[**Instructions**: Each data package must populate this readme template. **Delete the instructions (marked with brackets throughout) before saving.** If using the template in Git, save as .md. If starting with the SFA GitLab template, this template is pre-populated. **For inclusion in your ESS-DIVE data package, save as .pdf.**The readme is not limited to the information or sections described below. No matter what is included in the readme file, the documentation should be written for an audience that has never seen the data package contents before. Descriptions should be written as clearly and concisely as possible. We strongly recommend having someone review the content before finalizing. We strongly recommend looking at other RC SFA data packages on ESS-DIVE to see what they have written.]

[**Instructions for data packages**: Follow instructions for a [manuscript package](https://pnnl.sharepoint.com/:w:/r/teams/SubsurfaceBiogeochemicalResearchSFA/_layouts/15/Doc.aspx?sourcedoc=%7B3BAFC8FC-ACCA-4B2C-ACF3-D9E92A290A05%7D&file=How-to-publish-a-manuscript-associated-data-package.docx&action=default&mobileredirect=true) or talk to Amy Goldman.]

|  |
| --- |
| **Title** |

**[Section Required]**

[**Instruction**: Insert data package title here. Should be less than 20 words]

|  |
| --- |
| **Summary** |

**[Section Required]**

[**Instruction**: If the data package is associated with a manuscript, include the following sentence as the first sentence of the summary/abstract: “This data package is associated with the publication “[Insert title of manuscript]” published in [or submitted to] [insert name of journal] ([insert first author last name et al. and year; [insert pre-print link or manuscript doi).]

[**Instruction**: Insert abstract/summary/machine learning problem setup]

[**Instruction**: When publishing to ESS-DIVE include link to GitLab/GitHub if there is one: “This data package is associated with the [GitLab or GitHub] repository found at [insert link]”]

|  |
| --- |
| **Brief Overview of Methods** |

**[Section Required]**

[**Instruction**: Insert a very brief summary of the methods associated with your data package. This section should match the Methods section in the ESSDIVE\_Metadata]

[**Instruction**: In addition to the summary, you can point to methods information in your data package or in your manuscript.]

|  |
| --- |
| **Critical Details** |

**[Section Optional]**

[**Instructions**: This section is optional. It may contain a description of a labeling scheme and mapping across samples, replicates, sensors, experiments, sites, or files. It may contain clarifying information about differences among samples, replicates, sensors, experiments, sites, or files. It may contain clarifying information on how to work with the data, analyses, or models or what processing has been done. It may also contain the operating system if scripts are included and it is relevant]

|  |
| --- |
| **Data Package Structure** |

**[Section Required]**

[**Instructions**: SFA data packages are required to follow ESS-DIVE Reporting Formats. Follow the [step-by-step ‘how to’ instructions](https://pnnl.sharepoint.com/:w:/r/teams/SubsurfaceBiogeochemicalResearchSFA/Shared%20Documents/General/SFA%20Data%20and%20Software%20Management/How-to-Publish-Data-Package.docx?d=wced5db769f7e4d328434ea7ad7350d59&csf=1&web=1&e=Px7pHV).]

[**Instruction:** List and explain specific folders and/or files OR include the following sentence or something similar in your readme: “Please see [name of FLMD csv] for a list of all files contained in this data package and descriptions for each. Data dictionaries for csv files have \_dd” appended to the file names.”]

|  |
| --- |
| **Citations and Acknowledgements** |

**[Section Required]**

[**Instruction include the following text for SFA packages or replace with the text from your manuscript acknolwedgments:** This research was supported by the U.S. Department of Energy (DOE) Office of Science, Biological and Environmental Research (BER) Program, Environmental System Science (ESS) Program (<https://ess.science.energy.gov/>) through the Pacific Northwest National Laboratory (PNNL) River Corridor Science Focus Area (SFA). PNNL is operated by Battelle Memorial Institute for the DOE under Contract No. DE-AC05-76RL01830.

[**Instruction when publishing FTICR data generated at EMSL, include the statement:** FTICR-MS data were generated at the Environmental Molecular Sciences Laboratory, a DOE BER User Facility (EMSL; https://ror.org/04rc0xn13), under the EMSL User Proposal #####.]

[**Instruction when publishing to ESSDIVE and there is an associated manuscript, include the statement:** Cite this data package with the appropriate DOI. Cite the associated manuscript in any work that that uses analyses or conclusions presented in the manuscript. To cite the paper: [**Instruction**: Insert full paper citation]]

[**Instruction** **for software development, include the following statement verbatim:** “This material was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the United States Department of Energy, nor Battelle, nor any of their employees, nor any jurisdiction or organization that has cooperated in the development of these materials, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness or any information, apparatus, product, software, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. PACIFIC NORTHWEST NATIONAL LABORATORY operated by BATTELLE for the UNITED STATES DEPARTMENT OF ENERGY under Contract DE-AC05-76RL01830.”]

[**Instruction** **for WHONDRS data packages, include the following and modify as needed: “**- Please acknowledge the U.S. Department of Energy (DOE) Biological and Environmental Research (BER) Environmental System Science (ESS) program (<https://ess.science.energy.gov/>) — which generously provides funding to the WHONDRS — in work that uses this dataset. If using FTICR-MS, NMR, LC-MS, or GC-MS data, please also acknowledge the U.S. DOE Environmental Molecular Sciences Laboratory (EMSL; grid.436923.9).”]

[**Instructions for citations**: list citations as bullets below any reference. It is recommended to list them in alphabetical order]

|  |
| --- |
| **Contact** |

**[Section Required]**

[**Instruction**: Insert Contact name, Contact email]

|  |
| --- |
| **Change History** |

**[Section Required When on ESS-DIVE]**

**[Instruction**: **NOTE: this section only needs to be included in the ESS-DIVE data package readme. It does not need to be included in the GitLab/GitHub readme. The purpose is to indicate if you have revised a published data package after it has been published. Populate row 1 during first publication. Populate row 2 and beyond as needed if revisions are done.]**

Approach to change history and versioning:

Updates to **data package** version: When any file within a data package is updated, the data package version number is updated. The data package version number is indicated in the title of the data package, the data package folder name, and in the change history table below. You can access previous versions of the data package by sending a request to ESS-DIVE.

Updates to **individual file** versions: As files are changed, the file version number is also updated. The file version number is indicated in the file name, file level metadata (flmd) file, and the change history table below. The version number on an individual file may not match the version number of the data package. For example, v3 of a data package may include v2 of an individual file.

The change history below describes each file revised during versioning. If you are interested in seeing the exact cells within a file that have changed, you can utilize the daff package in R (<https://github.com/edwindj/daff>) to compare a previously downloaded file to a newly downloaded file.

In the change history table below, the sub-headers and bullets indicate the type of change in each file:

* New files: Describes new files added that were not present in previous data package versions
* Bulk changes to files: Describes a change to many files within the data package. The indicated superscript with be added to each file name that the change applies to.
* Modified files:
* Corrected: Describes existing information modified or removed to prevent sharing of incorrect information
* Added: Describes new information inserted into an existing file (e.g., appending new columns/rows)
* Updated: Describes modifying existing information to maintain accuracy though version changes. (e.g., changing version number to new version number)

Change history:

|  |  |
| --- | --- |
| **Data Package Version** | **Changes** |
| **Version 1**  *[MONTH YEAR]* | Original data package publication |