GRIIDC Experimental Lab Data Guidance

Rationale and Purpose

As part of GRIIDC's mission to ensure a data and information legacy to promote continual scientific discovery, GRIIDC accepts a breadth of data generated by a variety of methods. Data submitted to GRIIDC includes environmental data collected though field studies, large-scale climate and ocean models, and data generated solely through lab activities. GRIIDC requests that data be shared in the rawest format that permits the widest reuse; however, this format depends on the QA/QC performed at each level of processing, the size of data files, and the standards in the field of study. The purpose of this document is to outline general requirements for sharing data generated through experimental lab activities. This may include data generated from samples collected in the field and exposed to manipulated conditions in the lab (e.g., microcosms, mesocosms) and activities such as analyzing the structure of chemicals and/or spectral analysis in the laboratory that may not have a relevant field component.

Data Sharing Requirements for Experimental Data

Data File

1. Data should be in raw non-proprietary formats that permit the widest reuse. Please refer to the GRIIDC guidance document Preferred File Formats for additional information.
2. Data file should not contain graphs, figures, tables, or text published in peer review journals or any other copyright material.
3. Provide the data for each replicate in each treatment rather than averages and/or standard deviations.
4. Provide the data used to calculate values (e.g. percentages, LC50, LD50, proportions, ratios) rather than the final calculated values.
5. Provide the blanks information or calibration files use for spectral analysis, if applicable.
6. Provide enough information to understand the treatment and replicate that generated each data point. Additional information can be provided in a readme file or in the dataset description.
7. If a dataset includes images or videos, a readme file should be provided that describes the images and videos and explains the experimental conditions or sample reflected in each image.
8. If the data are organized into multiple folders, please provide a readme file that explains the organization and/or naming convention.

Dataset Description

The dataset description provided during submission should allow another person to understand the contents of the dataset without referencing an external resource, such as a publication.

1. In the dataset description the following information should be provided, if applicable:
   - Experimental setup:
       i. Initial concentration of chemicals used in different treatments
       ii. Volume of vessel
       iii. Number of organisms in each replicate
       iv. Dimensions of tank
       v. Wave production method
       vi. Standards used to spike samples
       vii. Number of replicates in each treatment
       viii. Velocity of jets
       ix. Method used to generate water accommodated fraction
   - Duration of experiment and timepoint when measurements were performed
   - Reference to specific EPA methods (e.g. EPA Method 537)
   - Brief description of methods (do not rely on a citation of publication to describe methods)
2. In the dataset description, define all parameters, variable names, acronyms, and units.
3. If a dataset was generated solely in the lab or through microcosm/mesocosm experiments where field samples are brought into the lab and then exposed to manipulated conditions:
   - Do not provide a place keyword
   - In the Data Extent tab select "Define Spatial Extent" and select "Non-Spatial" to provide a short sentence that explains why the data does not have a relevant geographic component. For example: "Dataset contains dynamic light scattering measurements of oil droplets dispersed with food-grade dispersants, no relevant geographic component".
4. Please include the UDI (or DOI) of any related datasets in the metadata abstract.