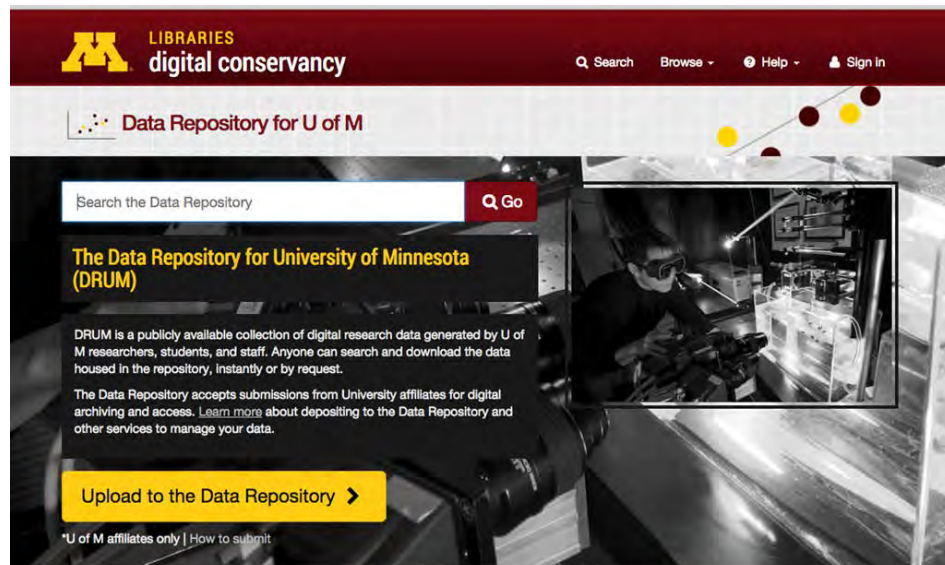
LIBRARIES

What kinds of data repositories exist?

Institutional Data Repositories

Running...

- Dataverse
- Digital Commons
- **DSpace**
- EPrints
- Fedora
- HubZero
- Hydra
- Islandora



How to Upload

1. Prepare Data

Data should be free of identifying or sensitive information and include adequate documentation. Not sure? [Contact us for help!](#)

2. Upload

Have your files ready (up to 2GB each) and use the upload form to fill out metadata about your data.

Features

Flexible Access Options

Choose to make your data immediately accessible to everyone, or moderate access to your data upon request.

Meet Grant Requirements

Comply with federal mandates for data management planning (DMP) and sharing. [Read more.](#)

Our Services

Data Management Plan Assistance

We offer personalized assistance for drafting your next grant's Data Management Plan. [Contact us](#) for assistance during your planning process.

Metadata Consultation

We can help structure your data using disciplinary best practices to ensure the best

What kinds of data repositories exist?

Institutional Data Repositories

Running...

- Dataverse
- Digital Commons
- DSpace
- **EPrints**
- Fedora
- HubZero
- Hydra
- Islandora

The screenshot shows the homepage of the University of Bath Research Data Archive. The header features the University of Bath logo and navigation links: Home, About, and Browse. Below the header is a search bar with a 'Search' button. The main content area has a large heading 'Welcome to University of Bath Research Data Archive' followed by the tagline 'Find data from our world class research'. Below this is a search bar with a 'Search' button. A 'Latest Additions' box is also visible. The footer contains links for 'About this Archive', 'Browse Archive', and 'Archive Policies', along with a footer note about OAI 2.0 support and a footer bar with the EPrints logo and copyright information.

UNIVERSITY OF BATH

Home About Browse -

Login Create Account

Search

Welcome to University of Bath Research Data Archive

Find data from our world class research

Search

Latest Additions

About this Archive

More information about this site.

Browse Archive

Browse the items in the repository by subject.

Archive Policies

Policy for use of material in the repository.

University of Bath Research Data Archive supports OAI 2.0 with a base URL of <http://researchdata.bath.ac.uk/cgi/oa2>

Atom RSS

University of Bath Research Data Archive is powered by [EPrints 3](#) which is developed by the School of Electronics and Computer Science at the University of Southampton. [More information and software credits.](#)

researchdata.bath.ac.uk

eprints

What kinds of data repositories exist?

Institutional Data Repositories

Running...

- Dataverse
- Digital Commons
- DSpace
- EPrints
- **Fedora**
- HubZero
- Hydra
- Islandora

The screenshot shows the RUcore website interface. At the top, the RUcore logo is displayed with the text 'Rutgers University Community Repository'. Navigation links for 'Help', 'Contact Us', and 'My Account' are in the top right. Below the logo, there are buttons for 'Search', 'Services', and 'Collections'. The main heading is 'RUresearch Data Portal'. A description states: 'The RUresearch Data Portal provides a place for Rutgers researchers to share their research data with the global scholarly community. RUresearch leverages all the capabilities of RUcore and adds additional tools and services specific to research data.' A 'Search' section includes a search bar and a 'Search' button. Below this are links for 'Browse All Projects' and 'Advanced Search'. A 'Learn More' section titled 'RUresearch Services' lists several bullet points: 'Free assistance and advice to design your data strategy and prepare the data Management component of a grant proposal.', 'A customized search and retrieval portal for your data.', 'Ongoing management and support for your data', 'Is there a fee for placing my data in RUresearch?', and 'Guidance for the NSF Data Management Grant Requirement'. A 'What does RUresearch Provide' section lists: 'Robust preservation and long term access', 'Metadata customized to your project's needs or community practice (e.g., DDI, Darwin Core, etc.)', 'Access control, including flexible embargo periods and statements of use', 'All associated information, presented in Context, such as codebooks, lab notes, Images, software, related publications, etc.', and 'Contact the RUresearch Data Team'. On the right side, there is a blue graphic of a head filled with icons, with text below it: 'About RUresearch', 'RUresearch Data Team', 'NSF Data Management', 'Data Management and RUresearch Presentation', and 'Highlighted RUresearch Collections'. At the bottom right, there are links for 'Video Mosaic Collaborative' and 'Ralston Horse Behavior Research'.

RUcore
Rutgers University Community Repository

Help Contact Us My Account

Search Services Collections

Home **RUresearch Data Portal**

RUresearch Data Portal

The **RUresearch** Data Portal provides a place for Rutgers researchers to share their research data with the global scholarly community. **RUresearch** leverages all the capabilities of RUcore and adds additional tools and services specific to research data.

Search [More...](#)

Search

[Browse All Projects](#) | [Advanced Search](#)

Learn More

RUresearch Services

- [Free assistance and advice](#) to design your data strategy and prepare the data Management component of a grant proposal.
- A customized search and retrieval portal for your data.
- Ongoing management and support for your data
- Is there a fee for placing my data in **RUresearch**?
- [Guidance for the NSF Data Management Grant Requirement](#)

What does **RUresearch** Provide

- [Robust preservation](#) and long term access
- [Metadata](#) customized to your project's needs or community practice (e.g., DDI, Darwin Core, etc.)
- [Access control](#), including flexible embargo periods and statements of use
- All associated information, presented in Context, such as codebooks, lab notes, Images, software, related publications, etc.
- [Contact the RUresearch Data Team](#)

About RUresearch

[RUresearch Data Team](#)

[NSF Data Management](#)

[Data Management and RUresearch Presentation](#)

Highlighted RUresearch Collections

[Video Mosaic Collaborative](#)

[Ralston Horse Behavior Research](#)

What kinds of data repositories exist?

Institutional Data Repositories

Running...

- Dataverse
- Digital Commons
- DSpace
- EPrints
- Fedora
- **HubZero**
- Hydra
- Islandora

The screenshot shows the Purdue University Research Repository homepage. At the top is a navigation bar with links for Home, Publications, Projects, Get Started, Policies, and Contact Us. A search bar is located on the right. Below the navigation bar, a section titled "Public Access to Federally Funded Research" features a paragraph about the Office of Science and Technology Policy's directive and a "Learn More" button. To the right of this section is a circular seal of the Executive Office of the President of the United States, Office of Science and Technology Policy. Below the "Public Access" section, there are three columns of content. The first column has three items: "Create a Data Management Plan" (with a clipboard icon), "Upload Research Data to Your Project" (with a gear icon), and "Publish your Dataset" (with a document icon). The second column features a box for "LARS - Vegetation and Soils Field Research Data" from the Laboratory for Applications of Remote Sensing (LARS) at Purdue University, which includes a grid of satellite images and a description of 127 downloadable datasets. The third column contains a login form with fields for Username (optional), Name, E-mail, and Question, along with a Submit button.

PURDUE UNIVERSITY Purdue University Research Repository

Login Register Have a Question?

Home Publications Projects Get Started Policies Contact Us

Search

Public Access to Federally Funded Research

The [Office of Science and Technology Policy](#) has directed federal agencies with more than \$100M in R&D expenditures to develop plans to make the published results of federally-funded research, including datasets, freely available to the public within one year of publication.

[Learn More](#)

EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES
OFFICE OF SCIENCE AND TECHNOLOGY POLICY

Create a Data Management Plan
Learn about the detailed requirements for your data management plan (DMP). Funding agency requirements are very specific and our DMP resources can help you to clear up any confusion. [Get Started](#)

Upload Research Data to Your Project
Create a project to upload and share your data with collaborators using our step-by-step form to guide you through the process. Invite collaborators from other institutions to join your project. [Create a Project](#)

Publish your Dataset
Package, describe, and publish your dataset with a Datacite DOI. Publishing will ensure your dataset is citable, reusable, and archived for the long-term. [See Published Datasets](#)

LARS - Vegetation and Soils Field Research Data
Laboratory for Applications of Remote Sensing (LARS)
Purdue University

The collection of 127 downloadable datasets, collected between 1975 and 1991, consists of calibrated spectral data and the associated biophysical and meteorological data.

Username: OPTIONAL
Name:
E-mail:
Question:
[Submit](#)

What kinds of data repositories exist?

Institutional Data Repositories

Running...

- Dataverse
- Digital Commons
- DSpace
- EPrints
- Fedora
- HubZero
- **Hydra**
- Islandora

The screenshot shows the Penn State ScholarSphere website. At the top is a blue header with the Penn State logo and the text "ScholarSphere". To the right of the header is a "Login" button. Below the header is a navigation bar with links for "Home", "About", "Help", and "Contact". A search bar is located on the right side of the navigation bar, with the text "Enter search terms" and a "Go" button. The main content area features a large banner with a background image of a globe and various data points. Overlaid on the banner are two white boxes: one asking "Need help with research data management?" and another encouraging users to "Share Your Work" with a "Terms of Use" link. Below the banner, there are three main sections: "Featured Works", "Recent Additions", and "Featured Researcher". The "Featured Works" section lists three items: "Analytical Comparisons of...", "4 1/2 stars: Creating a...", and "The Uptake of Vitamin A...". The "Recent Additions" section lists two items: "GSA 2015 Abstract &..." and "The Uptake of Vitamin A...". The "Featured Researcher" section features a photo of Dr. Laura Guertin, a Professor of Earth Science at Penn State Brandywine, and a brief biography of her research and teaching.

PENNSTATE ScholarSphere

Login

Home About Help Contact

Enter search terms Go All

Need help with research data management?

Consult the Libraries' guide on RDM services. Questions? Contact the Research Data Management Team.

Share Your Work

Terms of Use

Photo and image credits.

Featured Works

Recent Additions

Featured Researcher

View other featured researchers

Analytical Comparisons of...
Depositor: Mohannad Ahmed Abdulrahman
Keywords: radio, astronomy, telescope

4 1/2 stars: Creating a...
Depositor: Jennie Anne Levine Knies
Keywords: Libraries, Undergraduate research, Penn State Wilkes-Barre

The Uptake of Vitamin A...
Depositor: Joanna Kalina Urbanek
Keywords: vitamin A, brain, mathematical modeling

GSA 2015 Abstract &...
Depositor: Laura Guertin
Keywords: Geological Society of America, undergraduate education, undergraduate research, geoscience education, Google

Dr. Laura Guertin, Professor of Earth Science at Penn State Brandywine

I have a pedagogical research agenda, looking at how new and emerging technologies can enhance teaching and student learning in introductory-level geoscience courses.

Why does Dr. Guertin use ScholarSphere?

At scientific conferences, more and more presentations are being given in a poster format, on 4 foot by 8 foot boards. A poster is an excellent way for me to capture highlights of my work with text and images, but it is difficult for colleagues and other researchers to take away all of that information after viewing the poster.

What kinds of data repositories exist?

Institutional Data Repositories

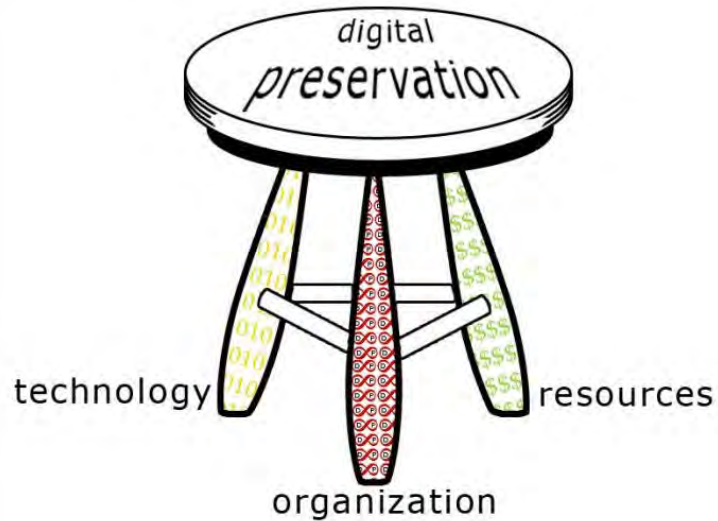
Running...

- Dataverse
- Digital Commons
- DSpace
- EPrints
- Fedora
- HubZero
- Hydra
- **Islandora**

The screenshot shows the SFU Research Data Repository (Radars) website. The header includes the SFU logo, the text "SIMON FRASER UNIVERSITY ENGAGING THE WORLD", the title "Research Data Repository", an "A-Z DIRECTORY" link, a search bar, and language selection options for "This site" and "SFU.ca". A left sidebar contains a menu with links: "BROWSE RESEARCH DATA", "PRIVATE DATA WORKSPACE", "DESKTOP SYNC", "DEPOSIT DATA", "REPOSITORY SEARCH", "DEPOSIT TERMS", "SUPPORT AND SERVICES", and an "ADMIN LOGIN" section with a "Log in" button. The main content area is titled "Radars: SFU's Research Data Repository" and includes a welcome message: "Welcome to Radars, SFU's Research Data Repository. Here, you can:". Below this are three large icons with corresponding links: a red icon with a downward arrow for "Deposit data", a blue icon with a magnifying glass for "Browse data", and a dark grey icon with a padlock for "Access your private data workspace". At the bottom, a paragraph states: "In addition to providing a place for you to deposit your research data, the SFU Library can assist in managing your data. Do you want to publish a completed data set alongside a research paper, need some help in formatting, or just looking for extra storage so that you can cope with the volume of your own backups? If you have any questions, we'd love to talk to you about your data needs. Please see our guides on [research data management](#) or contact us at lib-data@sfu.ca for more information."

How do I build an institutional data repository?

Blueprint for Institutional Data Repositories

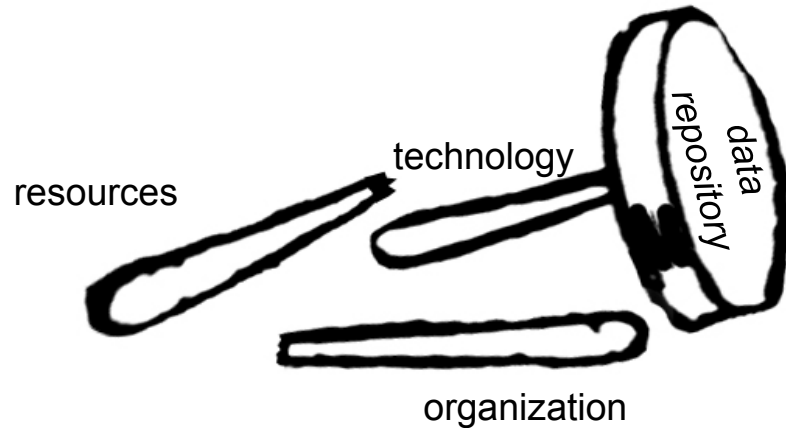


Nancy McGovern's Three-Legged Stool of building an organization's digital preservation program...

- technological infrastructure
- organizational infrastructure
- resources framework

How do I build an institutional data repository?

Blueprint for Institutional Data Repositories



Lisa's three-legged stool for building a data repository (basically the same...)

- technological infrastructure
- organizational infrastructure
- resources framework

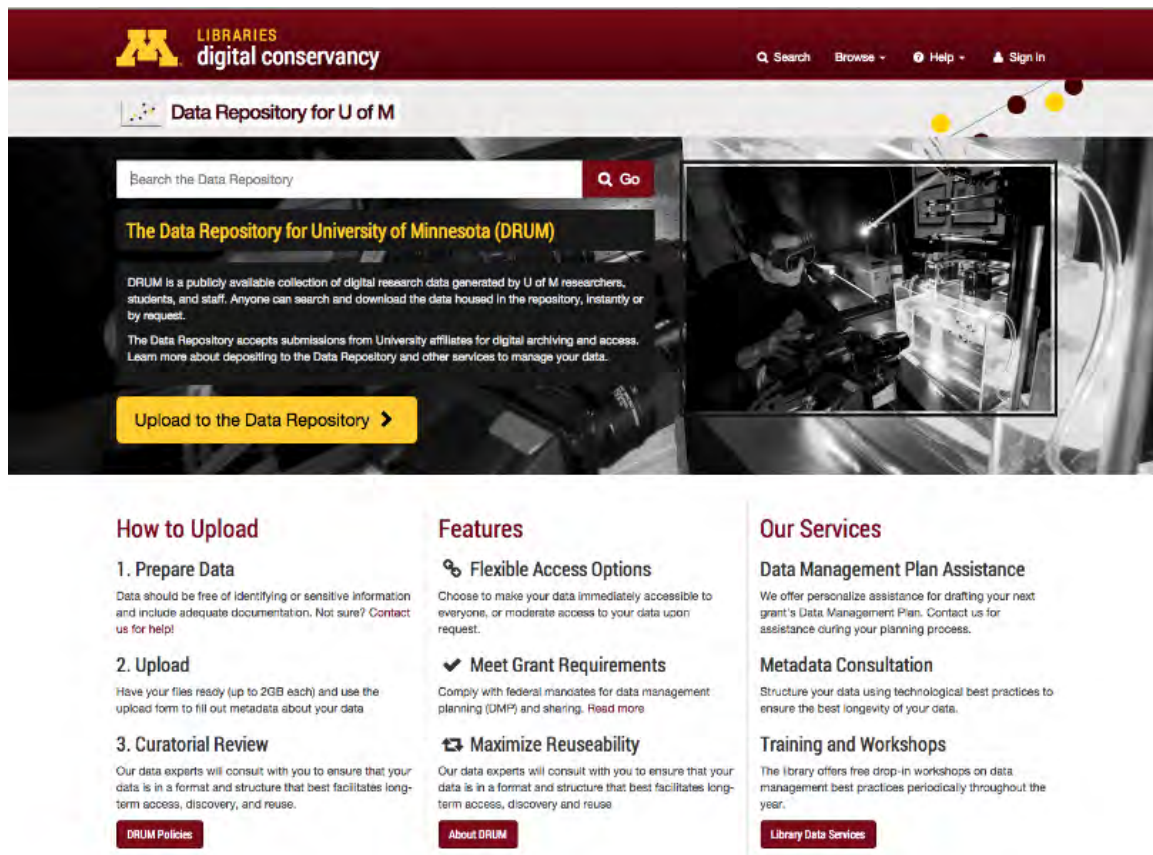
DRUM

<http://z.umn.edu/drum>

Launched Nov 2014

Available to U of M researchers and provides:

- Open access
- Curation services
- Permanent identifiers (DOI)
- Flexible Licenses
- File download analytics
- Preservation



The screenshot shows the DRUM website homepage. At the top is a dark red header with the 'LIBRARIES digital conservancy' logo on the left and search, browse, help, and sign-in links on the right. Below the header is a light gray banner with the text 'Data Repository for U of M' and a small icon. The main content area has a dark background with a search bar, a description of DRUM, and a yellow 'Upload to the Data Repository' button. To the right is a large image of a person using a microscope. Below the main content are three white boxes: 'How to Upload' with a 3-step process, 'Features' with three bullet points, and 'Our Services' with three service descriptions. Each box has a red button at the bottom: 'DRUM Policies', 'About DRUM', and 'Library Data Services' respectively.

LIBRARIES digital conservancy

Search Browse Help Sign in

Data Repository for U of M

Search the Data Repository **Go**

The Data Repository for University of Minnesota (DRUM)

DRUM is a publicly available collection of digital research data generated by U of M researchers, students, and staff. Anyone can search and download the data housed in the repository, instantly or by request.

The Data Repository accepts submissions from University affiliates for digital archiving and access. Learn more about depositing to the Data Repository and other services to manage your data.

Upload to the Data Repository >

How to Upload

- 1. Prepare Data**
Data should be free of identifying or sensitive information and include adequate documentation. Not sure? [Contact us for help!](#)
- 2. Upload**
Have your files ready (up to 2GB each) and use the upload form to fill out metadata about your data.
- 3. Curatorial Review**
Our data experts will consult with you to ensure that your data is in a format and structure that best facilitates long-term access, discovery, and reuse.

DRUM Policies

Features

- Flexible Access Options**
Choose to make your data immediately accessible to everyone, or moderate access to your data upon request.
- Meet Grant Requirements**
Comply with federal mandates for data management planning (DMP) and sharing. [Read more](#)
- Maximize Reuseability**
Our data experts will consult with you to ensure that your data is in a format and structure that best facilitates long-term access, discovery and reuse.

About DRUM

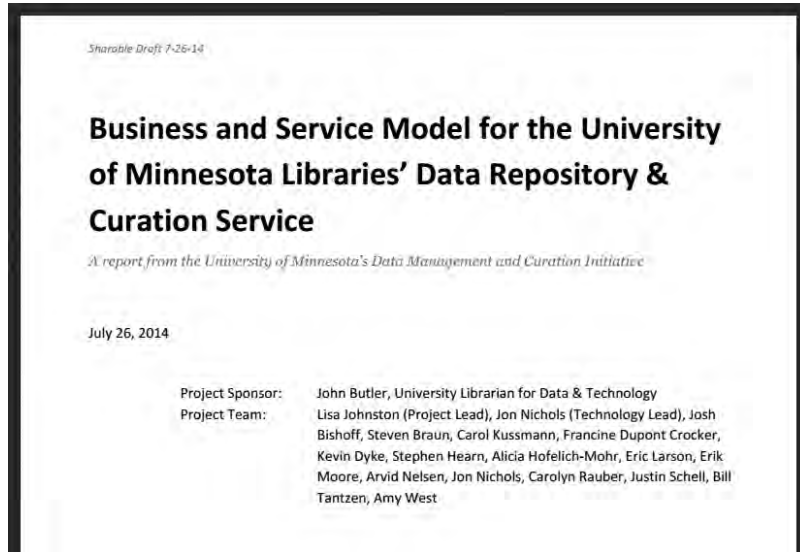
Our Services

- Data Management Plan Assistance**
We offer personalized assistance for drafting your next grant's Data Management Plan. Contact us for assistance during your planning process.
- Metadata Consultation**
Structure your data using technological best practices to ensure the best longevity of your data.
- Training and Workshops**
The library offers free drop-in workshops on data management best practices periodically throughout the year.

Library Data Services

How do I build my institutional data repositories?

Organizational Infrastructure



- Libraries DM+C Initiative 2014-2015 with dedicated resources and a 19-person team
 - **Business model**
 - Policies
 - Services
- Existing libraries digital preservation framework
- Campus: New research data management policy
- Marketing: Official launch in March 2015

How do I build my institutional data repositories?

Organizational Infrastructure

DRUM Policies and Terms of Use

Housed within the University of Minnesota Digital Conservancy, the Data Repository for U of M (DRUM) is subject to the [policies of the Digital Conservancy](#) and, in addition, include the following. Note: Policies and Terms are subject to change.

Policies related to DEPOSIT of Materials in DRUM

- [Data Collection Policy](#)
- [End-user Access Policy](#)
- [Preservation Policy](#)
- [Deposit License](#)

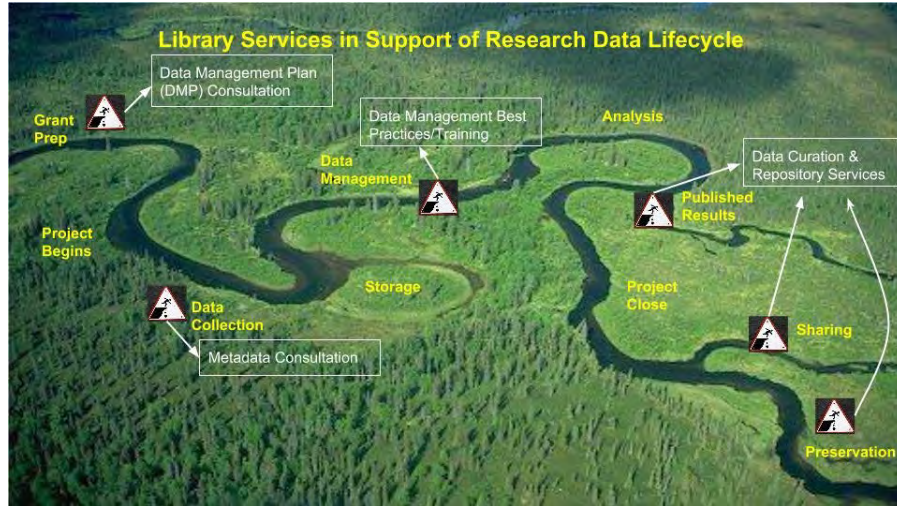
Policies related to USE of Materials Available in DRUM

- [Terms of Use](#)
- [End-user Access Policy](#)

- Libraries DM+C Initiative 2014-2015 with dedicated resources and a 19-person team
 - Business model
 - **Policies**
 - Services
- Existing libraries digital preservation framework
- Campus: New research data management policy
- Marketing: Official launch in March 2015

How do I build my institutional data repositories?

Organizational Infrastructure



- Libraries DM+C Initiative 2014-2015 with dedicated resources and a 19-person team
 - Business model
 - Policies
 - **Services**
- Existing libraries digital preservation framework
- Campus: New research data management policy
- Marketing: Official launch in March 2015

How do I build my institutional data repositories?

Organizational Infrastructure



- Libraries DM+C Initiative 2014-2015 with dedicated resources and a 19-person team
 - Business model
 - Policies
 - Services
- **Existing libraries digital preservation framework**
- Campus: New research data management policy
- Marketing: Official launch in March 2015

How do I build my institutional data repositories?

Organizational Infrastructure



The screenshot shows a webpage titled "ADMINISTRATIVE POLICY" with a red icon of two people. The main heading is "Research Data Management: Archiving, Ownership, Retention, Security, Storage, and Transfer". Below this is a yellow box containing a table of contents and metadata. The table of contents lists: Policy Statement, Reason for Policy, Procedures, Forms/Instructions, Appendices, FAQ, Additional Contacts, Definitions, Responsibilities, Related Information, and History. The metadata includes: Effective Date: Draft, Under review; Last Updated: Month, YYYY; Responsible University Officer: Vice President for Research; Policy Owner: Associate Vice President for Research; and Policy Contact: Frances Lawrenz. Below the yellow box is a section titled "POLICY STATEMENT" which describes the scope of the policy, covering the Office of the Vice President for Research (OVPR), the Office of the Vice President and Chief Information Officer (OVPCIO), the Academic Health Center (AHC), the University Libraries and system campus libraries, the colleges, schools and system campuses, and the Principal.

ADMINISTRATIVE POLICY

Research Data Management: Archiving, Ownership, Retention, Security, Storage, and Transfer

Policy Contents	Effective Date: Draft, Under review
<ul style="list-style-type: none">• Policy Statement• Reason for Policy• Procedures• Forms/Instructions• Appendices• FAQ• Additional Contacts• Definitions• Responsibilities• Related Information• History	Last Updated: Month, YYYY
Responsible University Officer: <ul style="list-style-type: none">• Vice President for Research	
Policy Owner: <ul style="list-style-type: none">• Associate Vice President for Research	
Policy Contact: <ul style="list-style-type: none">• Frances Lawrenz	

POLICY STATEMENT

The Office of the Vice President for Research (OVPR); the Office of the Vice President and Chief Information Officer (OVPCIO); the Academic Health Center (AHC); the University Libraries and system campus libraries; the colleges, schools and system campuses; and the Principal

- Libraries DM+C Initiative 2014-2015 with dedicated resources and a 19-person team
 - Business model
 - Policies
 - Services
- Existing libraries digital preservation framework
- **Campus: New research data management policy**
- Marketing: Official launch in March 2015

How do I build my institutional data repositories?

Organizational Infrastructure

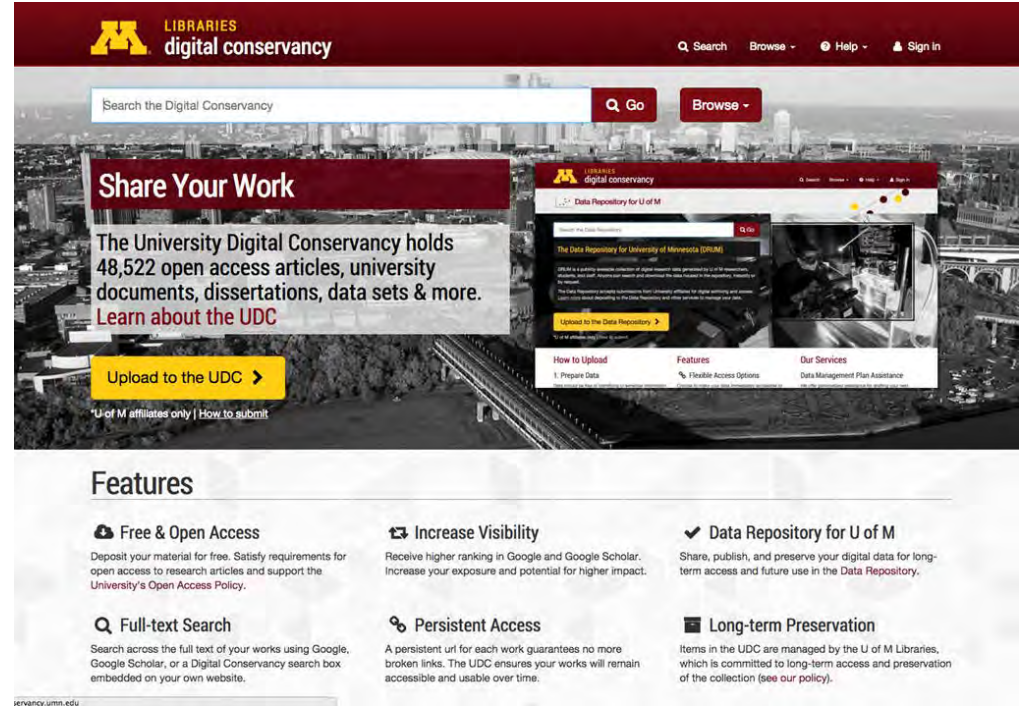


- Libraries DM+C Initiative 2014-2015 with dedicated resources and a 19-person team
 - Business model
 - Policies
 - Services
- Existing libraries digital preservation framework
- Campus: New research data management policy
- **Marketing: Official launch in March 2015**

Case Study: Building the Data Repo for the U of M

Technical Infrastructure

- **DRUM part of existing IR (DSpace):**
 - Metadata schema
 - Collection home page
 - Record view
- Meet federal funding requirements
- Curation procedures
 - Sensitive Data
 - Readme Template
 - Transform File Formats



Case Study: Building the Data Repo for the U of M

Technical Infrastructure

- DRUM part of existing IR (DSpace):
 - **Metadata schema**
 - Collection home page
 - Record view
- Meet federal funding requirements
- Curation procedures
 - Sensitive Data
 - Readme Template
 - Transform File Formats

University Digital Conservancy Item Submission

[Describe](#) [Describe](#) [Upload](#) [CC License](#) [License](#) [Confirm](#) [Complete](#)

Describe Item

Author(s)
Last name, e.g. Smith or University of Minnesota First, MI e.g. Barb H. or Robert Jr.
 [Add More](#)

List the author(s) of the data set. Use complete names eg. "James Ryan Swanson"

Group Author
 [Add More](#)

If applicable for datasets produced by a group. Ex. "University of Minnesota Department of Chemistry, Crystallography Research Lab"

Title (Required)

Titles should be descriptive and include the what, where, when and scale of your data.

Date Completed (Required)
Year (required) Month (optional) Day (optional)

Date the data was finalized for distribution.

Contact (Required)
Last name, e.g. Smith or University of Minnesota First, MI e.g. Barb H. or Robert Jr.

Name of the principal contact for the data.

Contact Email (Required)

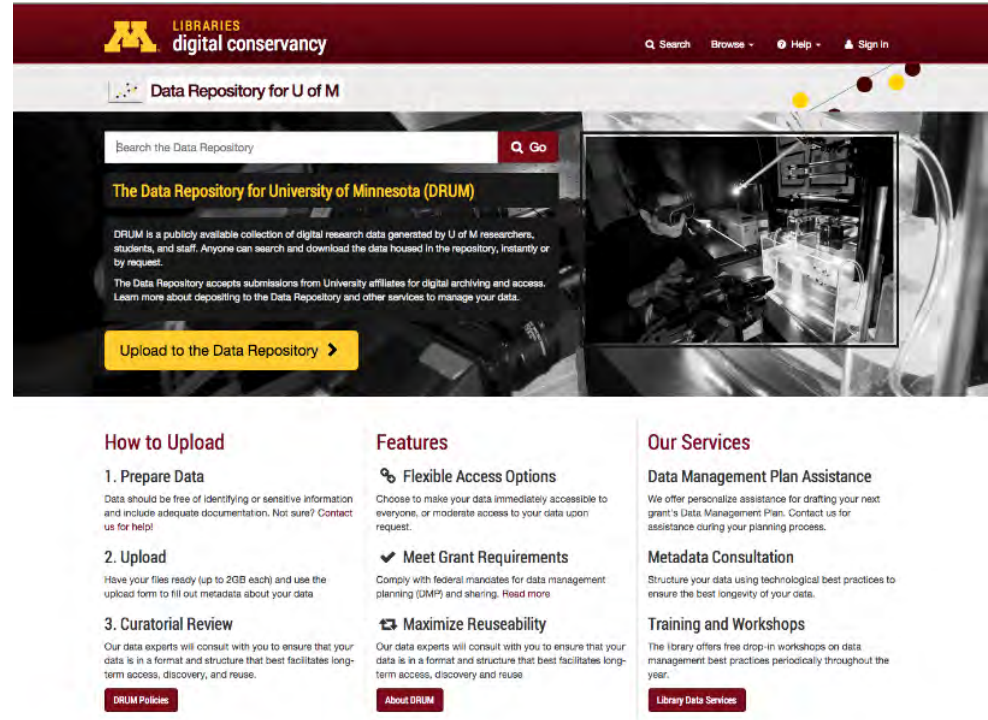
Email address of the principal contact for the data.

[Save & Exit](#) [Next >](#)

Case Study: Building the Data Repo for the U of M

Technical Infrastructure

- DRUM part of existing IR (DSpace):
 - Metadata schema
 - **Collection home page**
 - Record view
- Meet federal funding requirements
- Curation procedures
 - Sensitive Data
 - Readme Template
 - Transform File Formats



Schema published: "The Supporting Documentation for Implementing the Data Repository for the University of Minnesota (DRUM): A Business Model, Functional Requirements, and Metadata Schema" at <http://hdl.handle.net/11299/171761>.

Case Study: Building the Data Repo for the U of M

Technical Infrastructure

- DRUM part of existing IR (DSpace):
 - Metadata schema
 - Collection home page
 - **Record view**
- Meet federal funding requirements
- Curation procedures
 - Sensitive Data
 - Readme Template
 - Transform File Formats

The screenshot shows the DRUM record view for a dataset. The header includes the University of Minnesota Digital Conservancy logo and navigation links. The main content area displays the title, authors, published date, author contact, data type, abstract, license, and suggested citation. A map of the eastern United States is shown, color-coded by deer density. Below the map, there are persistent links to the item and a button to show the full item record. At the bottom, a table lists the files available for download.

LIBRARIES digital conservancy

Search Browse Help Sign In

Data Repository for U of M

University Digital Conservancy Home / University of Minnesota - Twin Cities / Data Repository for U of M (DRUM) / View Item

White-tailed deer density estimates across the eastern United States, 2008

Walters, Brian, F. Woodall, Christopher, W. Russell, Matthew, B. (2016)

Published Date
2016-03-13

Author Contact
Russell, Matthew, B. (russellm@umn.edu)

Data Type
Dataset, Map, Spatial Data

Abstract
In 2008, the Quality Deer Management Association (QDMA) developed a map of white-tailed deer density with information obtained from state wildlife agencies. The map contains information from 2001 to 2005, with noticeable changes since the development of the first deer density map made by QDMA in 2001. The University of Minnesota, Forest Ecosystem Health Lab and the US Department of Agriculture, Forest Service-Northern Research Station have digitized the deer density map to provide information on the status and trends of forest health across the eastern United States.

License: Attribution-NonCommercial-ShareAlike 3.0 United States

Suggested Citation
Walters, Brian, F. Woodall, Christopher, W. Russell, Matthew, B. (2016). White-tailed deer density estimates across the eastern United States, 2008 [dataset]. Retrieved from the Data Repository for the University of Minnesota, <http://dx.doi.org/10.13020/D6G014>.

[Show full item record](#)

Persistent link to this item
<http://dx.doi.org/10.13020/D6G014>
<http://hdl.handle.net/11299/178246>

Services
Full Metadata (XML)
[View Usage Statistics](#)

View/Download file

File View/Open	Description	Size	Format
deer_density_QDMA.zip	GIS data-single shapefile	3.792Mb	application/zip
README.txt	README text file with basic description of data	1.287Kb	text/plain

Case Study: Building the Data Repo for the U of M

Technical Infrastructure

- DRUM part of existing IR (DSpace):
 - Metadata schema
 - Collection home page
 - Record view
- **Meet federal funding requirements**
- Curation procedures
 - Sensitive Data
 - Readme Template
 - Transform File Formats

Resources and Examples

The Libraries have organized the following resources that may be useful for you as you write your DMP:

- [UMN DMP template \(google doc\): a template you can use to begin writing your DMP](#)
- [Data Management Checklist](#) (two page pdf)
- Boilerplate text that you can use in your DMP and project proposal:

A long-term data sharing and preservation plan will be used to store and make publicly accessible the data beyond the life of the project. The data will be deposited into the Data Repository for the University of Minnesota (DRUM), <http://hdl.handle.net/11299/166578>. The University Libraries' hosted institutional data repository is an open access platform for dissemination and archiving of university research data. Data files in DRUM are written to an isilon storage system with two copies, one local to each of the two geographically separated University of Minnesota Data Centers. The local Isilon cluster stores the data in such a way that the data can survive the loss of any two disks or any one node of the cluster. Within two hours of the initial write, data replication to the 2nd Isilon cluster commences. The 2nd cluster employs the same protections as the local cluster, and both verify with a checksum procedure that data has not altered on write. In addition, DRUM provides long-term preservation of digital data files for at least 10 years using services such as migration (limited format types), secure backup, bit-level checksums, and maintains a persistent DOI for data sets, facilitating data citations. In accordance to DRUM policies, the (deidentified, if applicable) data will be accompanied by the appropriate documentation, metadata, and code to facilitate reuse and provide the potential for interoperability with similar data sets.

Case Study: Building the Data Repo for the U of M

Technical Infrastructure

- DRUM part of existing IR (DSpace):
 - Metadata schema
 - Collection home page
 - Record view
- Meet federal funding requirements
- **Curation procedures**
 - Sensitive Data
 - Readme Template
 - Transform File Formats

Submission
under
curatorial
review

Case Study: Building the Data Repo for the U of M

Technical Infrastructure

- DRUM part of existing IR (DSpace):
 - Metadata schema
 - Collection home page
 - Record view
- Meet federal funding requirements
- Curation procedures
 - **Sensitive Data**
 - Readme Template
 - Transform File Formats

The image shows a 'Wizard' window for identifying sensitive data in DRUM. The window is titled 'Wizard' and has a 'Confirmation' section. It contains two main sections: 'Identities' and 'Locations'. The 'Identities' section lists various types of data with checkboxes indicating whether they are accepted or not. The 'Locations' section lists various file formats with checkboxes indicating whether they are accepted or not. The 'Confirmation' section includes a 'Back' button and a 'Finish' button. The background document is titled 'Procedure for Identifying Sensitive, Private, or Legally Protected Data in DRUM' and is from the University of Minnesota. It includes a 'Purpose' section, an 'Evaluation and Review of DRUM Submissions' section, and two steps: 'Step 1: U of M Researcher submits their data to DRUM at z.umn.edu/drum.' and 'Step 2: Data curators in the library receive notification that a new submission has occurred, but has not yet been reviewed.'

Procedure for Identifying Sensitive, Private, or Legally Protected Data in DRUM UNIVERSITY OF MINNESOTA Driven to Discover

Purpose: The Data Repository for the U of M (DRUM) has clear policies on what types of data may be accepted. DRUM policy states that "data must be non-restricted data that DO NOT contain any private, confidential, or otherwise legally protected information (e.g., personal identifiable information)." Therefore, to best support this policy, the following procedure outlines the steps that the library takes to prevent restricted data from being accepted in DRUM.

Evaluation and Review of DRUM Submissions:
Step 1: U of M Researcher submits their data to DRUM at z.umn.edu/drum.
Step 2: Data curators in the library receive notification that a new submission has occurred, but has not yet been reviewed.

Wizard

Confirmation
Please review your selections below and click Finish to begin your search. To make changes, click Back.

Identities

✓ Social Security Numbers	✗ Dates of Birth	✗ Australia: TFN
✓ Credit Card Numbers	✗ Phone Numbers	✗ Canada: SIN
✓ Password Entries	✗ E-Mail Addresses	✗ United Kingdom: NINO
✗ Bank Account Numbers	✗ Personal Addresses	✗ United Kingdom: NHS
✗ Driver Licenses	✗ Passport Numbers	

Locations

✓ E-Mails	✓ Browsers	✗ Windows Registry
✓ Files and Compressed Files	✓ My Documents and Settings	

< Back Finish Cancel Help

Case Study: Building the Data Repo for the U of M

Technical Infrastructure

- DRUM part of existing IR (DSpace):
 - Metadata schema
 - Collection home page
 - Record view
- Meet federal funding requirements
- Curation procedures
 - Sensitive Data
 - **Readme Template**
 - Transform File Formats

```
This readme.txt file was generated on <YYYYMMDD> by <Name>

GENERAL INFORMATION

1. Title of Dataset:
2. File Information:
  A. Filename:
  B. Short description:
  C. Filename:
  D. Short description
  E. Filename:
  F. Short description:
  G. If data set includes multiple files related to one another,
  include relationship here:

3. Principal Investigator Contact Information
  A. Name:
  B. Institution:
  C. Address:
  D. Email:

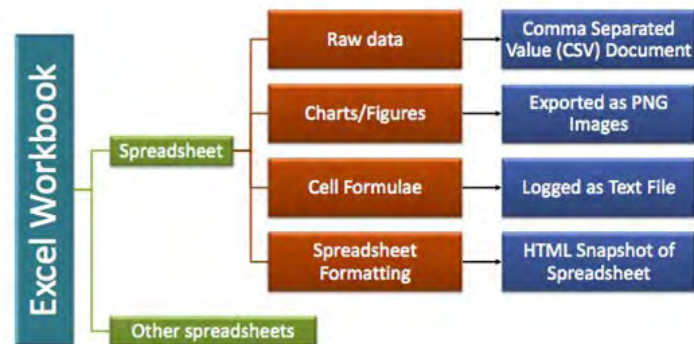
4. Associate or Co-investigator Contact Information
  A. Name:
  B. Institution:
  C. Address:
  D. Email:
```

Case Study: Building the Data Repo for the U of M

Technical Infrastructure

- DRUM part of existing IR (DSpace):
 - Metadata schema
 - Collection home page
 - Record view
- Meet federal funding requirements
- Curation procedures
 - Sensitive Data
 - Readme Template
 - **Transform File Formats**

Excel Archival Tool



File View/Open

Data For Gamma-toxin.xlsx

Archived_Data_for_Gamma_Toxin.zip

Combined Figures.pzf

Combined_Figures.xml

Description

Experiment Data Readings

Archival Version of Data for Gamma-toxin.xlsx

Data Analysis and Figures in Prism

Archival Version of Combined Figures.pzf

Case Study: Building the Data Repo for the U of M

Resources Infrastructure

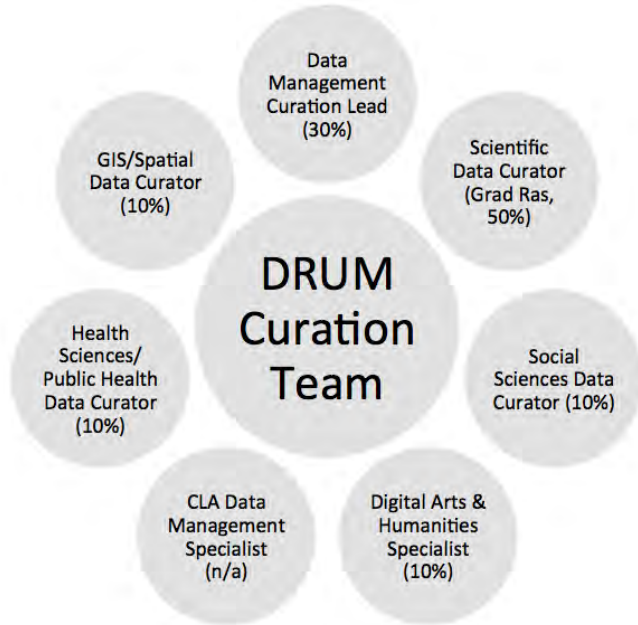


- Funding for DataCite DOIs
- Staffing Model
- Training for new curation staff



Case Study: Building the Data Repo for the U of M

Resources Infrastructure



- Funding for DataCite DOIs
- **Staffing Model**
- Training for new curation staff

Case Study: Building the Data Repo for the U of M

Resources Infrastructure

DMCI Scientific Data Curator Training Guide 2015

University Libraries Data Management and Curation Initiative (DMCI)

Please connect to and familiarize yourself with the [DMCI Staff Handbook](#) (2014-) where we will keep our shared documents. These include:

- DMCI Data Curation Procedures (Curator Handbook)
- Data Curators' Position Description
- This document

Table of Contents

[1.0 Responsibilities](#)

[1.1 Duties of the Scientific Data Curator](#)

[1.2 Daily Tasks of the Scientific Data Curator](#)

[1.3 Regular Meetings to Attend as the Scientific Data Curator](#)

[2.0 Research](#)

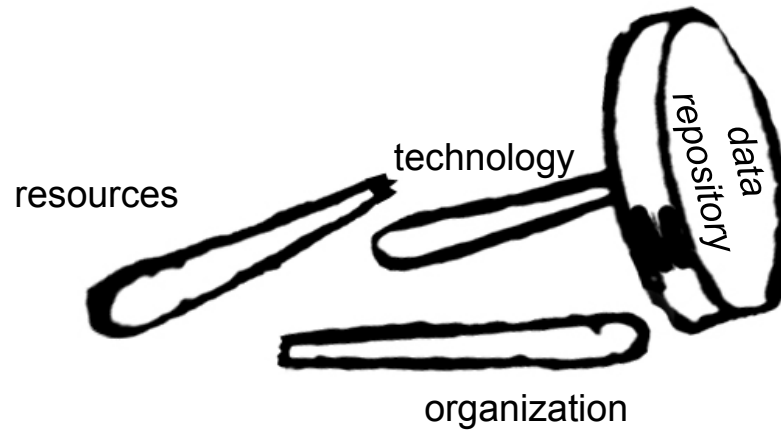
- Funding for DataCite DOIs
- Staffing Model
- Training for new curation staff

Conclusions



- Greater exposure of library services on campus
- Data repositories open up new conversations and opportunities (e.g., big data, campus networks, policies)
- Rewarding work to partner with researchers to publish their data sets.

Thanks and Questions



Keep building those stools...