Welcome!
ESS-DIVE Tutorial Preparation

If you want to participate in today’s hands on activities, please use this time to:

Create an ORCiD
https://orcid.org/register

Sign in to Sandbox

Type the following in the chat when you’re done!
“1 - <your ORCID>”
Create an ORCiD

The ESS-DIVE Sandbox and Main Portal both require an ORCiD account to login and submit or manage data.

https://orcid.org/register
Navigate to ESS-DIVE Sandbox

Sandbox is a test system where you can practice using any feature. This data is **NOT** preserved in ESS-DIVE.

Login with ORCiD account
Add email address to your profile
Project Data Management on ESS-DIVE

data.ess-dive.lbl.gov
Presenters

Joan Damerow
Community Engagement Lead Scientist

Charuleka Varadharajan
ESS-DIVE co-PI
Have you used ESS-DIVE before?

- Yes, I have stored data in ESS-DIVE before
- Yes, I have downloaded data from ESS-DIVE
- No, I have not logged into ESS-DIVE
What is your primary role in your research group?

Field Scientist/Data Collector
Lab Scientist/Data Analyzer
PI
Data User
Data Manager
What is a primary challenge that you have encountered in managing or publishing data?
Tutorial Overview

**Part 1:** About ESS-DIVE and Data Management Plans

**Part 2:** Learn how to use ESS-DIVE features for managing project data

Feedback on Project Data Management Features and Needs

*This session is designed for data managers and ESS Project PIs*
Lunch Tutorial and Posters

**UP NEXT: Creating and Publishing High-Quality Datasets**

1:00 PM – 2:30 PM ET / 10:00 AM – 11:30 AM PT

Learn how to upload and publish data with ESS-DIVE as a new contributor in this hands-on training session. **Presenter:** Emily Robles

**Posters**

ESS-DIVE Overview: A Scalable, User-focused Repository for Earth and Env Science Data **Presenter:** Shreyas Cholia

Community-developed (meta)data reporting formats to enable data reuse in ESS-DIVE, **Presenter:** Charu Varadharajan
Including ESS-DIVE in your Data Management Plan
What is ESS-DIVE?

Data Repository to preserve expand access to and improve usability of ESS data

data.ess-dive.lbl.gov

780 public datasets

Use of ESS-DIVE is mandatory for ESS project data management plans
## Who is eligible to store data with ESS-DIVE?

<table>
<thead>
<tr>
<th>Can I store data with ESS-DIVE?</th>
<th>ESS Program Funding</th>
<th>Other Funding Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Yes</strong>, you can start storing data right away</td>
<td></td>
<td><strong>No</strong>, you need to request permission to upload data to ESS-DIVE</td>
</tr>
</tbody>
</table>

### Rationale?

- **ESS Program Funding**
  - All projects funded by the ESS program should store or mirror data on ESS-DIVE
- **Other Funding Source**
  - Must provide rationale for why project data is relevant to ESS program

### How do I start storing data?

- **ESS Program Funding**
  - Fill out [New Data Contributor form](#)
- **Other Funding Source**
  - Fill out [New Project Request form](#)

### Proposals

- **ESS Program Funding**
  - Review terms of use and include in DMP. No need for a letter of support.
- **Other Funding Source**
  - Contact for permission to include in your proposal

*If your project request is approved we will follow up with instructions on how to start uploading data.*
Using ESS-DIVE in a Data Management Plan (DMP)

- Data Collected, Generated, or Used
- \textbf{Standards and reporting formats}
- \textbf{Data Sharing and Preservation}
- Related tools, software or code
- Data protection: Security and Integrity
- DMP Oversight
- Rationale for the proposed data management plan

ESS-DIVE is DOE ESS’s central location for long-term data preservation
DMP: Data Sharing and Preservation

● **ESS-DIVE provides**
  - Storage of *diverse data* types
  - Public release of data with digital object identifier (*DOI*)
    - Pre-publish DOIs may be available on request
    - **Citation** text for published data
  - Long-term data *search and download* availability
  - **Curation** of dataset metadata and basic file checks

● **ESS-DIVE does not provide**
  - Enforcement of data publication timelines
  - Guidance/obtaining permissions on what project data can be shared
Archiving data in multiple data repositories

At minimum ESS projects need to store metadata on ESS-DIVE

Options if you have data in other repositories:

- Replicate the **data and metadata** at ESS-DIVE and the other repository
- Replicate the **metadata** at ESS-DIVE and **link to the data** at the other repository
- Store the **data and metadata** in ESS-DIVE and **link to the data** ESS-DIVE from the other repository
- Use the same or a different **DOIs**
DMP: Data and metadata standards

- **ESS-DIVE provides**
  - Metadata standards for data packages
  - Community reporting formats for several ESS data and metadata types
  - Basic guidance for organizing data packages and authorship
  - Checks of metadata quality (FAIR scores) and that files are readable

- **ESS-DIVE does not provide**
  - Project is expected to provide data formatting and quality checking.
    Include data reporting formats early in your scientific workflows
Using ESS-DIVE for your project: Terms of Use

**Step 1:** Register as a new data contributor (https://docs.ess-dive.lbl.gov/new-user-registration)

**Step 2:** Submit data in relevant community standards and apply QA/QC

**Step 3:** Provide data package documentation & appropriate standardized metadata

**Step 4:** Agree to the data contributor license (Part B) for serving data to the public.

ESS-DIVE supports Creative Commons Attribution 4.0 and Creative Commons Public Domain policies for data usage rights
NEW Project Data Management Features
Overview of Data Management Features

*NEW* Project Data Management Features

- Manage project information, data managers, and teams

Dataset Sharing
- Collaborate on a single dataset

Data Portals
- Create a collection of datasets

External Linking
- Link to other repos and data systems
Overview of **NEW** Data Management Features

These project management features are **not set up by default** for projects. PIs must initiate the process.

We will describe what these features are and learn hands-on how to set them up:

1. ESS-DIVE’s Project List
2. Project Data Manager Role
3. My Projects
4. Teams

**Additional Resources:**

Webinar June 26, 10-11 am PT

[Project Data Management documentation]
ESS-DIVE’s Project List

Is a searchable list of all project’s approved to publish data on ESS-DIVE.

- Search for project’s by title, PI, or identifier
  - Must login to view and search by PI name
- Look-up project identifiers when submitting datasets with the API

https://data.ess-dive.lbl.gov/projects
ESS-DIVE’s Project List

This list does away with inconsistent project titles and provides a controlled list with unique identifiers to ensure your dataset citations are consistent within ESS-DIVE.

Your project must be on this list before you can submit data

- **ESS** Projects are approved by default
- **Non-ESS** projects are added once approved by ESS program
A project data manager **is a role** in ESS-DIVE that enables you to **access** project information and teams.

- **Become a project manager**
- **Gain access to your project page**
- **Create teams for your project**
Project Data Managers

Assigning project data managers for your project is an important starting point to leverage ESS-DIVE’s growing suite of project management tools.

Who can become a project manager?

- **PIs** have automatic approval to become project managers.
- **Any other project member** must first be approved by their PI.

If you are on multiple projects, you will need to be assigned project manager for each project.
Last Call

ESS-DIVE Tutorial Preparation

If you want to participate in today’s hands on activities, please:

Create an ORCiD
https://orcid.org/register

Sign in to Sandbox

Type the following in the chat when you’re done!
“1 - <your ORCID>”
My Projects

Lists **all projects you can access** as a project data manager

- The “My Projects” page is only visible to project managers
- Access additional features from this page by selecting a project

This is a quick way to check if you are a project manager on any projects!

---

<table>
<thead>
<tr>
<th>Project Title</th>
<th>PI(s)</th>
<th>ESS-DIVE Project Identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental Systems Science Data Infrastructure for a Virtual Ecosystem (ESS-DIVE)</td>
<td>Deborah Agarwala, Charchitava Vardharajap, Shreyasi Cholia</td>
<td>153be88f-224c-4d9a-9f40-1e0bafeda25bde6</td>
</tr>
</tbody>
</table>
After selecting a project, you’ll be directed to the project information page.

This page lists **important metadata** about your project:

- Principal Investigator(s)
- Project Data Managers
- Sponsor Program

Use this to review who has project manager access to your project!
My Teams

Select the My Teams tab to:

- **View** teams you are a member of
- **Create** teams for your project
- **Manage** team members

Projects can have multiple teams. If you are **not a member** of a team, you **won’t be able to view** that team.

Project managers cannot view all project teams unless they are added to all teams.
What are teams?

Teams are groups of ESS-DIVE users.

The purpose of teams is to provide a way to share data permissions with a group of people at once and support collaboration.

What can teams be used for?

- **Data managers** who need to manage project dataset permissions and publication
- **Collaborators** who need to edit the same datasets together
- **Reviewers** who need to view private datasets

Example of a team

Project managers is not a team and does not behave like a team.
Overview of Roles on ESS-DIVE

**ESS-DIVE User**
Someone who logged into ESS-DIVE using their ORCID.

**Data Contributor**
An ESS-DIVE user who was approved by ESS-DIVE to upload data; they can create, edit, and publish datasets.

**NEW: Project Data Manager**
A data contributor approved by a project PI to create and manage teams for that project.

**NEW: Team**
A group of ESS-DIVE users who share access permissions to datasets and data portals.

Icons from [https://www.flaticon.com/authors/juicy-fish](https://www.flaticon.com/authors/juicy-fish) and [https://www.flaticon.com/authors/freepik](https://www.flaticon.com/authors/freepik)
Time to practice!
Set up Project Management

Project management features are not set up by default for any project. PIs must initiate the process.

The hands-on tutorial section is designed to help PIs get started. We will learn how to set up and use each of the four features we discussed.
Time to practice!
Set up Project Management

On ESS-DIVE Sandbox

Sandbox is a test system where you can practice using any ESS-DIVE feature. Data on Sandbox is NOT preserved in ESS-DIVE.

Step 1: Ensure project information is accurate

Project information may change over time. The default project information provided to ESS-DIVE is when the funding was awarded.

Use the project list to:

✓ Verify that your project is in ESS-DIVE’s Project List and approved to store data
✓ Review the project title and PI info used in dataset citations
✓ Contact ESS-DIVE if you need to update any project information ess-dive-support@lbl.gov

https://data.ess-dive.lbl.gov/projects
Step 2: Become a Project Data Manager

During the tutorial:
- We’re using **Sandbox**
- Quick and **approval isn’t necessary**!
- **Same process** for everyone
- We will be assigned to a **test project** called “ESS PI Meeting 2023 Project”
- Everyone who entered a “1” in the chat has already been assigned as a project data manager **to the test project**

After the tutorial:
- You’ll use **ESS-DIVE’s main repository**
- **PI approval required** and may take time
- PIs have a **different process** than other project members
- You’ll be assigned to your real project

Step 2: Become a Project Data Manager on ESS-DIVE

Let the ESS-DIVE Team know who should have permission to manage your project.

- **PIs and Co-PIs:**
  - ✓ Register as a data contributor
  - ✓ Notify ESS-DIVE Support

- **All other members of project:**
  - ✓ PI must fill out the Add Project Manager form to grant you approval
  - ✓ Register as a data contributor

https://data.ess-dive.lbl.gov
Step 2: Become a Project Data Manager on ESS-DIVE

Let the ESS-DIVE Team know who should have permission to manage your project.

- **PIs and Co-PIs:**
  
  **Good News!** Any PI who registered as a data contributor by May 5th is **already a Project Data Manager** on their project.

  We recommend continuing the tutorial on Sandbox, but you have the option to use our main data repository: [https://data.ess-dive.lbl.gov](https://data.ess-dive.lbl.gov)
Step 2: Become a Project Data Manager on Sandbox

Confirm that you are a project manager on the tutorial test project:

✓ Login to Sandbox
✓ Navigate to the project list
✓ Select “My Projects” and look for “ESS PI Meeting 2023 Project”

Don’t see the “My Projects” tab? Type a “0” in the chat and our team will try to add you if there’s time

Step 3: Visit your Project Page

From the “My Projects” page, select “ESS PI Meeting 2023 Project”

Use your project page to:

✓ **Confirm** who your project data managers are

✓ **Access** teams page
Step 4: Create a Team for your Project

Creating a team involves 3 important components:

✓ Consider **who you want** to be on the team and **what the team for**

✓ Carefully choose a **team name**

✓ Choose which team members should **co-own your team**

**Using Sandbox to create test teams first** may help your project figure out how you want to use teams!
Step 4: Create a Team for your Project

What can teams be used for?

You may be interested in creating teams if you have ever wanted to have groups of:

- **Data managers** who need to *manage* project dataset permissions and publication
- **Collaborators** who need to *edit* the same datasets together
- **Reviewers** who need to *view* private datasets

**Who can I add to a team?** Anyone who has logged into ESS-DIVE with their ORCID
Step 4: Create a Team for your Project

Why do we need to be careful with team names?

Teams cannot be deleted or renamed later

ESS-DIVE’s recommendations for naming teams:

- Use common **prefix** for all teams in your project
- Use the intended purpose or access level* as **suffix**
- Have a team just for project data managers

*Access levels will be discussed in the Sharing Data section of this tutorial

Team Name Examples

- ess-dive-view
- ess-dive-edit
- ess-dive-reviewers
- ess-dive-admin
- ess-dive-data-managers
Step 4: Create a Team for your Project

Why choose team co-owners?

Co-owners can add and remove team members over time as people enter or leave your project to determine who should access team datasets.

- The project manager who created the team is the only team owner at first.
- Add at least one co-owner after creating a team by starring their name.
Step 4: Create a Team for your Project

All teams have a profile

Profiles are used to view:

- **Team members**
- **Datasets shared** with that team
- **Cumulative metrics** of datasets shared with that team
Step 4: Create a Team for your Project

Time to create a team!

1. Select the My Teams page
2. Click the “+ Add Team” button at the bottom of the page
3. Carefully choose a team name
   *Reminder: teams cannot be deleted or renamed later*
4. Search for team members by name or ORCID to add them
5. Select “Create Team”
6. A green pop-up message will appear, click the **View it Now** link to visit the team profile and ensure your team was successfully created
7. **Return** to your My Teams page
8. Add at least one **co-owner**

To enable full visibility and access to teams for project managers:

**Add and star all project managers** on every team you create
Review the project management set up steps

You can review these steps at any time in our documentation

1. Ensure project information is accurate
2. Assign project data managers
3. Visit your project page
4. Create a team for your project

https://docs.ess-dive.lbl.gov/manage-data/manage-project-data
Review the project management set up steps

You can review these steps at any time in our documentation

Please keep in mind that if you used Sandbox for this tutorial, you will need to repeat these steps on ESS-DIVE’s main data repository

https://docs.ess-dive.lbl.gov/manage-data/manage-project-data
What’s Next?

What can we do now that we have teams?

- **Project data managers** can encourage data contributors to start sharing datasets and portals with teams
- **Data contributors** can share their datasets with their project teams

Eventually, your teams will have access to a list of datasets. Project managers can update team members over time as people enter or leave a project.
Share Datasets
Assign and View Dataset Permissions
What is Dataset Sharing on ESS-DIVE?

Dataset creators can choose who should view, edit, and manage their datasets.
## Permission Types

<table>
<thead>
<tr>
<th>Role</th>
<th>Permissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viewer</td>
<td>• view private datasets</td>
</tr>
<tr>
<td>Editor</td>
<td>• view private datasets • edit metadata • add/remove files • publish datasets</td>
</tr>
<tr>
<td>Manager</td>
<td>• view private datasets • edit metadata • add/remove files • publish datasets • share datasets with others</td>
</tr>
</tbody>
</table>
Who can I share datasets with?

Data contributors (i.e. the person who created the dataset) must initially share their dataset.

Share data permissions with:

- Individual ESS-DIVE users (e.g. Co-authors, Collaborators, PIs, Data Managers)
- Teams (e.g. a team of project data managers)

Project data managers do not have default permissions to any datasets
Why share datasets?

- **Collaborate** with team members and co-authors
- Provide PIs and data managers **access to project datasets**
  - Manually share project datasets to view all public and private datasets together

**Benefits of sharing with teams:**

- Easy to share with a **consistent group, instead of** one person at a time
- **Update team members** over time, **instead of** individual dataset permissions
  - As people enter or leave a project, manage the team members who can access project datasets
Sharing datasets with teams

**Data contributors** should start sharing with data managers at the time they submit a new dataset.

**Data managers** will then be able to access shared datasets from the team profile* right away.

*If you share view access, the dataset will not show up on the team profile*

Consider sharing manage access with data managers to allow data managers to independently update data permissions in the future.
Time to Practice!
Assign dataset permissions

https://data-sandbox.ess-dive.lbl.gov/data

1. Submit new or open existing dataset
   *new datasets cannot be shared until after initial submission*
2. Select the “Edit” button then the “Share” button
3. Grant permission to one team and one individual
4. Click Save
5. Click Submit Dataset

Additional Resources:
Data Sharing Tutorial Webinar
Time to Practice!
Find your datasets

Datasets you can edit, manage or have created:
1. Hover over your name in the top-right corner and select “My Datasets” from dropdown

Datasets you can view:
1. Clear any search filters
2. Expand the “Access” search filter
3. Select “Private datasets only”

Datasets you have view permission to will not show up under “My Datasets” or on team profiles

https://data-sandbox.ess-dive.lbl.gov/data
Feedback - What do you think?
Policy question: Who should have access to project datasets by default?

- Only the dataset creator
- Co-authors
- PIs
- Project Data Managers
- Other
Discussion on project management workflow

Does the project managers, teams, default sharing system work for you?

Was anything confusing or unexpected?

Summary of how project data access works currently:

- PIs must approve project data managers
- Project data managers are not automatically on a team
- Can’t view all project teams until you are on all teams
- Datasets must be shared by the dataset creator
  - PIs and data managers do not have automatic access to datasets
Create a Data Portal
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.
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**
Data Portals: provide custom view to showcase your project data

Data tab:

This data portal creates a view of 226 datasets related to a specific research topic

Can see public and private datasets when logged in
What’s inside a Data Portal?

Collection of Datasets

Custom Descriptive Pages

Metrics: Summary of Holdings, View/Download/Citation Metrics...
Explore Data Portal Metrics

- Total number of datasets
- Volume (bytes) of metadata records, data files
- Most recent date updated
- File types
- Years data collected
- Views, Downloads, Citations

https://data.ess-dive.lbl.gov/portals/ngeetropics/Metrics
Sharing Data Portals

- Share permission to view, edit, or manage a portal
  - **Edit page** markdown,
  - **Add/remove pages**, and
  - **Add/remove data filters**

- **Does NOT** give access to **datasets** in portal

- Share data portal with individuals or team(s)
Public Listing of Data Portals

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
Time to Practice!
Create and share a portal

1. Select “Portals” and “New Portal” button
2. On the Settings tab: add url, logo(s), title, and description
3. Share the portal with individual or team
   a. Only people who have logged into ESS-DIVE using their ORCID can be found in the Sharing Options

https://data-sandbox.ess-dive.lbl.gov/portals
Time to Practice!
Add datasets to portal using data filters

Create portals based on research theme, data type, project, etc. using rules

You try: Create an ESS Project portal. Add datasets to your portal using any combination of 2-3 filters. TIP: NGEE Tropics test datasets

1. In Data tab - scroll to “Change the data in your collection”
2. Add a Rule
   a. Under Awards and Funding select “Project” as one of the rules
   b. Enter project name exactly and select “equals” operator
Portal publication

- Portals are automatically set to private
- *To release your portal to the public, contact* ess-dive-support@lbl.gov
- After it’s public, you can provide a link to your portal in publications, etc.
Linking to External Project Data
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 is “External Linking”?

- Store your data where it makes the most sense scientifically or for practical reasons
- **Metadata** will be on ESS-DIVE
- Original dataset will be **searchable** on ESS-DIVE using metadata
- Data can only be externally linked if stored on an approved repository
- External dataset metadata can be bundled into a **Data Portal**
What does external linking look like?

External Links to Data or Metadata

<table>
<thead>
<tr>
<th>Description</th>
<th>Relationship</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>doi:10.6073/pasta/388c70359c8d0cd79dc3875d6510071d</td>
<td>[same as] Original publication of this dataset (where the data+metadata can be found)</td>
<td><a href="https://doi.org/10.6073/pasta/388c70359c8d0cd79dc3875d6510071d">https://doi.org/10.6073/pasta/388c70359c8d0cd79dc3875d6510071d</a></td>
</tr>
</tbody>
</table>
Why include external links?

The ESS publication policy requires replicating datasets on ESS-DIVE even if the data are already published with a DOI.

Consider using this feature if:

- Your data is spread across established repositories (e.g. NMDC, EDI, Zenodo) and you don’t want to replicate existing public datasets.
- You want to establish links between data on different repositories.
- Individual files in a dataset are more useful when viewed from a particular data or analysis platform (e.g. google earth engine).
Use Cases to Store your Data Externally

Scientific, practical, or other funding compliance reasons

1) Original Dataset Publication
   - Approved Data Repositories
     - EDI
     - USGS ScienceBase

3) Individual Files
   - Data Analysis Platform
     - Github
     - Google Earth Engine
   - Large Data
     - File size
     - Number of files

1) Complete Copy of Data Files
   - Large Project Data Archives
     - SPRUCE
     - Watershed Function SFA
     - NGEE Tropics

same as
archived at
Documentation Provides How-To Instructions

Manage Project Data
https://docs.ess-dive.lbl.gov/manage-data/manage-project-data

Share Data Permissions
https://docs.ess-dive.lbl.gov/manage-data/share-data-permissions

Portals
docs.ess-dive.lbl.gov/portals/why-use-data-portals

External Linking
docs.ess-dive.lbl.gov/contributing-data/link-to-external-data-sources
Summary: Evolving Project Data Management Tools

- Project-specific data search
- Enable project-centric data portals
- Project data citations and metrics
- Assign project data managers and create teams that can change over time
- Share datasets and portals with project members and teams to enable collaboration
Questions?
Lunch Tutorial and Posters

UP NEXT: Creating and Publishing High-Quality Datasets
1:00 PM – 2:30 PM ET / 10:00 AM – 11:30 AM PT

Learn how to upload and publish data with ESS-DIVE as a new contributor in this hands-on training session. **Presenter:** Emily Robles

**Posters**

ESS-DIVE Overview: A Scalable, User-focused Repository for Earth and Env Science Data **Presenter:** Shreyas Cholia

Community-developed (meta)data reporting formats to enable data reuse in ESS-DIVE, **Presenter:** Charu Varadharajan
Connect With Our Team!

To get help:
ess-dive.lbl.gov
ess-dive-support@lbl.gov
docs.ess-dive.lbl.gov

To stay updated:
ess-dive-community@lbl.gov
@essdive

Acknowledgements
Advisory Groups: ESS-DIVE Archive Partnership Board, ESS Cyberinfrastructure Working Groups
Funding: EESSD Data Management