

Palmyra Atoll Data Library

Table of contents

Welcome!	3
Why PADL?	3
About this guide	4
What to expect	4
Basic Concepts	6
Data Workflow	8
1 Metadata Components	9
1.1 Title	9
1.2 Data Tables	10
1.3 Project Participants and Personnel	10
1.3.1 Creators	10
1.3.2 Contacts	10
1.3.3 Associated Parties	10
1.4 Abstract	10
1.5 Keywords	10
1.6 Intellectual Rights	10
1.7 Geographic Coverage	10
1.8 Temporal Coverage	10
1.9 Taxonomic Coverage	10
1.10 Maintenance	10
1.11 Methods	10
1.12 Project	10
1.13 Other Entities	10
Using ezEML	11
Metadata Templates	12
Alternative options	13
Additional Resources	14

Welcome!

The Palmyra Atoll Data Library (PADL) aims to publish as much data collected at Palmyra Atoll as possible.

We believe that facilitating access to research and its associated data will amplify academic discussions, help other researchers, and further inform management and policy decisions being made in Palmyra and other tropical ecosystems.

We are working with the [Environmental Data Initiative \(EDI\)](#), an NSF-supported initiative that provides a reliable, registered and certified trustworthy data repository with an open-access API. EDI follows a high metadata standard that ensures that your future self and others can utilize data for scientific and other forms of inquiry.

Why PADL?

We want to have a centralized place where scientist and managers have access to all data collected at Palmyra Atoll. We believe that data are among the most valuable outputs of research. If no-one can access these data, all the data collection is a shocking waste of resources. Having an open data library allows for aggregating and synthesizing data from different contexts. This is essential to establishing broader ecological knowledge and informing conservation management. Long-term data are crucial to understand historical patterns and baselines in a changing world. PADL wants to make it easy for scientist to know and have access to research done at Palmyra in the past.

Note

This guide borrows heavily from [EDI resources](#), [CAP LTER Getting Started Guide](#) and [NCEAS coreR curriculum](#). Check out the links and see complete citation on the reference section.

About this guide

We created this guide to help you document your data in a way that facilitates the process of publishing your data or metadata to EDI. Here we describe what you need to do to make your data publicly available through the EDI repository.

! Important

We understand that EDI might not be suitable to all data. If you are planning to publish your data elsewhere we ask to please collect the metadata following EDI metadata standards presented in this guide and submit your metadata to TNC Palmyra.

What to expect

- The goal of this guide is for you to know what you need to do to be in compliance with TNC Palmyra Data Sharing Agreement.
- We provide some orientation and tips for when planning your data collection, together with resources to learn more about tidy data.
- We describe each of the elements needed to best document your data.
- This guide goes in-depth on how to publish your data using the [ezEML](#) tool suite developed by the Environmental Data Initiative ([EDI](#)).
- We also provide metadata templates for you to know and plan on documenting your data from the get go.

💡 Tip

We highly recommend to have a plan for collecting your metadata since the beginning of your research life cycle. According to EDI experience “continuous creation of metadata during the research life cycle greatly benefits data management during a project and when it comes time to publish data when the project concludes”.

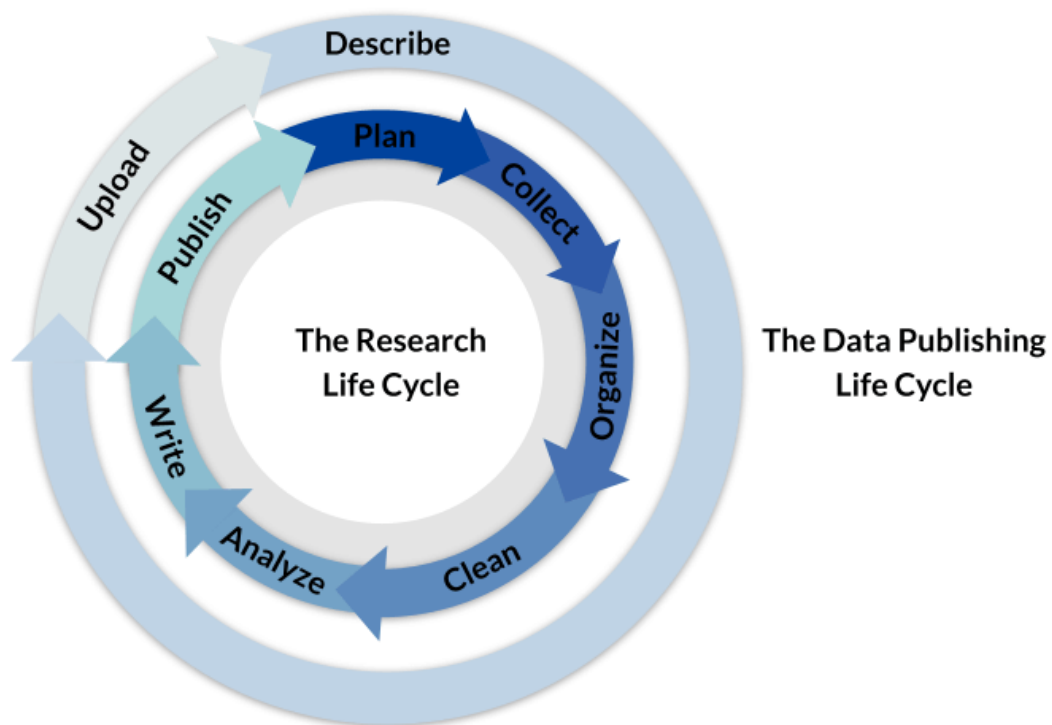


Figure 1: from edirepository.org, Creating Metadata During the Research Life Cycle

Basic Concepts

Before we start we want to introduce to some relevant concepts related to data publishing. This sets the ground to navigate the environmental/ecological data publishing world.

Data Package

- EDI publishes data from the ecological and environmental sciences irrespective of funding origin.
- The *unit of publication* is called a data package.
- This is an assemblage of science metadata, the data it self, and any other file you want to publish with your data.
- The basic rule of thumb is to package your data in the form you would like to receive it.
- Each data package is assigned a Digital Object Identifier (DOI) and published in the repository for future use.
- EDI resources on [data packages](#)

Metadata

- The metadata is the data that provides information about your data. It describe the structure, content, and context of your data.
- They are vital to the discovery and reuse of data, and are a required element of a data package.
- EDI resources on [Creating Metadata for a Data Package](#)

Ecological Metadata Language (EML)

- EDI uses the Ecological Metadata Language (EML) format to document the metadata in each data package.
- “EML provides a comprehensive vocabulary and a readable XML markup syntax for documenting research data. It is a community-maintained specification, and evolves to meet the data documentation needs of researchers who want to openly document, preserve, and share data and outputs” ([Jones et al, 2019](#)).

ezEML

- EDI created a user-friendly tool that streamlined the process of the creating metadata in the Ecological Metadata Language (EML).
- ezEML is a form-based online application that leads the user through EML document creation step by step.
- Check out EDI [ezEML User Guide](#) for more information.

Data Workflow

1. Organize
2. Clean
3. Describe
4. Upload
5. Cite

1 Metadata Components

This section describes of the most relevant components of the EML. The goal is for you to learn about all information you need to document from your data from the get go. The next sections show formats that can help you do this along the way.

1.1 Title

Every data package needs a **Title**. Data packages title should be informative and provide helpful context about the dataset. Title should inform about the temporal and geographic extent of the data.

Example

Activity budgets and space use for two common Pacific parrotfish (*Chlorurus* Sp.) at Palmyra Atoll National Wildlife Refuge 2014.

1.2 Data Tables

1.3 Project Participants and Personnel

1.3.1 Creators

1.3.2 Contacts

1.3.3 Associated Parties

1.4 Abstract

1.5 Keywords

1.6 Intellectual Rights

1.7 Geographic Coverage

1.8 Temporal Coverage

1.9 Taxonomic Coverage

1.10 Maintenance

1.11 Methods

1.12 Project

1.13 Other Entities

Using ezEML

Metadata Templates

Alternative options

Additional Resources