

# Community Resources for Data Management



Megan Mach and Amber Budden  
DataONE

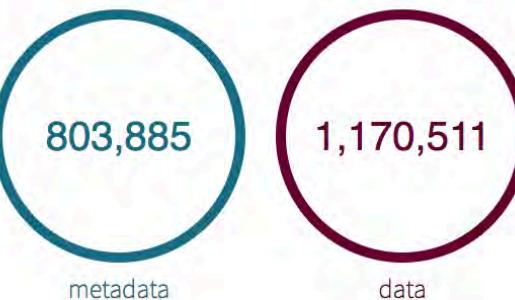
Nancy Hoebelheinrich  
Knowledge Motifs

# DataONE Cyberinfrastructure



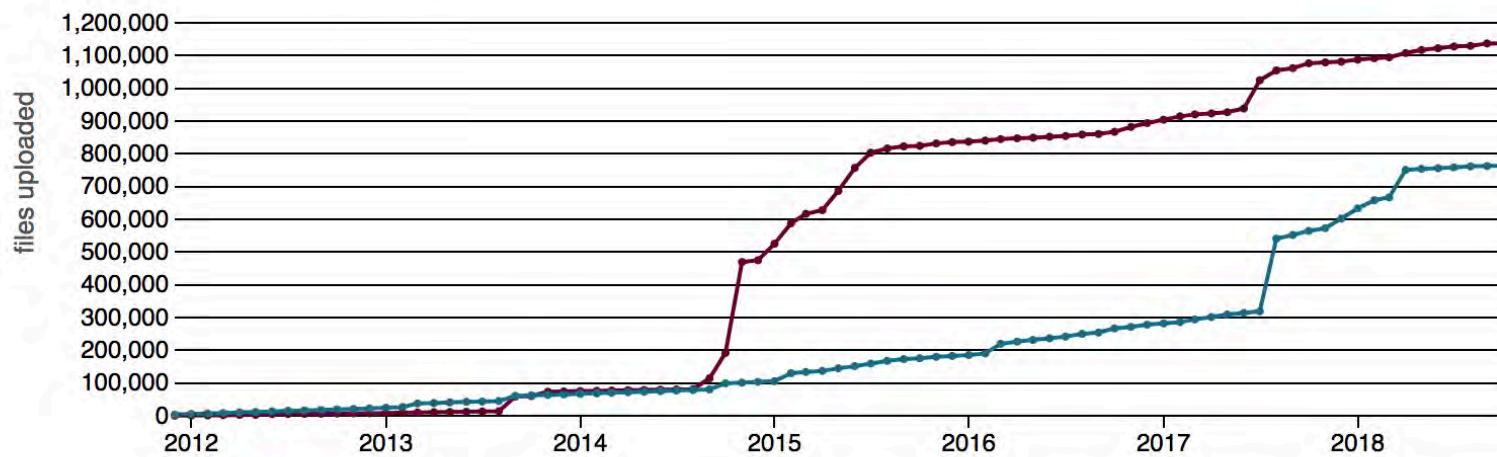
# Data Holdings

[dataone.org/current-member-nodes](http://dataone.org/current-member-nodes)



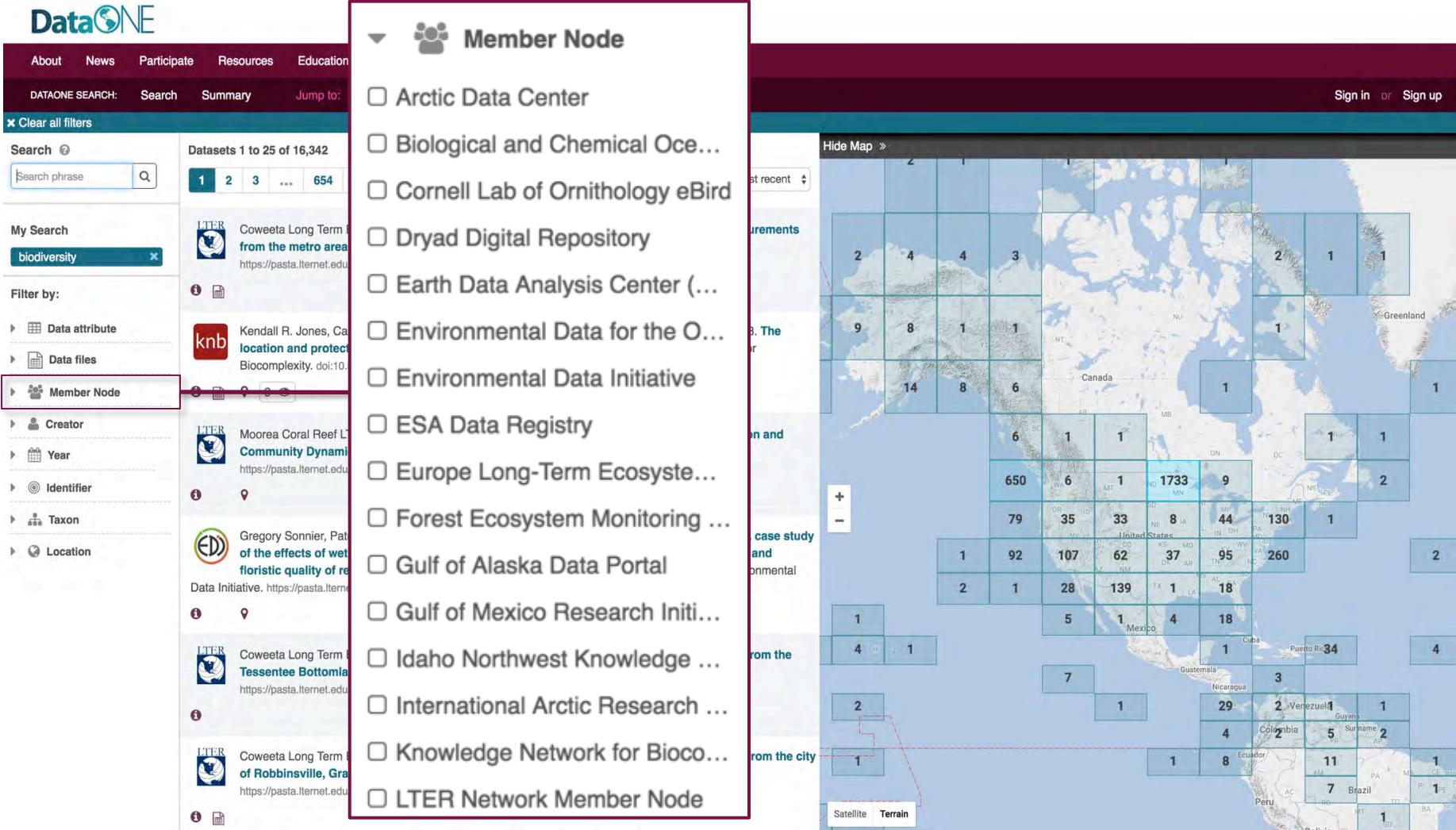
## Uploads

The number of individual metadata and data files uploaded over time. Only the first version of each file is counted.



# DataONE Search

*search.dataone.org*



# Data Management Education

**DataONE**

**Primer on Data Management: What you always wanted to know but were afraid to ask**

Carly Strasser, Robert Cook, William Michener, Amber Budden

**Contents**

1. Objective of This Primer
2. Why Manage Data?
  - 2.1. It will benefit you and your collaborators
  - 2.2. It will benefit the scientific community
  - 2.3. Journals and sponsors want you to share your data
3. How To Use This Primer
4. The Data Life Cycle: An Overview
5. Data Management Throughout the Data Life Cycle.
  - 5.1. Plan
  - 5.2. Collect
  - 5.3. Assure
  - 5.4. Describe: Data Documentation
  - 5.5. Preserve
  - 5.6. Discover, Integrate, and Analyze
6. Conclusion
7. Acknowledgements
8. References
9. Glossary

**1. Objective of This Primer**  
 The goal of data management is to produce self-describing data sets. If you give your data colleague who has not been involved with your project, will they be able to make sense of it effectively and properly? This primer describes a few fundamental data management practices that will enable you to develop a data management plan, as well as how to effectively organize, manage, describe, preserve and share data.

**2. Why Manage Data?**  
 The goal of data management is to produce self-describing data sets. If you give your data colleague who has not been involved with your project, will they be able to make sense of it effectively and properly? This primer describes a few fundamental data management practices that will enable you to develop a data management plan, as well as how to effectively organize, manage, describe, preserve and share data.

**DataONE Best Practices Primer**

**DataONE**

**Typical data analyses**

**Lesson 10: Analysis and Workflows**

View all Education Modules at <https://www.datatone.org/education-modules>

**Formal Workflow**  
 Analytical pipeline where each step can be implemented in different software systems. Parameters and requirements for each step are formally recorded.

- Single access point for multiple analyses across software packages
- Keeps track of data and provenance to better enable reproducibility
- Workflow can be stored
- Allows sharing and reuse of individual steps or overall workflow

**Formal workflow example: Kepler software**

**Reproducibility.**  
 ...is at the core of the scientific process. If results are not reproducible, they lose credibility. Good documentation of the data and the analysis are essential!

**Workflows**  
**Definition:** Precise description of the procedures used in a project. Can be formal or informal.

**Informal workflow**  
 No special software is needed to create workflow diagrams. Workflow diagrams include:

- Inputs and outputs
- Transformation rules or analytical processes
- Decision points
- Arrows indicating direction of process flow

**Informal Workflow Example**

**Best practices for data analysis**  
 Formal documentation document the workflows used to create results. Include:

- Data provenance
- Analyses and parameters used
- Connections between analyses via inputs and outputs

Document the code you write for analyses.

- Well-documented code is easier to review and share and enables repeated analyses
- Include project level information; script dependencies, inputs, and outputs; parameters; and what happens in individual sections

Construct end-to-end scripts that run the entire process from start to finish without intervention.

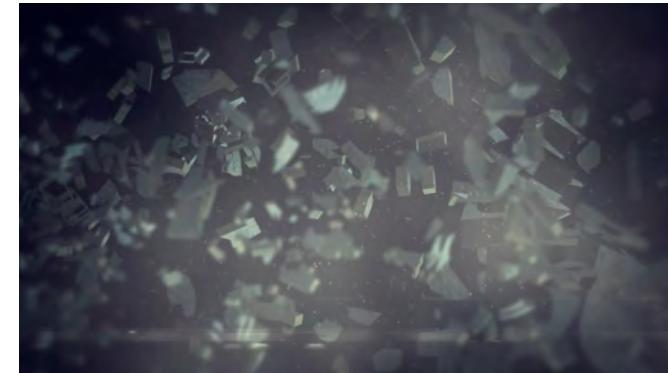
**Local contact information**

**DataONE**

**Tutorials on Data Management**

Lesson 5: Data Quality Control and Assurance

DataONE Education Module 05: Data Quality, Control and Assurance 624 views



**DataONE**

About News Participate Resources Education Data

DATATONE SEARCH: Search Summary Jump to DOI Go

**Clear filters**

Search

My Search  
 soil   
 water

Filter by:

- Data attribute
- Data files
- Member Node
- Creator
- Year
- Identifier

Datasets 1 to 25 of 11,077

Sort by: Most recent

1	2	3	...	444	Next
1	47	98	80		
201	484	7	98	1109	
72	211	148	233	167	
7	259	102	137	1	
				472	
				6	5
				1	

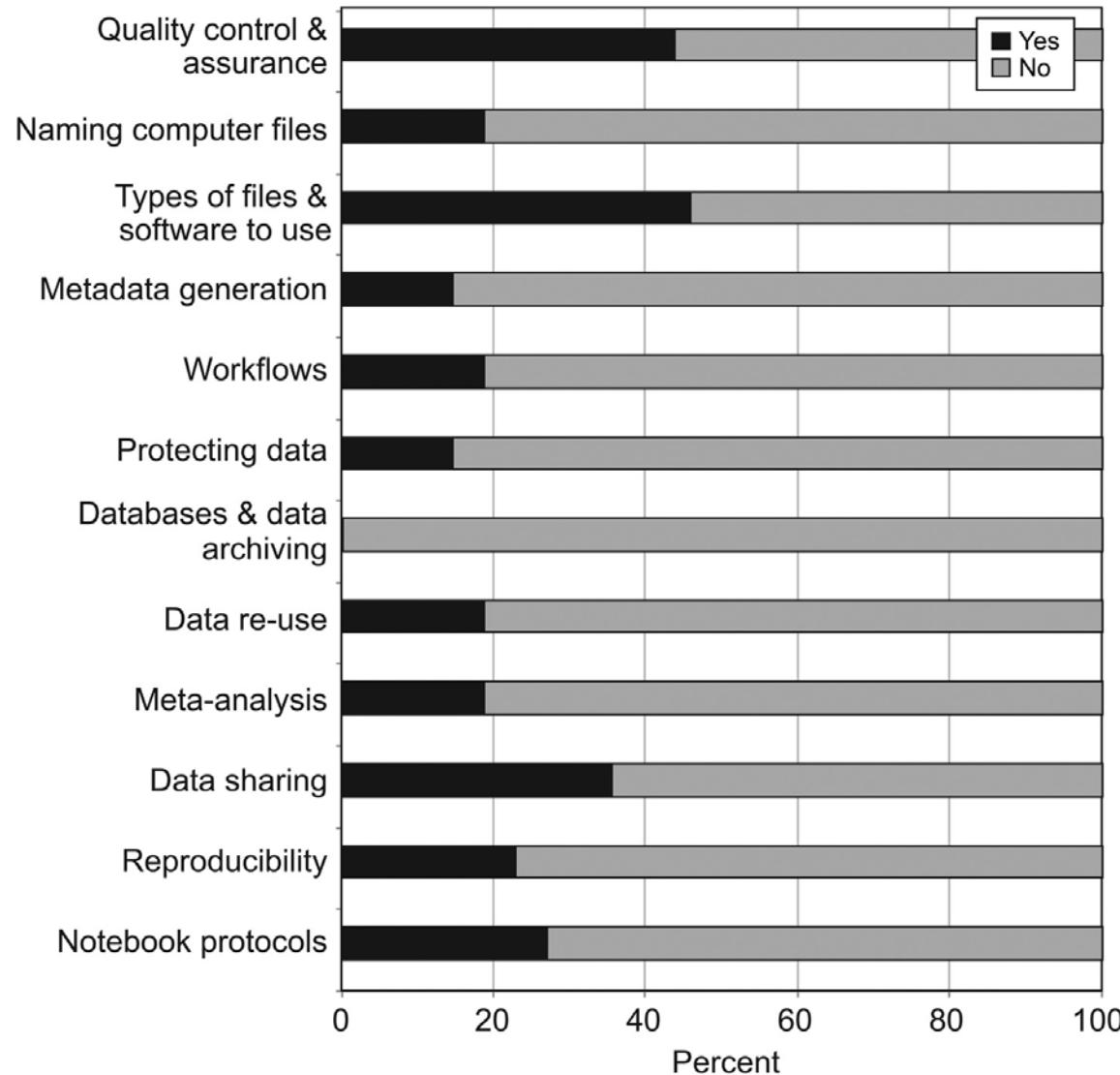
1 LTER John Yarie, and Bonanza Creek LTER. 2016. Soil Temperature at LTER Moisture Manipulation Treatments. U.S. LTER Network. <https://pasta.bremec.edu/package/metadata/rrn/rrn-bnz48f13>.

2 LTER John Yarie, and Bonanza Creek LTER. 2016. Soil Moisture (VWC) at LTER Moisture Manipulation Treatments. U.S. LTER Network. <https://pasta.bremec.edu/package/metadata/rrn/rrn-bnz48f13>.

3 LTER Jeff Walker, and Paddy Sullivan. 2015. Walker IPY. <https://pasta.bremec.edu/package/metadata/rrn/rrn-wlk15>.

4 Google Maps ©2016 Google. All rights reserved. Terms of Use

# The Fractured Lab Notebook



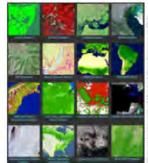
Strasser & Hampton (2012) *Ecosphere* 3:12 DOI: 10.1890/ES12-00139.1

# DataONE Education Modules

*[dataone.org/education-modules](http://dataone.org/education-modules)*

**Tutorials on Data Management**

Lesson 1: Introduction to Data Management  
Why Data Management?



Why Data Management

**DataONE**

**Tutorials on Data Management**

Lesson 2: Data Sharing



Data Sharing

**DataONE**

**Tutorials on Data Management**

Lesson 3: Data Management Planning



Data Management Planning

**DataONE**

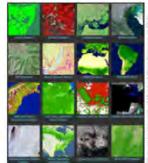


# DataONE Education Modules

[dataone.org/education-modules](http://dataone.org/education-modules)

## Tutorials on Data Management

Lesson 1: Introduction to Data Management  
Why Data Management?



Why Data Management

**DataONE**

## Tutorials on Data Management

Lesson 2: Data Sharing



Hands-on Exercises for <http://www.datatone.org>

**DataONE**

## Tutorials on Data Management

Lesson 3: Data Management Planning



Data Management Planning

**DataONE**

## Hands-on Activity 3: Data Management Planning

**Associated Data**

**Objectives:** Students identify the basic components of the DMPTool.

**Outcomes:** Students will learn how to use the DMPTool.

**Time Needed:** One hour

**URLs:** DMPTool: <http://www.datatone.org/dmp/>

**Additional Files:** None

**Key Reading:**

- Digital Data Management: A How-to Guide
- DCC How-to: <http://www.dcc.ac.uk/resources/guides/digital-data-management>
- NSF requirements: [http://www.nsf.gov/funding/pgm\\_index.jsp?pgm\\_id=430](http://www.nsf.gov/funding/pgm_index.jsp?pgm_id=430)
- DMPTool: <http://www.datatone.org/dmp/>
- Example DMP: <http://www.datatone.org/dmp/>

**Notes and Instructions:**

This exercise will help students understand the terms used in DMPTool. The tool terms used include: data, metadata, and data management plan. The tool to familiarize students with these terms.

- Decide on an exercise. "How can we ensure that our data is accessible?"
- Allow students to discuss what they have learned about data management.
- Have students reason about why data management is important.
- There is a good discussion on the importance of data management in the DMPTool FAQ site.

**Hands-on Exercises for <http://www.datatone.org>**

## Hands-on Activity 1: Accessing Data in the Literature

**DataONE**

View all Education Modules at <http://www.datatone.org/education-modules>

## Lesson 6: Protecting Your Data

**Backups** vs. **Archiving**

<b>Backups:</b> Periodic snapshots of current version	<b>Archiving:</b> Final version for historical reference or disasters
• Standard for short or near-term	• Often done on a somewhat frequent schedule
• Limit or negate data loss, particularly if data is not reproducible	• Save time and money

**Why perform backups?**

• Limit or negate data loss, particularly if data is not reproducible

• Save time and money

**DataONE**

View all Education Modules at <http://www.datatone.org/education-modules>

## Lesson 2: Data Sharing

Address data sharing throughout the data lifecycle

Describe data content, character, and process.

Describe a location from which it can be accessed.

Preserve it in formats on media good for long term.

Publish information about the data so that others can discover it easily.

**Concern**

Inappropriate use due to misunderstanding of research purpose or parameters	<b>Solution</b>
Security and confidentiality of sensitive data	• Provide rich abstract, purpose, constraints of use, & supplemental information as needed
Lack of credit or acknowledgement	• Publish data with author info
Loss of advantage when competing for research dollars	• Use Constraints can be used to say who may access & how

**The value of data sharing**

To the scientist

- Receive researcher recognition as an authoritative source
- Improved data quality
- Greater opportunity for data exchange
- Improved scientific network connections & potential collaboration

To the research sponsor/funder

- Enhanced value of investments by maximizing the return on research dollars spent.

To the research community

- Better ability to build trust rather than reputation among other researchers
- Ability to perform meta analyses
- Increased transparency, reproducibility, and comparability of results
- Ability to conduct methodology assessment, recommendations, & improvements
- Better education for new researchers on the most current and significant findings

To the public

- Better access to information leads to better understanding and contributions toward effective public & personal decision making

**Methods for making data sharable**

• Create a reusable, robust template.

• Include unique IDs & citation information.

• Have contributors review metadata for accuracy.

• Publish metadata via a portal or clearinghouse.

**Data sharing to understand Alzheimer's Disease**

"It's not science the way most of us have practiced in in our careers. But we all realized that we would never get biomarkers or treatments for Alzheimer's disease and intellectual-property issues outside the door and agreed that all of our data would be made public immediately." - John Trojanowski, U. Penn

**More at the NY Times: <http://nyti.ms/1pvKte44>**

**Local contact information**



# Community Use

Page	Total Events
	<b>232,749</b> % of Total: 100.00% (232,749)
1. <a href="/education-modules">/education-modules</a>	<b>30,854</b> (13.24%)
2. <a href="/best-practices/create-and-document-data-back-up-policy">/best-practices/create-and-document-data-back-up-policy</a>	<b>23,989</b> (10.30%)
3. <a href="/data-management-planning">/data-management-planning</a>	<b>23,604</b> (10.13%)
4. <a href="/best-practices">/best-practices</a>	<b>10,892</b> (4.67%)
5. <a href="/software-tools/fusion-lidar-software">/software-tools/fusion-lidar-software</a>	<b>8,562</b> (3.67%)
6. <a href="/current-member-nodes">/current-member-nodes</a>	<b>8,198</b> (3.52%)
7. <a href="/find-data">/find-data</a>	<b>7,635</b> (3.28%)
8. <a href="/">/</a>	<b>6,504</b> (2.79%)
9. <a href="/software-tools/dbdesigner-4">/software-tools/dbdesigner-4</a>	<b>6,409</b> (2.75%)
10. <a href="/investigator-toolkit">/investigator-toolkit</a>	<b>5,756</b> (2.47%)

# Community Use

Page	Event Label	Total Events
		<b>14,947</b> % of Total: 6.42% (232,749)
1. <a href="#">/education-module</a>	1. /sites/all/documents/L01_DataManagement.pptx	<b>2,115</b> (14.15%)
2. <a href="#">/best-practices/cre up-policy</a>	2. /sites/all/documents/education- modules/pptx/L01_DataManagement.pptx	<b>1,364</b> (9.12%)
3. <a href="#">/data-managemen</a>	3. /sites/all/documents/L07_Metadata.pptx	<b>1,029</b> (6.88%)
4. <a href="#">/best-practices</a>	4. /sites/all/documents/DataONE_Education_Modules_Full_Set.p ptx	<b>913</b> (6.11%)
5. <a href="#">/software-tools/fu</a>	5. /sites/all/documents/L03_DataManagementPlanning.pptx	<b>793</b> (5.30%)
6. <a href="#">/current-member-r</a>	6. /sites/all/documents/education- modules/pptx/L03_DataManagementPlanning.pptx	<b>762</b> (5.10%)
7. <a href="#">/find-data</a>	7. /sites/all/documents/L02_DataSharing.pptx	<b>740</b> (4.95%)
8. <a href="#">/</a>	8. /sites/all/documents/L04_DataEntryManipulation.pptx	<b>725</b> (4.85%)
9. <a href="#">/software-tools/db</a>	9. /sites/all/documents/education- modules/pptx/L07_Metadata.pptx	<b>634</b> (4.24%)
10. <a href="#">/investigator-toolk</a>	10. /sites/all/documents/L08_WriteQualityMetadata.pptx	<b>600</b> (4.01%)

# Maintenance

## Journal of eScience Librarianship

Volume 6 | Issue 2

Article 1

2017-09-08

### Using Peer Review to Support Development of Community Resources for Research Data Management

Heather Soyka

*Kent State University*

Amber Budden

*DataONE/University of New Mexico*

Viv Hutchison

*US Geological Survey*

# Challenges



# Solutions

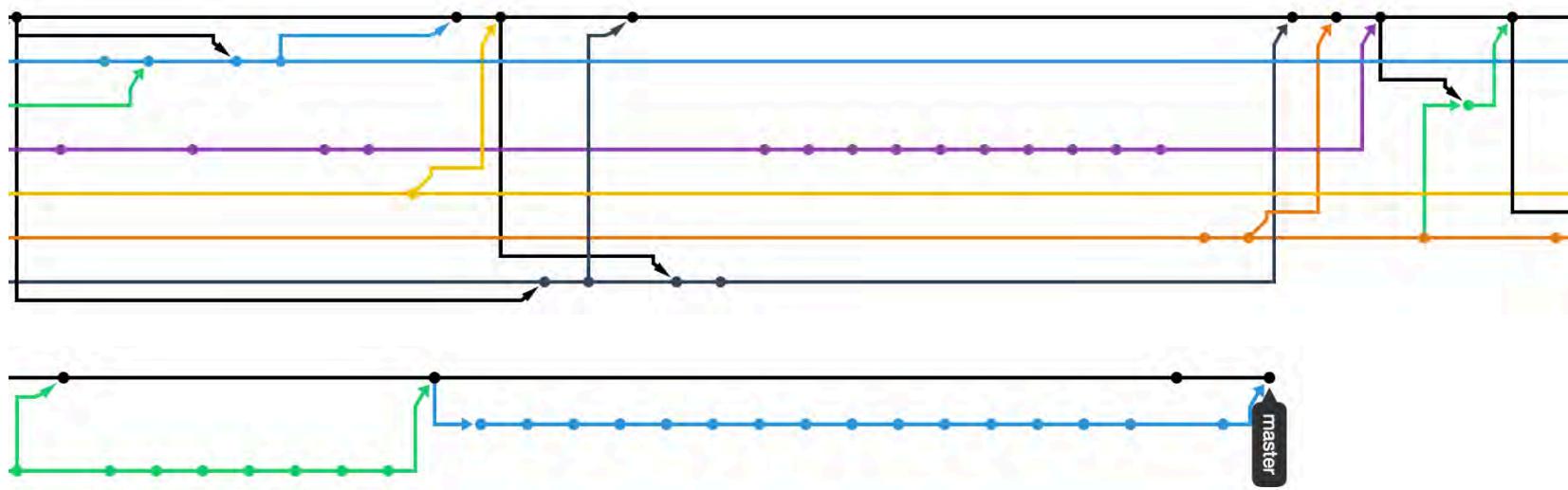


# Versioned work space

Clone or download ▾

Fork

11



# Versioned work space

Clone or download ▾

Fork 11

DataONEorg / Education

- aebudden / dataone\_lessons
- amandawhitmire / dataone\_lessons
- andosmith / Education
- datadavev / d1\_Education\_fork
- dbloom / dataone\_lessons
- hsoyka / dataone\_lessons
- jduckles / dataone\_lessons
- lwasser / dataone\_lessons
- mayernik / dataone\_lessons
- vhutch / dataone\_lessons
- whshannon / dataone\_lessons

# Structured repository

The screenshot shows a GitHub repository page for 'DataONEorg / hub\_lessons'. The repository has 11 watchers, 5 stars, and 18 forks. The main navigation tabs include Code, Issues (0), Pull requests (0), Projects (1), Wiki, Insights, and Settings. Below the tabs, there's a breadcrumb navigation showing 'Tree: 23581c038b' and the path 'hub\_lessons / lessons /'. There are buttons for Create new file, Upload files, Find file, and History.

A commit by user 'meg3mach' is displayed, titled 'update org info and title/author in front matter'. The commit was made on Jul 23, 2019, with the latest commit hash being f82ccb0. The commit message also includes 'Latest commit f82ccb0 on Jul 23'.

The repository contains a folder structure under 'lessons':

- ..
- 00\_markdown
- 01\_management
- 02\_datasharing
- 03\_planning
- 04\_entry
- 05\_qaqc
- 06\_protect
- 07\_metadata
- 08\_citation
- 09\_analysis
- 10\_policy

Each item in the list has a description: 'update org info and title/author in front matter' and a timestamp: '3 months ago'.

# Structured repository

The screenshot shows a GitHub repository interface for the DataONEorg/hub\_lessons repository. The repository has 11 stars, 5 forks, and 18 issues. It contains one project and one wiki page. The current branch is master, and the user is viewing the hub\_lessons/lessons/01\_management directory. The latest commit was made by meg3mach on July 23, 2018. The commit message is "update org info and title/author in front matter". Below the commit, there is a list of files and their changes:

File	Description	Time Ago
images	Changes to M1	2 years ago
01_management.pdf	Regenerated PDFs for lessons	2 years ago
L01_DataManagement_Handout.pdf	upload handout and exercise lesson 1	a year ago
L01_Exercise.pdf	upload handout and exercise lesson 1	a year ago
index.md	update org info and title/author in front matter	3 months ago
index.png	change all cover.png to index.png in lessons	5 months ago
slides.md	switch index to cover and vis versa in lessons	5 months ago

# Data Management Skillbuilding Hub



Home Contribute FAQ GitHub



The Data Management Skillbuilding Hub contains resources for better data management and is open to community input and update. These resources are adaptable across a range of contexts and intended for use by researchers, teachers, librarians, or anyone who wants to learn better data management practices. Each tile below links to community contributed education materials, such as best practices and lesson plans.

The resources presented on the Data Management Skillbuilding Hub can be updated by users to promote a current, well-maintained, and sustainable educational tool. Learn more about how you can [contribute](#).

## *Using This Resource*

Click individual tiles to learn more and use each resource. You can limit resources by content type and [Data Life Cycle](#) stage. Comprehensive information is available in the [FAQ](#).

» Filter by content type:

**ALL**

TEACHING MODULE

BEST PRACTICE

VIDEO

» Filter by stage of the Data Life Cycle

All



01.  
Why Data  
Management



02.  
Data  
Sharing



03.  
Data Management  
Planning



04.  
Data Entry  
and Manipulation



05.  
Data Quality  
Control and Assurance



06.  
Protecting  
Your Data



07.  
Metadata



08.  
Data Citation

# Data Management Skillbuilding Hub

- Structure: Data life-cycle
- Current holdings: Education modules and best practices
- Citation: Credit where credit is due
- Editing content (forking!)
- Creating new content (in the works)
- Future holdings

The Data Management Skillbuilding Hub contains resources for better data management and is open to community input and update. These resources are adaptable across a range of contexts and intended for use by researchers, teachers, librarians, or anyone who wants to learn better data management practices. Each tile below links to community contributed education materials, such as best practices and lesson plans

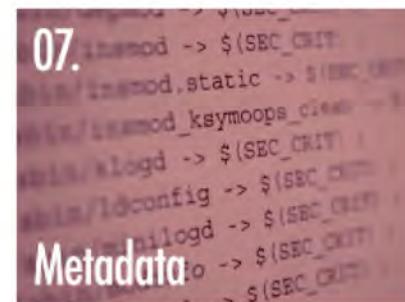
The resources presented on the Data Management Skillbuilding Hub can be updated by users to promote a current, well-maintained, and sustainable educational tool. Learn more about how you can [contribute](#).

## Using This Resource

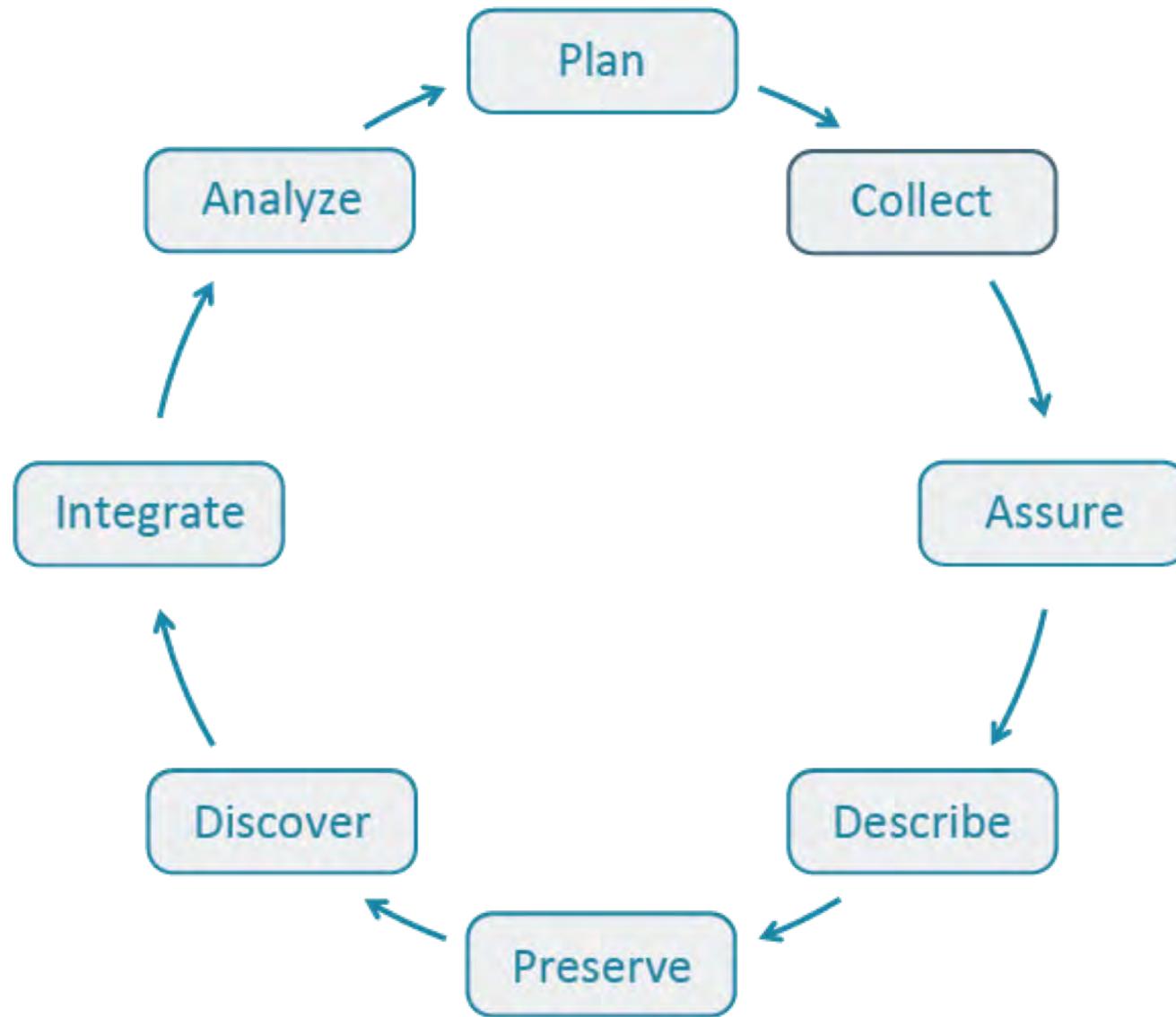
Click individual tiles to learn more and use each resource. You can limit resources by content type and [Data Life Cycle](#) stage. Comprehensive information is available in the [FAQ](#).

» Filter by content type: [ALL](#) [TEACHING MODULE](#) [BEST PRACTICE](#) [VIDEO](#)

» Filter by stage of the Data Life Cycle  



# Data life-cycle stages



# Data Management Skillbuilding Hub



Home Contribute FAQ GitHub



The Data Management Skillbuilding Hub contains resources for better data management and is open to community input and update. These resources are adaptable across a range of contexts and intended for use by researchers, teachers, librarians, or anyone who wants to learn better data management practices. Each tile below links to community contributed education materials, such as best practices and lesson plans

The resources presented on the Data Management Skillbuilding Hub can be updated by users to promote a current, well-maintained, and sustainable educational tool. Learn more about how you can [contribute](#).

## Using This Resource

Click individual tiles to learn more and use each resource. You can limit resources by content type and [Data Life Cycle](#) stage. Comprehensive information is available in the [FAQ](#).

» Filter by content type:

**ALL**

All  
Plan  
Collect  
Assure  
✓ Describe  
Preserve  
Discover  
Integrate  
Analyze

**BEST PRACTICE**

**VIDEO**

» Filter by stage of the Data Life Cycle



01.  
**Why Data Management**



02.  
**Data Sharing**



07.  
**Metadata**



08.  
**Data Citation**

## Hosted by DataONE

In collaboration with the community, DataONE has developed high quality resources for helping educators and librarians with training in data management, including teaching materials, webinars and a database of best-practices to improve methods for data sharing and management.

ⓘ If you have a question or concern, please open an [Issue](#) in this repository on GitHub.

## TEACHING MODULE

02.

### Data Sharing



### Presentation View

Quick tips: Press **p** for presentation; **f** for full screen

#### Supporting downloads:

[PDF Download](#)[PPT Download](#)[Handout](#)[Hands-on Exercise](#)

When first sharing research data, researchers often raise questions about the value, benefits, and mechanisms for sharing. Many stakeholders and interested parties, such as funding agencies, communities, other researchers, or members of the public may be interested in research, results and related data. This lesson addresses data sharing in the context of the data life cycle, the value of sharing data, concerns about sharing data, and methods and best practices for sharing data.

#### Cite this lesson:

DataONE Community Engagement & Outreach Working Group (2017) "Data Sharing". Accessed through the Data Management Skillbuilding Hub at [https://dataoneorg.github.io/Education/lessons/02\\_datasharing/index](https://dataoneorg.github.io/Education/lessons/02_datasharing/index) on Oct 04, 2018



< [Home](#) >

## The Data Lifecycle

Several stages require critical attention to ensure effective data sharing

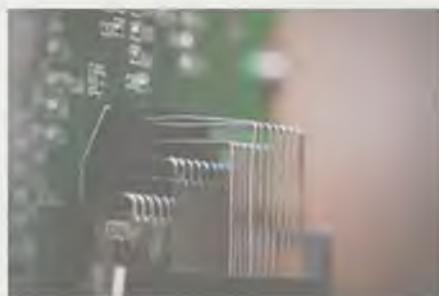
Step	Action
<b>Describe</b>	document the data content, character and process
<b>Deposit</b>	store the data in a location from which it can be accessed
<b>Preserve</b>	select storage formats and media with long term use in mind
<b>Discover</b>	publish information about the data so that others can find it

## Why share data

Data sharing requires effort, resources, and faith in others. Why do it?

For the benefit of:

- the public
- the research sponsor
- the research community
- the researcher



CC image by Jessica Lucia on Flickr

### NOTES FOR CURRENT SLIDE

Effective data sharing requires careful thought during each stage of the data development process including:

- description and documentation of the data process, content, and character;
- deposition and storage of the data in a location from which it can be accessed or shared;
- preservation of the data using a format and media that enable long term reuse; and
- making the data discoverable by publishing information about the data in research publications, data clearinghouses and data distribution portals.

### NOTES FOR NEXT SLIDE

Why expend the extra effort to share data? Because it benefits the public, the research sponsor, the research community and, perhaps most importantly, the researcher.

## TEACHING MODULE

02.

### Data Sharing



## Presentation View

Quick tips: Press **p** for presentation; **f** for full screen

### Supporting downloads:

[PDF Download](#)[PPT Download](#)[Handout](#)[Hands-on Exercise](#)

When first sharing research data, researchers often raise questions about the value, benefits, and mechanisms for sharing. Many stakeholders and interested parties, such as funding agencies, communities, other researchers, or members of the public may be interested in research, results and related data. This lesson addresses data sharing in the context of the data life cycle, the value of sharing data, concerns about sharing data, and methods and best practices for sharing data.

### Cite this lesson:

DataONE Community Engagement & Outreach Working Group (2017) "Data Sharing". Accessed through the Data Management Skillbuilding Hub at [https://dataoneorg.github.io/Education/lessons/02\\_datasharing/index](https://dataoneorg.github.io/Education/lessons/02_datasharing/index) on Oct 04, 2018



< [Home](#) >

# Data Management Skillbuilding Hub



[Home](#)   [Contribute](#)   [FAQ](#)   [GitHub](#)



The Data Management Skillbuilding Hub contains resources for better data management and is open to community input and update. These resources are adaptable across a range of contexts and intended for use by researchers, teachers, librarians, or anyone who wants to learn better data management practices. Each tile below links to community contributed education materials, such as best practices and lesson plans

The resources presented on the Data Management Skillbuilding Hub can be updated by users to promote a current, well-maintained, and sustainable educational tool. Learn more about how you can [contribute](#).

## Using This Resource

Click individual tiles to learn more and use each resource. You can limit resources by content type and [Data Life Cycle](#) stage. Comprehensive information is available in the [FAQ](#).

» Filter by content type: [ALL](#) [TEACHING MODULE](#) [BEST PRACTICE](#) [VIDEO](#)

» Filter by stage of the Data Life Cycle



## BEST PRACTICE

### Best Practice: Assure

Select a Best Practice below to learn more about the "Assure" stage in the [Data Life Cycle](#).

#### What is the “Assure” stage?

Employ quality assurance and quality control procedures that enhance the quality of data (e.g., training participants, routine instrument calibration) and identify potential errors and techniques to address them.

More information can be found in the [Best Practices Primer](#).



#### Best Practices by Data Life Cycle

All  
Plan  
Collect  
Assure  
Describe  
Preserve  
Discover  
Integrate  
Analyze

Learn more:  
[BP Primer](#)

## BEST PRACTICE

### Best Practice: Assure

Select a Best Practice below to learn more about the "Assure" stage in the [Data Life Cycle](#).

#### What is the “Assure” stage?

Employ quality assurance and quality control procedures that enhance the quality of data (e.g., training participants, routine instrument calibration) and identify potential errors and techniques to address them.

More information can be found in the [Best Practices Primer](#).



#### Communicate data quality

Information about quality control and quality assurance are important components of the metadata: [\(click for more\)](#)

Tags: assure flag qualify

#### Confirm a match between data and their description in metadata

To assure that metadata correctly describes what is actually in a data file, visual inspection or analysis should be done by someone not otherwise familiar with the data and its format. This will assure that the metadata is sufficient to describe the data... [\(click for more\)](#)

Tags: assure data consistency describe documentation metadata quality

#### Consider the compatibility of the data you are integrating

The integration of multiple data sets from different sources requires that they be compatible. Methods used to create the data should be considered early in the process, to avoid problems later during attempts to integrate data sets. Note that just because... [\(click for more\)](#)

Tags: analyze assure database integrate quality tabular

#### Develop a quality assurance and quality control plan

Just as data checking and review are important components of data management, so is the step of documenting how these tasks were accomplished. Creating a plan for how to review the data before it is collected or compiled allows a researcher to

#### Best Practices by Data Life Cycle

All  
Plan  
Collect  
Assure  
Describe  
Preserve  
Discover  
Integrate  
Analyze

#### Learn more:

BP Primer

## BEST PRACTICE

### Communicate data quality

Data Life Cycle stage(s): [Assure](#)

Information about quality control and quality assurance are important components of the metadata:

- Qualify (flag) data that have been identified as questionable by including a flagging\_column next to the column of data values. The two columns should be properly associated through a naming convention such as Temperature, flag\_Temperature.
- Describe the quality control methods applied and their assumptions in the metadata. Describe any software used when performing the quality analysis, including code where practical. Include in the metadata who did the quality control analysis, when it was done, and what changes were made to the dataset.
- Describe standards or test data used for the quality analysis. For instance, include, when practical, the data used to make a calibration curve.
- If data with qualifier flags are summarized to create a derived data set, include the percent flagged data and percent missing data in the metadata of the derived data file. High frequency observations are often downsampled, and it is critical to know how much of the data were rejected in the primary data.

#### Best Practices by Data Life Cycle

All  
Plan  
Collect  
Assure  
Describe  
Preserve  
Discover  
Integrate  
Analyze

#### Learn more:

[BP Primer](#)

### Description Rationale

Data quality and any methods used for quality control should be communicated so others can assess the data independently.

### Additional Information

Hook, L.A., Beaty, T.W., Santhana-Vannan, S., Baskaran, L. and Cook, R.B. 2007. Best practices for preparing environmental data sets to share and archive. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. ([daac.ornl.gov/PI/bestprac.html](http://daac.ornl.gov/PI/bestprac.html))

Sheldon, W., Henshaw, D. and Ramsey, K. 2007. Final Report: Workshop to define quality management standards for data completeness in derived data products. Long Term Ecological Research Network Document Archive, University of New Mexico, Albuquerque, NM.

### Additional Information (Biblio)

[Best Practices for Preparing Ecological and Ground-Based Data Sets to Share and Archive](#)

## Search the Data Management Skill Building Hub

Results are listed in order of find, not by best match to search word, and will be alphabetically ordered by result type.

data backup

### **Teaching Module: Protecting Your Data: Backups, Archives & Data Preservation**

Life Cycle Step(s): preserve, assure

Authoring Organization: DataONE

### **Best Practice: Create and document a data backup policy**

Life Cycle Step(s): plan, preserve

Authoring Organization: DataONE

### **Best Practice: Ensure integrity and accessibility when making backups of data**

Life Cycle Step(s): preserve

Authoring Organization: DataONE

### **Hosted by DataONE**

In collaboration with the community, DataONE has developed high quality resources for helping educators and librarians with training in data management, including teaching materials, webinars and a database of best-practices to improve methods for data sharing and management.

?

If you have a question or concern, please open an [Issue](#) in this repository on GitHub.

## TEACHING MODULE

02.

### Data Sharing



### Presentation View

Quick tips: Press **p** for presentation; **f** for full screen

#### Supporting downloads:

[PDF Download](#)[PPT Download](#)[Handout](#)[Hands-on Exercise](#)

When first sharing research data, researchers often raise questions about the value, benefits, and mechanisms for sharing. Many stakeholders and interested parties, such as funding agencies, communities, other researchers, or members of the public may be interested in research, results and related data. This lesson addresses data sharing in the context of the data life cycle, the value of sharing data, concerns about sharing data, and methods and best practices for sharing data.

#### Cite this lesson:

DataONE Community Engagement & Outreach Working Group (2017) "Data Sharing". Accessed through the Data Management Skillbuilding Hub at [https://dataoneorg.github.io/Education/lessons/02\\_datasharing/index](https://dataoneorg.github.io/Education/lessons/02_datasharing/index) on Oct 04, 2018



< [Home](#) >

# Data Management Skillbuilding Hub

- Structure: Data life-cycle stage
- Current holdings: Education modules and best practices
- Citation: Credit where credit is due
- **Editing content (forking!)**
- Creating new content (in the works)
- Future holdings

## Guidelines for contributors and content editors

This document details our recommended processes to [update current content](#), [suggest changes to content](#), and [fork content for your own use](#), as well as an introduction to [how the content is organized](#) and the [tools we use to display content](#).

This repository was developed by the DataONE Community Engagement and Outreach Working Group and continues to be maintained by members of this team. Thank you for your interest in contributing to these educational materials.

## Guidelines for contributors and content editors

This document details our recommended processes to [update current content](#), [suggest changes to content](#), and [fork content for your own use](#), as well as an introduction to [how the content is organized](#) and the [tools we use to display content](#).

This repository was developed by the DataONE Community Engagement and Outreach Working Group and continues to be maintained by members of this team. Thank you for your interest in contributing to these educational materials.

### Update current content

Want to update a link or method? See a spelling error? Changes can be easily proposed by opening the [GitHub Education page](#) and editing content directly. For help, try this brief [GitHub tutorial](#) on forking and editing content.

### Edit content

1. Create a fork of the [lessons](#) or [best practices](#) repository into your github account, depending on which content you wish to edit.
2. Modify the files that you want to change ([See "Structure" below for tips on making changes](#)).
3. Submit a pull-request against the `master` branch of this repository.
4. Your changes will be reviewed by the repository admins.

## Guidelines for contributors and content editors

This document details our recommended processes to [update current content](#), [suggest changes to content](#), and [fork content for your own use](#), as well as an introduction to [how the content is organized](#) and the [tools we use to display content](#).

This repository was developed by the DataONE Community Engagement and Outreach Working Group and continues to be maintained by members of this team. Thank you for your interest in contributing to these educational materials.

### Update current content

Want to update a link or method? See a spelling error? Changes can be easily proposed by opening the [GitHub Education page](#) and editing content directly. For help, try this brief [GitHub tutorial](#) on forking and editing content.

### Edit content

1. Create a fork of the [lessons](#) or [best practices](#) repository into your github account, depending on which content you wish to edit.
2. Modify the files that you want to change ([See "Structure" below](#) for tips on making changes).
3. Submit a pull-request against the `master` branch of this repository.
4. Your changes will be reviewed by the repository admins.

### Page not rendering?

Check that the `title` field of the YAML header (the first line of each lesson) is in quotes.

### Suggest changes to content

1. Open an [Issue](#) on this repository.
2. Provide your suggested changes with as much detail and guidance as possible. Be specific.
3. Your suggestions will be reviewed by the repository admins.
4. Changes will be pushed to the repository by the repository admins regularly/as needed.



## DataONEorg / Education

Watch 1

Star 0

Fork 1

Code

Issues 0

Pull requests 0

Projects 0

Insights

is:issue is:open

Labels

Milestones

New Issue

① 0 Open ✓ 0 Closed

Author ▾

Labels ▾

Projects ▾

Milestones ▾

Assignee ▾

Sort ▾

**There aren't any open issues.**You could search [all of GitHub](#) or try an [advanced search](#).💡 **ProTip!** Type on any issue or pull request to go back to the issue listing page.

## Guidelines for contributors and content editors

This document details our recommended processes to [update current content](#), [suggest changes to content](#), and [fork content for your own use](#), as well as an introduction to [how the content is organized](#) and the [tools we use to display content](#).

This repository was developed by the DataONE Community Engagement and Outreach Working Group and continues to be maintained by members of this team. Thank you for your interest in contributing to these educational materials.

### Update current content

Want to update a link or method? See a spelling error? Changes can be easily proposed by opening the [GitHub Education page](#) and editing content directly. For help, try this brief [GitHub tutorial](#) on forking and editing content.

### Edit content

1. Create a fork of the [lessons](#) or [best practices](#) repository into your github account, depending on which content you wish to edit.
2. Modify the files that you want to change ([See "Structure" below for tips on making changes](#)).
3. Submit a pull-request against the `master` branch of this repository.
4. Your changes will be reviewed by the repository admins.

### Page not rendering?

Check that the `title` field of the YAML header (the first line of each lesson) is in quotes.

### Suggest changes to content

1. Open an [Issue](#) on this repository.
2. Provide your suggested changes with as much detail and guidance as possible. Be specific.
3. Your suggestions will be reviewed by the repository admins.
4. Changes will be pushed to the repository by the repository admins regularly/as needed.

### Fork content for your own use

Fork and edit content through GitHub, rather than editing privately, to enable others to use your edited content and to track how these materials are used.

1. Create a fork of the [lessons](#) or [best practices](#) repository into your github account
2. Modify the files that you want to change ([See "Structure" below for tips on making changes](#))

Edit or fork content for your own use

Select one of the buttons below to open the appropriate GitHub repository:

[Teaching Module](#)[Best Practice](#)

***Hosted by DataONE***

In collaboration with the community, DataONE has developed high quality resources for helping educators and librarians with training in data management, including teaching materials, webinars and a database of best-practices to improve methods for data sharing and management.

 If you have a question or concern, please open an [Issue](#) in this repository on GitHub.

[DataONEorg / hub\\_bestpractices](#)[Watch 6](#)[Star 0](#)[Fork 8](#)[Code](#)[Issues 0](#)[Pull requests 0](#)[Projects 1](#)[Insights](#)

Tree: b10d396c32 ▾

[hub\\_bestpractices / bestpractices /](#)[Create new file](#)[Find file](#)[History](#)

meg3mach updates to frontmatter and stage blurbs

Latest commit 407423d on Aug 30

..

<a href="#">images</a>	adding organization info to bp front matter	3 months ago
<a href="#">logos</a>	bestpractices citation	3 months ago
<a href="#">advertise-your-data.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">assign-descriptive-file.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">backup-your-data.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">check-data-and.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">choose-and-use.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">communicate-data-quality.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">confirm-a-match.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">consider-the-compatibility.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">create-a-data.md</a>	updates to frontmatter and stage blurbs	a month ago
<a href="#">create-and-document.md</a>	updates to frontmatter and stage blurbs	a month ago
<a href="#">create-manage-and.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">decide-what-data.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">define-expected-data.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">define-roles-and.md</a>	Add authorship and update date information to BP	2 months ago
<a href="#">define-the-data.md</a>	Add authorship and update date information to BP	2 months ago



Search or jump to...

Pull requests Issues Marketplace Explore

Bell icon + P icon

meg3mach / BestPractices

forked from DataONEorg/hub\_bestpractices

Watch 0

Star 0

Fork 8

Code

Pull requests 1

Projects 0

Wiki

Insights

Settings

BestPractices / bestpractices /

confirm-a-match.md

or cancel

Edit file

Preview changes

Spaces

2

Soft wrap

```
1  ---
2  title: Confirm a match between data and their description in metadata
3  layout: bestpractice_cover
4  tags:
5    - assure
6    - data consistency
7    - describe
8    - documentation
9    - metadata
10   - quality
11  step:
12    - assure
13    - describe
14  related:
15    - consider-the-compatibility
16    - describe-the-contents
17    - define-the-data
18  update:
19    - May 11, 2011
20  author:
21    - Eric Lind
22    - John Porter
23    - Michael Grady
24  organization: DataONE
25  org_url: http://www.dataone.org
26  org_logo: DataONE.png
27  resource: true
28  categories: ["Best Practice"]
29  ---
30
31 To assure that metadata correctly describes what is actually in a data file, visual inspection or analysis should be done by someone not otherwise familiar with the data and its format. This will assure that the metadata is sufficient to describe the data. For example,
```

# Data Management Skillbuilding Hub

- Structure: Data life-cycle stage
- Current holdings: Education modules and best practices
- Citation: Credit where credit is due
- Editing content (forking!)
- **Creating new content (in the works)**
- Future holdings

# Data Management Skillbuilding Hub

- Structure: Data life-cycle stage
- Current holdings: Education modules and best practices
- Citation: Credit where credit is due
- Editing content (forking!)
- Creating new content (in the works)
- **Future holdings**

# Data Management Skillbuilding Hub

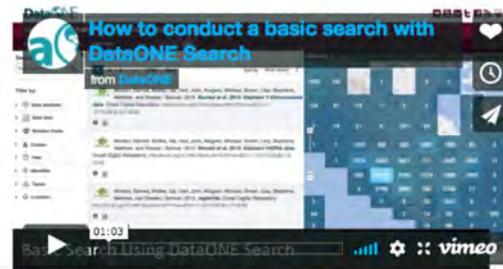
- Structure: Data life-cycle stage
- Current holdings: Education modules and best practices
- Citation: Credit where credit is due
- Editing content (forking!)
- Creating new content (in the works)
- **Future holdings**

## Upcoming Webinar:



## Tutorials:

### Conducting a Basic Search



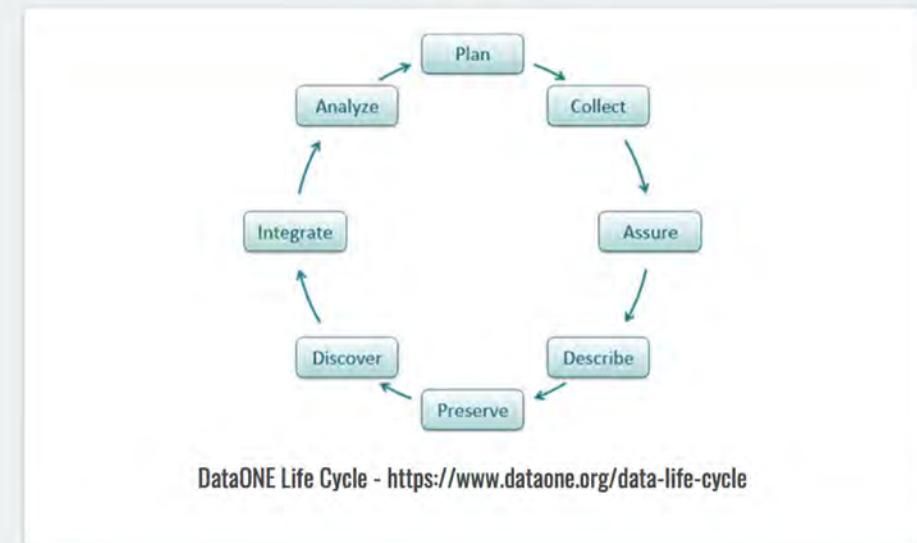
DataONE Search: Basic Search from DataONE on Vimeo.

## Welcome to the DMT Clearinghouse

The Data Management Training (DMT) Clearinghouse is a registry for online learning resources focusing on research data management.

It was created in a collaboration between the [U.S. Geological Survey's Community for Data Integration](#), the [Earth Sciences Information Partnership \(ESIP\)](#), and [DataONE](#).

For questions or feedback, please contact  
[clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

[Read More](#)

### Search

Find learning resources by keyword, name, date, license and cost

[Search](#)

### Browse

See a list of learning resources by educational framework

[Browse](#)

### Submit

Submit your learning resources to the Clearinghouse

[Submit](#)

# Data Management Training (DMT) Clearinghouse:



**A Convenient and Curated Source  
for Finding Educational Resources  
on Research Data Management**

**(RDM)**

DataONE Webinar

October 9, 2018

Nancy J. Hoebelheinrich

Knowledge Motifs LLC



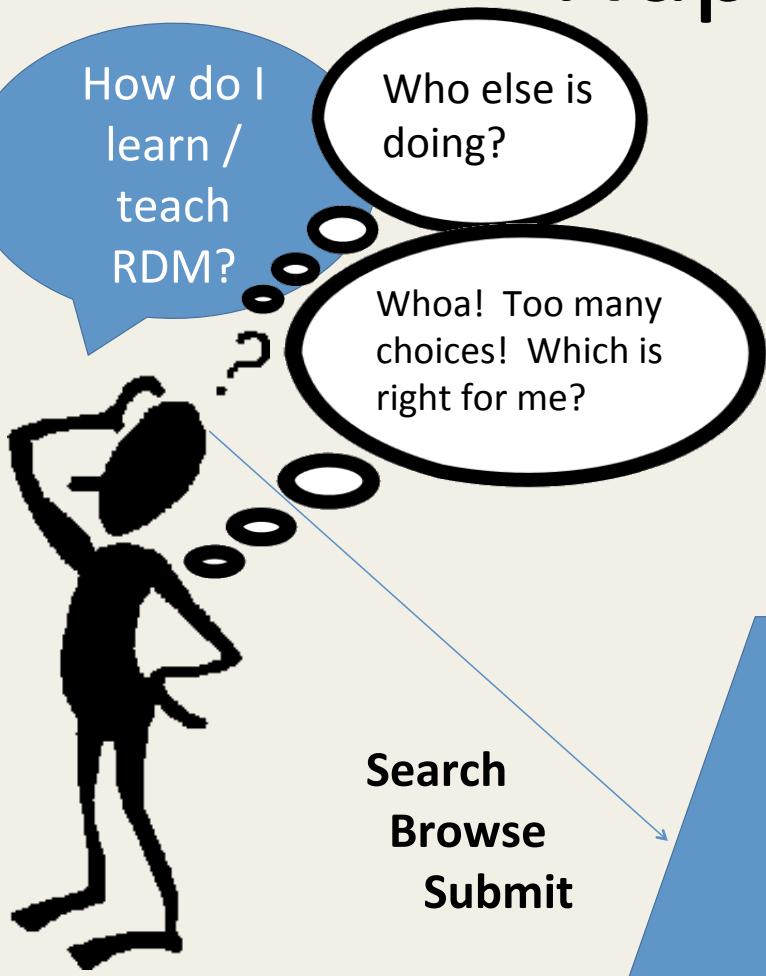
*Mapping sensible data relationships*

# Research Data Management Training – Sometimes you or your research team need it,

- Why is training needed?
  - Motivators:
    - Funders require
    - Publishers are beginning to require
    - For scientific reproducibility
    - For data re-use by colleagues & collaborators
    - Community culture beginning to expect open data
    - Others...?

but  
where to find  
training  
resources?

# Napkin Drawing



Lab / Classroom / 1 on 1



Self – taught  
Researcher

Excerpted from the SGCI BootCamp “Pitch Deck”

# Introducing the ESIP-hosted Data Management Training Clearinghouse!

 **ESIP** Data Management Training

Home Browse Search Submit Help About Log In

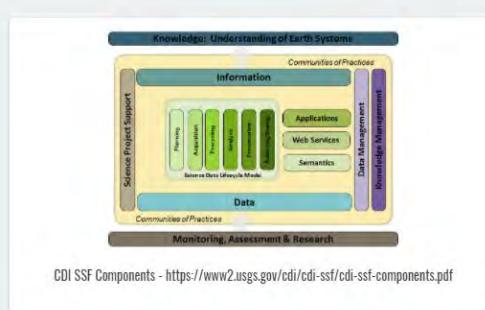
### Welcome to the DMT Clearinghouse

The Data Management Training (DMT) Clearinghouse is a registry for online learning resources focusing on research data management.

It was created in a collaboration between the U.S. Geological Survey's Community for Data Integration, the Earth Sciences Information Partnership (ESIP), and DataONE.

For questions or feedback, please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org).

[Read More](#)



CDI SSF Components - <https://www2.usgs.gov/cdi/cdi-ssf/cdi-ssf-components.pdf>

**Search**

Find learning resources by keyword, name, date, license and cost

 **Browse****Submit****CONNECT WITH US**  
on social networks**SUBSCRIBE**  
to the monday updates**SEARCH**  
this esip site

Questions or issues with the website? Please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)  
ESIP is a collaboration among many partner organizations, activities are sponsored by [NASA](#) and [NOAA](#) and managed by the [Foundation for Earth Science](#).  
appropriate reward and career opportunities for RSES. We organise regular events to

<http://dmtclearinghouse.esipfed.org/>

# What is the Data Management Training (DMT) Clearinghouse??

## What?

- *Metadata registry* for educational resources on research data management
- Capabilities include:
  - *Search*
  - *Browse*
  - *Submit*
- Collaboratively developed & maintained

## What kind of training resources?

- Short courses ala “Kahn Academy” (7 – 15 min. modules)
- Videos
- Learning activities to supplement courses
- Presentations & webinars
- Data “recipes”
- Syllabi & curricula

<http://dmtclearinghouse.esipfed.org/>

# Why use the DMT Clearinghouse??

## Answers to these Researcher questions:

- What kind of resources are available on RDM & where did they come from?
- Do they pertain to my subject domain?
- Do they fit the data management framework of my organization?
- Are they appropriate for my role on my research team?
- What do they cost?

## Answers to these Data Specialist questions:

- I've been asked to provide some training to a research team on RDM . What have others done that I can adapt?
- Where can I find practical, subject-domain targeted exercises for my generic tutorials on RDM?
- Colleagues keep asking me for the training resources that I've created. Where can I share them easily?

# Where, how is the DMT Clearinghouse maintained??

## Currently, hosted & maintained

- On & by the **ESIP Federation Commons** – a Drupal based content management system
- Don't need to register for Search / Browse / Suggest a Resource to add
- Do not need an ESIP Acct to "submit" a resource, unless planning to create full description

## Sustainability Plan

- Crowd-sourced submissions
- Domain-knowledgeable reviewers & editors to maintain quality & currency of resources
- Always seeking user interface & functionality feedback
- After initial seed \$\$, have been funded by IMLS for 3 year **National Leadership Grant**
- Exploring options for longer term Sustainability from NSF

# Let's take a look!

 **ESIP** Data Management Training

Home Browse Search Submit Help About Log in

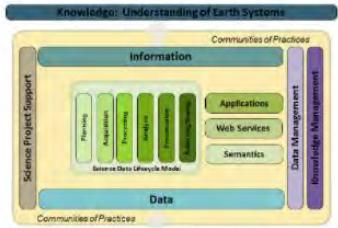
## Welcome to the DMT Clearinghouse

The Data Management Training (DMT) Clearinghouse is a registry for online learning resources focusing on research data management.

It was created in a collaboration between the U.S. Geological Survey's Community for Data Integration, the Earth Sciences Information Partnership (ESIP), and DataONE.

For questions or feedback, please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

[Read More](#)



CDI SSF Components - <https://www2.usgs.gov/cdi/cdi-ssf/cdi-ssf-components.pdf>

**Search** 

Find learning resources by keyword, name, date, license and cost

**Browse** 

See a list of learning resources by educational framework

**Submit** 

Submit your learning resources to the Clearinghouse

**CONNECT WITH US**  
on social networks

**SUBSCRIBE**  
to the monday updates

**SEARCH**  
this esip site

Questions or issues with the website? Please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

ESIP is a collaboration among many partner organizations, activities are sponsored by [NASA](#) and [NOAA](#) and managed by the Foundation for Earth Science. appropriate reward and career opportunities for RSFs. We organise regular events to

<http://dmtclearinghouse.esipfed.org/>

# Browse Function

The screenshot shows the DMT Clearinghouse website interface. A red circle highlights the 'Browse' link in the top navigation bar. Another red circle highlights the 'Browse' button in the main content area.

**Welcome to the DMT Clearinghouse**

The Data Management Training (DMT) Clearinghouse is a registry for online learning resources focusing on research data management.

It was created in a collaboration between the U.S. Geological Survey's Community for Data Integration, the Earth Sciences Information Partnership (ESIP), and DataONE.

For questions or feedback, please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

[Read More](#)

**Search**  
Find learning resources by keyword, name, date, license and cost  
 [Search](#)

**Browse**  
See a list of learning resources by educational framework  
[Browse](#)

**Submit**  
Submit your learning resources to the Clearinghouse (login required)  
[Submit](#)

**CONNECT WITH US**  
on social networks

**SUBSCRIBE**  
to the monday updates  
 [join](#)

**SEARCH**  
this esip site  
 [search](#)

Questions or Issues with the website? Please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

ESIP is a collaboration among many partner organizations, activities are sponsored by [NASA](#) and [NOAA](#) and managed by the [Foundation for Earth Science](#).

Copyright © 2016 Federation of Earth Science Information Partners. All rights reserved.

# Browse Function

**Browse**

**All Learning Resources**

**DataONE Data Management Module 02: Data Sharing** More ▾ May 2012

  DataONE Education Modules

Analyze Assure Collect Describe Discover Integrate Plan Preserve

**DataONE Data Management Module 03: Data Management Planning** More ▾ May 2012

  DataONE Education Modules Plan

**DataONE Data Management Module 01: Why Data Management** More ▾ May 2012

  DataONE Education Modules

Analyze Assure Collect Describe Discover Integrate Plan Preserve

**DataONE Data Management Module 05: Data Quality Control and Assurance** More ▾ May 2012

  DataONE Education Modules Assure Collect

**DataONE Data Management Module 07: Metadata** More ▾ May 2012

  DataONE Education Modules Describe

**Mozilla Science Lab's Open Data Primers** More ▾

**Mozilla Science Lab Open Data Instructor Guides** More ▾

# Browse Function

**Browse**

Framework

- DataONE Education Modules
- ESIP Data Management for Scientists Short Course
- ICSU - World Data System Training Resources Guide
- The Digital Preservation Network
- USGS Science Support Framework

All Learning Resources

**DataONE Data Management Module 02: Data Sharing** More ▾ May 2012

  DataONE Education Modules

Analyze Assure Collect Describe Discover Integrate Plan Preserve

**DataONE Data Management Module 03: Data Management Planning** More ▾ May 2012

  DataONE Education Modules Plan

**DataONE Data Management Module 01: Why Data Management** More ▾ May 2012

  DataONE Education Modules

Analyze Assure Collect Describe Discover Integrate Plan Preserve

**DataONE Data Management Module 05: Data Quality Control and Assurance** More ▾ May 2012

  DataONE Education Modules Assure Collect

**DataONE Data Management Module 07: Metadata** More ▾ May 2012

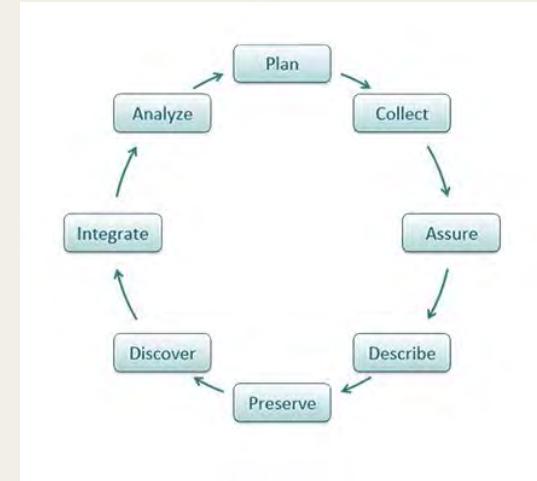
  DataONE Education Modules Describe

**Mozilla Science Lab's Open Data Primers** More ▾

**Mozilla Science Lab Open Data Instructor Guides** More ▾

# What is an educational framework?

- An [educational] framework is a plan or set of steps that defines or collects the content using clear, definable standards about what the student should know and understand.
- For purposes of the DMT Clearinghouse, a given learning resource may be associated with a community-defined standard for data management.
- For example, the DataONE framework represents the DataONE's "*Data Life Cycle*".



<https://www.dataone.org/data-life-cycle>

The DataONE data life cycle was developed ...in collaboration with the broader DataONE community ... [and] serves as an underlying framework for the development of tools, services and education materials by DataONE.

# Browse Function

From ~250 'published' resources → ~10

## Browse

- Framework
- DataONE Education Modules
- ESIP Data Management for Scientists Short Course (33)
- FAIR Data Principles (12)
- ICSU - World Data System Training Resources Guide (76)
- The Digital Preservation Network (7)
- USGS Science Support Framework (19)

### Steps for DataONE Education Modules

- Analyze (3)
- Assure (4)
- Collect (4)
- Describe (4)
- Discover (3)
- Integrate (3)
- Plan (3)
- Preserve (5)

## DataONE Education Modules

- DataONE Data Management Module 01: Why Data Management**  
More ▾  
May 2012  
PUBLIC DOMAIN DataONE Education Modules
- DataONE Data Management Module 02: Data Sharing**  
More ▾  
May 2012  
PUBLIC DOMAIN DataONE Education Modules
- DataONE Data Management Module 03: Data Management Planning**  
More ▾  
May 2012  
PUBLIC DOMAIN DataONE Education Modules
- DataONE Data Management Module 04: Data Entry and Manipulation**  
More ▾  
May 2012  
PUBLIC DOMAIN DataONE Education Modules
- DataONE Data Management Module 05: Data Quality Control and Assurance**  
More ▾  
May 2012  
PUBLIC DOMAIN DataONE Education Modules
- DataONE Data Management Module 06: Data Protection and Backups**  
More ▾  
May 2012  
PUBLIC DOMAIN DataONE Education Modules

# Browse Function -- you found one that looks useful! What next?

You can either click on the Title or the More button to get more info

Resource Title	Publication Date
Open Teacher Information Management Courses	
Essentials 4 Data Support	August 2017
Open Teacher Data Management Courses	
ISRIC Spring School	
ISRIC - World Soil Information Educational Videos	
NASA Earthdata Webinar Series	
ORNL DAAC Data Management Workshops	
Environmental Data Management Best Practices Part 2: Geospatial Data	May 2014

# Browse Function -- you found one that looks useful from the brief description!

## What next?

### ORNL DAAC Data Management Workshops

More ▾

Educational workshops on various scientific data management best practices designed to (1) introduce new data collectors to best practices in data curation and (2) enhance the skillsets of experienced data providers. New workshops are added as they are made available.

[View full description](#)

[View resource ↗](#)



ICSU - World Data System Training Resources Guide

# Browse Function -- you want to see more about this one! What next?

You can either ...

Click on the ***View full description*** to look at the full metadata ... or the ***View resource*** button to go directly to the “landing page” of the resource.

The screenshot shows a section of a website titled "ORNL DAAC Data Management Workshops". Below the title, there is a brief description of the workshops: "Educational workshops on various scientific data management best practices designed to (1) introduce new data collectors to best practices in data curation and (2) enhance the skillsets of experienced data providers. New workshops are added as they are made available." At the bottom of this text block are two buttons: a black one labeled "View full description" and a blue one labeled "View resource". A red star is placed over the "View full description" button, and another red star is placed over the "View resource" button. At the very bottom of the screenshot, there are three small icons: a document, a public domain logo, and the ICSU - World Data System Training Resources Guide logo.

# Search Function

dmclearinghouse.esipfed.org

ESIP Data Management Training

Home Browse Search Help About Log in

Welcome to the DMT Clearinghouse

The Data Management Training (DMT) Clearinghouse is a registry for online learning resources focusing on research data management.

It was created in a collaboration between the U.S. Geological Survey's Community for Data Integration, the Earth Sciences Information Partnership (ESIP), and DataONE.

For questions or feedback, please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

[Read More](#)

DOC Lifecycle - <http://www.dcc.ac.uk/resources/curation-lifecycle-model>

Search

Find learning resources by keyword, name, date, license and cost

Your search terms  Search

Browse

See a list of learning resources by educational framework

[Browse](#)

Submit

Submit your learning resources to the Clearinghouse (login required)

[Submit](#)

CONNECT WITH US  
on social networks

SUBSCRIBE  
to the monday updates

your email [join](#)

SEARCH  
this esip site

[search](#)

Questions or issues with the website? Please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

ESIP is a collaboration among many partner organizations, activities are sponsored by NASA and NOAA and managed by the Foundation for Earth Science.

Copyright © 2016 Federation of Earth Science Information Partners. All rights reserved.

# Search Function: 3 approaches

You filter  
by  
pre-set  
categories  
("filters")



Your search terms

You enter your own  
search terms

**Filter**

**Framework**

- ESIP Data Management for Scientists Short Course (33)
- USGS Science Support Framework (18)
- DataONE Education Modules (10)
- ICSU - World Data System Training Resources Guide (8)
- The Digital Preservation Network (7)

**Keywords**

- Data management (32)
- Data sharing (27)
- Data management planning (23)
- Data preservation (15)
- Data life cycle (11)
- Data citation (10)
- Data reuse - Core Trustworthy Data Repositories Requirements (10)
- Community standards (9)
- Data access methods (9)
- Metadata (9)

Show more

**Organizations**

- Federation of Earth Science Information Partners (ESIP Federation) (34)
- U.S. Geological Survey (16)
- DataONE (10)

**Your search** 104 results

**DataONE Data Management Module 02: Data Sharing** May 2012

**DataONE Data Management Module 03: Data Management Planning** May 2012

**DataONE Data Management Module 01: Why Data Management** May 2012

**DataONE Data Management Module 05: Data Quality Control and Assurance** May 2012

**DataONE Data Management Module 07: Metadata** May 2012

Mozilla Science Lab's Open Data Primers More ▾

# Search Function: Approach 3: Combining Approach 1 & Approach 2 ...



# Search Function: Approach 1 = You search using your own terms

**Filter**

**Framework**

- ICSU - World Data System
- Training Resources Guide (76)
- FAIR Data Principles (70)
- ESIP Data Management for Scientists Short Course (33)
- USGS Science Support Framework (19)
- DataONE Education Modules (10)
- The Digital Preservation Network (7)

**Keywords**

- Data management (88)
- Data management planning (63)
- Data sharing (59)
- Data access (54)
- Data reuse - Core Trustworthy Data Repositories Requirements (44)
- Data publication (36)
- Accessible data - FAIR Data Principle (33)
- Open data (33)
- Data collection (32)
- Data preservation (31)

Show more

**Organizations**

- Federation of Earth Science Information Partners (ESIP Federation) (34)
- Facilitate Open Science

Your search terms

Search

Your search

250 results

Clear all

Creating Documentation and Metadata: Creating a Citation for Your Data

More

September 2012

Local Data Management - Data Formats: Using Self-describing Data Formats

More

January 2013

Responsible Data Use: Data Restrictions

More

February 2013

Preserving the Scientific Record: Establishing Relationships with Archives

More

October 2012

Preserving the Scientific Record: Preserving a Record of Environmental Change

More

October 2012

USGS Data Management Training Modules—Metadata for Research Data

More

# Search Function: Approach 2 = You start by checking boxes within the pre-set categories (“filters”)

The screenshot shows a search interface with a sidebar on the left containing filter categories and a main search results area on the right.

**Filter Sidebar:**

- Framework**
  - ICSU - World Data System
  - Training Resources Guide (76)
  - FAIR Data Principles (70)
  - ESIP Data Management for Scientists Short Course (33)
  - USGS Science Support Framework (19)
  - DataONE Education Modules (10)
  - The Digital Preservation Network (7)
- Keywords**
  - Data management (88)
  - Data management planning (63)
  - Data sharing (59)
  - Data access (54)
  - Data reuse - Core Trustworthy Data Repositories Requirements (44)
  - Data publication (36)
  - Accessible data - FAIR Data Principle (33)
  - Open data (33)
  - Data collection (32)
  - Data preservation (31)
- Show more**
- Organizations**
  - Federation of Earth Science Information Partners (ESIP Federation) (34)
  - Facilitate Open Science

**Search Results Area:**

Your search terms:  Search

**Your search** 250 results Clear all

Creating Documentation and Metadata: Creating a Citation for Your Data More ▾  
[CC BY] ESIP Data Management for Scientists Short Course September 2012

Local Data Management - Data Formats: Using Self-describing Data Formats More ▾  
[CC BY] ESIP Data Management for Scientists Short Course January 2013

Responsible Data Use: Data Restrictions More ▾  
[CC BY] ESIP Data Management for Scientists Short Course February 2013

Preserving the Scientific Record: Establishing Relationships with Archives More ▾  
[CC BY] ESIP Data Management for Scientists Short Course October 2012

Preserving the Scientific Record: Preserving a Record of Environmental Change More ▾  
[CC BY] ESIP Data Management for Scientists Short Course October 2012

USGS Data Management Training Modules—Metadata for Research Data More ▾  
[CC BY] USGS Science Support Framework

# Built-in Search Filters

**Filter**

**Framework** (circled)  
■ ESIP Data Management for Scientists Short Course (33)  
■ USGS Science Support Framework (18)  
■ DataONE Education Modules (10)  
■ ICSU - World Data System Training Resources Guide (8)  
■ The Digital Preservation Network (7)

**Keywords** (circled)  
■ Data management (32)  
■ Data sharing (27)  
■ Data management planning (23)  
■ Data preservation (15)  
■ Data life cycle (11)  
■ Data citation (10)  
■ Data reuse - Core Trustworthy Data Repositories Requirements (10)  
■ Community standards (9)  
■ Data access methods (9)  
■ Metadata (9)  
[Show more](#)

**Organizations** (circled)  
■ Federation of Earth Science Information Partners (ESIP Federation) (34)  
■ U.S. Geological Survey (16)  
■ DataONE (10)

**People** (circled)  
■ Nancy J. Hoebelheinrich (33)  
■ Ruth E. Duerr (33)  
■ Matthew Mayernik (9)  
■ Robert R. Downs (7)  
■ Robert Cook (5)  
■ Chen Chiu (2)  
■ Curt Tilmes (2)  
■ Jason Kudulis (2)  
■ Kathy Martinolich (2)  
■ Lola Olsen (2)  
■ Mimi Tzeng (2)  
■ Suresh K.S. Vannan (2)  
■ Tyler Stevens (2)  
■ Zannah Marsh (2)  
■ Ben Wheeler (1)  
■ Dave Fearon (1)  
■ Dorothea Salo (1)  
■ Drew Ignizio (1)  
■ Elliot Metsger (1)  
■ Emily Fort (1)  
[Show more](#)

**Publication Date** (circled)  
■ 2012 (31)  
■ 2017 (15)  
■ 2013 (13)  
■ 2015 (9)  
■ 2016 (5)  
■ 2014 (2)  
■ 2010 (1)

**License** (circled)  
■ Creative Commons Attribution 3.0 United States - CC BY 3.0 US (35)  
■ Creative Commons 0 - CC0 "No Rights Reserved" (Public Domain) (31)  
■ Creative Commons Attribution 4.0 International - CC BY 4.0 (15)  
■ Creative Commons Attribution-NonCommercial 4.0 International - CC BY-NC 4.0 (5)  
■ Creative Commons 1.0 Universal (Public Domain Dedication) (3)  
■ Creative Commons Attribution-ShareAlike 4.0 International License - CC BY-SA 4.0 (3)  
■ Creative Commons Attribution 3.0 Unported - CC BY 3.0 (1)

**Cost** (circled)  
■ No fee (101)  
■ Fee (3)

# Demonstrating Approach 3: Start by entering your own search term...

A screenshot of a search interface. At the top, there is a search bar containing the text "sharing practices". To the right of the search bar is a blue "Search" button. Below the search bar, a green header bar displays the text "Your search sharing practices" and "89 results". A red oval highlights the search term "sharing practices" in the search bar. Another red oval highlights the number "89 results" in the green header. The word "Yikes!" is written in blue text above the results area.

... then limiting by Keyword filter...

A screenshot of a search interface. At the top, there is a search bar containing the text "sharing practices". To the right of the search bar is a blue "Search" button. Below the search bar, a green header bar displays the text "Your search sharing practices" and "26 results". A red oval highlights the search term "sharing practices" in the search bar. Another red oval highlights the number "26 results" in the green header. The word "More precise!" is written in blue text above the results area. On the left side of the interface, there is a sidebar with a "Framework" section listing various resources and a "Keywords" section with a list of checked items. The checked items in the "Keywords" section include "Data sharing" (with a checked checkbox), "Data management (17)" (with an unchecked checkbox), "Data management planning (12)" (with an unchecked checkbox), and "Data access (11)" (with an unchecked checkbox). The main content area shows a list of search results. The first result is titled "Best practices for preparing data to share and preserve" and includes a brief description and a "More" link. The second result is titled "Research Data Management and Sharing MOOC" and includes a brief description and a "More" link. Both results have a "USGS Science Support Framework" badge and a "November 2015" date stamp.

# Demonstrating Search: finally, limiting even more by the Framework filter...

Filter

sharing practices

Still more precise! Search

2 results Clear all

Your search sharing practices

Framework - A community-based organization plan or set of steps for education or training DataONE Education Modules [X]

Keywords - short phrases describing what the learning resource is about Data sharing [X]

DataONE Data Management Module 02: Data Sharing More ▾

When first sharing research data, researchers often raise questions about the value, benefits, and mechanisms for sharing. Many stakeholders and interested parties, such as funding agencies, ... data, concerns about sharing data, and methods and best practices for sharing data and includes a downloadable presentation (PPT or ...)

PUBLIC DOMAIN DataONE Education Modules FAIR Data Principles May 2012

DataONE Data Management Module 08: Data Citation More ▾

... the process for obtaining one, and to summarize best practices for supporting data citation. This 30-40 minute lesson includes a ...

PUBLIC DOMAIN DataONE Education Modules FAIR Data Principles May 2012

Framework

- (-) DataONE Education Modules
- ICSU - World Data System
- Training Resources Guide (11)
- FAIR Data Principles (10)
- ESIP Data Management for Scientists Short Course (2)
- USGS Science Support Framework (1)

DataONE Framework Steps

- Describe (2)
- Discover (2)
- Preserve (2)
- Analyze (1)
- Assure (1)
- Collect (1)
- Integrate (1)
- Plan (1)

Keywords

- Data sharing

# Demonstrating Search: however, if too precise, you can either *Clear all* to start over ... or uncheck filters / facets

Filter

sharing practices Search

Your search sharing practices 2 results

**Framework - A community-based organization plan or set of steps for education or training** DataONE Education Modules [X] Clear all

**Keywords - short phrases describing what the learning resource is about** Data sharing [X]

**DataONE Data Management Module 02: Data Sharing** More ▾

When first sharing research data, researchers often raise questions about the value, benefits, and mechanisms for sharing. Many stakeholders and interested parties, such as funding agencies, ... data, concerns about sharing data, and methods and best practices for sharing data and includes a downloadable presentation (PPT or ...)

May 2012

**DataONE Data Management Module 08: Data Citation** More ▾

... the process for obtaining one, and to summarize best practices for supporting data citation. This 30-40 minute lesson includes a ...

May 2012

Keywords

Data sharing

# Full Description – 1 of 2

## ORNL DAAC Data Management Workshops

### Key Info

**URL - the landing page for the learning resource:**

<http://daac.ornl.gov/workshops/workshops.shtml>



**Description - a brief synopsis, abstract or summary of what the learning resource is about:**

Educational workshops on various scientific data management best practices designed to (1) introduce new data collectors to best practices in data curation and (2) enhance the skillsets of experienced data providers. New workshops are added as they are made available.



**Authoring Organization(s) Name:**

NASA ORNL DAAC (Oak Ridge National Laboratory Distributed Active Archive Center)



**License - link to legal statement specifying the copyright status of the learning resource:**

Creative Commons 1.0 Universal (Public Domain Dedication)

**Access Cost:**

No fee



**Primary language(s) in which the learning resource was originally published or made available:**

English

### More info about

**Keywords - short phrases describing what the learning resource is about:**

Appraisal - Core Trustworthy Data Repositories Requirements

Data discovery and identification - Core Trustworthy Data Repositories Requirements

Data integrity and authenticity - Core Trustworthy Data Repositories Requirements

Data quality - Core Trustworthy Data Repositories Requirements

Data reuse - Core Trustworthy Data Repositories Requirements

Expert guidance - Core Trustworthy Data Repositories Requirements

Preservation plan - Core Trustworthy Data Repositories Requirements



# Full Description – 2 of 2

**Subject Discipline - subject domain(s) toward which the learning resource is targeted:**

Engineering: Aerospace Engineering

Physical Sciences and Mathematics: Earth Sciences

Physical Sciences and Mathematics: Environmental Sciences

**Publisher - organization credited with publishing or broadcasting the learning resource:**

NASA ORNL DAAC (Oak Ridge National Laboratory Distributed Active Archive Center)

**Media Type - designation of the form in which the content of the learning resource is represented, e.g., moving image:**

Collection - a group or set of items that comprise a single learning resource, e.g., a PDF version of a slide presentation, an audio file of the presentation and a textual representation of the oral transcription of the presentation.

**Contributor Organization(s):****Name:**

ICSU - World Data System (WDS)

**Type:**

Endorser

**Contact Organization(s):**

NASA ORNL DAAC (Oak Ridge National Laboratory Distributed Active Archive Center)

## Educational Info

**Purpose - primary educational reason for which the learning resource was created:**

Professional Development - increasing knowledge and capabilities related to managing the data produced, used or re-used, curated and/or archived.

**Learning Resource Type - category of the learning resource from the point of view of a professional educator:**

Lesson - detailed description of an element of instruction in a course, contained in a unit of one or more lessons, and used by a teacher to guide class instruction.

**Target Audience - intended audience for which the learning resource was created:**

Data supporter

Early-career Research Scientist

Research Scientist

**Intended time to complete - approximate amount of time the average student will take to complete the learning resource:**

More than 1 hour (but less than 1 day)

**Framework - A community-based organization plan or set of steps for education or training:**

ICSU - World Data System Training Resources Guide

# Submit Function, briefly

 **ESIP** Data Management Training

Home Browse Search **Submit** Help ▾ About  Log in

## Welcome to the DMT Clearinghouse

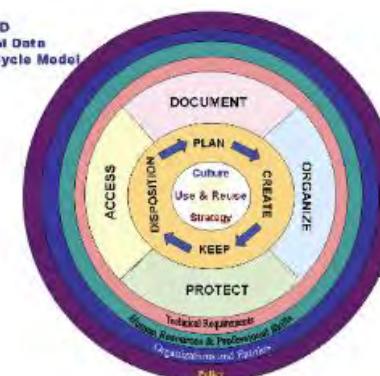
The Data Management Training (DMT) Clearinghouse is a registry for online learning resources focusing on research data management.

It was created in a collaboration between the U.S. Geological Survey's Community for Data Integration, the Earth Sciences Information Partnership (ESIP), and DataONE.

For questions or feedback, please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

[Read More](#)

IWGDD  
Digital Data  
Life Cycle Model



The diagram illustrates the IWGDD Digital Data Life Cycle Model as a circular process. The outer ring is divided into four quadrants: ACCESS (top-left), DOCUMENT (top-right), PROTECT (bottom-right), and DISPOSITION (bottom-left). The inner circle contains the word "CULTURE" at the top and "USE & REUSE STRATEGY" below it. Arrows indicate a clockwise flow from DISPOSITION through PROTECT, then moving inward to DOCUMENT, then outward to ACCESS, and finally back to DISPOSITION. At the bottom of the inner circle, there is a reference to "Technical Requirements, Business Requirements & Professional Practice".

IWGDD Lifecycle - [https://www.nitrd.gov/About/Harnessing\\_Power\\_Web.pdf](https://www.nitrd.gov/About/Harnessing_Power_Web.pdf)

## Search

Find learning resources by keyword, name, date, license and cost

sharing practices

Search

## Browse

See a list of learning resources by educational framework

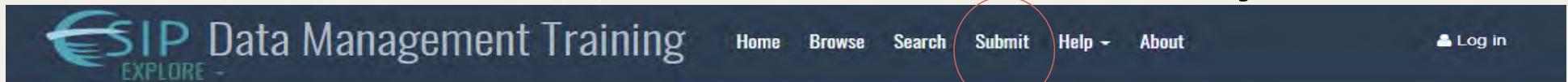
Browse

## Submit

Submit your learning resources to the Clearinghouse

Submit

# Submit Function, briefly



## Submit new Learning Resource

Your learning resource suggestion will be submitted to the Clearinghouse, but will not be published immediately as the information needs to be reviewed for quality control and relevancy.

If you would like to submit more information, please log in and return to this page. If you don't yet have an ESIP account, you can create one by clicking "Log in" above, then "I want to create an account."

From the Help pages, you can also find out more about how and what kind of information to submit. We do require that you give us your name and email address if you submit without having a user login in case reviewers or editors have questions. Rest assured that your contact information will not be shared publicly without your permission.

Thank you for your interest in making data management training resources widely available!

**Title \***

**URL - the landing page for the learning resource**

**Access Cost \***

- No fee
- Fee

**Submission Contact Name \***

**Submission Contact Email Address \***

Address will not be shared without permission.

**CAPTCHA**

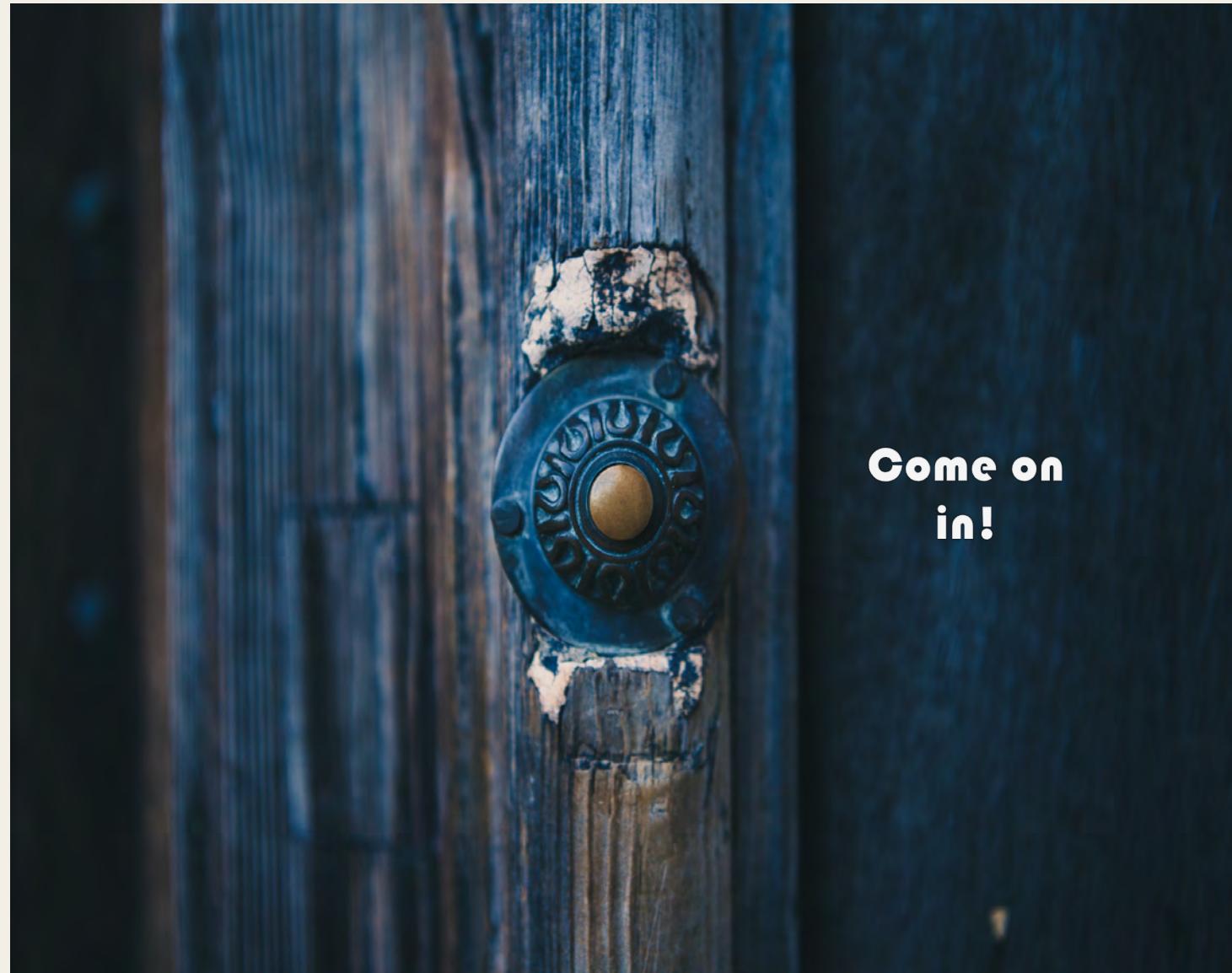
This question is for testing whether you are a human visitor and to prevent automated spam submissions.

 I'm not a robot

Privacy • Terms

# As a community supported resource we'd love to have you to join us by...

- Submitting your learning resources
- Joining our Working Groups on
  - Assessment Framework
  - Metadata Enhancement
  - Content Diversification
- Editorial assistance
- Usability testing
- Spreading the word
- Jumping in on our crowdsourcing events



# Join us!

Join the ESIP Research Data Management Cluster at: [esip\\_dmtraining@lists.esipfed.org](mailto:esip_dmtraining@lists.esipfed.org)

Contact:

Nancy Hoebelheinrich ([nhoebel@kmotifs.com](mailto:nhoebel@kmotifs.com)) or  
[clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

 Data Management Training

Home Browse Search Submit Help About Log in

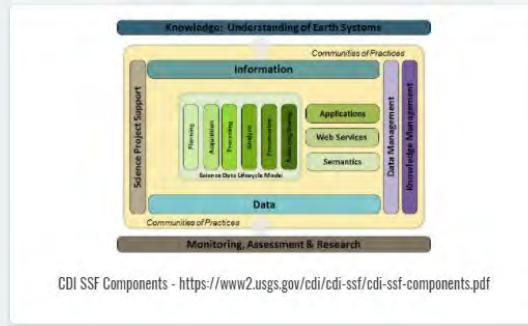
## Welcome to the DMT Clearinghouse

The Data Management Training (DMT) Clearinghouse is a registry for online learning resources focusing on research data management.

It was created in a collaboration between the U.S. Geological Survey's Community for Data Integration, the Earth Sciences Information Partnership (ESIP), and DataONE.

For questions or feedback, please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

[Read More](#)



CDI SSF Components - <https://www2.usgs.gov/cdi/cdi-ssf/cdi-ssf-components.pdf>

### Search

Find learning resources by keyword, name, date, license and cost

### Browse

See a list of learning resources by educational framework

### Submit

Submit your learning resources to the Clearinghouse

#### CONNECT WITH US

on social networks

#### SUBSCRIBE

to the monday updates

your email

#### SEARCH

this esip site

your search terms

Questions or issues with the website? Please contact [clearinghouseEd@esipfed.org](mailto:clearinghouseEd@esipfed.org)

ESIP is a collaboration among many partner organizations, activities are sponsored by [NASA](#) and [NOAA](#) and managed by the [Foundation for Earth Science](#). We organise regular events to appropriate reward and career opportunities for RSFs.

<http://dmtclearinghouse.esipfed.org/>

## Data Management Skillbuilding Hub

- One off lessons
- Host/store here
- Indexed at DMT Clearinghouse

<http://dataoneorg.github.io/Education>

## DMT Clearinghouse

- Index here
- Metadata
- This is a registry

<http://dmtclearinghouse.esipfed.org>



**FALL MEETING**  
Washington, D.C. | 10-14 Dec 2018

*Please take some time and come ask us questions at the help desk*

Megan Mach [mach@unm.edu](mailto:mach@unm.edu)

Nancy Hoebelheinrich [nhoebel@kmotifs.com](mailto:nhoebel@kmotifs.com)

Amber Budden [aebudden@dataone.unm.edu](mailto:aebudden@dataone.unm.edu)

*DataONE Webinar Series*

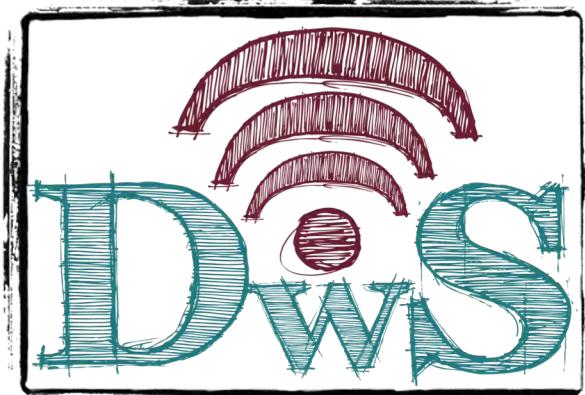
**[www.dataone.org/webinars](http://www.dataone.org/webinars)**

*Upcoming Webinar Event*

**[www.dataone.org/upcoming-webinar](http://www.dataone.org/upcoming-webinar)**

*Previous Webinar Events (Recording and Discussion)*

**[www.dataone.org/previous-webinars](http://www.dataone.org/previous-webinars)**



#DWS2018  
@DataONEorg

## *Upcoming Webinar Event*

***[www.dataone.org/upcoming-webinar](http://www.dataone.org/upcoming-webinar)***

### **Schema.org: Improving access to data through a standardized language**

November 13, 2018

**Bryce Mecum**, Scientific Software Engineer, National Center  
for Ecological Analysis and Synthesis (NCEAS)

**Doug Fils**, Data Management Technical Expert, Consortium  
for Ocean Leadership (COL)

