

# Introduction to Data Management and Publication

By Gabe Kamener  
FCE Information Manager, Florida International University



2026 Florida Coastal Everglades LTER Information Management  
*January 22, 2026*



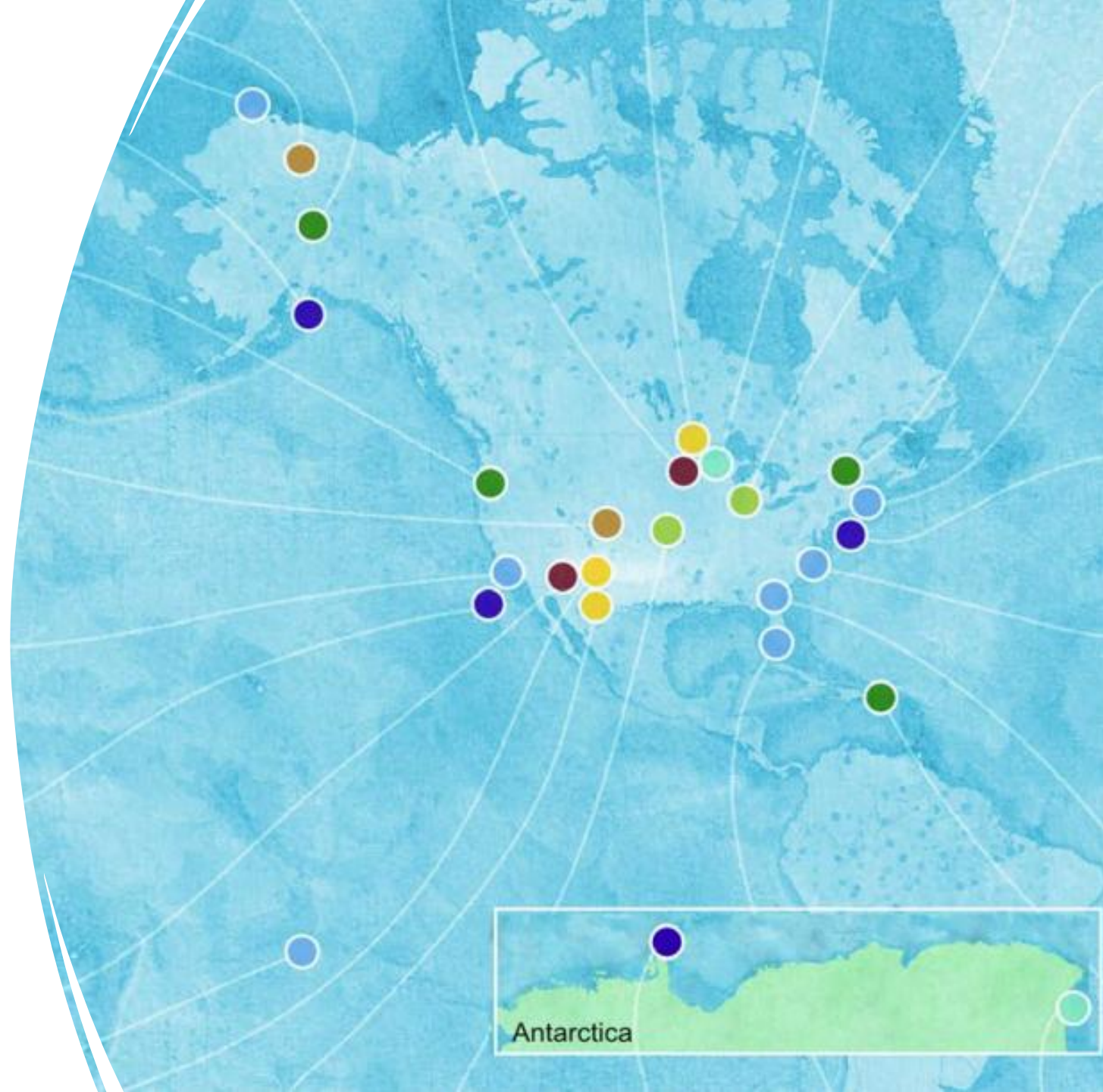
# Overview

- Review
  - Sharing of LTER Network data
  - EDI data repository and FCE data catalog
  - Citing data
- Best practices
  - Data and project management
  - Formatting data
  - Describing metadata
- Preparing for publication with ezEML

# Sharing of LTER Network Data

---

- Ongoing since NSF funded first LTER sites in 1980
- Enables new science!
- Supports open science and reproducibility
- Funders and journal publishers often require datasets in a recognized data repository

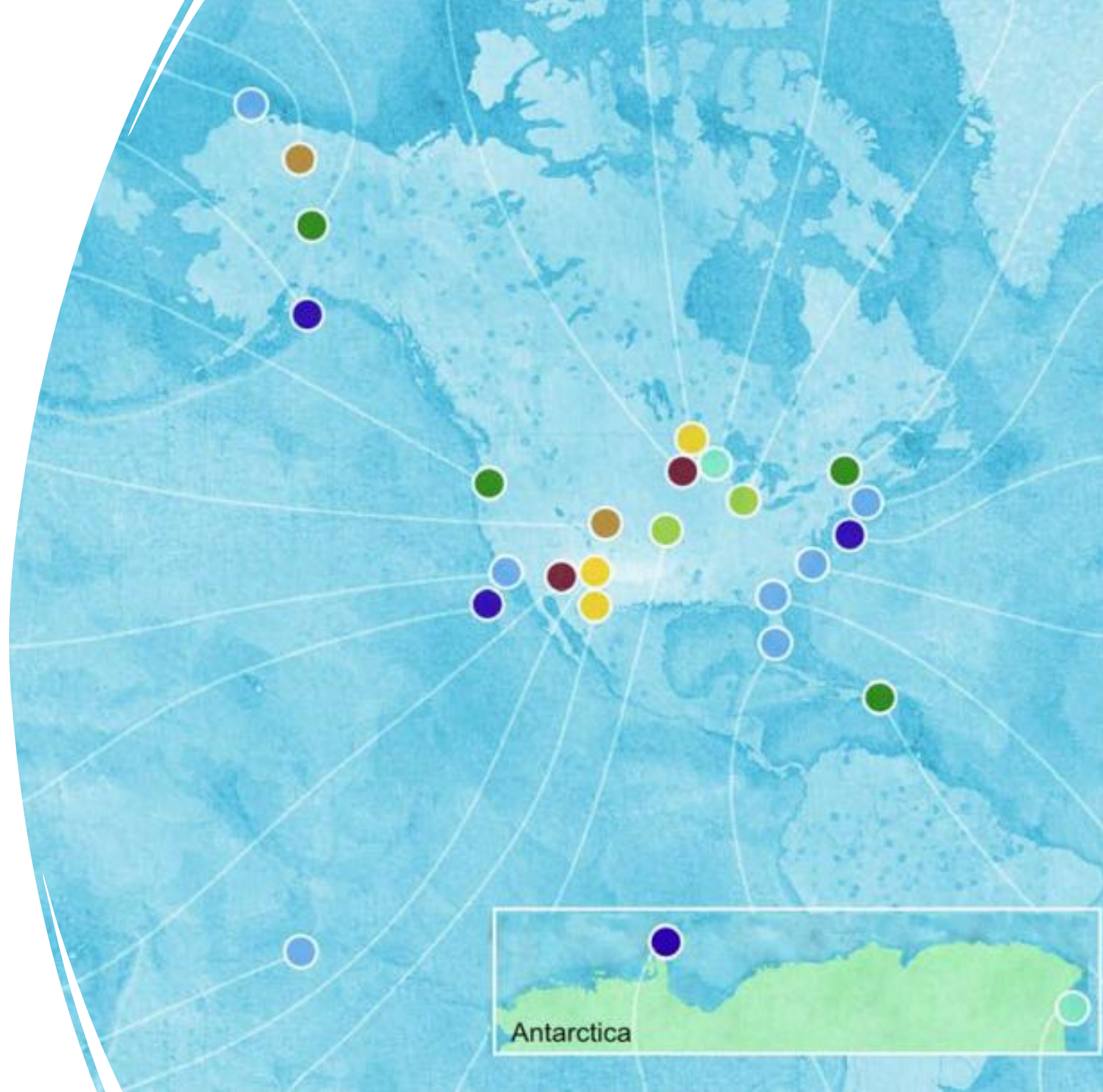




# LTER Network Data Release Policy

---

- Data and information derived from publicly funded research in the U.S. LTER Network, totally or partially from LTER funds from NSF... [must be].. made available in a community accepted data repository... with as few restrictions as possible, on a nondiscriminatory basis.

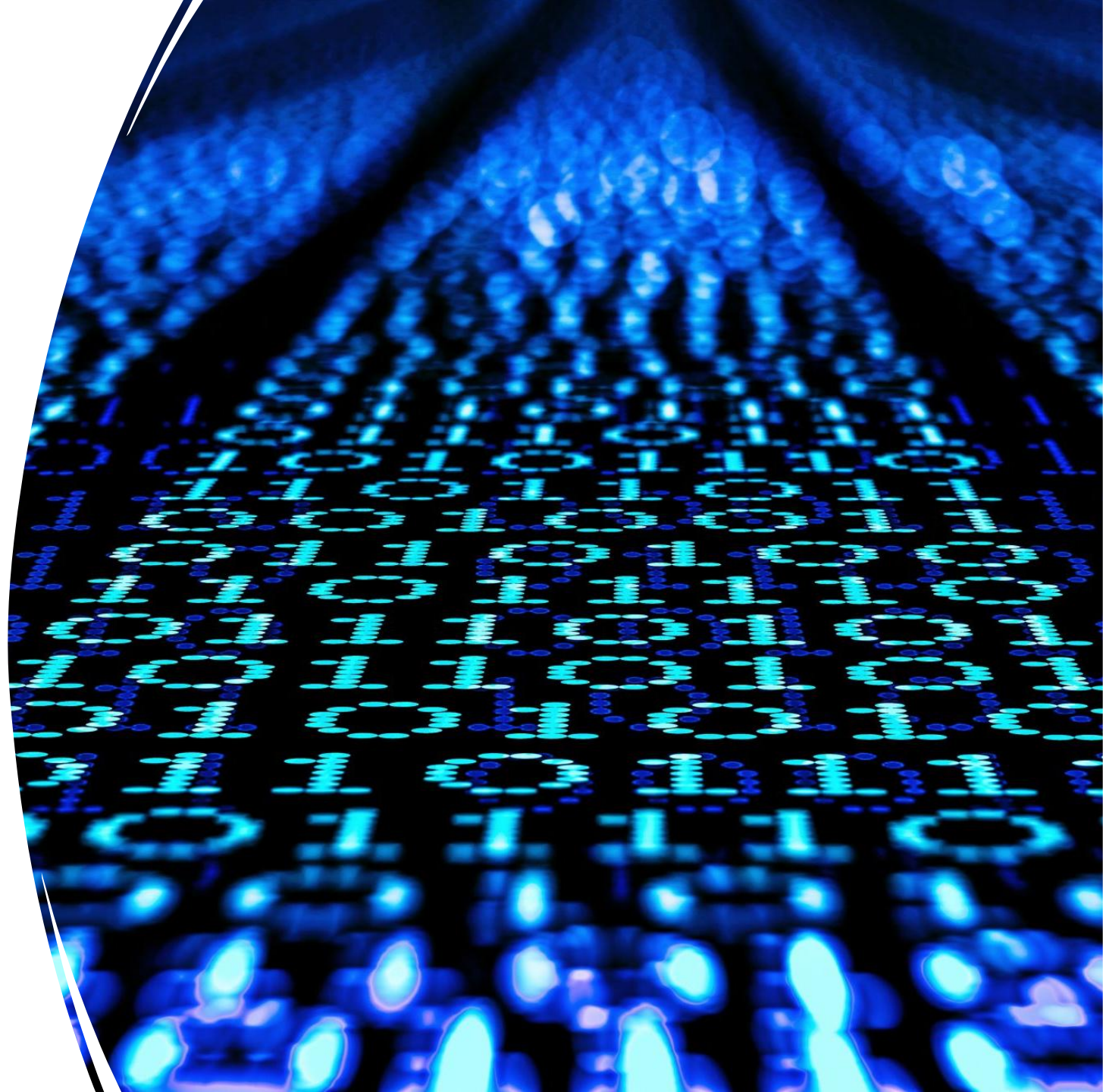




# Two Types of LTER Data

---

- Type I – data are to be released to the general public according to the terms of the general data use agreement **within 2 years from collection** and no later than the publication of the main findings from the dataset
- Type II - data are to be released to restricted audiences according to terms specified by the owners of the data. **Type II data are considered to be exceptional and should be rare in occurrence.**

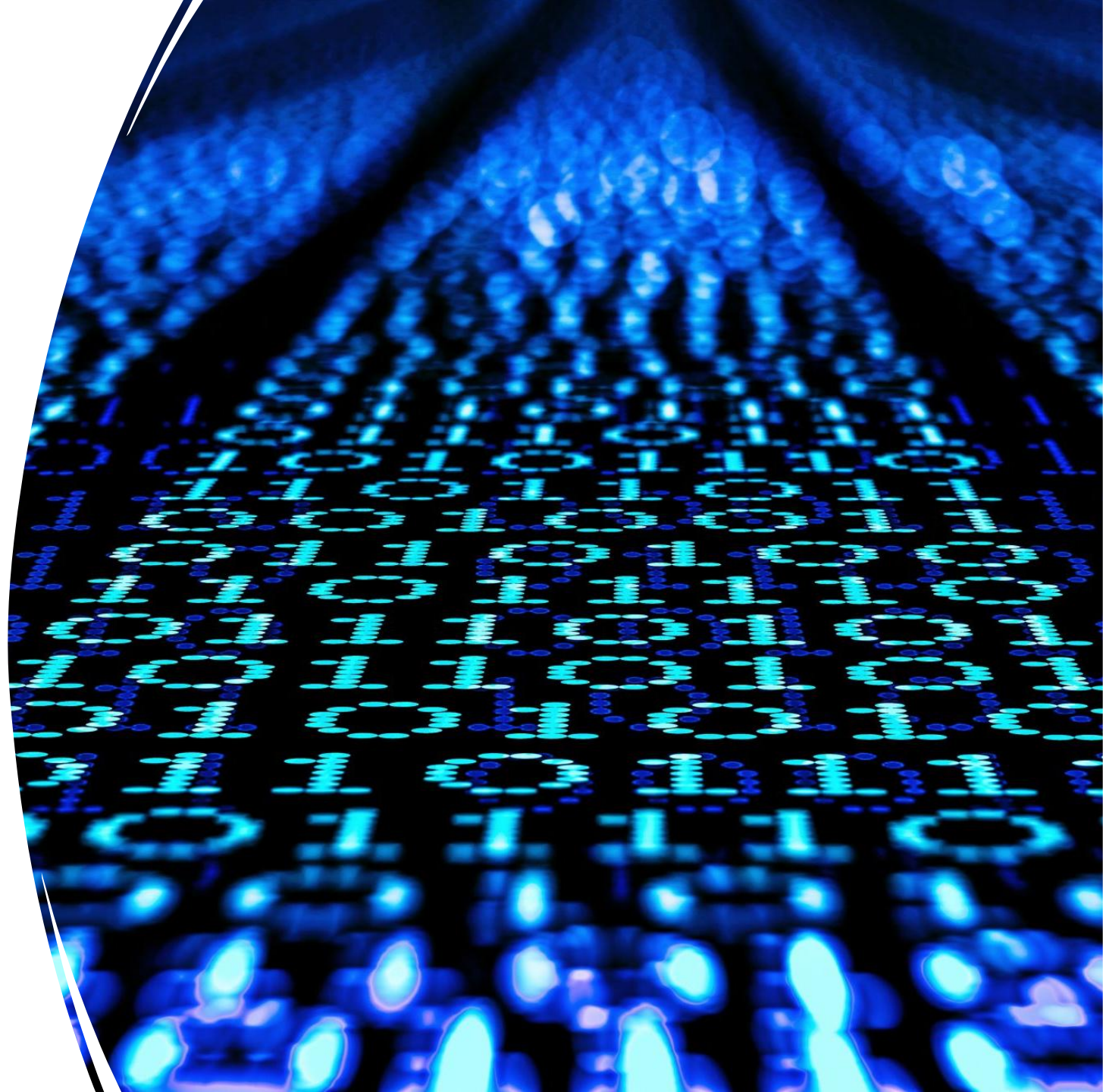




# FCE LTER Graduate Student Data

---

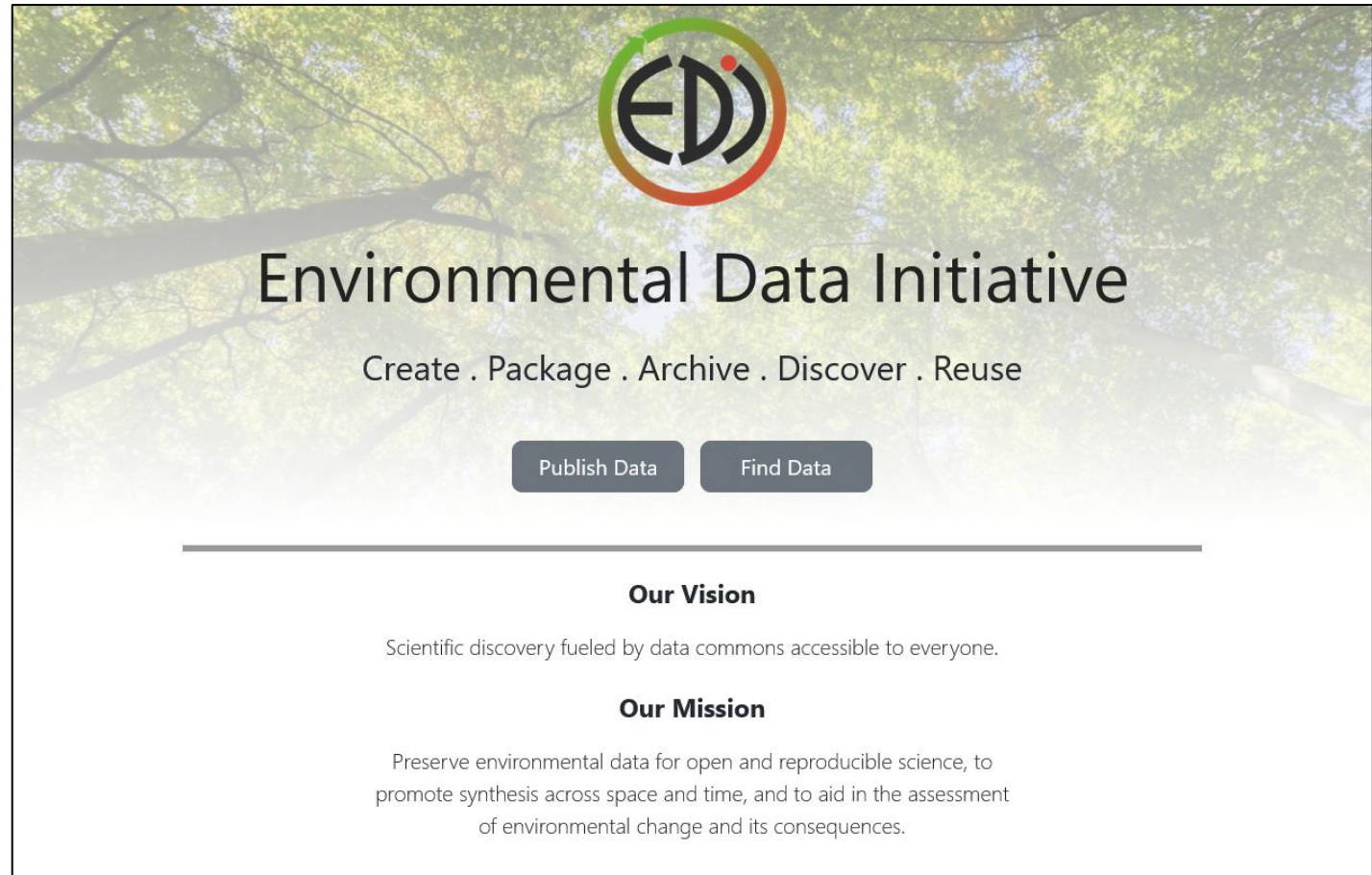
- Submit complete dataset and metadata to FCE Information Manager (IM) before graduation
- Meet with IM **at least 2 months before** you need to publish data



# The Environmental Data Initiative (EDI) Data Repository

---

- Funded by NSF
- Provides:
  - Long-term data security
  - Long-term data accessibility
  - Data integrity
  - Data discovery
  - Citable digital object identifiers (DOIs) for datasets

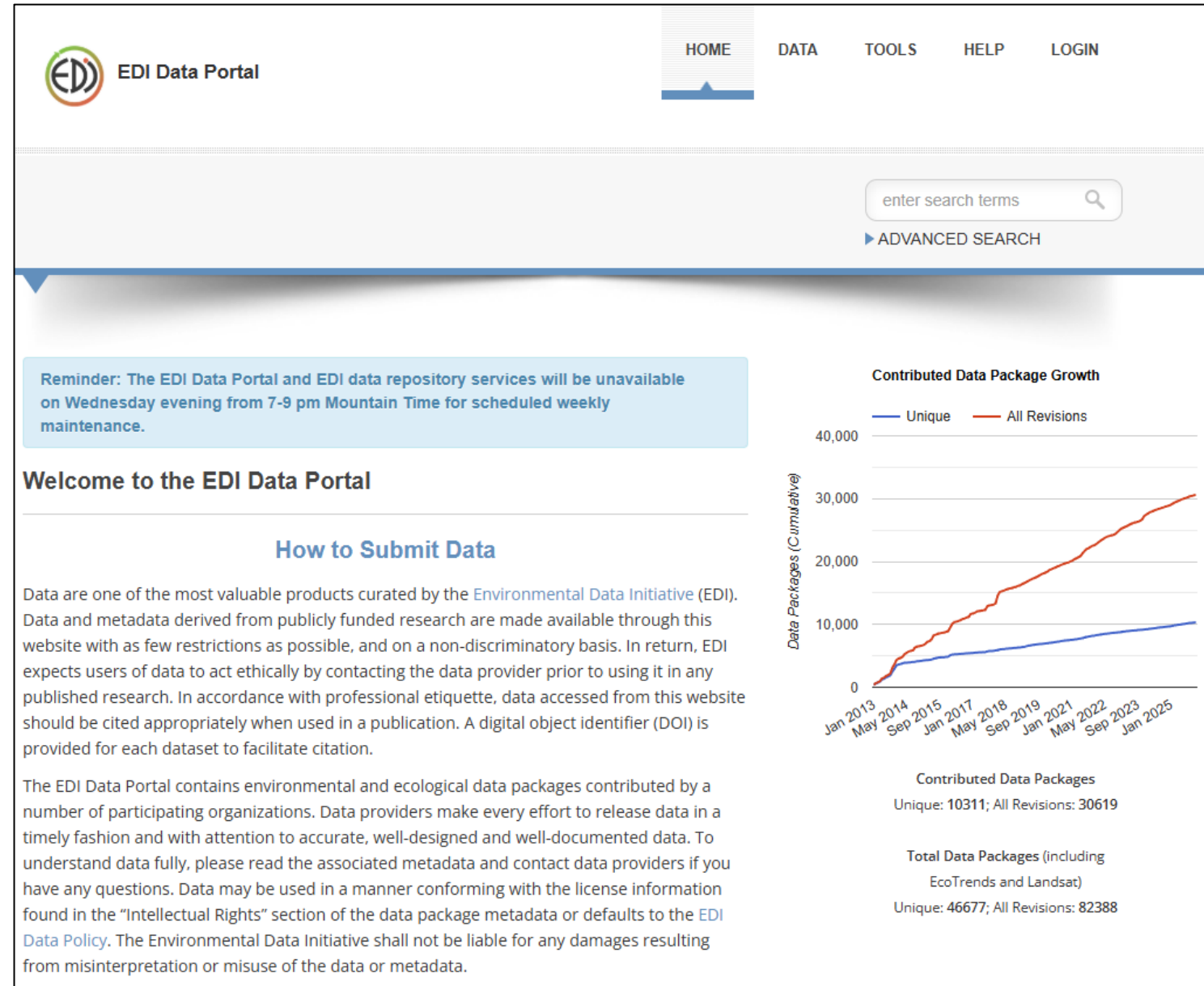


<https://edirepository.org>



# EDI Data Portal

- User-friendly interface of EDI repository
- More than 10,300 searchable, unique data packages
- Advanced search functionality



<https://portal.edirepository.org/nis/home.jsp>



# FCE Data Catalog

- Lists FCE datasets published in EDI
- Easily search the catalog from the FCE website!

[Home](#) / [Data](#) / [Core](#)

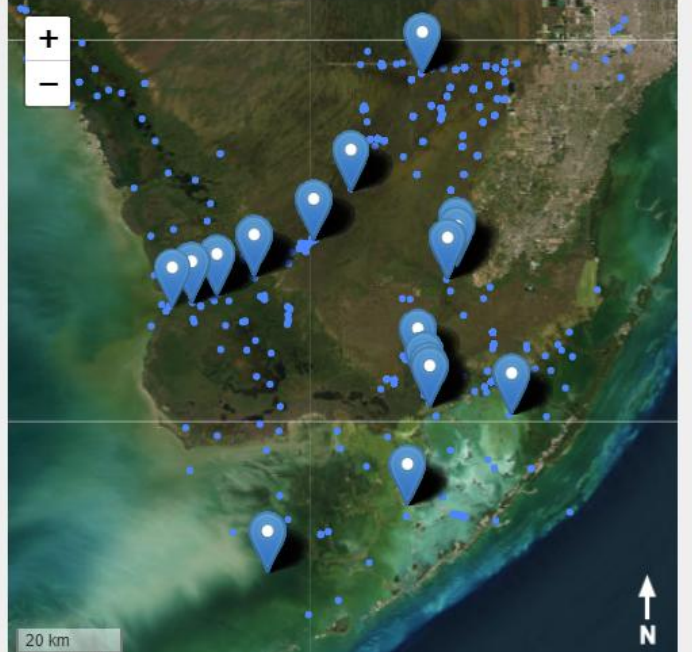
## FCE Data Catalog

### Search FCE Datasets

Keywords

Originator

LTER Core Area  
All core areas



20 km  
10 mi

Leaflet | Powered by Esri | DigitalGlobe, GeoEye, I-cubed, USDA, USGS, AE...

FCE Sites      Satellite Sites

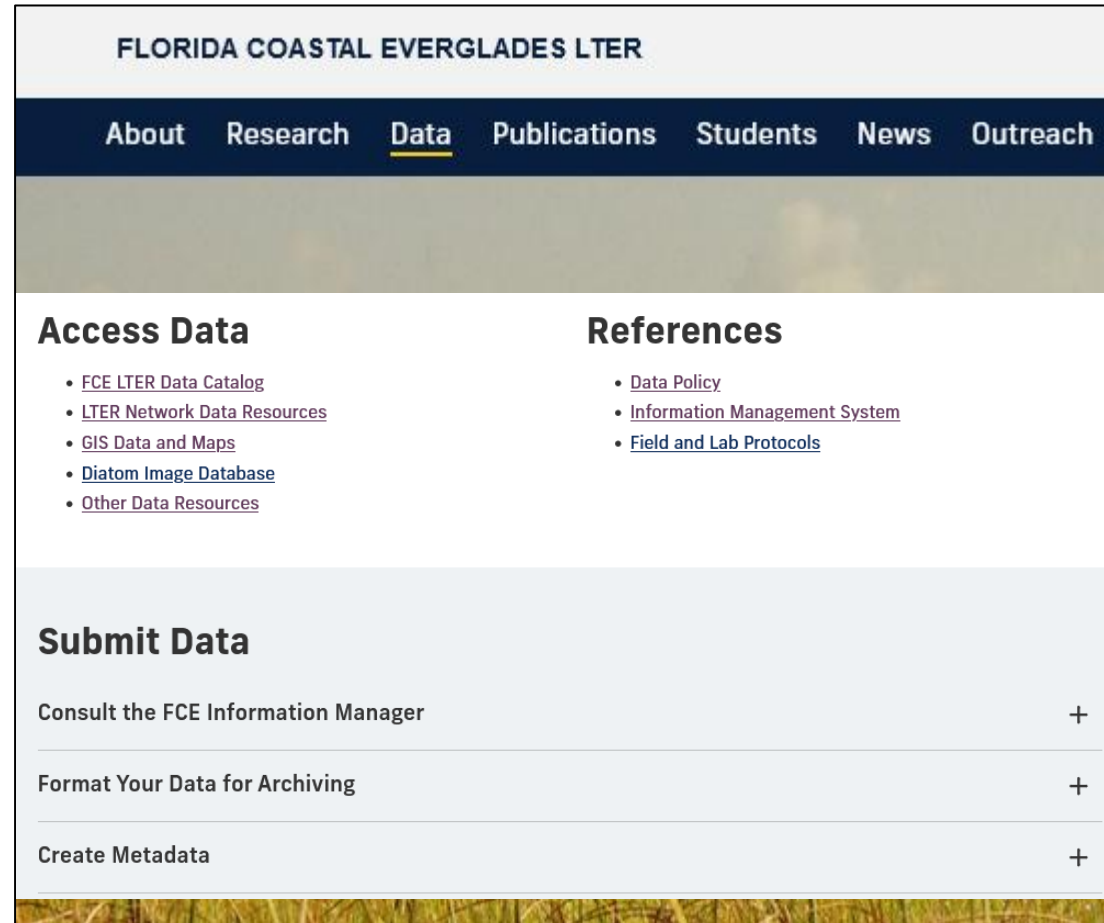
To search for data from specific FCE sites, zoom or pan the map to change the area inside a fixed yellow box, which will be included in the search.

Total number of data sets found = 228

<https://fce-lter.fiu.edu/data/core/index.php>

# Steps to Publish Data in the EDI Repository

1. Review the FCE Data page!
2. Contact the FCE IM about:
  - Data (tabular, model code, imagery, etc.)
  - Required metadata
  - How to format your dataset
  - Getting an FCE dataset ID
3. Enter data and metadata into ezEML
4. Review package with FCE IM and publish!



<https://fcelter.fiu.edu/data>

# Citing Data

Kleindl, P.M., E. Gaiser, A. Wachnicka, J. Sah, and M. Ross. 2024. Macrophyte and microbial mat biomass co-variation along a hydrologic gradient and response to a removal experiment in temporary wetlands Everglades, FL, USA, February 2003 – November 2006 ver 3. Environmental Data Initiative. <https://doi.org/10.6073/pasta/01ef607bc7090c3488e5bcb7475847e9> (Accessed 2026-01-22)

**Title:** Macrophyte and microbial mat biomass co-variation along a hydrologic gradient and response to a removal experiment in temporary wetlands Everglades, FL, USA, February 2003 – November 2006

**Creators:** Kleindl, Paige Marie; Ph.D. Candidate - Periphyton Laboratory; Florida Coastal Everglades LTER  
Gaiser, Evelyn; Principal Investigator; Florida Coastal Everglades LTER  
Wachnicka, Anna; Principal Scientist; South Florida Water Management District...

Show more >

**Publication Date:** 2024-12-04

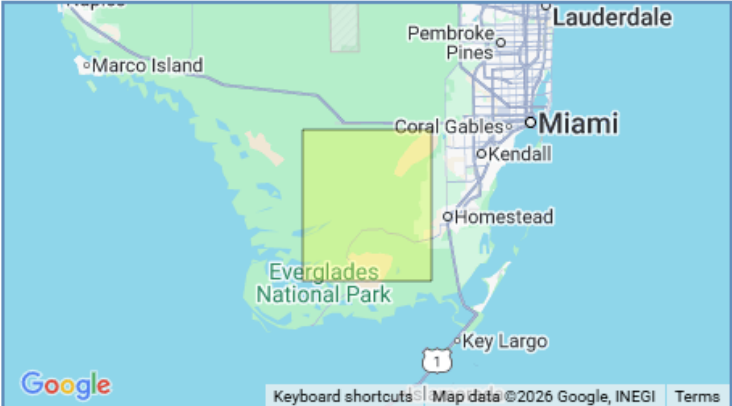
**Citation:** Kleindl, P.M., E. Gaiser, A. Wachnicka, J. Sah, and M. Ross. 2024. Macrophyte and microbial mat biomass co-variation along a hydrologic gradient and response to a removal experiment in temporary wetlands Everglades, FL, USA, February 2003 – November 2006 ver 3. Environmental Data Initiative. <https://doi.org/10.6073/pasta/01ef607bc7090c3488e5bcb7475847e9> (Accessed 2026-01-22).

Copy Citation

**Abstract:** This data package encompasses hydrologic variables, soil depth, hydrologically-regulated macrophyte community types, macrophyte biomass and community structure, and microbial mat biomass that was collected in two observational surveys and one in-situ experimental manipulation in six temporary wetland regions located in the...

Show more >

**Spatial Coverage:**





# Citing Data

- Include full citations (with DOI) in References and then cite those references in Data Availability Statement (also include DOI links there)

## DATA AVAILABILITY STATEMENT

Data and R code (Kleindl et al., 2024) are available from the Environmental Data Initiative (EDI) Data Portal: <https://doi.org/10.6073/pasta/01ef607bc7090c3488e5bcb7475847e9>.

### Supporting Information

### REFERENCES

Adams, A. E., E. M. Besozzi, G. Shahrokhi, and M. A. Patten. 2022. "A Case for Associational Resistance: Apparent Support for the Stress Gradient Hypothesis Varies with Study System." *Ecology Letters* 25(1): 202–217. <https://doi.org/10.1111/ele.13917>.

[View](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#) | [find it@FIU](#)

Kleindl, P. 2025b. "Figure 8: Structural equation models (SEMs) for microbial mat biomass in the a) census and b) transect surveys." <https://BioRender.com/o08j359>.

[Google Scholar](#) | [find it@FIU](#)

Kleindl, P. M., E. Gaiser, A. Wachnicka, J. Sah, and M. Ross. 2024. "Macrophyte and Microbial Mat Biomass Co-Variation along a Hydrologic Gradient and Response to a Removal Experiment in Temporary Wetlands Everglades, FL, USA, February 2003–November 2006 ver 3." Environmental Data Initiative. <https://doi.org/10.6073/pasta/01ef607bc7090>

[View](#) | [Google Scholar](#) | [find it@FIU](#)

Citation use in Kleindl et al. (2025). <https://doi.org/10.1002/ecs2.70384>

# Best Practices in Data Management

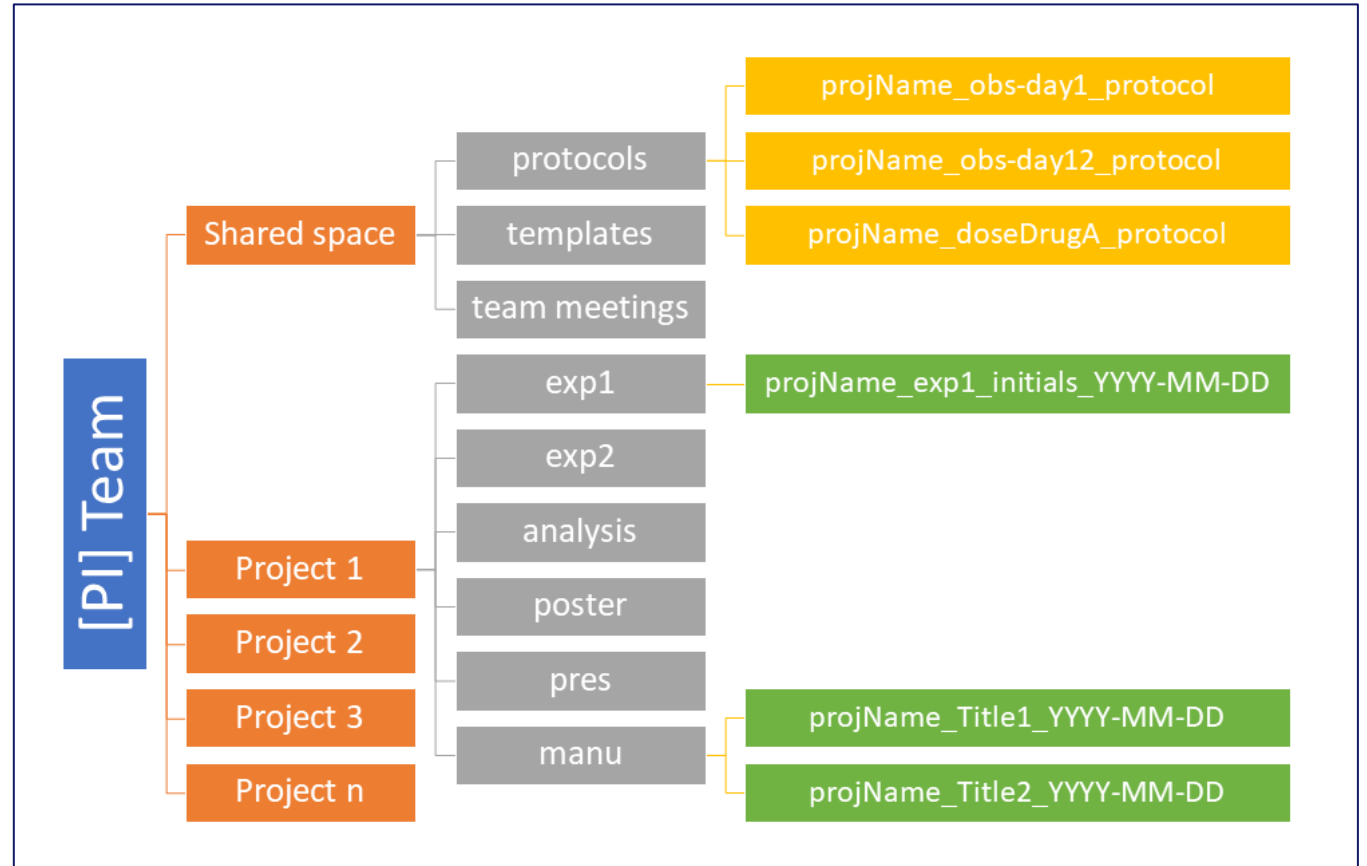
“Well managed data is a benefit to any researcher as it requires less digging to find, less effort to understand, and less processing to prepare for collaboration, reuse, and sharing.” -Briney et al. 2020.



Artwork by @allison\_horst (CC BY 4.0)

# File Organization

- Use folder structures
- Use consistent file names
- Keep raw data separate from analysis
- Use file versioning





# Backup Your Data

- 3-2-1 rule
  - **Three** copies of the data
  - **Two** geographically separate locations
  - More than **one** type of storage device



Artwork by @allison\_horst (CC BY 4.0)

@allison\_horst

# Write a Living Data Management Plan

- Document important details (e.g. file organization and storage, backup plan, etc.) in one place
- Can be relatively short
- Update as research project evolves



Artwork by @allison\_horst (CC BY 4.0)

@allison\_horst

# Sharpen Your Data Management Skills

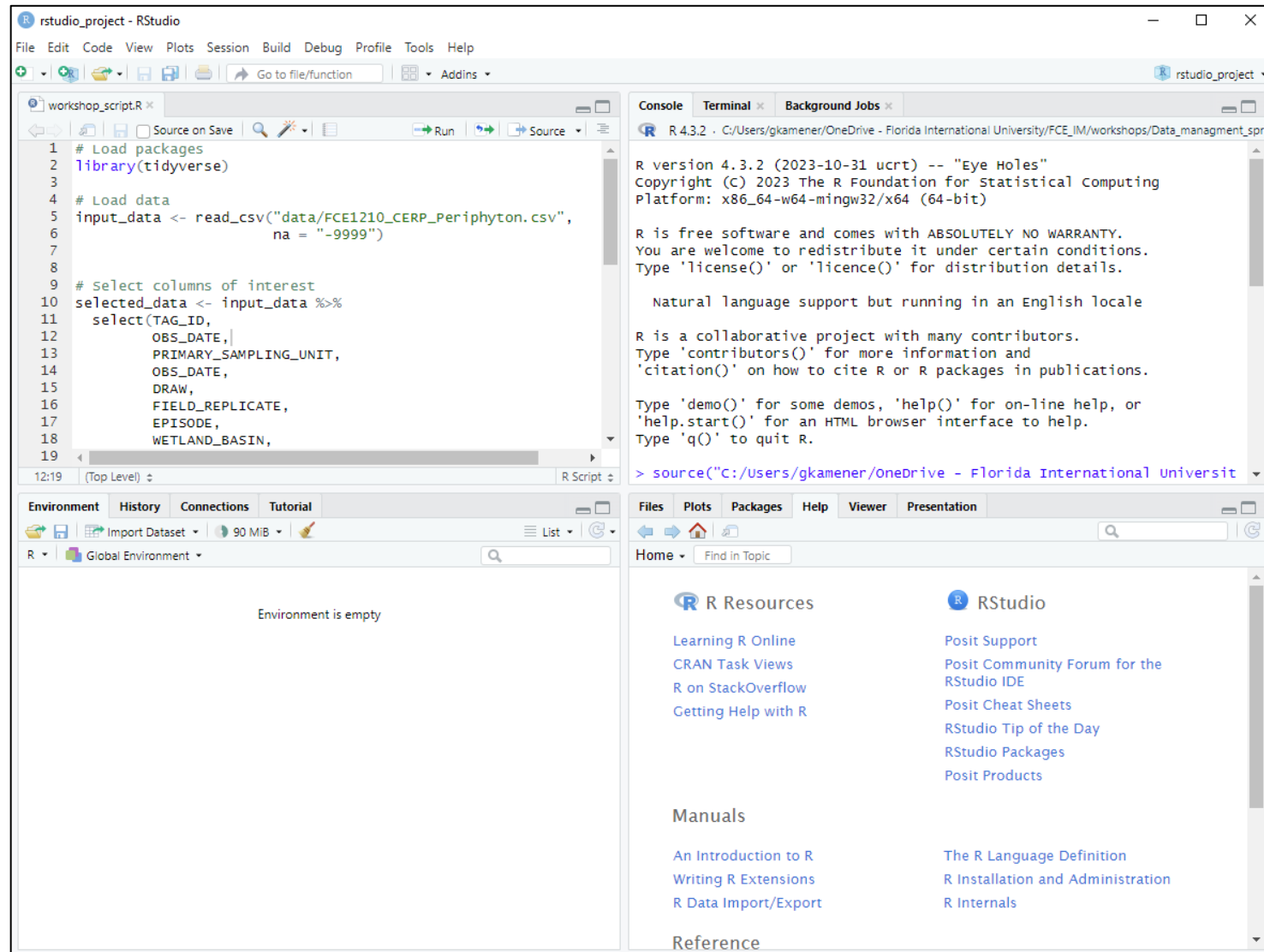
- Briney, K. A., Coates, H. L., & Goben, A. (2020). Foundational practices of research data management. Research Ideas and Outcomes 6: e56508.  
<https://doi.org/10.3897/rio.6.e56508>
- Briney, K. (2023). The Research Data Management Workbook.  
Caltech Library.  
<https://doi.org/10.7907/z6czt-7zx60>



Briney 2023



# Project Management in RStudio



# RStudio Projects

Advantages of RStudio Projects:

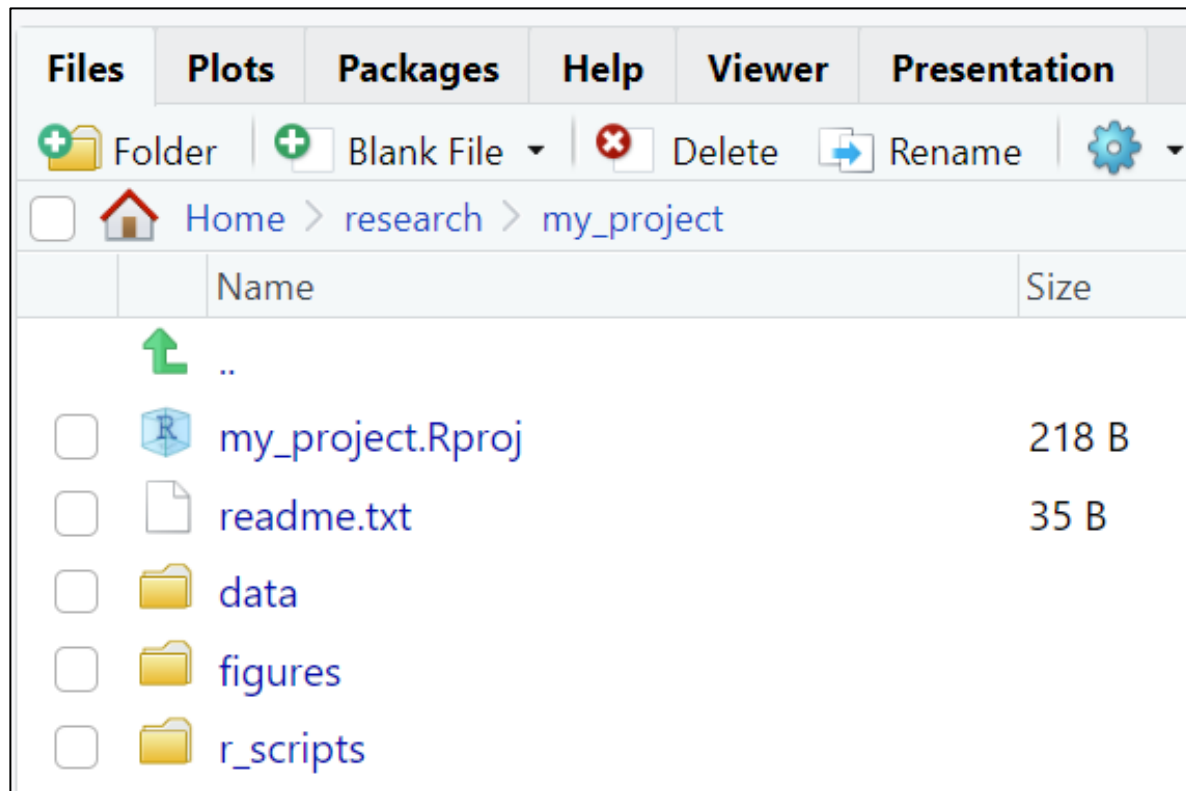
- **Automatically set working directory**
- **Portable**
- **Collaborator friendly**
- Supports version control with Git/GitHub
- Can aid reproducibility with renv package
- **Can integrate with GitHub Copilot**

# Setting up an RStudio Project

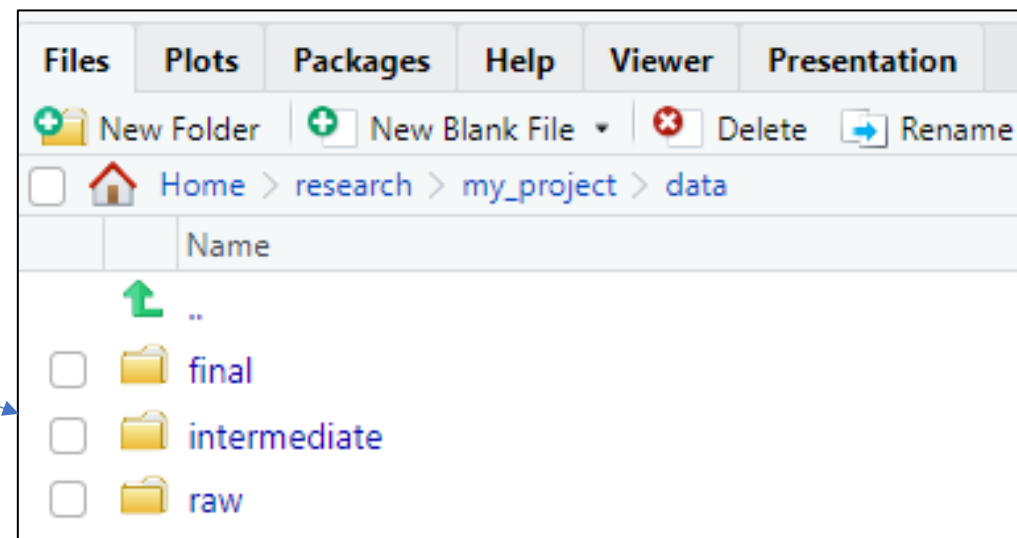
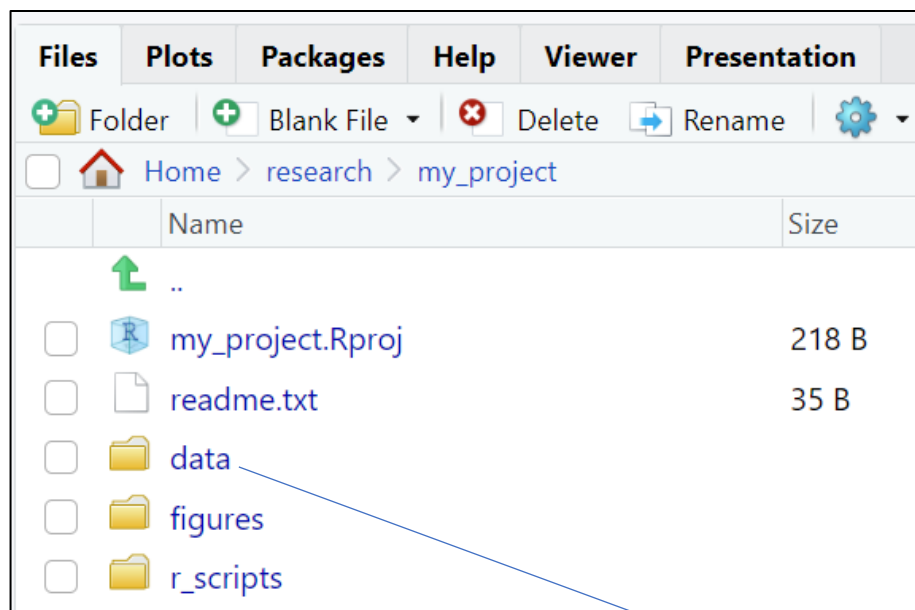
1. Start RStudio.
2. Under the File menu, click on New Project. Choose **New Directory**, then **New Project**.
3. Enter a name for this new folder (or “directory”) and choose a convenient location for it. This will be your working directory.
4. Click on Create Project.
5. Create folders for data (with “raw”, “intermediate”, and “final” subfolders), r scripts, and figures in your working directory.
6. Place raw data files and script files into respective folders.



# Organizing Your Working Directory



# Organizing Your Working Directory



# RStudio Projects: Portability

- R script without RStudio project:

```
df <- read_csv("C:/Users/gkamener/Documents/research/my_project/  
data/raw/my_project_data_GK_2024_11_13.csv")
```

- Often requires specifying working directory or full file path
- Script may break if “my\_project” folder is moved or shared

- With RStudio project:

```
df <- read_csv("data/raw/my_project_data_GK_2024_11_13.csv")
```

- Path always relative to “my\_project” folder
- Project becomes portable

# GitHub Copilot in RStudio

- “Any use of generative ‘AI’ should be **transparent**, **accountable** and **acknowledged**. Check the editorial policies of journals before submission to ensure you are using the most up-to-date guidance” – Cooper et al. 2025
- Check with your advisor and institution before using generative AI for anything research related.



# GitHub Copilot in RStudio

- “You need a fundamental understanding of coding to be able to use Copilot effectively.” – O’Hara and Kui 2025
- Following data-management best practices will enable better use of GitHub Copilot

# GitHub Copilot in RStudio

- GenAI programming resources recently featured during LTER IM monthly meeting
  - Casey O'Hara and Li Kui (2025), AI for Ecologists. URL [https://nceas-learning-hub.github.io/ai\\_for\\_nceas](https://nceas-learning-hub.github.io/ai_for_nceas).
  - Elmendorf (2026), NWT LTER AI Tools Introduction. URL [https://nwtlter.github.io/data\\_skills/AI\\_intro\\_to\\_AI.html](https://nwtlter.github.io/data_skills/AI_intro_to_AI.html)

# Formatting Data: Tidy is the Goal

“**TIDY DATA** is a standard way of mapping the meaning of a dataset to its structure.”  
—HADLEY WICKHAM

In tidy data:

- each variable forms a column
- each observation forms a row
- each cell is a single measurement

each column a variable

id	name	color
1	floof	gray
2	max	black
3	cat	orange
4	donut	gray
5	merlin	black
6	panda	calico

each row an observation

Wickham, H. (2014). Tidy Data. Journal of Statistical Software 59 (10). DOI: 10.18637/jss.v059.i10

# Formatting Data

- One rectangle of data
  - No blank columns or rows
  - No blank cells
- One data type per column
- One value per cell
- Use simple headers
  - No spaces
  - No special characters



Plot #	Subject ID	Weight1207	Stage1207	Sex
1	1	33.2, 1		m
five	2? Not 100% sure	44.0	2	F
#1	3		three	male



Date	Plot	Subject_ID	Weight	Stage	Sex	Comment
2012-07-01	1	1	33.2	1	m	NA
2012-07-01	5	2	44.0	2	f	Unsure of id
2012-07-01	1	3	-9999.0	3	m	Forgot to measure



# Formatting Data

- Format dates as YYYY-MM-DD
- Use consistent categorical variable codes
- Avoid calculations or graphs



Plot #	Subject ID	Weight1207	Stage1207	Sex
1	1	33.2, 1		m
five	2? Not 100% sure	44.0	2	F
#1	3		three	male



Date	Plot	Subject_ID	Weight	Stage	Sex	Comment
2012-07-01	1	1	33.2	1	m	NA
2012-07-01	5	2	44.0	2	f	Unsure of id
2012-07-01	1	3	-9999.0	3	m	Forgot to measure

# Formatting Data

- Don't highlight cells or embed comments
- Document any changes you make
- Store metadata in separate sheet or file



Plot #	Subject ID	Weight1207	Stage1207	Sex
1	1	33.2, 1		m
five	2? Not 100% sure	44.0	2	F
#1	3		three	male



Date	Plot	Subject_ID	Weight	Stage	Sex	Comment
2012-07-01	1	1	33.2	1	m	NA
2012-07-01	5	2	44.0	2	f	Unsure of id
2012-07-01	1	3	-9999.0	3	m	Forgot to measure

# Metadata

- What are metadata?
  - Metadata = data about data
  - Document **who**, **what**, **why**, **where**, and **when**
- Why use metadata?
  - Keep track of important details about data
  - Required to publish datasets
  - Increase findability and usability of published data

# Documenting Metadata

- Dataset metadata should include:
  - Data table information (i.e. data dictionary)
  - Title
  - Abstract
  - Parties responsible for dataset
  - Methods
  - Intellectual rights
  - Keywords
  - Geographic, temporal, and taxonomic coverage
  - Information about non-tabular data (if applicable)
  - Project funding
  - Permits (if applicable)

# Documenting Metadata: Data Dictionary

- Describes data table variables  
(Broman and Woo 2018)
- Create as early as possible!

Column Name	Definition	Variable Type	Units	Precision	Codes	Date Time Format String	Missing Value Code	Missing Value Code Explanation
SITENAME	Name of site	Categorical			SRS2 = Shark River Slough 2			
Date	Date of sample collection	DateTime				YYYY-MM-DD		
Time	Time of sample collection	DateTime				hh:mm	NA	Not recorded
Salinity	Concentration of salinity	Numerical	PSU	0.1			-9999.0	Not recorded
TN	Concentration of total nitrogen	Numerical	micromolePerLiter	0.001			-9999.000	Not recorded
TP	Concentration of total phosphorus	Numerical	micromolePerLiter	0.01			-9999.00	Not recorded
Comment	Field comments	Text					NA	Not recorded



# Documenting Metadata

- Title
  - Should be descriptive, including what, where, and when
  - **Do not** use manuscript title
- Abstract
  - Describe who, what, why, where, and when of dataset in more detail
  - **Do not** include results or conclusions from study
- Methods
  - Detail to answer any questions someone might have
  - Include list of cited references (if applicable)
- Intellectual rights
  - FCE uses CC-BY Creative Commons license

# Documenting Metadata

- Keywords
  - Source from LTER Controlled Vocabulary\* when possible
- Geographic coverage
  - Coordinates in decimal degrees for publishing
- Temporal coverage
  - Documents when data were collected
  - Format as YYYY-MM-DD or YYYY
- Taxonomic coverage
  - Important to check spelling of names

\*LTER Controlled Vocabulary link: <https://vocab.lternet.edu/vocab/vocab/index.php>

# Documenting Metadata

- Information about non-tabular data entities
  - Includes metadata about model code, geospatial, imagery, or other data entities
- Project funding
  - Information about funding for data collection
- Permits
  - Permits that were required for data collection

# Timing and Importance of Quality Metadata

- Important for future you!
- Important for discovery and re-use of published data!
- LTER Network and EDI use Ecological Metadata Language (EML) to document quality metadata

Metadata: Why are they important?

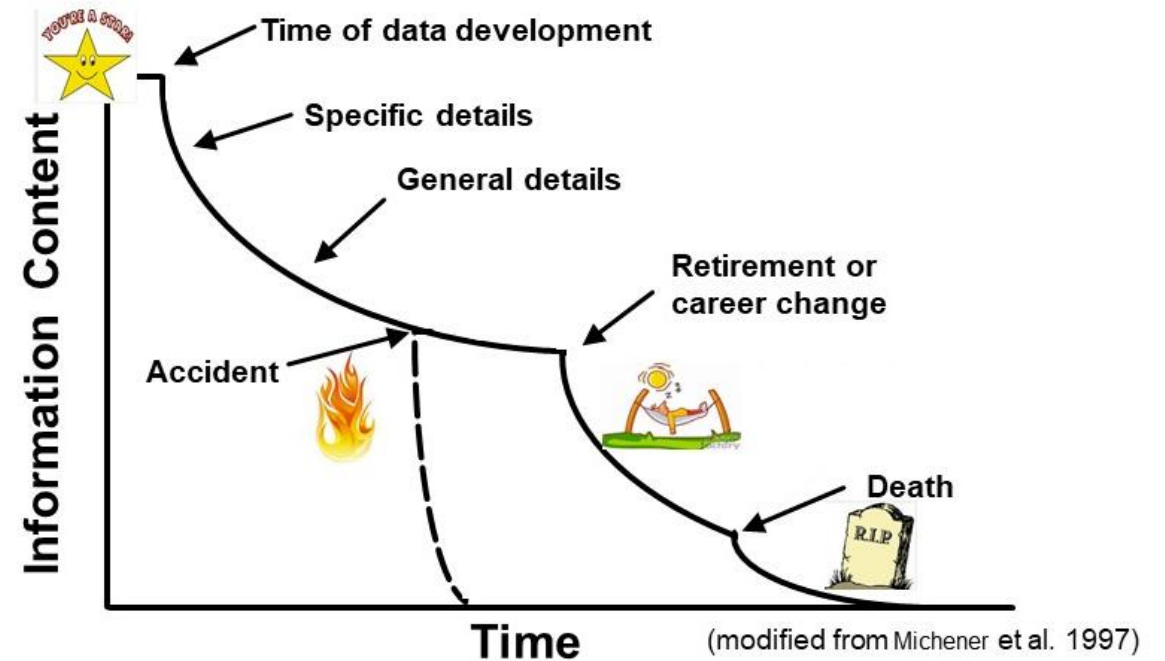


Image courtesy of DataONE

# Metadata Creation During Research Life Cycle

- Metadata should be described throughout the research life cycle
  - Describe as much as possible in early stages
  - Regularly add/update throughout life cycle

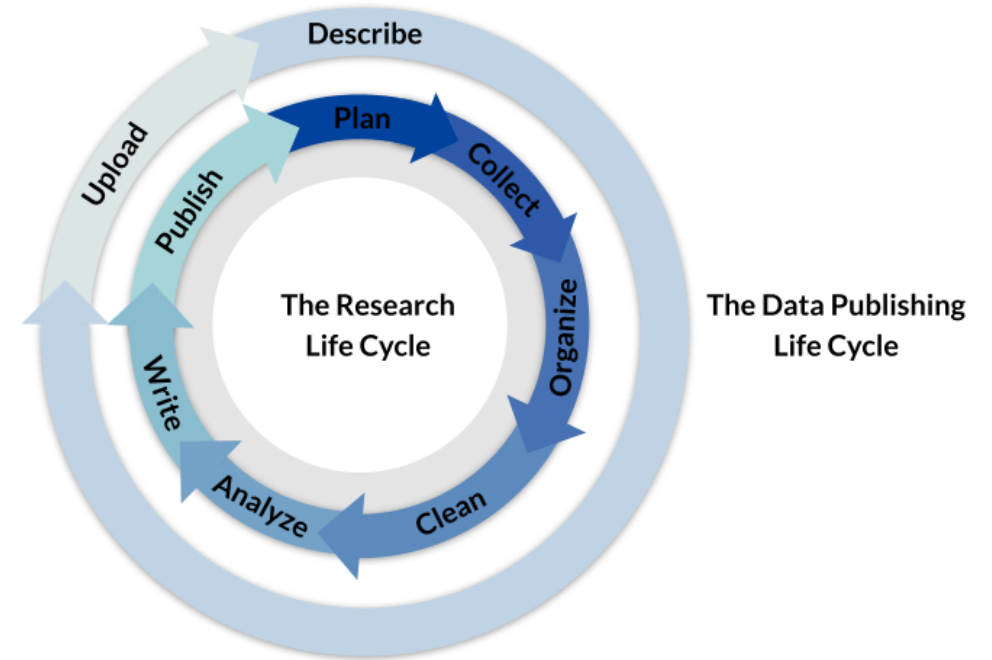
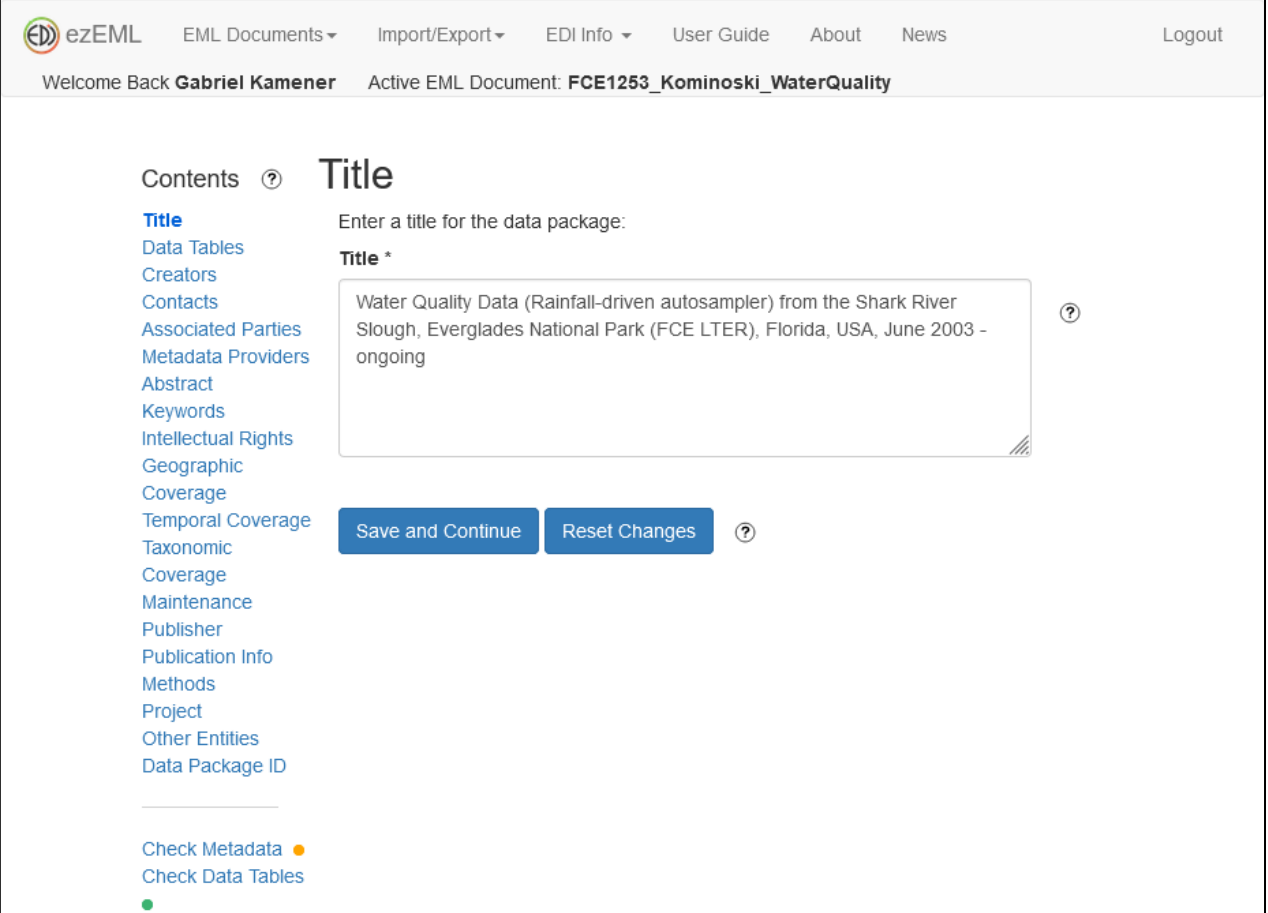


Image courtesy of Environmental data Initiative



# ezEML's Role in Data Publishing Life Cycle

- Form-based web interface to create and edit structured metadata in EML
- Can utilize pre-filled FCE templates
- Saves as you go
- Can quality check data and metadata
- Produces data package (data + metadata)



The screenshot displays the ezEML web interface. At the top, there is a navigation bar with links: ezEML, EML Documents, Import/Export, EDI Info, User Guide, About, News, and Logout. Below this, a status bar reads "Welcome Back Gabriel Kamener" and "Active EML Document: FCE1253\_Kominoski\_WaterQuality". The main content area is divided into two sections. On the left, a "Contents" sidebar lists various metadata fields: Title, Data Tables, Creators, Contacts, Associated Parties, Metadata Providers, Abstract, Keywords, Intellectual Rights, Geographic, Coverage, Temporal Coverage, Taxonomic, Coverage, Maintenance, Publisher, Publication Info, Methods, Project, Other Entities, and Data Package ID. The "Title" field is currently selected. On the right, the "Title" form is shown with the instruction "Enter a title for the data package:". The "Title \*" field contains the text: "Water Quality Data (Rainfall-driven autosampler) from the Shark River Slough, Everglades National Park (FCE LTER), Florida, USA, June 2003 - ongoing". Below the form are two buttons: "Save and Continue" and "Reset Changes". At the bottom left, there are two status indicators: "Check Metadata" with a yellow dot and "Check Data Tables" with a green dot.

<https://ezeml.edirepository.org/eml>

# ezEML's Role in Data Publishing Life Cycle

- Can create metadata for tables and other data entities (e.g. model code, imagery, and other nontabular data files)

The screenshot displays the ezEML web interface. At the top, a navigation bar includes the ezEML logo, links for EML Documents, Import/Export, EDI Info, User Guide, About, News, and a Logout button. Below this, a status bar shows 'Welcome Back Gabriel Kamener' and 'Active EML Document: knb-lter-fce.1250.1'.

The main content area is titled 'Data Tables' with a help icon. On the left, a sidebar lists various metadata fields: Title, Data Tables (highlighted), Creators, Contacts, Associated Parties, Metadata Providers, Abstract, Keywords, Intellectual Rights, Geographic Coverage, Temporal Coverage, Taxonomic Coverage, Maintenance, Publisher, Publication Info, Methods, Project, Other Entities, and Data Package ID.

The 'Data Tables' section contains a table with two entries:

Data Table Name	
FCE1250_Lamb_AltStableState_HydroGeo	<div><div>▼</div><div>Edit</div><div>Remove</div><div>Re-upload</div></div>
FCE1250_Lamb_AltStableState_Bio	<div><div>▲</div><div>Edit</div><div>Remove</div><div>Re-upload</div></div>

Below the table, there are two buttons: 'Load Data Table from CSV File' and 'Add Data Table from Scratch' (with a help icon). A 'Save and Continue' button is also present.

At the bottom, there are two status indicators: 'Check Metadata' with a green dot and 'Check Data Tables' with an orange dot.

# ezEML's Role in Data Publishing Life Cycle

- Quality checks data and metadata
- Provides feedback on potential showstoppers (errors) and points to investigate (warnings)

[Contents](#) [Check Data Table: Results](#)

[Title](#)  
[Data Tables](#)  
[Creators](#)  
[Contacts](#)  
[Associated Parties](#)  
[Metadata Providers](#)  
[Abstract](#)  
[Keywords](#)  
[Intellectual Rights](#)  
[Geographic Coverage](#)  
[Temporal Coverage](#)  
[Taxonomic Coverage](#)  
[Maintenance](#)  
[Publisher](#)  
[Publication Info](#)  
[Methods](#)  
[Project](#)  
[Other Entities](#)  
[Data Package ID](#)

[Check Metadata](#) ●  
[Check Data Tables](#) ●

Please note: When data packages are submitted to EDI's data repository, data table error checking is performed there as well. Experienced users of the repository may recognize that the repository's error checking is more permissive than the checking being done here in ezEML. ezEML's error checking is intended to reflect best practices and help data providers minimize the data cleaning burden that will be passed on to consumers of their data.

[Back](#)

**Data Table:** SRS\_Rain\_Water\_Chemistry

Column: **Date**    Type: DATETIME

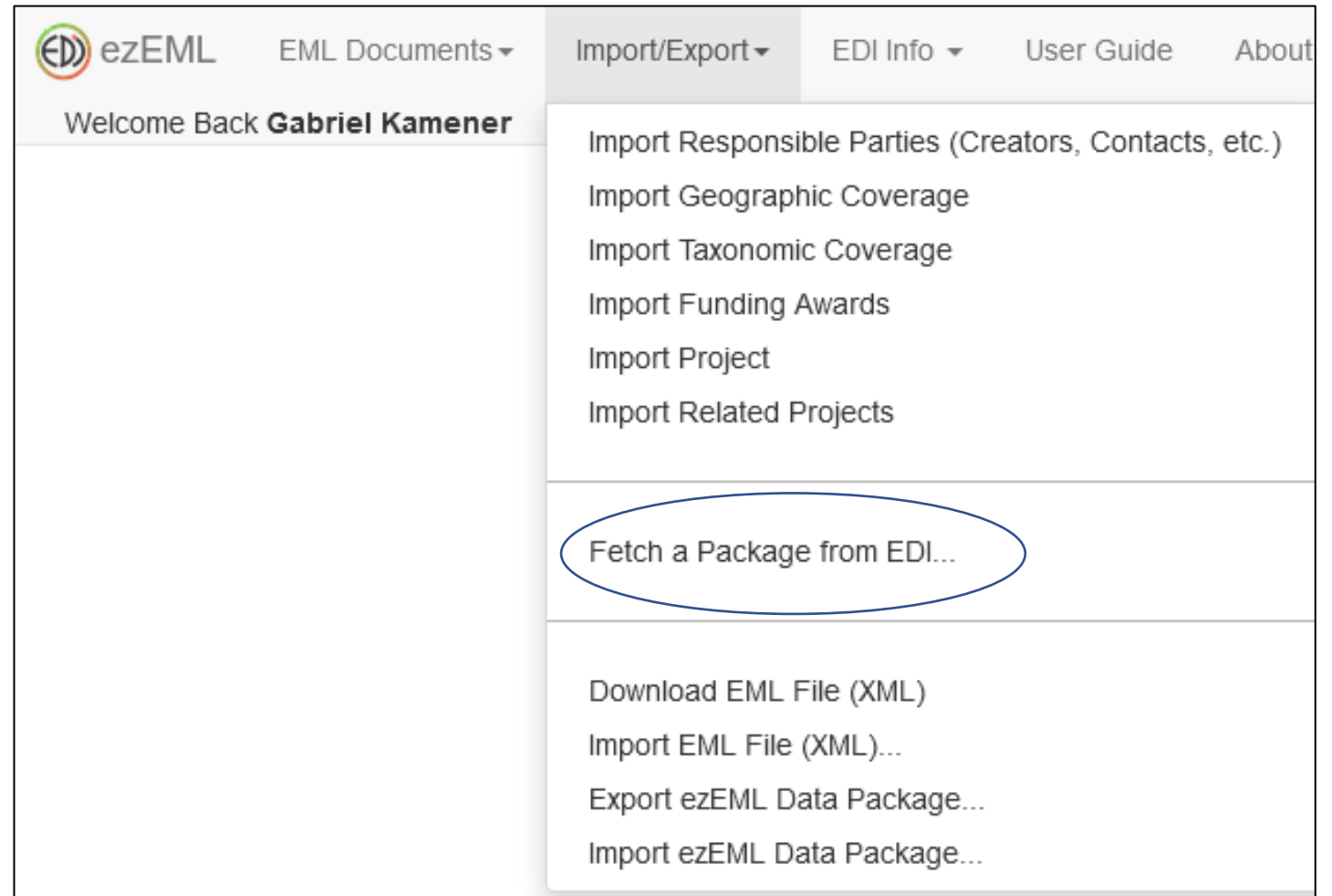
Row	Error	Expected	Found
	The specified DateTime Format String is not supported.	A <a href="#">supported</a> format	mm/dd/yyyy

Column: **Time**    Type: DATETIME

Row	Error	Expected	Found
4	DateTime element does not have expected format	hh:mm	9:05
25	DateTime element does not have expected format	hh:mm	6:25
26	DateTime element does not have expected format	hh:mm	6:40
45	DateTime element does not have expected format	hh:mm	9:29

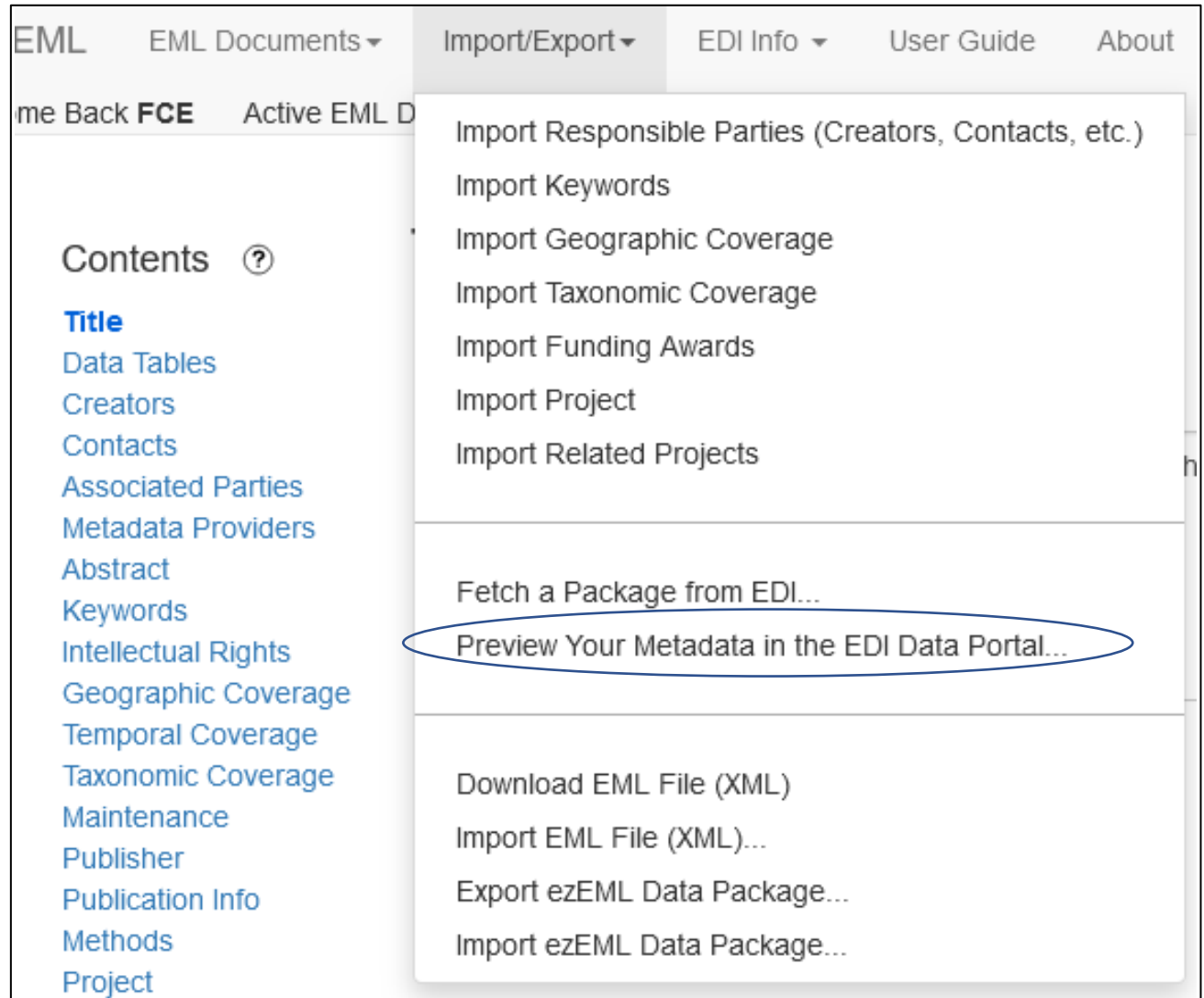
# ezEML's Role in Data Publishing Life Cycle

- Can also “fetch” existing packages from EDI to facilitate updates




# ezEML's Role in Data Publishing Life Cycle

- Preview how your package will look on the EDI Data Portal



# ezEML's Role in Data Publishing Life Cycle

- Preview how your package will look on the EDI Data Portal



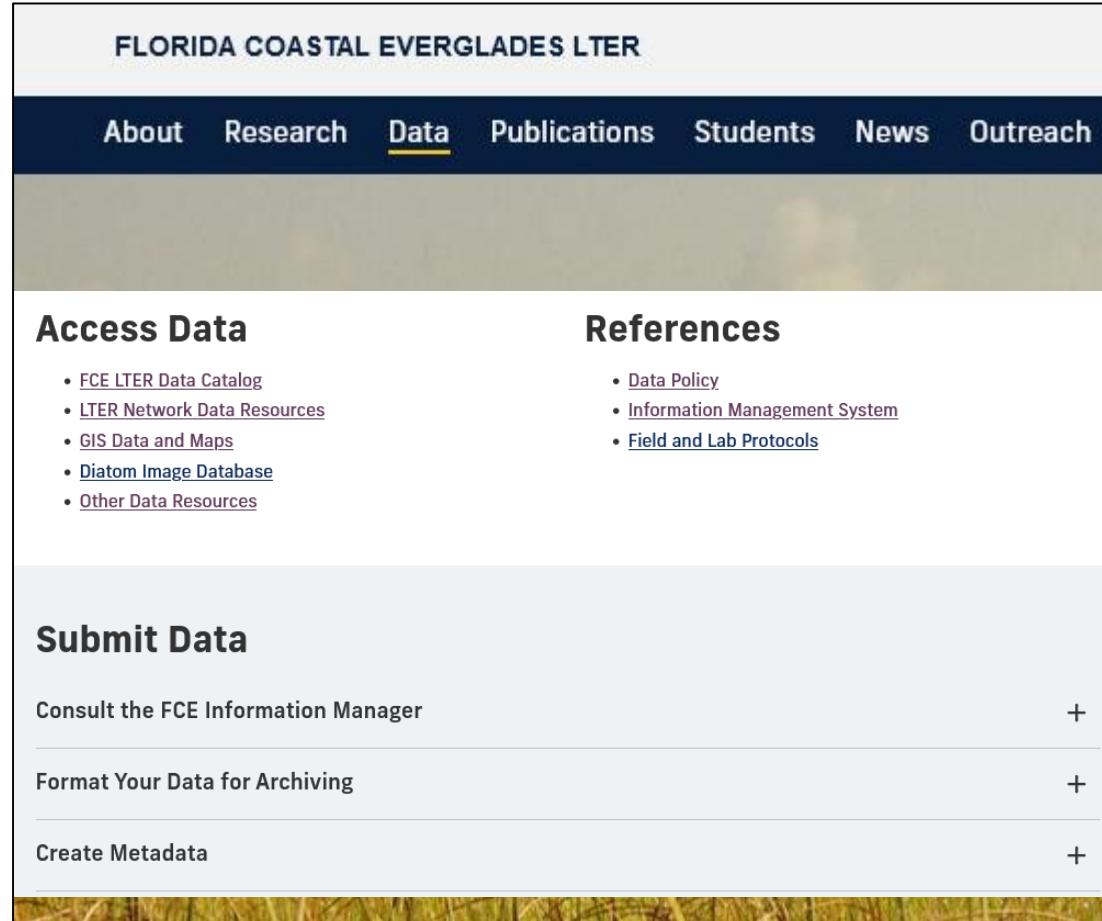
The screenshot displays the EDI Data Portal interface. At the top, there is a navigation bar with links for HOME, DATA, TOOLS, HELP, and LOGIN. Below this is a search bar with the placeholder text "enter search terms" and a magnifying glass icon, followed by a link to "ADVANCED SEARCH". The main content area features a blue header for "Data Package Metadata" with a "View Summary" link. The title of the data package is "Water Quality Data (Rainfall-driven autosampler) from the Shark River Slough, Everglades National Park (FCE LTER), Florida, USA, June 2003 - ongoing". Under the "General Information" section, the "Data Package:" details are listed in a table-like format.

Data Package:	
Local Identifier:	knb-lter-fce.1253.5
Title:	Water Quality Data (Rainfall-driven autosampler) from the Shark River Slough, Everglades National Park (FCE LTER), Florida, USA, June 2003 - ongoing
Abstract:	Water quality samples are being collected using ISCO autosamplers at all freshwater sites: SRS1a (not active), SRS1c (not active), SRS1d, SRS2, and SRS3. Rain level actuators are used at the sites to trigger water sampling after rain events exceed a given threshold of duration and/or intensity. As currently programmed, when a rain event at a site exceeds the threshold of 2.5 cm per hour, the autosampler at that site collects a 1000mL sample 30 minutes later. The samples are retrieved from the site every 3-4 weeks and analyzed for total phosphorus (TP), total nitrogen (TN), and salinity. Salinity values were not taken consistently from 2000 to mid-2017; those values were replaced by -9999 in the data. See also Shark River Slough precipitation data package (knb-lter-fce.1092) and Shark River Slough extensive water quality data (knb-lter-fce.1072) on the FCE LTER website's data catalog or in the EDI repository.
Publication Date:	2023-09-18



# Submitting Your Data to the FCE IM

1. Review the FCE Data page!
2. Contact the FCE IM to review your data and to receive an FCE dataset ID
3. Accept ezEML collaboration invite from FCE IM
4. Enter data and metadata into ezEML
5. Review with FCE IM



The screenshot shows the 'Data' page of the Florida Coastal Everglades LTER website. The header includes the site name and a navigation menu with 'About', 'Research', 'Data' (highlighted), 'Publications', 'Students', 'News', and 'Outreach'. The main content area is divided into two columns: 'Access Data' and 'References'. The 'Access Data' column lists links to the FCE LTER Data Catalog, LTER Network Data Resources, GIS Data and Maps, Diatom Image Database, and Other Data Resources. The 'References' column lists links to Data Policy, Information Management System, and Field and Lab Protocols. Below these columns is a 'Submit Data' section with a light blue background, containing three expandable items: 'Consult the FCE Information Manager', 'Format Your Data for Archiving', and 'Create Metadata', each with a plus sign icon to its right.

FLORIDA COASTAL EVERGLADES LTER

About Research Data Publications Students News Outreach

**Access Data**

- [FCE LTER Data Catalog](#)
- [LTER Network Data Resources](#)
- [GIS Data and Maps](#)
- [Diatom Image Database](#)
- [Other Data Resources](#)

**References**

- [Data Policy](#)
- [Information Management System](#)
- [Field and Lab Protocols](#)

**Submit Data**

Consult the FCE Information Manager +

Format Your Data for Archiving +

Create Metadata +

<https://fcelter.fiu.edu/data>

# IM Support at FCE

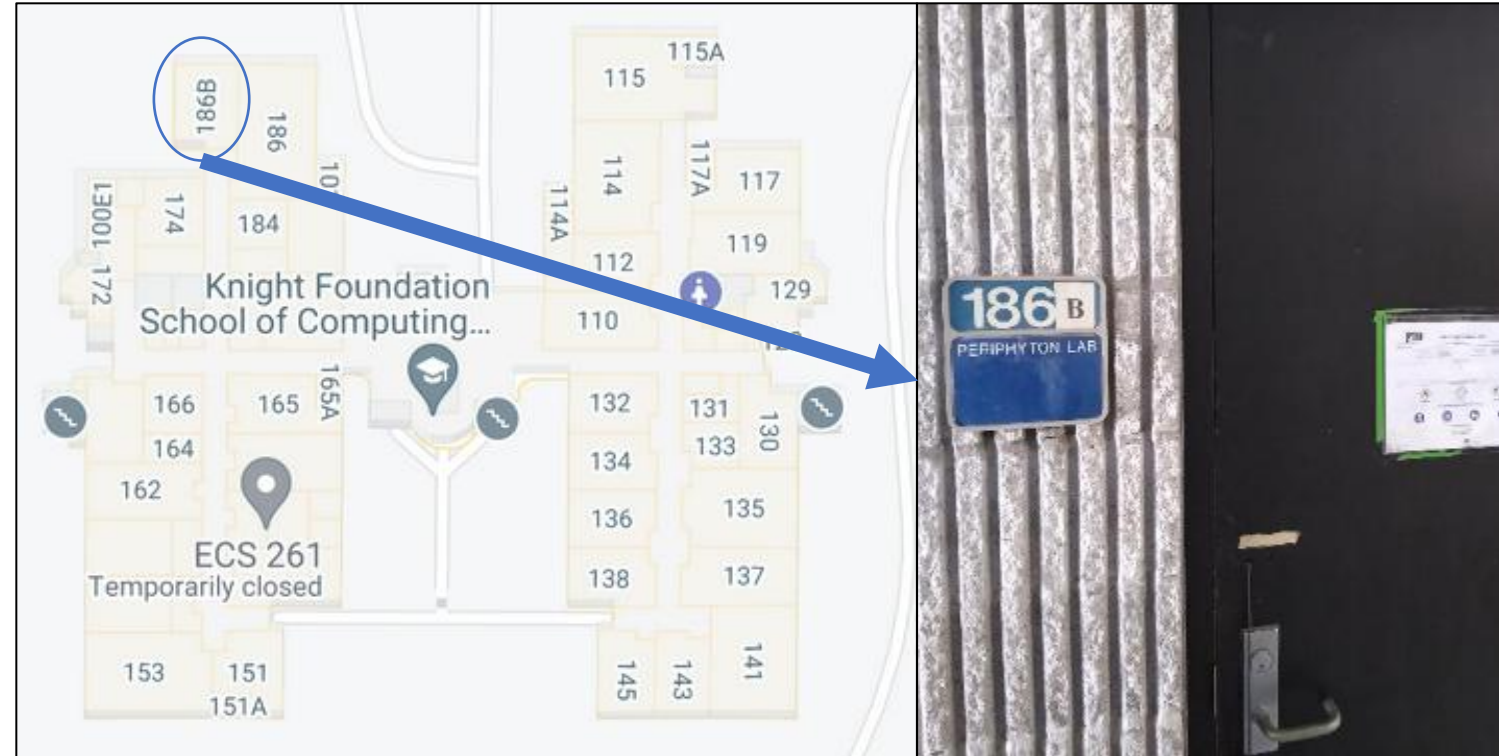
- Section on FCE Data page!
- Contains IM support information and a growing collection of links to resources on best practices

Access Data	
<ul style="list-style-type: none"><li>• <a href="#">FCE LTER Data Catalog</a></li><li>• <a href="#">LTER Network Data Resources</a></li><li>• <a href="#">GIS Data and Maps</a></li><li>• <a href="#">Diatom Image Database</a></li><li>• <a href="#">Other Data Resources</a></li></ul>	
References	
<ul style="list-style-type: none"><li>• <a href="#">Data Policy</a></li><li>• <a href="#">Information Management System</a></li><li>• <a href="#">Field and Lab Protocols</a></li></ul>	
Submit Data	
Consult the FCE Information Manager	+
Format Your Data for Archiving	+
Create Metadata	+
Information Management Support and Best Practices	
Information Management Support	+
Additional Resources	+

<https://fcelter.fiu.edu/data>

# IM Support at FCE

- Weekly IM office hours at MMC-CASE 186B, Thursdays 2:00-7:00 pm
- IM support at FCE student Think Tank events
- Available by email or appointment (in-person or Zoom)



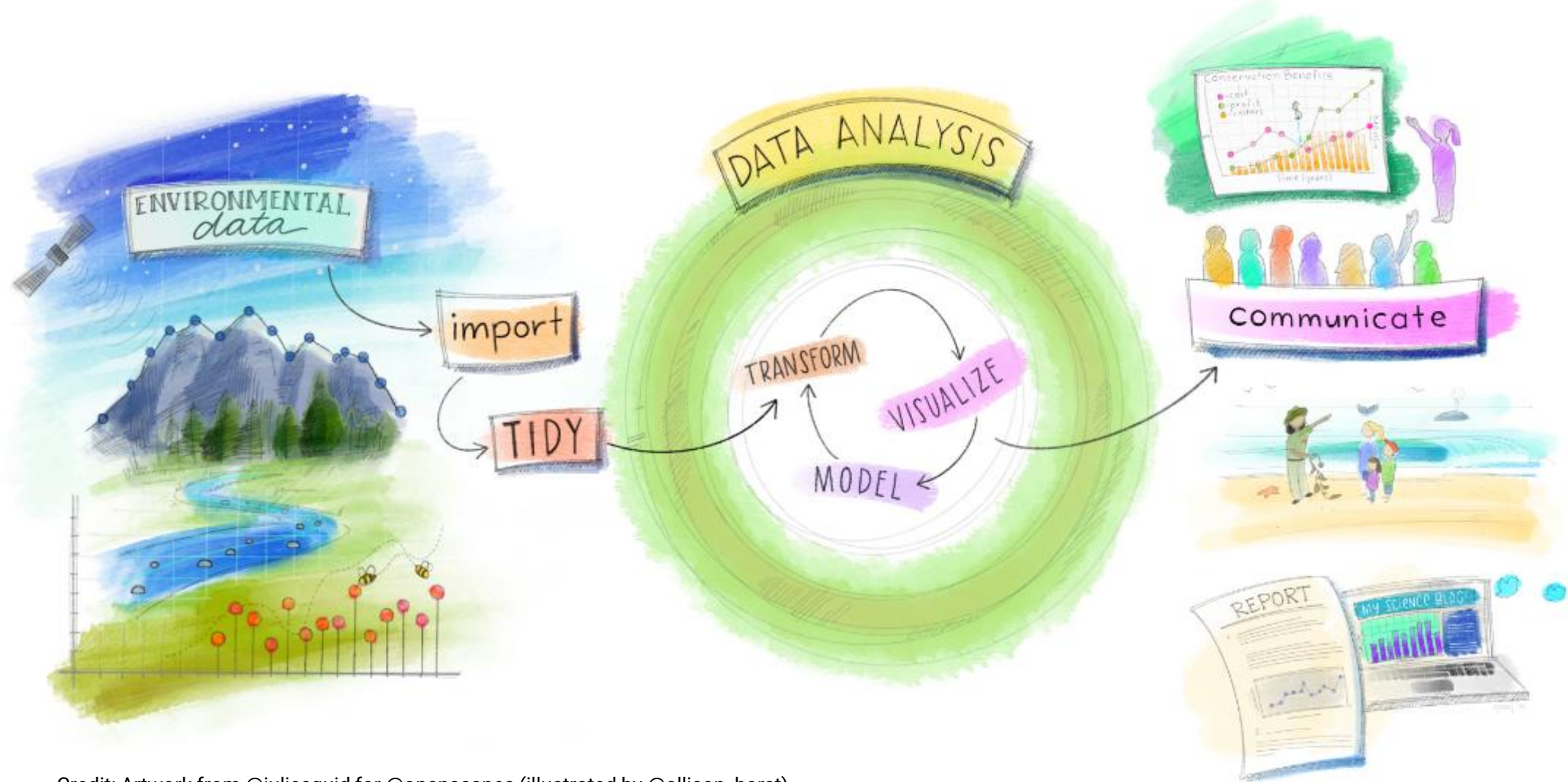
# Resources and References

- FCE LTER data page - <https://fcelter.fiu.edu/data>
- [FCE ezEML instructions](#) (PDF file)
- [Tips for submitting FCE data and metadata](#) (PDF file)
- FCE GitHub repository containing this presentation and example metadata documents: [https://github.com/FCE-LTER/intro\\_to\\_fce\\_data\\_management\\_and\\_publication](https://github.com/FCE-LTER/intro_to_fce_data_management_and_publication)
- Briney, K. A., Coates, H. L., & Goben, A. (2020). Foundational practices of research data management. Research Ideas and Outcomes 6: e56508. <https://doi.org/10.3897/rio.6.e56508>
- Briney, K. (2023). The Research Data Management Workbook. Caltech Library. <https://doi.org/10.7907/z6czh-7zx60>
- Broman, K. W., & Woo, K. H. (2018). Data organization in spreadsheets. The American Statistician, 72(1), 2-10. <https://doi.org/10.1080/00031305.2017.1375989>

# Resources and References

- Casey O'Hara and Li Kui (2025), AI for Ecologists. [https://nceas-learning-hub.github.io/ai\\_for\\_nceas](https://nceas-learning-hub.github.io/ai_for_nceas).
- Cooper, N., Hsing, P.-Y., Almarzouq, B., Baldauf, S., Chatterjee, N., Plomp, E., Strydom, T., Takola, E., & Zagrodzka, Z. (2025). Guide to Reproducible Code (v2.0.1). Zenodo. <https://doi.org/10.5281/zenodo.17855982>
- Data Carpentry course episode: [Introduction to R and RStudio](#)
- Elmendorf (2026), NWT LTER AI Tools Introduction. [https://nwltter.github.io/data\\_skills/AI\\_intro\\_to\\_AI.html](https://nwltter.github.io/data_skills/AI_intro_to_AI.html)
- Kleindl, P. M., Wachnicka, A., Sah, J. P., Ross, M. S., & Gaiser, E. E. (2025). Hydrology drives facilitative and competitive strategies in freshwater macrophyte and microbial communities. *Ecosphere*, 16(9), e70384. <https://doi.org/10.1002/ecs2.70384>

# Questions and Discussion



Credit: Artwork from @juliesquid for @openscapes (illustrated by @allison\_horst).

[https://github.com/allisonhorst/stats-illustrations/tree/master/openscapes?fbclid=IwAR0M\\_Bksd0Hr5IL4Jla7l8m1azBojU8acN50mZ4QIGv20xcVbBhbkvpgBlw](https://github.com/allisonhorst/stats-illustrations/tree/master/openscapes?fbclid=IwAR0M_Bksd0Hr5IL4Jla7l8m1azBojU8acN50mZ4QIGv20xcVbBhbkvpgBlw)