

Publishing IEP Datasets to the Environmental Data Initiative

Adopting an Open Science Framework, the Interagency Ecological Program (IEP) Data Utilization Workgroup has started to archive long-term monitoring and special study datasets to the Environmental Data Initiative (EDI) data repository. EDI supports permanent citations of datasets using digital object identifiers (DOI), data versions, and tracking. EDI serves as a means for IEP to have more transparent data practices, discoverability, and defensibility as well as increase the opportunity for synthesis. Archived and open data can also more easily be shared and linked to other platforms and portals.

For IEP investigators and data managers interested in archiving open data on the EDI data repository, there are three ways to submit and publish datasets to the EDI data repository.

1. EDI Data Curation Services – Word Template

EDI's Data Curation Team will do the data packaging and upload to the repository for you. Simply fill out [EDI's Metadata Template](#) for each dataset you intend to archive and send it with your data to info@environmentaldatainitiative.org.

- Modifications to EDI's Metadata Template for IEP datasets can be found at [the IEP GitHub page](#).
- Sent data must be cleaned and ready to publish. This includes:
 - Perform any standard QAQC procedures and document your procedures in the metadata.
 - Making sure all tables have a single header row, a single data type per column, and a single value per cell. Use .csv files when possible.

2. EDI Data Curation Services – ezEML

EDI recently developed a graphical user interface as an alternative to the Word template for entering your metadata information. You can log in to the [ezEML web application](#) and create an ezEML data package by manually entering your values for each section in the metadata.

- EDI has [step-by-step instructions](#) for uploading data using ezEML.
- You can save your ezEML file and modify it when you update your package with additional data.

3. Do-it-Yourself - EMLAssemblyline

Package and upload data to the repository yourself. Data packaging involves organizing, cleaning, and describing your data with information rich EML metadata. Uploading involves a repository user account and some familiarity with EDI's Data Portal. Reference the [Five Phases of Data Publishing](#) for details on each of these steps. Contact info@environmentaldatainitiative.org if you have any questions.

- *IEP's [GitHub](#) page has instructional documents and templates for "Do-it-Yourself" publishing.*
 - IEP-EDI metadata template
 - Procedures for creating EML metadata in R programming (EMLassemblyLine package)
 - Examples of R code from other IEP long-term monitoring datasets
 - Staging and Production Instructions for final upload and publication
 - EDI repository account information IEP
- *For an alternate EML R package, see [EMLaide](#)*

Additional contacts and links for EDI that IEP has been directly working with are:

- Corinna Gries; cgries@wisc.edu (lead PI)
- Colin Smith; colin.smith@wisc.edu (Data curation lead)
- Mark Servilla; mark.servilla@gmail.com (PASTA software and application developer)
- [EDI Website Link](#)
- [EDI Data Portal Link](#)