



ESS-DIVE and Urban Integrated Field Laboratories Data

Presenters



Joan Damerow

Project Scientist



Emily Robles

*Data Management
Research Associate*

Presentation Overview



- Overview of ESS-DIVE and our data;
- ESS-DIVE project data management features;
- Resources to organize and curate high-quality datasets;
 - Dataset organization
 - Overview of reporting formats and templates
- Discussion on Urban IFL Data Management Needs

Key Takeaway: Learn how you can use ESS-DIVE features and reporting formats for your project data management and publication

What Urban IFL project or ESS project do you work on?



In relation to publishing data, which of the following terms are you familiar with?

Data package

Metadata

Dataset

Data repository

DOI (Digital Object Identifier)

FAIR



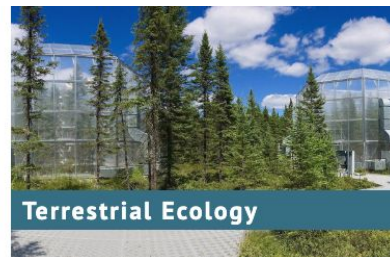
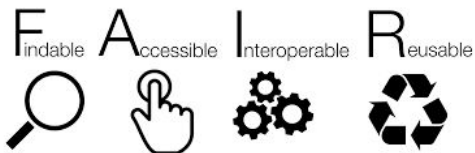
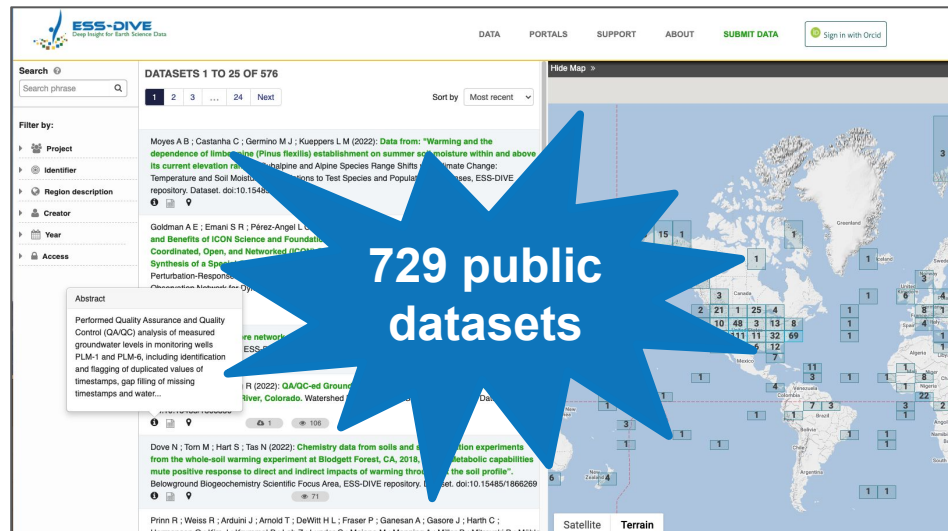
Overview of ESS-DIVE

What is ESS-DIVE?

data.ess-dive.lbl.gov

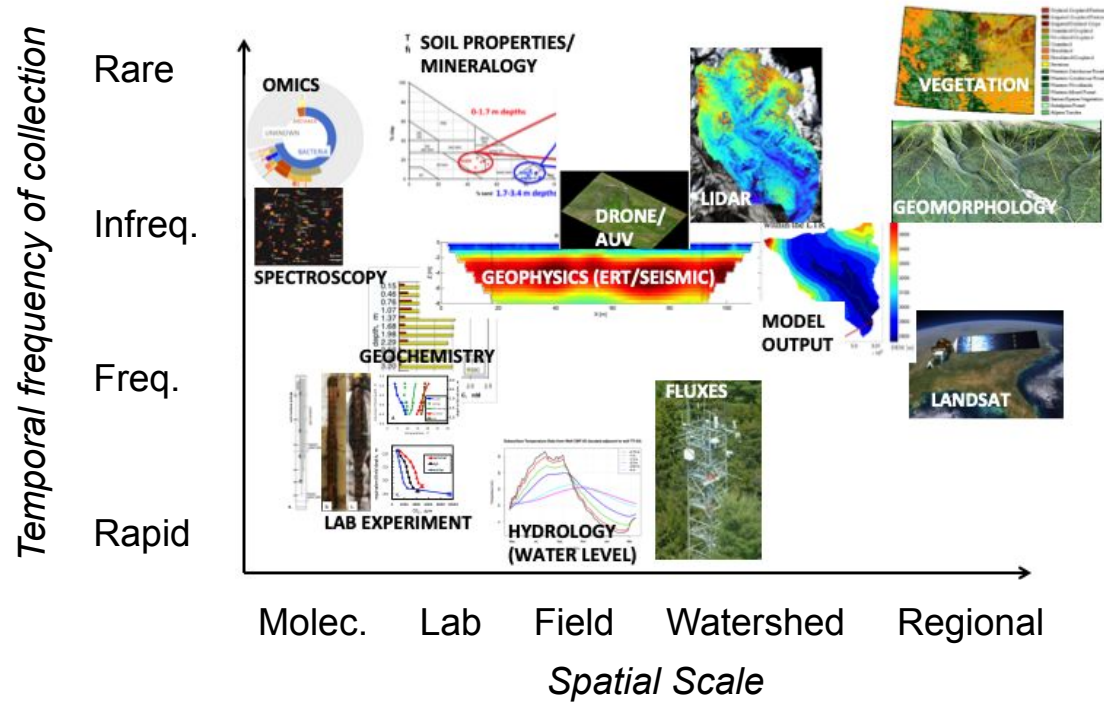


Data Repository to **preserve**
expand **access** to and
improve **usability** of
Environmental Systems
Science data



Complex data in Earth and Environmental Science

- The data are **diverse** and interdisciplinary
- Work with Urban-IFL projects to understand data types and support publication



How to upload and publish data on ESS-DIVE



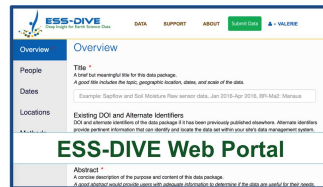
Register w/ ORCID.

Organize **data packages**.



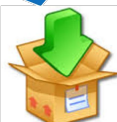
Revision

Revise/Create



Define **metadata** for data package.

Upload

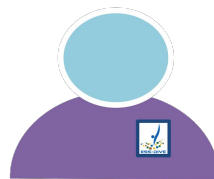


Data Package Metadata

Upload **files** for data package.

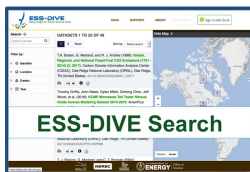
Revise

Submit



ESS-DIVE Project User

Review



Available for **public search/download**.

Catalog

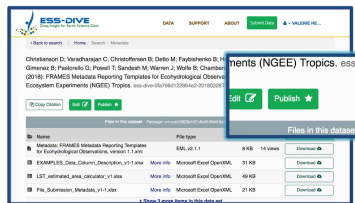
ESS-DIVE **review**.

DOI assigned.



Citation Information

Publish



Submit to save privately.
When ready click **Publish**.

ess-dive.lbl.gov/archive/

Tools and resources for dataset creation



Help Documentation

Refer to **ESS-DIVE's website and Gitbooks** for detailed information on data package requirements

ess-dive.lbl.gov/
docs.ess-dive.lbl.gov/



Offline Metadata Guide

Collaborate on metadata with co-authors before working on ESS-DIVE

docs.ess-dive.lbl.gov/



Sandbox Testing Server

Practice uploading datasets to our sandbox test server, which does not permanently save data

[data-sandbox.ess-dive.l
bl.gov/](https://data-sandbox.ess-dive.lbl.gov/)



Support Email Service

Feel free to **contact the ESS-DIVE support team** through email for any questions

[ess-dive.lbl.gov/contact/
ess-dive-support@lbl.gov](https://ess-dive.lbl.gov/contact/ess-dive-support@lbl.gov)

Web Interface Submission and Automated Review



Overview *	<h2>Overview</h2> <p>Title * A brief but meaningful title for this data package. <i>A good title includes the topic, geographic location, dates, and scale of the data.</i></p> <p>Example: Sapflow and Soil Moisture Raw sensor data, Jan 2016-Apr 2016, BR-Ma2: Manaus</p> <p>Existing DOI and Alternate Identifiers DOI and alternate identifiers of the data package if it has been previously published elsewhere. Alternate identifier Enter as many identifiers as needed below.</p> <p>Example: http://dx.doi.org/10.15486/NGT/XXXXXX</p> <p>Abstract * A concise description of the purpose and content of this data package. <i>A good abstract would provide users with adequate information to determine if the data are useful for their needs.</i></p> <p>Example: Raw output from the data logger connected to 9 sapflow and 5 soil moisture sensors where the sensors were installed, and other installation/maintenance details. No data processing.</p>
Overview *	<h2>Creators</h2> <p>The main researchers involved in producing the data. These include authors, owners, originators and principal investigators. One or more creators is required. If none are entered, you will be set as the creator of this document. List creators in the order you want them to appear.</p> <p>Creator #1</p> <p>First name: <input type="text"/> Last name: <input type="text"/></p> <p>Organization name: <input type="text"/></p> <p><i>All information entered in this section will be made public once the data package is published.</i></p>

<https://ess-dive.lbl.gov/webinars/>

Metadata Assessment Report

Perzan Z ; Maher K ; Boye K ; Bargar J (2023): Meteorological, reference evapotranspiration, and groundwater quality data from the Crested Butte, Colorado. Groundwater Quality SFA, ESS-DIVE repository. Dataset. [divelbl.gov/datasets/doi:10.15485/1958211](https://ess-dive.lbl.gov/datasets/doi:10.15485/1958211) on 2023-02-26

After running your metadata against our standard set of metadata, data, and congruency checks, we have found the following:



Assessment suite: ESS-DIVE Conformance

Findable: 100% complete

Accessible: 100% complete

Reusable: 100% complete

Interoperable: 100% complete

▶ Passed 19 checks out of 19 (informational checks not included).

▶ Warning for 0 checks.

▼ Failed 0 checks.

Web Search



ESS-DIVE
Only insight for Earth Science Data

DATA SUPPORT ABOUT Submit Data Sign in with ORCID

Search

DATASETS 1 TO 25 OF 441

Sort by: Most recent

Filter by:

- Identifier
- Region description
- Creator
- Year

Chadwick K D ; Brodrick P ; Grant K ; Henderson A ; Bill M ; Breckheimer I ; Williams C F R ; Goulden T ; Falco N ; McCormick M ; Musinsky J ; Pierce S ; Hastings Pono M ; Scott A ; Brodie E ; Hancher M ; Staltzer H ; Wainwright H ; Williams K ; Maher K (2020): NEON AOP foliar trait maps, maps of model uncertainty estimates, and conifer map, East River, CO 2018. A Multiscale Approach to Modeling Carbon and Nitrogen Cycling within a High Elevation Watershed. doi:10.15485/1509695

Damerow J ; Varadarajan G ; Boye K ; Brodie E ; Burrus M ; Chadwick D ; Cholla S ; Crystal-Omelas R ; Elbaashandy H ; Eloy Alves R ; Ely K ; Goldman A ; Hendrix V ; Jones C ; Jones M ; Kakalia Z ; Kemner K ; Kersting A ; Maher K ; Merino N ; O'Brien F ; Perzan Z ; Robles E ; Shavelev C ; Sorensen P ; Stagen J ; Waisenhorn P ; Whelanack K ; Zavarzin M ; Agarwal D (2020): ESS-DIVE Global Sample Numbers and Metadata Reporting Format for Environmental Systems Science (IGSN-ESS). Environmental Systems Science Data Infrastructure for a Virtual Ecosystem (ESS-DIVE). doi:10.15485/1500470

<https://data.ess-dive.lbl.gov>

DataONE

Making data more discoverable, accessible, & usable

What data are you looking for?

<https://search.dataone.org/>

WHONDRS 48 Hour Diel Cycling Study at HJ Andrews Experimental Forest Watershed 1 (WS1)

Explore at ESS-DIVE Explore at www.osti.gov Explore at search.datacite.org

2 scholarly articles cite this dataset (View in Google Scholar)

Unique identifier
<https://doi.org/10.15485/1509695>

Dataset updated 2019

Dataset provided by
Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems (WHONDRS)

Authors
James C Stegen; Amy E Goldman; Rosalie K Chu; Robert E Danczak; Vanessa A Garayburu-Caruso; Emily B Graham; Xinming Lin; Joseph W Morad; Huiying Ren; Lupita Renteria; Charles T Resch; Malak Tfaily; Nikola Tolici; Jason G Toyoda; Jacqueline R Wells

License
CC0 1.0 Universal Public Domain Dedication
License information was derived automatically

Area covered

Map Satellite

Not seeing a result you expected?
[Learn](#) how you can add new datasets to our index.

[Google Dataset Search](#)

Downloads, Citing and Reusing Data

Goldman A E ; Chu R K ; Danczak R E ; Daly R A ; Fansler S ; Garayburu-Caruso V A ; Graham E B ; McCall M L ; Ren H ; Renteria L ; Resch C T ; Tfaily M ; Tolic N ; Torgeson J M ; Toyoda J G ; Wells J ; Wrighton K C ; Stegen J C ; WHONDRS Consortium T (2020): WHONDRS Summer 2019 Sampling Campaign: Global River Corridor Sediment FTICR-MS, NPOC, and Aerobic Respiration. River Corridor and Watershed Biogeochemistry SFA Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems (WHONDRS). doi:10.15485/1729719



Citations

0

Downloads

0

Views

12

Copy Citation

Assessment report

Files in this dataset Package: ess-dive-915b1859fbd5ec6-20210430T040533424808

Name	File type	Size	Login to Download
Metadata: WHONDRS_Summer_2019_Sampling_Campaign_Global_River_Corridor_Sediment_FTICR_MS_NPOC_and_Aerobic_Respiration.xml	EML v2.1.1	64 KB	Download
WHONDRS_S19S_SitePhotos.zip	More info ZIP file	2 GB	Download
WHONDRS_S19S_Sediment_v2.zip	More info ZIP file	68 MB	Download

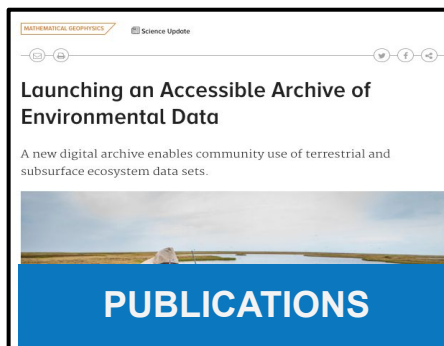
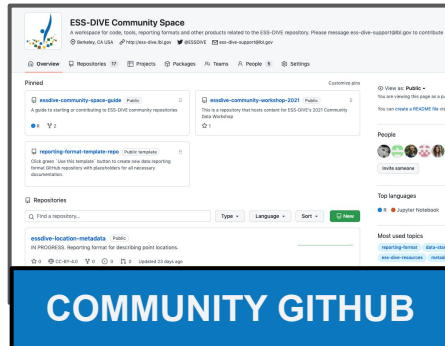
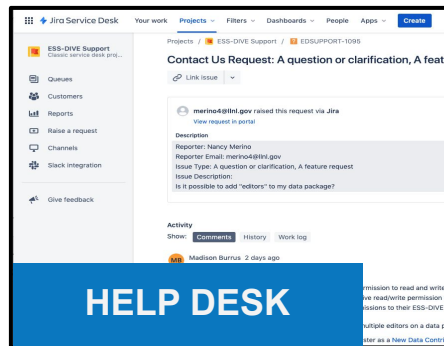
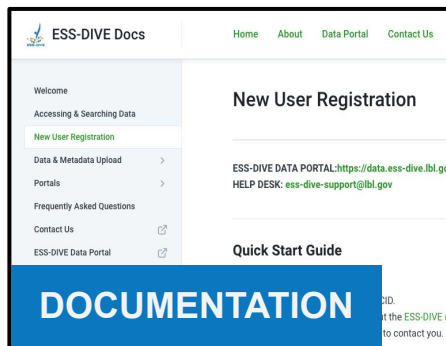
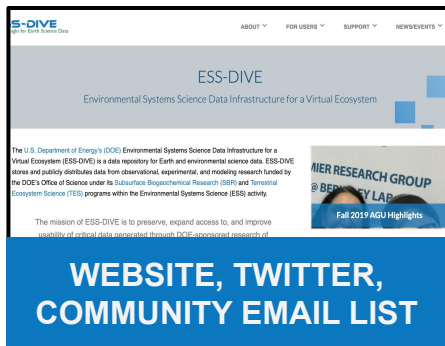
General

Identifier ess-dive-87e6b9a956fc602-20210430T040525320818

Abstract The WHONDRS Summer 2019 Sampling (S19S) Campaign collected samples in 97 globally distributed river corridor systems between July and September 2019. Surficial streambed sediments were collected at three locations within each site (upstream, midstream, and downstream). Surface water was collected at the downstream site. This dataset includes a portion of the data types produced from the sediment samples and does not include any results from the surface water. S19S surface water data can be found at <https://data.ess-dive.lbl.gov/view/doi:10.15485/1603775>. Future datasets from this study will include geochemical, hydrologic, and microbial data from the surface water and sediment. The S19S campaign was designed with the science community to ask questions associated with links among core/transient metabolomes, microbial metabolism, biogeochemical function, and physical properties of watershed and river corridor systems.

This dataset contains (1) high resolution characterization of dissolved organic matter from sediment via 12 Tesla Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) through the Environmental Molecular Sciences Laboratory (EMSL; <https://www.pnnl.gov/environmental-molecular-sciences-laboratory>); (2) non-purgeable organic carbon (NPOC); (3) respiration rates calculated from laboratory incubations; (4) site photos; and (5) metadata. The field metadata file (WHONDRS_S19S_metadata_v3.csv) is the same as the field metadata included in the S19S surface water data package but has had stream order added and is listed as version 3. The

Supporting and Engaging the ESS Community



Overview of Data Management Features



Overview of Data Management Features



**Dataset
Sharing**

Collaborate

**Data
Portals**

Create a
collection of
datasets

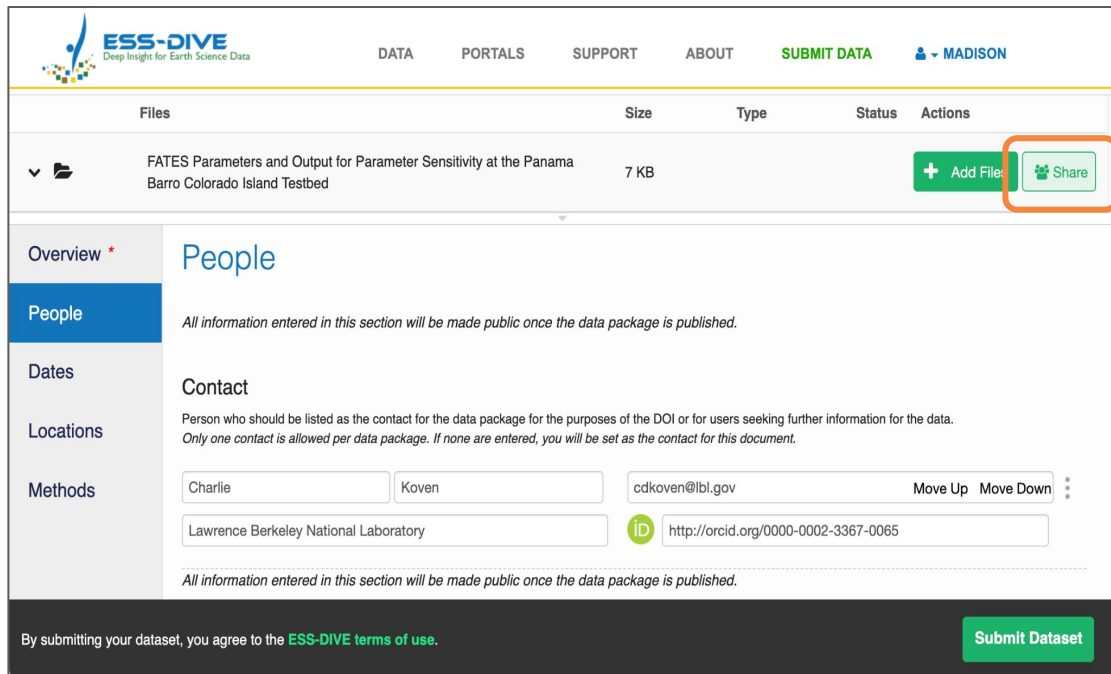
**External
Linking**

Link to other
repos and data
systems

**Project Data
Management**

Additional
project features

What is dataset sharing?



The screenshot shows the ESS-DIVE web interface. At the top, there is a navigation bar with links: DATA, PORTALS, SUPPORT, ABOUT, SUBMIT DATA, and a user profile icon labeled "MADISON". Below this is a table of datasets. The first dataset is "FATES Parameters and Output for Parameter Sensitivity at the Panama Barro Colorado Island Testbed", which is 7 KB in size. In the "Actions" column for this dataset, there is a green button labeled "Add File" and a green button labeled "Share". The "Share" button is highlighted with an orange rectangle. An orange arrow points from this button to a larger, magnified version of the "Share" button on the right. Below the table, there is a section titled "People" with a sub-section "Contact". This section contains a form for entering contact information, including a name field (with "Charlie" and "Koven" as examples), an email field (with "cdkoven@lbl.gov" as an example), and an ORCID iD field (with "http://orcid.org/0000-0002-3367-0065" as an example). At the bottom of the page, there is a "Submit Dataset" button.

Files	Size	Type	Status	Actions
▼ FATES Parameters and Output for Parameter Sensitivity at the Panama Barro Colorado Island Testbed	7 KB			Add File Share

People

All information entered in this section will be made public once the data package is published.

Contact

Person who should be listed as the contact for the data package for the purposes of the DOI or for users seeking further information for the data. Only one contact is allowed per data package. If none are entered, you will be set as the contact for this document.

Charlie Koven cdkoven@lbl.gov Move Up Move Down

Lawrence Berkeley National Laboratory http://orcid.org/0000-0002-3367-0065

All information entered in this section will be made public once the data package is published.

By submitting your dataset, you agree to the [ESS-DIVE terms of use](#).



Review and Edit
datasets with registered
team members

Dataset Sharing Permission Types

Viewer

- **view** private datasets

Editor

- **view** private datasets
- **edit** metadata
- **add/remove** files
- **publish** datasets

Manager

- **view** private datasets
- **edit** metadata
- **add/remove** files
- **publish** datasets
- **share** datasets with others

Sharing datasets - PIs and data managers

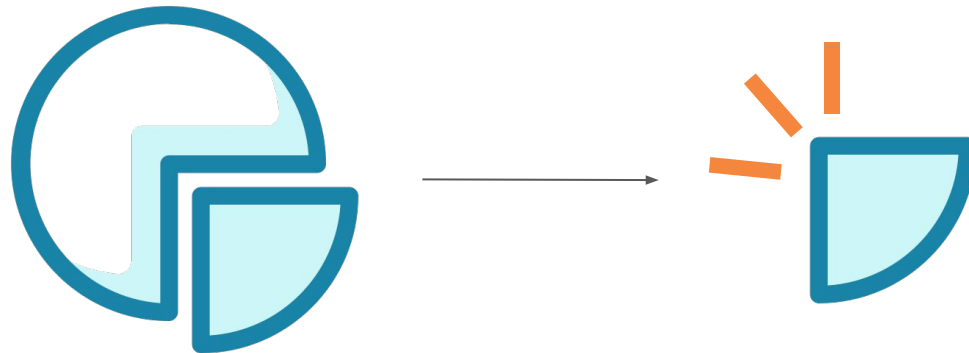
Access and report on project publication progress

- **View public and private datasets** generated by the project
 - Verify datasets are being published on time
 - Track how many datasets are on ESS-DIVE
 - Ability to review dataset content before publication
- **Manage datasets** from your project that were originally created or edited by team members

What is a data portal?



A data portal is a **collection** of any ESS-DIVE **datasets**. Data portals allow you to easily **highlight and share** related datasets.



What's inside a Data Portal?



WHONDRS Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems

Discover datasets associated with the Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems (WHONDRS)

About Data Metrics

Search

Search these datasets

DATASETS 1 TO 18 OF 18

Sort by: Most recent

Torgeson J M, Bambalkide T, Bates T L, Chu R, Crump B C, Danczak R E, Forbes B, Garayburu-Canuso V A, Goldman A E, Logozzo L, Maavara T, Martin E W, McKee S A, Raymond P A, Renteria L, Toyoda J O, Stegen J C (2022): **WHONDRS Surface Water Dissolved Organic Carbon and FTICR-MS across Stream Orders in Four United States Watersheds in 2019 and 2020**. River Corridor and Watershed Hydrochemistry SFA, ESS-DIVE repository. Dataset. doi:10.15485/1891155

Garayburu-Canuso V A, Goldman A E, Toyoda J O, Chu R, Renteria L, Stegen J C, Bengtson A, Torgeson J M, Wili K, Ross M (2022): **FTICR-MS Data from Global Inland River Water and Sediment and from Coastal River Fresh and Saline Sediment Associated with "Functional Trait Relationships in Organic Matter are Conserved in River Corridors Across Continents"**. Early Career Research Program: Watershed Perturbation-Response Traits Derived Through Ecological Theory, ESS-DIVE repository. Dataset. doi:10.15485/182422

Map showing the locations of the datasets across the world.

Collection of Datasets

WHONDRS Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems

https://whondrs.pnnl.gov

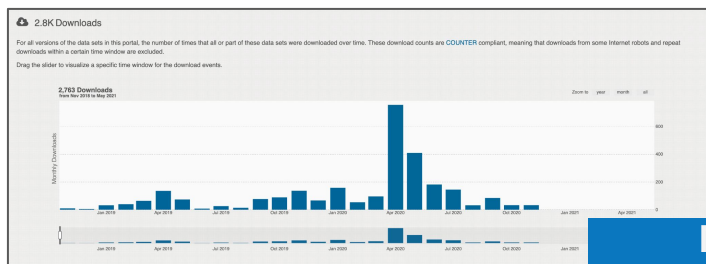
About Data Metrics

Data from the Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems

For more information about WHONDRS, go to <https://whondrs.pnnl.gov>

The Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems (WHONDRS) is a research consortium that aims to understand coupled hydrologic, biogeochemical, and microbial function within river corridors, with an emphasis on increasing accessibility and awareness and knowledge about the research life cycle. WHONDRS seeks to incentivize a global community around understanding these coupled systems from local to global scales and ultimately to

Custom Descriptive Pages



234 Citations

For all versions of the data sets in this repository, the number of citations reported to DataONE.

Christian Schöner and Michael Rother. 2018. Methanogenesis from Carbon Monoxide. Biogenesis of Hydrocarbons. pp. 1-29. https://doi.org/10.1007/978-3-319-53114-4_4-1. Cites Data: (T. J. Blasing 2013).

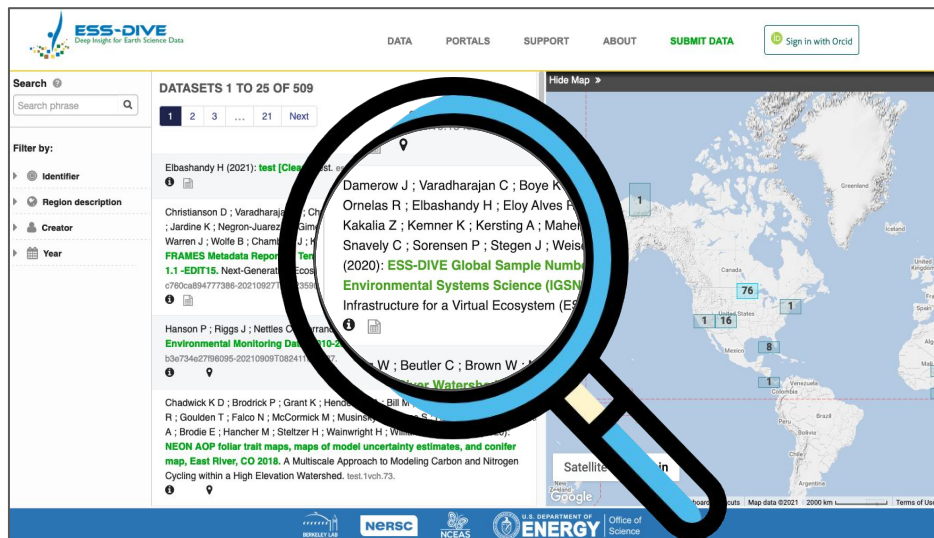
George Brett RUNION, Stephen A. PRIOR, Tyler A. MONDAY, and Janice RYAN-BOHAC. 2018. Effects of Elevated CO₂ on Growth of the Industrial Sweetpotato Cultivar CX-1. Environment Control in Biology. Vol. 56. pp. 89-92. <https://doi.org/10.2525/evb.56.89>. Cites Data: (Ralph F. Keeling et al. 2009).

Kieran M. Stanley, Aoife Grant, Simon O'Doherty, Dickon Young, Alistair J. Manning, et al. 2018. Greenhouse gas measurements from a UK network of tall towers: technical description and first results. Atmospheric Measurement Techniques. Vol. 11. pp. 1437-1458. <https://doi.org/10.5194/amt-11-1437-2018>. Cites Data: (R. Pinn et al. 2001).

R. Lakshaj-Sownder and SW. Grab. 2018. Temperature trends for coastal and adjacent higher lying interior regions of KwaZulu-Natal, South Africa. Theoretical and Applied Climatology. Vol. 38. pp. 4384-4394. <https://doi.org/10.1002/poc.5675>.

Metrics: Summary of Holdings, View/Download/Citation Metrics...

Why use data portals?



- Increase **Findability** of project datasets
- Elevate **public project identity** in ESS-DIVE
- Collect **cumulative metrics** across all project datasets
- Link to project data portals in **annual reports** and **paper publications**

Public Listing of Data Portals

A screenshot of the ESS-DIVE website's "PORTALS" page. The header includes the ESS-DIVE logo, navigation links (DATA, PORTALS, SUPPORT, ABOUT, SUBMIT DATA), and a "Sign in with Orcid" button. The main content area is titled "All portals" and lists five data portals, each with a logo, title, and brief description.

Portal Logo	Portal Title	Description
	Alpine Treeline Warming Experiment Data Portal	The Alpine Treeline Warming Experiment (ATWE) ran from 2008-2016 on Nivot Ridge, in the Front Range of the Colorado Rocky Mountains. It combined common gardens with climate manipulations, using infrared heaters to warm soil and plant surfaces by an amount comparable to current average projections...
	Carbon Dioxide Information Analysis Center	The Carbon Dioxide Information Analysis Center (CDIAC) had the primary responsibility for providing the US government and research community with global warming data and analysis as it pertains to energy issues. The CDIAC archive data holdings have been transferred predominantly to the U.S....
	Data from the East River Watershed, Colorado	This is a data portal to collect all data collected from the East River Watershed, Colorado.
	Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems	https://whondrs.pnnl.gov
	SPRUCE Experiment Data Portal	TBD - description

Your project data portal will be listed on **ESS-DIVE's portal page**

Anyone from the public can access this page and locate your data portal

Portal Sharing

- When granting a registered team member access to edit your portal, the same permission types as dataset sharing are available:
view, edit, and manage

⚙️ Settings

Portal URL *

This is a unique short name that is used in the URL so others can easily get to this portal. The URL is currently set to:
<https://data.ess-dive.lbl.gov/portals/portal-tutorial>

[Change URL](#)

Description

Describe your portal in one brief paragraph. This description will appear in the portal header and in search summaries.

This is a portal tutorial


Sharing options

Permissions

Allow specific users or groups to view or edit this portal, or change these permissions.

- View** - your team member can see this dataset when it's private. This dataset will appear on the search page when they login to ESS-DIVE.
- Edit** - your registered* team member can see this dataset when it's private, change the dataset metadata, add/remove files, publish the dataset, and edit the dataset after its published. This dataset will appear in their "My Data Packages" list.
- Manage** - your registered* team member has all the same access abilities as an editor and in addition can share this dataset with other people or remove people from the dataset permissions. This dataset will appear in their "My Data Packages" list.

*this person completed the [New Data Contributor form](#) and was approved to edit and submit metadata to ESS-DIVE

Name	ORCID	Access
Madison Burrus (You)	 http://orcid.org/0000-0002-0439-9173	Manage




⚙️ Sharing options

Permissions

Give registered ESS-DIVE users permissions to view, edit, or manage the metadata record and data files for this dataset publication.

- View** - your team member can see this dataset when it's private. This dataset will appear on the search page when they login to ESS-DIVE.
- Edit** - your registered* team member can see this dataset when it's private, change the dataset metadata, add/remove files, publish the dataset, and edit the dataset after its published. This dataset will appear in their "My Data Packages" list.
- Manage** - your registered* team member has all the same access abilities as an editor and in addition can share this dataset with other people or remove people from the dataset permissions. This dataset will appear in their "My Data Packages" list.

*this person completed the [New Data Contributor form](#) and was approved to edit and submit metadata to ESS-DIVE

Name	ORCID	Access
Madison Burrus (You)	 http://orcid.org/0000-0002-0439-9173	Manage

Search by name, ORCID, or group name

Hesham [

Hesham Elbashandy (<http://orcid.org/0000-0002-7535-8045>)

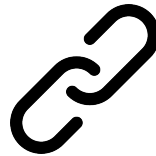
View +

[Cancel](#) [Save](#)

What is External Linking?



Data **stored on other repositories** can be connected to ESS-DIVE metadata **without needing to upload** it directly to the dataset



What does external linking look like?



Downloads 29

Citations 0

Views 24

Copy Citation

Assessment report

Files in this dataset Package: ess-dive-44b1b49c5d7a7ff-20220414T225815865

Name	File type	Size	Download All
Metadata: Snow_depth_sensor_measurement_data_for_Upper_Sub.xml	EML v2.2.0	15 KB	Download

External Links to Data or Metadata

External links for this dataset

Description	Relationship	URL
doi:10.6073/pasta/388c70359c8d0cd79dc3875d6510071d	[same as] Original publication of this dataset (where the data+metadata can be found)	https://doi.org/10.6073/pasta/388c70359c8d0cd79dc3875d6510071d

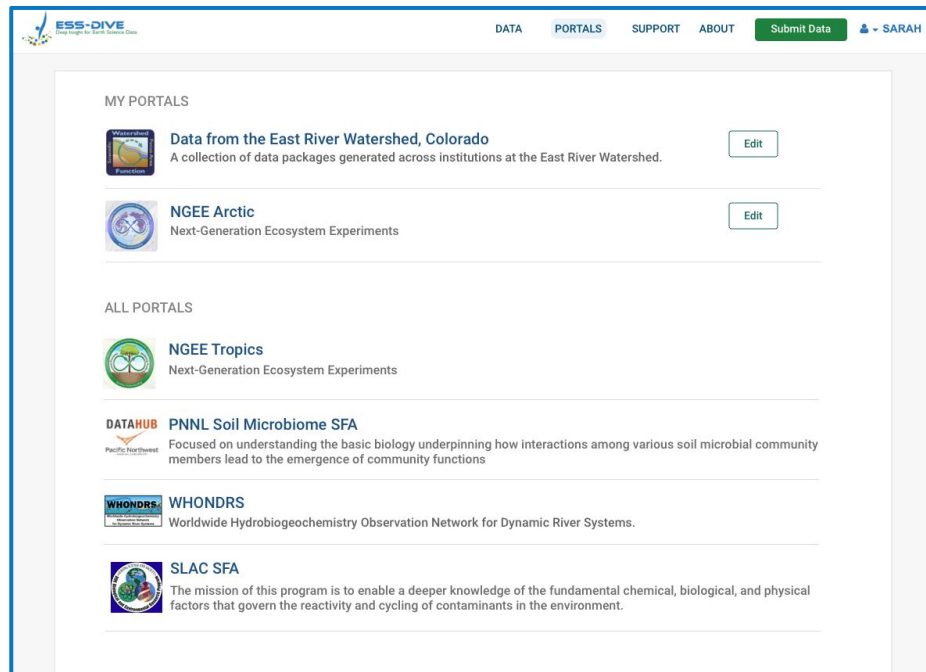
Upcoming: Project Spaces to Enable Project Data and User Management, Team Collaboration



Desired Features from User Interviews

- Creating project data **collections**
- Project **administration** of users and **curation** of data packages
- **Internal sharing within project teams** to promote collaboration

Expected Early 2023!



It's important to keep project information updated on ESS-DIVE

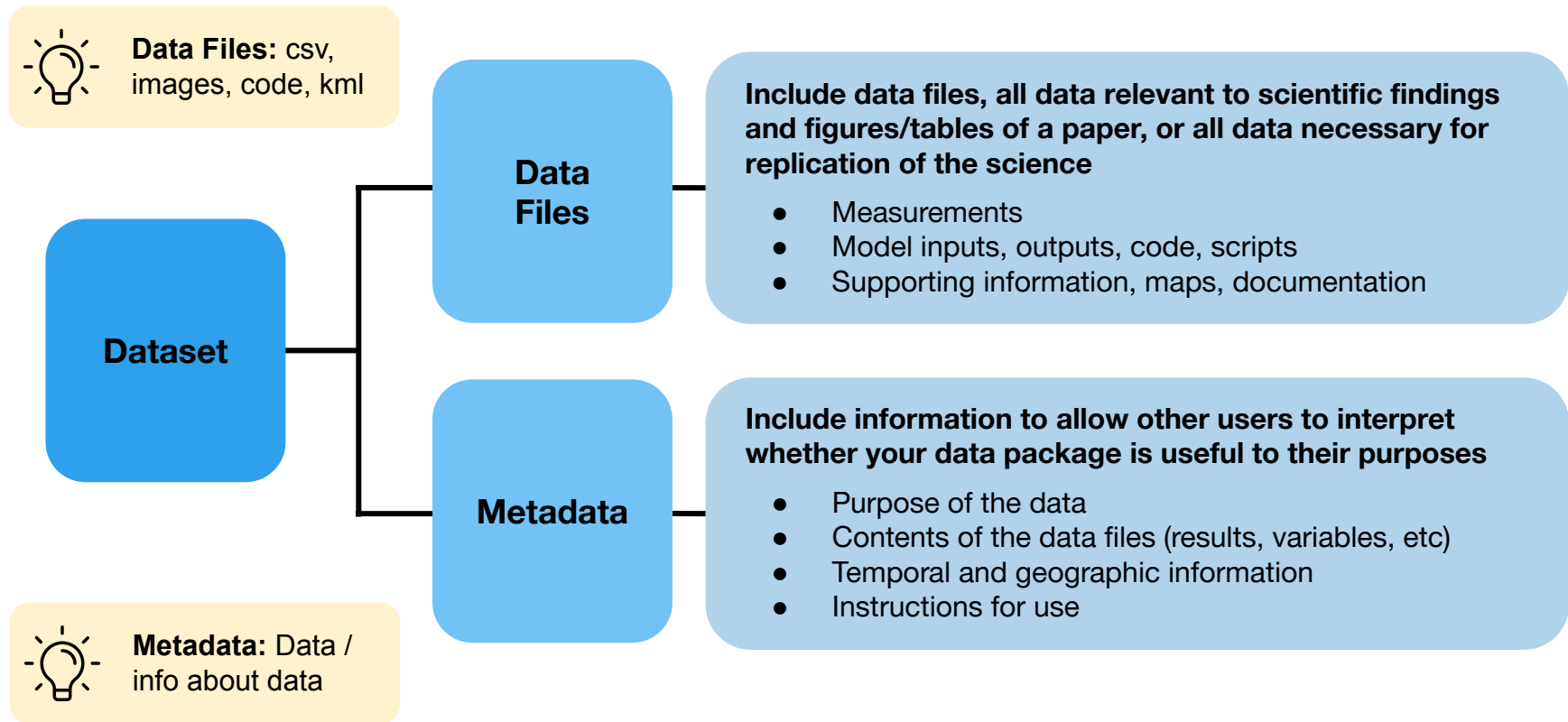


- **Project information may change over time.** The default project information provided to ESS-DIVE is when the funding was awarded.
- Don't let this affect your data! Inconsistent **project titles** can lead to incorrect data citations and make it difficult to manage project data.
- **We can help!** PIs or project managers can **send a request to ess-dive-support@lbl.gov to update project information.** Look out for further announcements from ESS-DIVE asking projects to update information.

ORGANIZING YOUR DATA

Best practices, considerations, & requirements

Components of a dataset



Choosing data to publish

Data related to paper publications

- Data that are used to create figures/tables and data that support the conclusions of paper and validate results

Data collected as part of a project

- Meet BER and other funding requirements by making data publicly available in ESS-DIVE

Publishing *scientifically useful data*

- Raw data, processed data, code, model inputs/outputs, or products built off of data

Considerations to divide data into datasets

Author contributions

Based level of contributor effort for portions of data - affects author order

Size of Data

Could impact upload method, usability/accessibility

Campaign / Time Period

Data from a field campaign or season that need to be viewed together

Data Type & Use

Particular data type from a project - e.g. continuously generated sensor data, sample data, data synthesis product
(raw data vs processed)



Reporting formats enable discovery and reuse



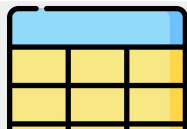
ESS-DIVE Reporting Formats ensure that

- Researchers can **find** data more easily in ESS-DIVE
- Data can be **reused** in synthesis research
- **Integrated search** that facilitates data interoperability

Benefits for data contributors, collaborators and project members, and data users.

13 ESS-DIVE Reporting Formats

Data Files



CSV Metadata

Velliquette, Heinz,
Devarakonda (ORNL)



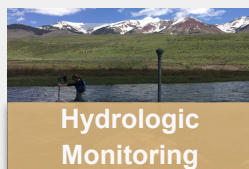
**Model Data
Archiving**
Simmonds (LBNL)



Soil Respiration
Bond-Lamberty,
Pennington (PNNL)



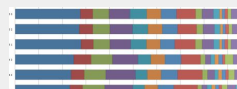
Leaf Physiology
Rogers, Ely (BNL)



**Hydrologic
Monitoring**
Goldman (PNNL)



**Water/Soil
Chemistry**
Boye (SLAC)

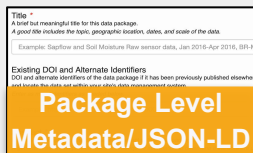


**Amplicon
Sequencing**
Weisenhorn (ANL)



Drone/UAV
Serbin, Ely (BNL)

Metadata



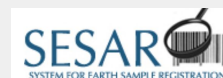
**Package Level
Metadata/JSON-LD**
Agarwal,
Hendrix (LBNL)



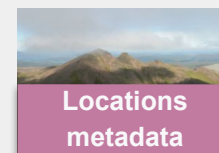
**Package Metadata
Quality**
Damerow (LBNL)



**File-level
Metadata**
Velliquette, Heinz,
Devarakonda (ORNL)



**Sample
IDs/Metadata**
Damerow (LBNL)



**Locations
metadata**
Crystal-Ornelas
(LBNL)

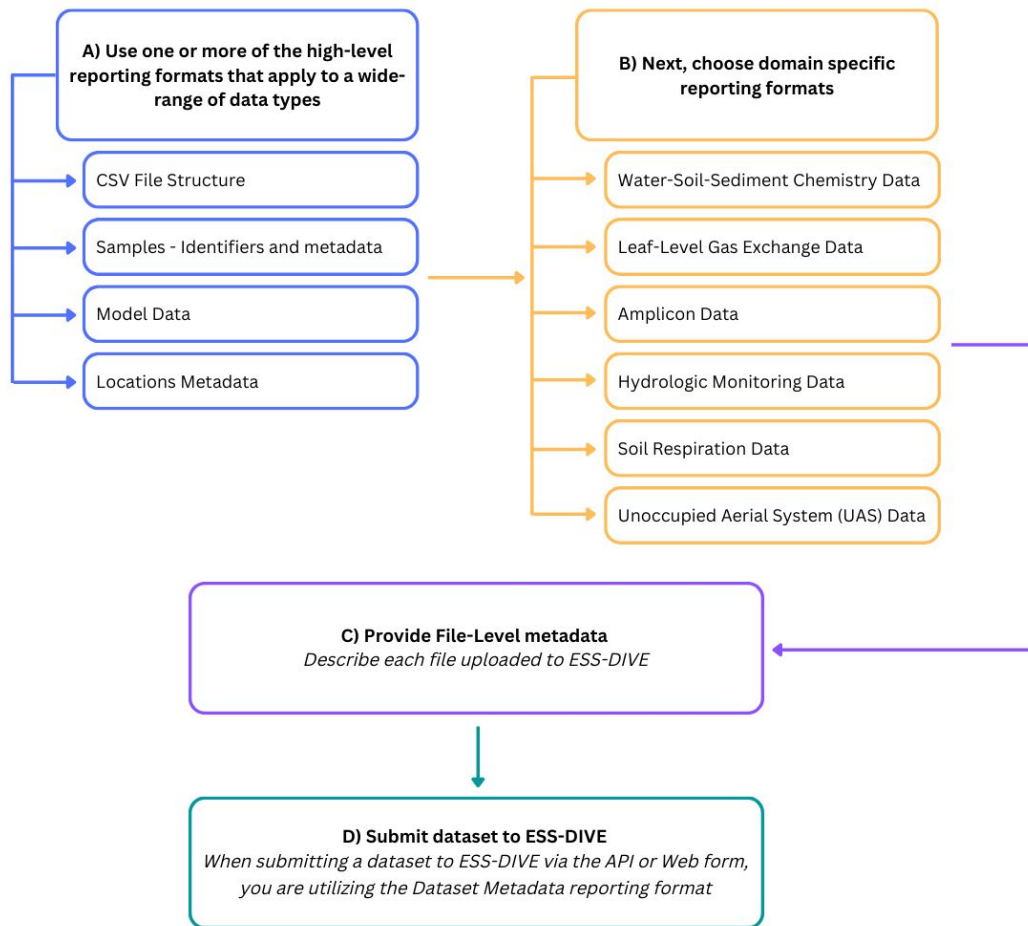
Crystal-Ornelas, R. et al. Enabling FAIR data in Earth and environmental science with community-centric (meta)data reporting formats. *Sci Data* 9, 700 (2022). <https://doi.org/10.1038/s41597-022-01606-w>

Choices for describing data



Choose high-level reporting formats, *then* domain-specific where applicable.

Package-level metadata is **required** for all ESS-DIVE datasets.



File-level metadata (FLMD)



What are file-level metadata?

- Granular information at the data file level (e.g., file name & description, start and end dates)

Why provide file-level metadata?

- Data users will have general understanding of info contained within each file
- FLMD can enable automatic parsing of data files so that users can eventually search & locate files across data collections



**Terri Veliquette,
ORNL**

File-level metadata example

Can use wildcard
* to indicate entry
that applies to
multiple files

Required information

Optional information

	A	B	C	D
1	File_Name	File_Description	Standard	UTC_offset
2	soil_samples_*.csv	15 soil samples taken in the summer of 2019 using small hand trowel and soil probe.	csv v1.0	- 5 hours
3	SoilPoreWaterHillslope2019.csv	50 soil pore water samples taken from the hillslope at the site over a one year period.	EPA	- 5 hours

CSV Reporting Format

What is the CSV reporting format?

- The CSV file is a non-proprietary format for tabular data
- Archives tabular data in its simplest form
- Defines structure and some content

Why use the CSV reporting format?

- Specifies common format for elements within your CSV files (e.g., missing values) which make CSVs easier to read
- Reduces inconsistencies (e.g., 2021-04-26 vs. 4/26/2021)



**Terri Veliquette,
ORNL**

Documentation and Community Space



A screenshot of the GitHub organization page for ESS-DIVE. The header shows the organization's name, a description, location (Berkeley, CA USA), website (http://ess-dive.lbl.gov), Twitter handle (@ESSDIVE), and email (ess-dive-support@lbl.gov). Below the header is a navigation bar with tabs: Overview, Repositories (17), Packages, People (5), Teams, Projects, and Settings. The main content area displays three pinned repositories: 'essdive-community-space-guide' (Public), 'essdive-community-workshop-2021' (Public), and 'reporting-format-template-repo' (Public template). Each repository card includes a brief description and a star count (2, 1, and 0 respectively).

View all reporting format **templates** and **instructions** for use

GitHub is primary location for **providing feedback**

GitHub: <https://github.com/ess-dive-community>

Website version: <https://docs.ess-dive.lbl.gov/contributing-data/data-reporting-formats>

How to Connect With Our Team



To get help:

Webinars: <https://ess-dive.lbl.gov/webinars/>

Guide to using ESS-DIVE: docs.ess-dive.lbl.gov



Support email: ess-dive-support@lbl.gov

To stay updated:

ess-dive-community@lbl.gov



[@essdive](https://twitter.com/essdive)

<https://bit.ly/essdiveMailingList>



Acknowledgements

Advisory Groups: ESS-DIVE Archive Partnership Board, ESS Cyberinfrastructure Working Groups

Funding: EESSD Data Management



Questions?

Urban IFL Data Management Needs

What data types are you planning to collect?



Approximately how many datasets do you anticipate publishing across the project in a year?



**What challenges do you face when creating datasets, or
what challenges do you anticipate as a new user?**



Urban IFL Data Management Needs

1. Are you interested in using existing standards or formats for your data (e.g. any of the ESS-DIVE reporting formats)? If yes, which ones?
2. Considering the data type responses, is there need to develop standard formats for any specific data type, particularly relevant across the Urban IFL projects?
3. Who manages and publishes data on your project (e.g. how many team members, and what is their role on the project)?
4. What tools are you planning to use for data management (e.g. any of the ESS-DIVE features mentioned in the presentation, or others)?
5. What challenges have you encountered or anticipate in managing data for your project, team or individual research?