Jornada Metadata Standards for EML Creation

The Information Management Team

27 June 2019

Table of Contents

**If you edit this document please track your changes and send to Greg Maurer (gmaurer.jrn.lter@gmail.com).**

# Introduction

This document defines metadata standards to use when creating EML files for Jornada Basin LTER research data packages. These standards are under active development as we revamp our data catalog, but they are essentially derived this reference document:

<https://environmentaldatainitiative.files.wordpress.com/2017/11/emlbestpractices-v3.pdf>

with modifications to suit the Jornada Basin LTER and its researchers.

# Metadata standards by EML section

## Personnel

At least one contact entity must be supplied in every EML document. Almost all EML documents should have at least one creator entity.

### creator entities

Creator entities can consist of one or more persons, who must include the original PI for the project AND and original co-PIs on the project.

No other creator entities are allowed.

### contact entities

Contact entities can consist of one or more persons who should be contacted for access to, or information about, the data package. These must include the Data Manager (listed by position and email - datamanager.jrn.lter@gmail.com - , not name) and the “currently responsible PI.” If there is no “currently responsible PI” this contact may be left out of the EML document, leaving only the Data Manager

JRN defines the “currently responsible PI” as the LTER principal investigator who curates a study and its associated data package(s). For data packages without a “currently responsible PI”, John Anderson has historically been listed in this role in the list of contact entities. When updating one of these packages, ask John if he would like to remain in this position. If not he may be removed, leaving only the Data Manager.

No other contacts are defined.

# Information management, metadata, and EML resources

The LTER network (funded by NSF) maintains a [Data Access policy](https://lternet.edu/data-access-policy/) that forms the basis for how Information Managment (IM) systems should function at sites in the LTER network. LTER maintains an [IM web portal](https://im.lternet.edu) with some resources on IM practices, guidelines, and training materials. The [Guidelines for LTER IM Systems](https://im.lternet.edu/im_requirements/ims_guidelines) document encapsulates the guidelines and best practices for what an LTER IM system should be.

Recently, many of the resources on curating and publishing data and metadata packages has been pushed to [EDI](https://environmentaldatainitiative.org). EDI’s “5 Phases of data publishing” is a large collection of resources, but in particular, [Phase 3](https://environmentaldatainitiative.org/resources/five-phases-of-data-publishing/phase-3/) provides useful information on EML metadata best practices like the [EML Best Practices V3 document](https://environmentaldatainitiative.files.wordpress.com/2017/11/emlbestpractices-v3.pdf).

The EML specification is defined and maintained by the [Knowledge Network for Biocomplexity](https://knb.ecoinformatics.org). They provide the EML Schema and additional (human readable) information here:

<https://knb.ecoinformatics.org/external//emlparser/docs/index.html>