

Shrubwise Data Challenge

Winter/Spring 2026



Agenda

- Welcome – Dr. Ilkay Altintas
- Data Challenge Background – Dr. Leticia Lee
- Perspective from Chris Anthony
- Perspective from Dr. Adam Atchley
- Data Challenge Schedule – Pedro Ramonetti
- Onboarding Overview – Pedro Ramonetti
- Next steps / Q&A – Claire Stirm

Welcome to the **Shrubwise Data Challenge!**

The largest and most contiguous patches of shrublands are found in the Southern California region.

Shrubs are also hidden in higher canopy fuel areas which makes prescribed burns and modeling wildfire behavior challenging.

We need your help identifying ways to map shrub distributions.

Background

- Low-lying woody vegetation (shrubs) is an essential driver of fire behavior
 - Of great importance in Southern California, shrub-dominated
 - Spatial arrangement of shrubs play a critical role in how fire spreads
- Traditional fuel models (2D models) have focused on forest ecosystems and have ignored shrubs
- New fire models (3D models) can more readily take into account shrubs
 - There is a need for high resolution 3D vegetation inputs with a focus on shrubs (i.e. the spatial and vertical distribution of shrubs)

Challenge Description

- Goal: Map shrub distributions at a high spatial resolution
 - Ideally 1 m resolution
- Build a machine learning pipeline that accurately characterizes shrub ecosystems

Example AI Workflow

Overview of steps in a machine learning pipeline

Training Labels

Define the target outputs your model will predict (e.g., shrub cover and height) and develop reliable label products to train against.

Training Data

Prepare and curate your inputs by integrating multiple data sources, cleaning artifacts and noise, and validating data quality and alignment.

Model Development

Train and refine your ML/AI model using the curated dataset, iterating on features, architecture, and hyperparameters to improve performance on target metrics.

Model Application

Apply the finalized model to generate predictions at scale, producing a new high-resolution shrub data product ready for analysis, visualization, and sharing.

Chris Anthony

Fellow at the Societal Computing and Innovation Lab (SCIL) at UC San Diego.

- Former Chief Deputy Director of CAL FIRE
- Interested in connecting innovative technologies to the wildfire community to protect lives, communities, ensure resilient landscapes



Adam Atchley

Research Scientist at
Los Alamos National
Laboratory

- Research interests in wildland fire, hydrology, land surface energy balances, and ecosystem response to climate change
- Developed physics-based representation of wildland fire disturbances imbedded within ecosystem process models.



Schedule - Enrollment

- **Friday, Feb 20:** Register your team to receive priority support
 - All team members must register with Wildfire Commons; otherwise, attempting to add a non-registered email will result in an error.
 - Registration should be completed using an academic email (i.e., a @university.edu domain) or OrcID. Other users will need additional support to access computing resources.
 - Anyone can participate
 - No participant can be part of more than 1 team
- **Friday, Apr 3:** Last day to register team

Phases

Suggested pace for data challenge

Onboarding

You will have completed onboarding in FireForge, learned how to launch the JupyterHub service, and practiced working collaboratively in shared workspaces.

Introductory Data

You will have explored core datasets and foundational concepts needed to understand the challenge goals.

Relevant Data Products

You will have identified additional relevant data products and learned how to incorporate them into your pipeline.

Workflow Development

You will have built and documented an end-to-end ML/AI pipeline to characterize shrubs.

Provided Datasets

- 10m and 30m products from the Rangeland Analysis Platform (RAP)
- Sentinel 2 Imagery
- USGS 3DEP Aerial Lidar (ALS)
- InteliMon Terrestrial Laser Scanning (TLS)
- National Agriculture Imagery Program (NAIP)
- BurnPro3D UC Climate Action Seed Grant field data
- 1m, 10m, 30m NASA WERK California Lidar-Derived Product

You are encouraged to use additional open-source data in your pipeline.

Schedule – Module Releases

The Data Challenge consists of module releases, each with specific tasks to complete. The tasks for each sprint will be detailed in the instructions provided on the following dates:

- Thursday, January 22, 2026 – Onboarding Workspace
- Thursday, February 19, 2026 – Sprint 1
- Thursday, March 5, 2026 – Sprint 2
- Thursday, March 19, 2026 – Sprint 3 and Sprint 4

Final Deliverable and Judging

Deadlines:

- Each team can request feedback or assistance on sprint 1-3 modules during office hours
 - Sprint 1-3 Modules due Sunday, April 5, 2026 by 10 pm Pacific
- Final deliverables will be evaluated by a panel of judges against a set of criteria.
 - Sprint 4 Final deliverable due Sunday, April 19, 2026 by 10 pm Pacific

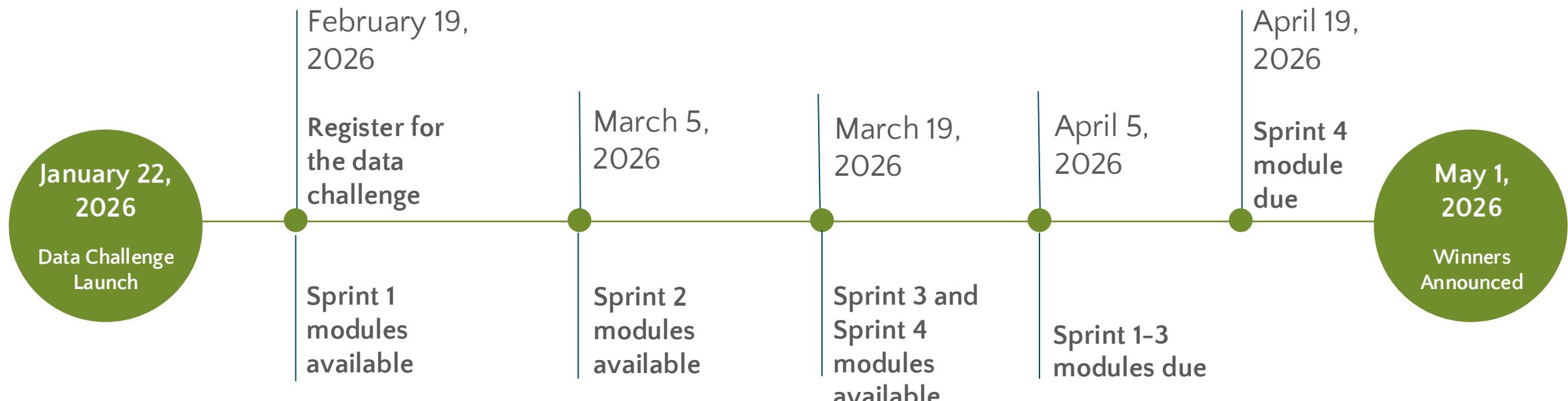
Schedule – Virutal Office Hours

Participants can join us for office hours (1:30-3:00 pm Pacific) during the data challenge:

- Friday, February 20, 2026
- Friday, February 27, 2026
- Friday, March 6, 2026
- Friday, March 13, 2026
- Friday, March 20, 2026
- Friday, March 27, 2026 (NO OFFICE HOURS)
- Friday, April 3, 2026
- Friday, April 10, 2026
- Friday, April 17, 2026

The link to join the office hours can be found [HERE](#).

Data Challenge Timeline



Onboarding

- [Wildfire Commons Documentation](#)
- **Onboarding Module:** This onboarding module is designed to familiarize Data Challenge participants with the user experience of working with workspaces.

Participants are encouraged to complete this onboarding module after registering for the Shrubwise Data Challenge.

Onboarding

The following slides will show you the onboarding process of FireForge, including how to join and start working on the Onboarding Module.

Register with your institution/ university account

Wildfire Science & Technology Commons

HOME ABOUT COMMUNITY FIREFORGE MARKETPLACE HELP

Log In/Register

Open Collaborations for a Wildfire Resilient Future

Introducing the Wildfire Scie

We believe that avoiding devastating wildfires requires a collaborative, open, and transparent approach. The Wildfire Commons is a platform designed to accelerate technological innovations for wildfire preparedness, response, and recovery. It brings together cutting-edge science, AI, and shared knowledge.

Welcome back

Login Staff

Log in using your institution credentials.
Do not use your personal email.

Sign in with CI Logon

By using the Wildfire Commons, you are agreeing with our Code of Conduct. Please email <https://wildfirecommons.org> with any problems logging in.

Register with your institution/university account

Wildfire Science & Technology Commons

Log in/Register

Welcome back

Login Staff

Log in using your institution credentials.
Do not use your personal email.

1 Sign in with CI Logon

By using the Wildfire Commons, you are agreeing with our Code of Conduct. Please email info@wildfirecommons.org with any problems logging in.

CI-Logon

Consent to Attribute Release

NSE National Data Platform (NDP) requests access to the following information. If you do not approve this request, do not proceed.

- Your CI-Logon user identifier
- Your name
- Your email address
- Your username and affiliation from your identity provider

2

Select an Identity Provider

- University of California, San Diego
- University of Cal
- University of California, San Diego
- University of Calabria
- University of Calcutta
- University of Calgary
- University of Calicut
- University of California, Agriculture and Natural Resources
- University of California, Berkeley
- University of California, College of the Law San Francisco
- University of California, Davis
- University of California, Irvine
- University of California, Los Angeles
- University of California, Merced
- University of California, Office of the President
- University of California, Riverside
- University of California, San Francisco
- University of California, Santa Barbara
- University of California, Santa Cruz

Search for your institution/university

Set-up your profile

1

PROFILE
LOG OUT

2

View Edit

Email: tutor@ndp.com

3

Wildfire Science & Technology Commons

Picture

Choose File No file chosen

Please upload a headshot
One file only:
40 MB limit.
Allowed types: png gif jpg jpeg webp.

First Name *
Tutor

Last Name *
Admin

Expertise (start typing, then click on existing expertise or press return to add your own) *

Wildland Fire Solution Interest *

Choose some options

Please identify all solutions areas you are interested in

Organization *

Please enter the name of the primary organization that you are currently affiliated with. If you are not affiliated with an organization, please enter 'Community Member'

Department or Unit

WSTC EMAIL LIST

Subscribe

Receive email about WSTC events and news. Uncheck box above to opt out of WSTC's mailing list.

4

Save

Navigate to data challenge

Once your profile is set-up, navigate to
fireforge.wildfirecommons.org



Navigate to data challenge and join challenge

The image shows the Wildfire Commons FireForge interface with three numbered steps indicating the navigation process:

- 1** Click on **Data Challenges** in the sidebar.
- 2** Click on **Explore** in the sidebar.
- 3** Click on the **View More →** button to see more challenges.

The interface includes a header with the Wildfire Commons FireForge logo and navigation links for Back to Wildfire Commons, My Dashboard, My Hub, Catalogs, and Data Challenges. The main content area displays a list of challenges, with the "Shrubwise Data Challenge" by the University of California, San Diego, highlighted. A "Join Challenge" button is visible on the right side of the challenge details page.

Explore Community Resources

Data Challenges **Workspaces** **Projects**

Live Challenges **Past Challenges** **Published Challenges**

Shrubwise Data Challenge **7 participants**

University of California, San Diego

View More →

My Dashboard

Shrubwise Data Challenge

University of California, San Diego

Join Challenge

Overview **Rules and Eligibility** **FAQ** **Modules** **Dataset Details**

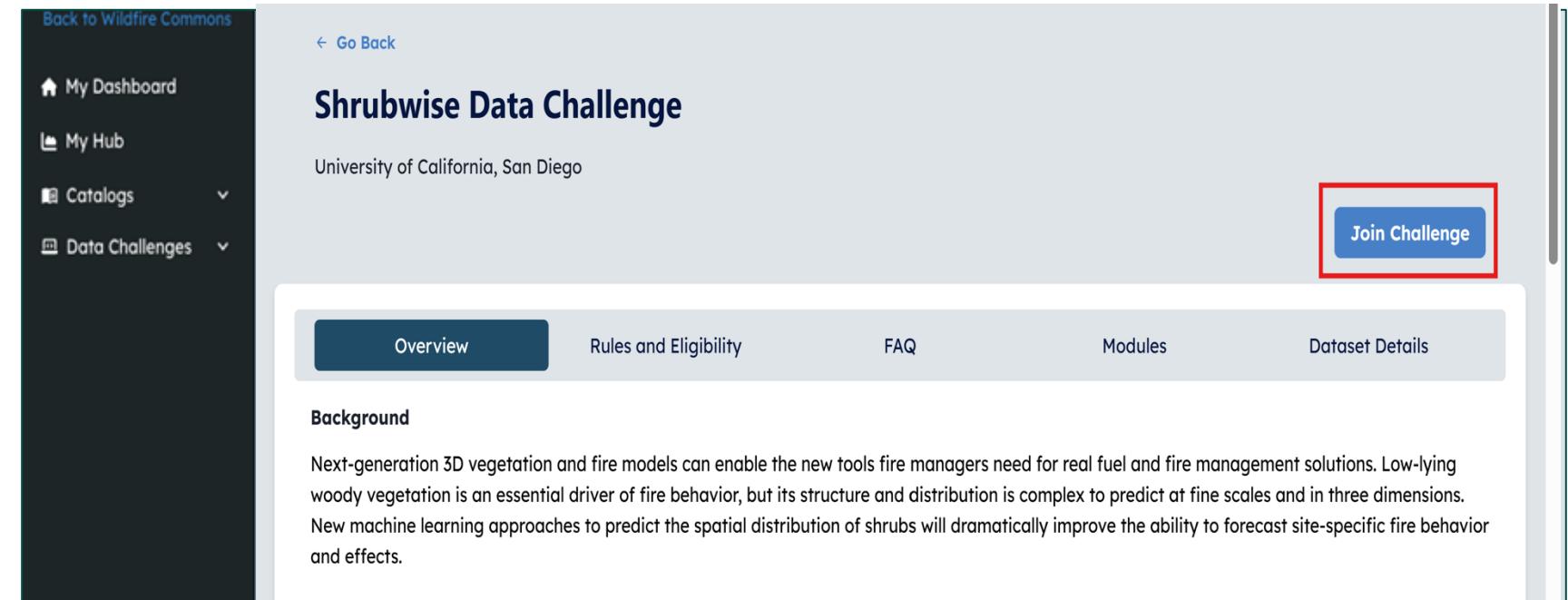
Background

Next-generation 3D vegetation and fire models can enable the new tools fire managers need for real fuel and fire management solutions. Low-lying woody vegetation is an essential driver of fire behavior, but its structure and distribution is complex to predict at fine scales and in three dimensions. New machine learning approaches to predict the spatial distribution of shrubs will dramatically improve the ability to forecast site-specific fire behavior and effects.

Traditional 2D fuel models are heavily focused on parametrizing forested ecosystems and largely ignore the influence of shrubs on fire behavior, often treating rangelands and shrub dominated ecosystems as uniform across large areas. In reality, shrubs are highly variable in both density and spatial configuration. This variability strongly influences fire behavior—shrubs can act as heat sinks, fuel breaks, or ladder fuels depending on their arrangement. When fuel models fail to represent this structure, fire behavior models inherit those inaccuracies, limiting their ability to simulate fire spread and intensity. This limitation is particularly consequential in Southern California, where shrub-dominated ecosystems are widespread and where spatial patterns of shrubs play a critical role in modulating fire spread under fluctuating and extreme weather conditions.

The fuel beds and forest structure generated by vegetation modeling tools like FastFuels (USFS) are used as inputs to fire models like QUIC-Fire (LANL), which uses coupled-fire atmosphere dynamics to produce simulations grounded in the underlying physics that drive fire behavior. Multimodal datasets such as airborne laser scanning, high resolution imagery, and multispectral satellite data offer new insights for describing the 3D arrangement of shrubs across ecosystems. Combined with 3D fuel modeling approaches, these data can become inputs to 3D physics-based fire models that account

Navigate to data challenge and join challenge



Back to Wildfire Commons

← Go Back

Shrubwise Data Challenge

University of California, San Diego

Join Challenge

Overview Rules and Eligibility FAQ Modules Dataset Details

Background

Next-generation 3D vegetation and fire models can enable the new tools fire managers need for real fuel and fire management solutions. Low-lying woody vegetation is an essential driver of fire behavior, but its structure and distribution is complex to predict at fine scales and in three dimensions. New machine learning approaches to predict the spatial distribution of shrubs will dramatically improve the ability to forecast site-specific fire behavior and effects.

Assign a name to your team and type the email addresses of your team members. **Make sure to separate them with a comma and leave no blank spaces.**

If you receive an error message while adding team members, it may mean that they haven't registered with the Wildfire Commons.

2

Join Challenge

Team Name

Team Commons

Team Member Emails

VALID emails separated by comma
Leave the below field empty to join this data challenge as a solo participant!

info@wildfirecommons.org

3

Join Challenge

Data challenge modules

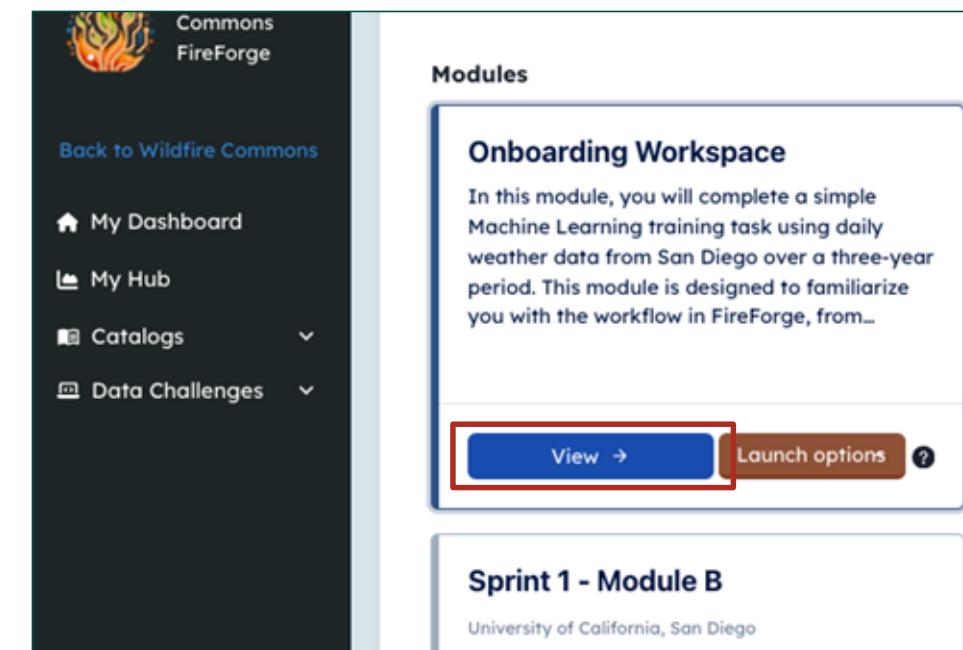
Once you join a Data Challenge, your modules will appear. To view the content of the onboarding module, click on the “View” button.

Guidance: Read modules carefully

Read carefully the instructions of each module prior to launching them on JupyterHub.

Each module will come with guidance on how to:

- Reserve your JupyterHub server computing resources
- Clone the attached repository (if provided)
- Install dependencies
- Download your data to JupyterHub



JupyterHub Guidance

Next we will be providing a quick overview of JupyterHub. The following slides are for reviewing first.

If you need guidance seek us at office hours.

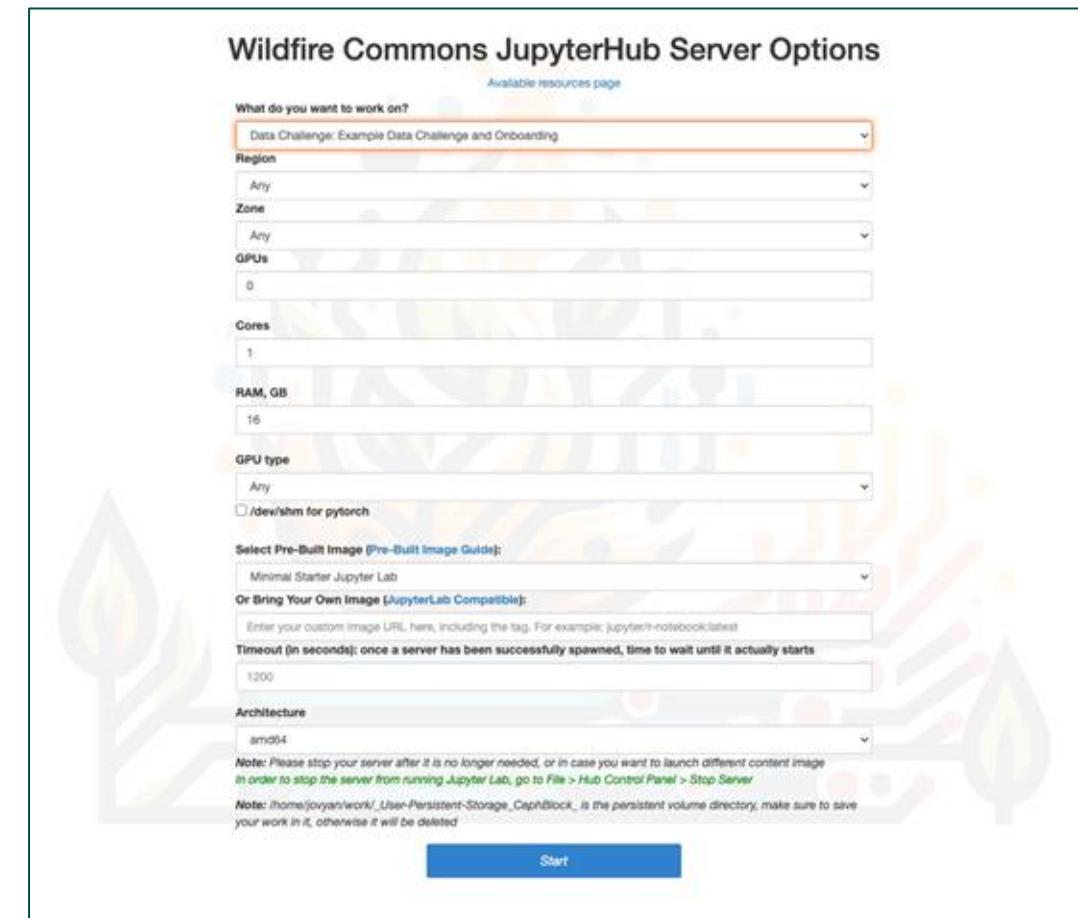
Launching JupyterHub

Once you have read the module's instructions, navigate to JupyterHub by clicking on the JupyterHub button at the bottom of the module.

Guidance:
Selecting your
compute resources

There will be multiple offerings for compute resources.

Rely on the guidance in the module instructions **before** selecting your compute.



Wildfire Commons JupyterHub Server Options
Available resources page

What do you want to work on?
Data Challenge: Example Data Challenge and Onboarding

Region: Any

Zone: Any

GPUs: 0

Cores: 1

RAM, GB: 16

GPU type: Any
 /dev/shm for pytorch

Select Pre-Built Image (Pre-Built Image Guide):
Minimal Starter Jupyter Lab

Or Bring Your Own Image (JupyterLab Compatibile):
Enter your custom image URL here, including the tag. For example: jupyter/t-notebook:latest

Timeout (in seconds): once a server has been successfully spawned, time to wait until it actually starts
1200

Architecture: amd64

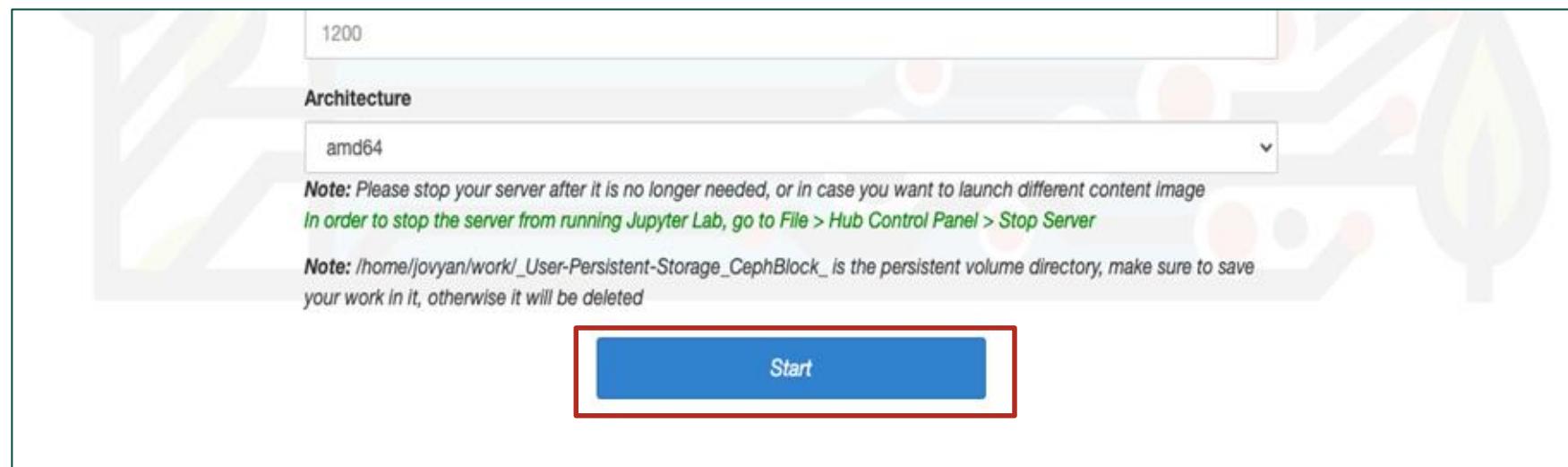
Note: Please stop your server after it is no longer needed, or in case you want to launch different content image in order to stop the server from running Jupyter Lab, go to File > Hub Control Panel > Stop Server

Note: /home/joyan/work/_User-Persistent-Storage_CephBlock_ is the persistent volume directory, make sure to save your work in it, otherwise it will be deleted

Start

Launching JupyterHub

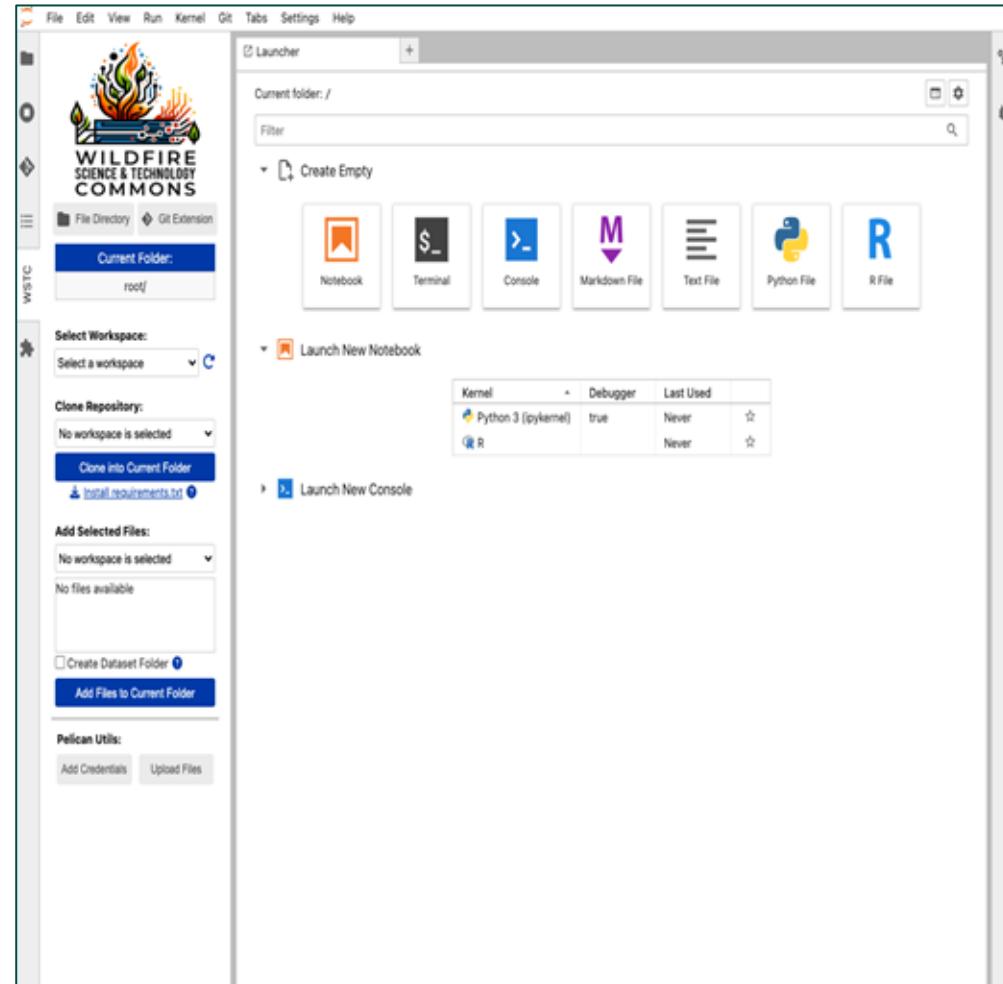
After setting up your resources, click on “Start” to launch your server.



Using JupyterHub

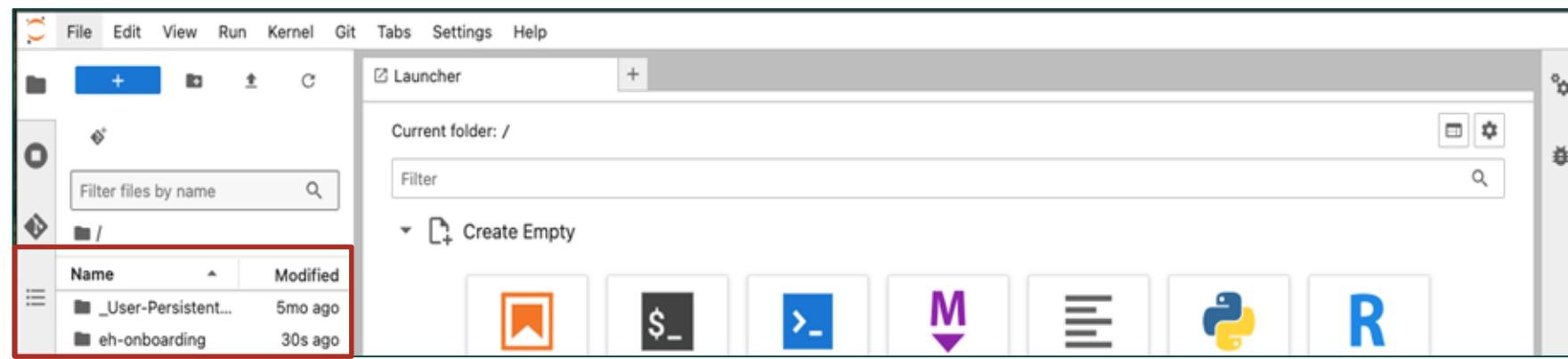
Once your server initiates, it will look like this screen →

Notice on the left side of the screen, you have options to connect in your module resources with your JupyterHub session.



Using JupyterHub

At this point, you should have two folders:



The first one corresponds to your user storage, while the second one to the team shared storage. **Files that are not saved in any of these folders will be lost when you stop your server.**

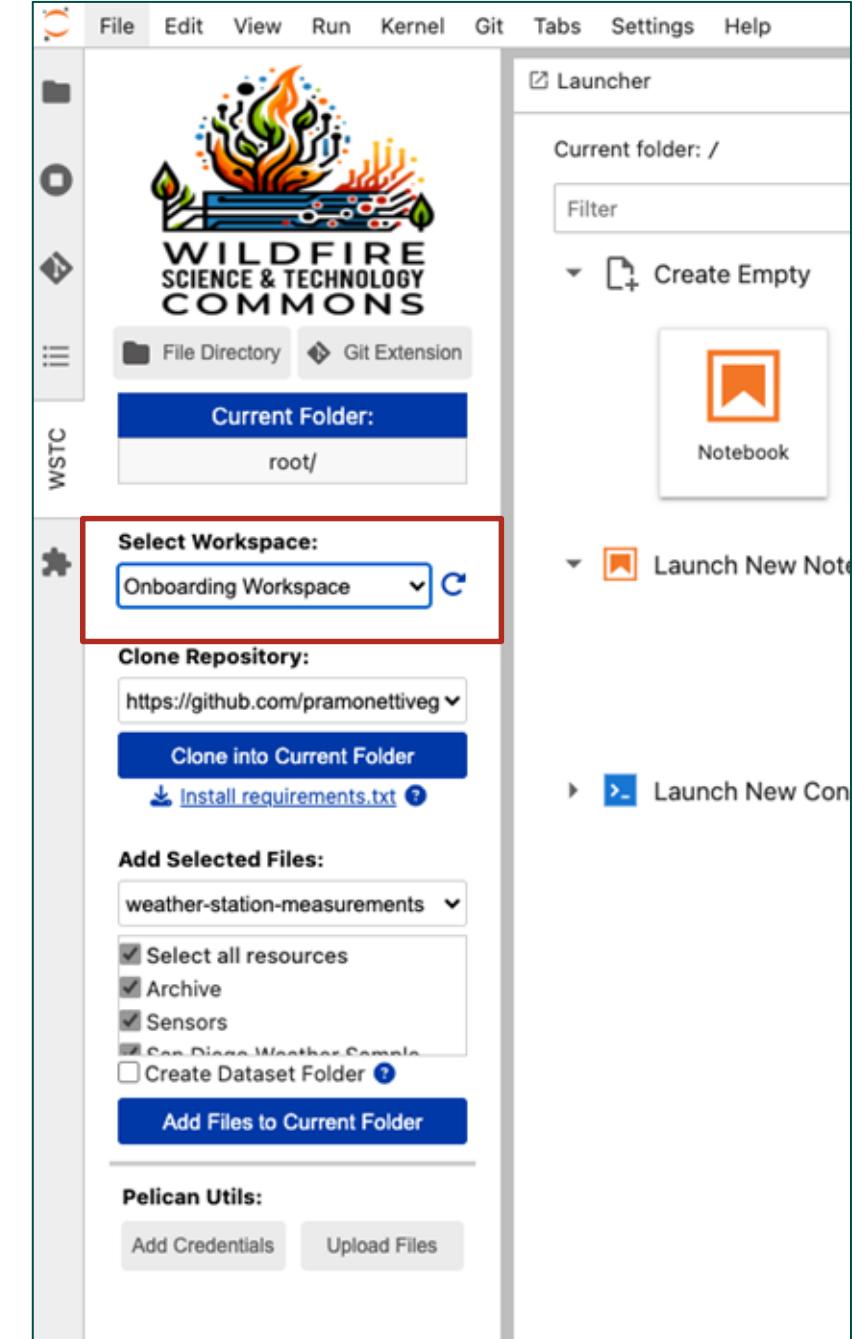
Good practices about shared storage

- Always communicate with your team members about the files they place/remove from this folder.
- Do not work on the same file at the same time or you will run into an overwrite conflict.
- Shared storage is limited, so use it to share frequently accessed files or data derived from your workflow.

Using JupyterHub

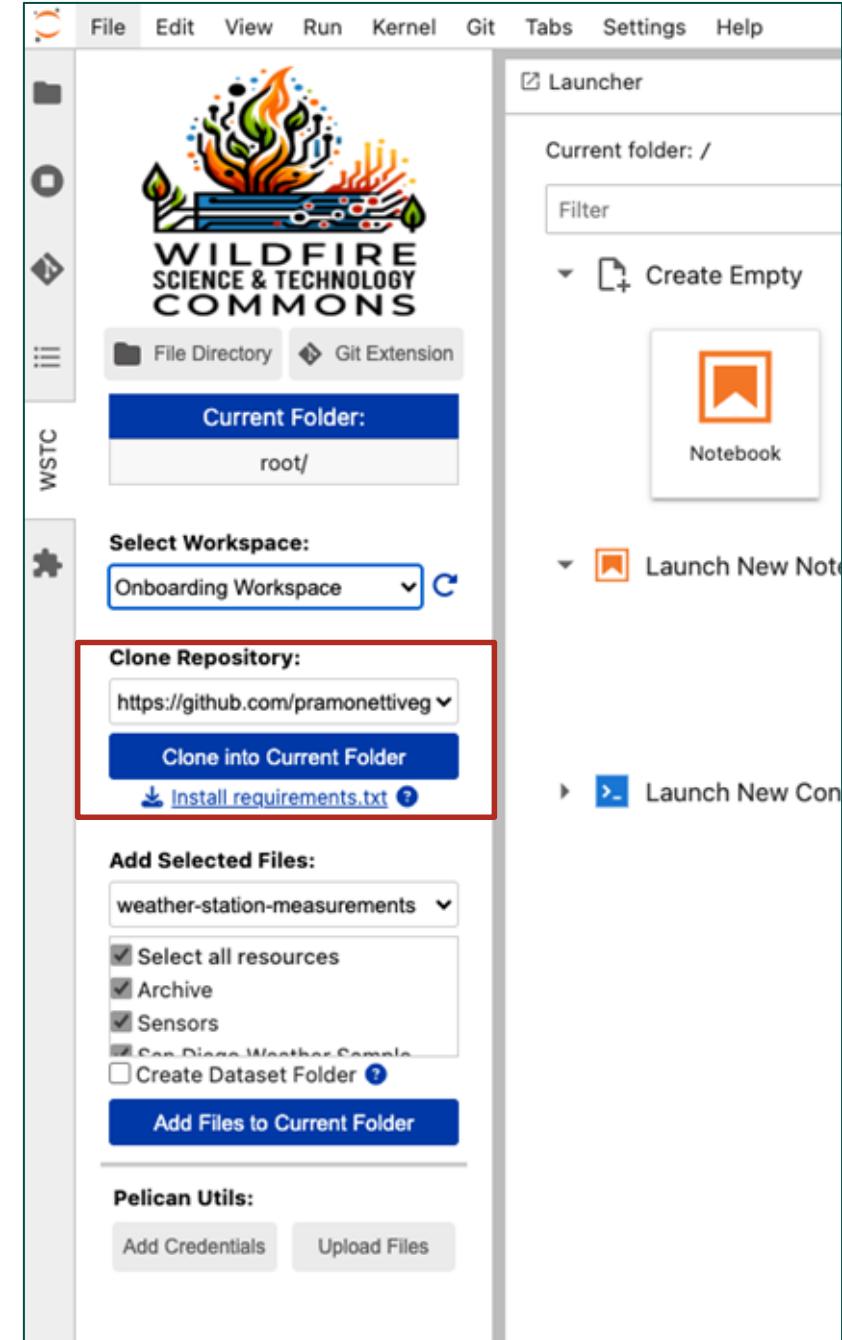
In this dropdown menu →

You can select the module or workspace you plan to work on.



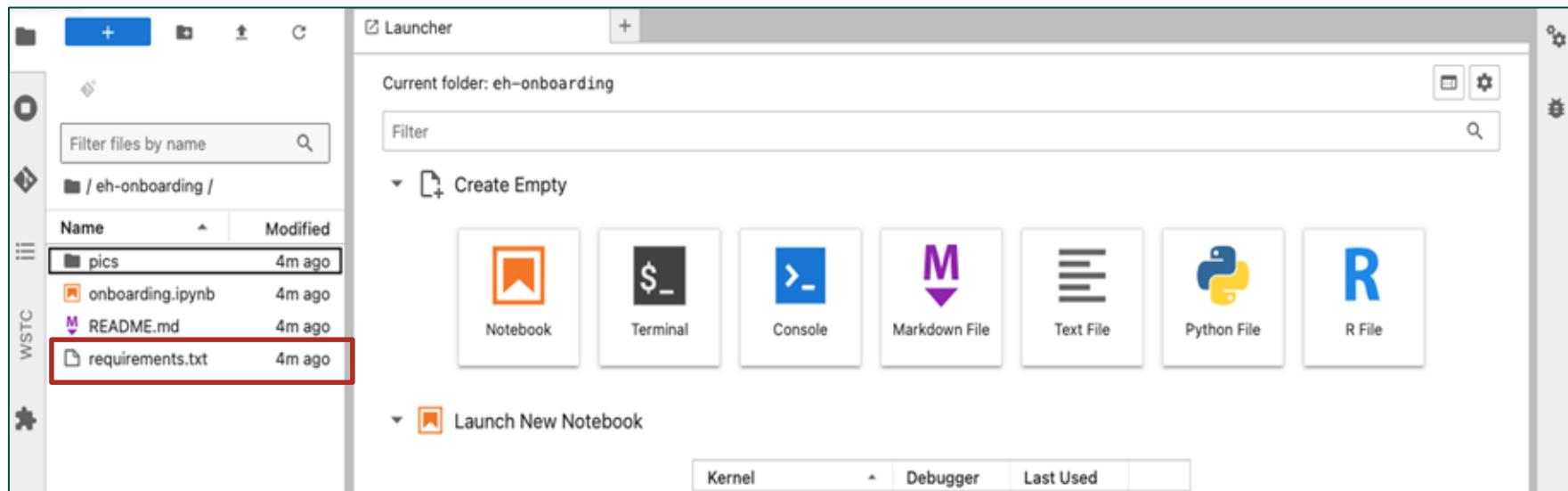
Using JupyterHub

You can clone the module's repositories using this section.



Using JupyterHub

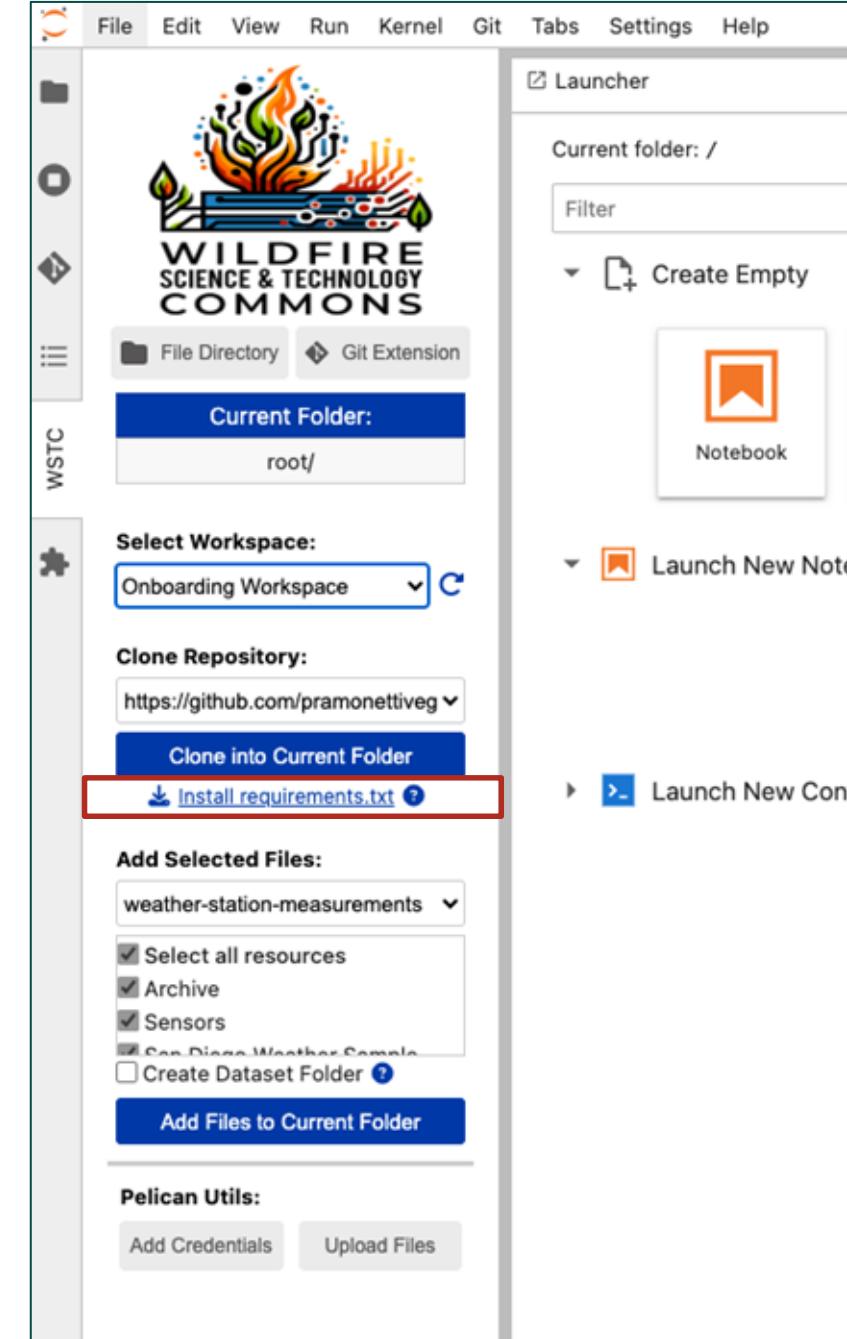
After cloning the repository and moving the cloned folder, the repository may contain a requirements.txt file.



Using JupyterHub

If the repository contains a requirements.txt file, you can install the libraries listed in it by clicking the “Install Requirements” button.

You can only take this action in the repository’s folder.

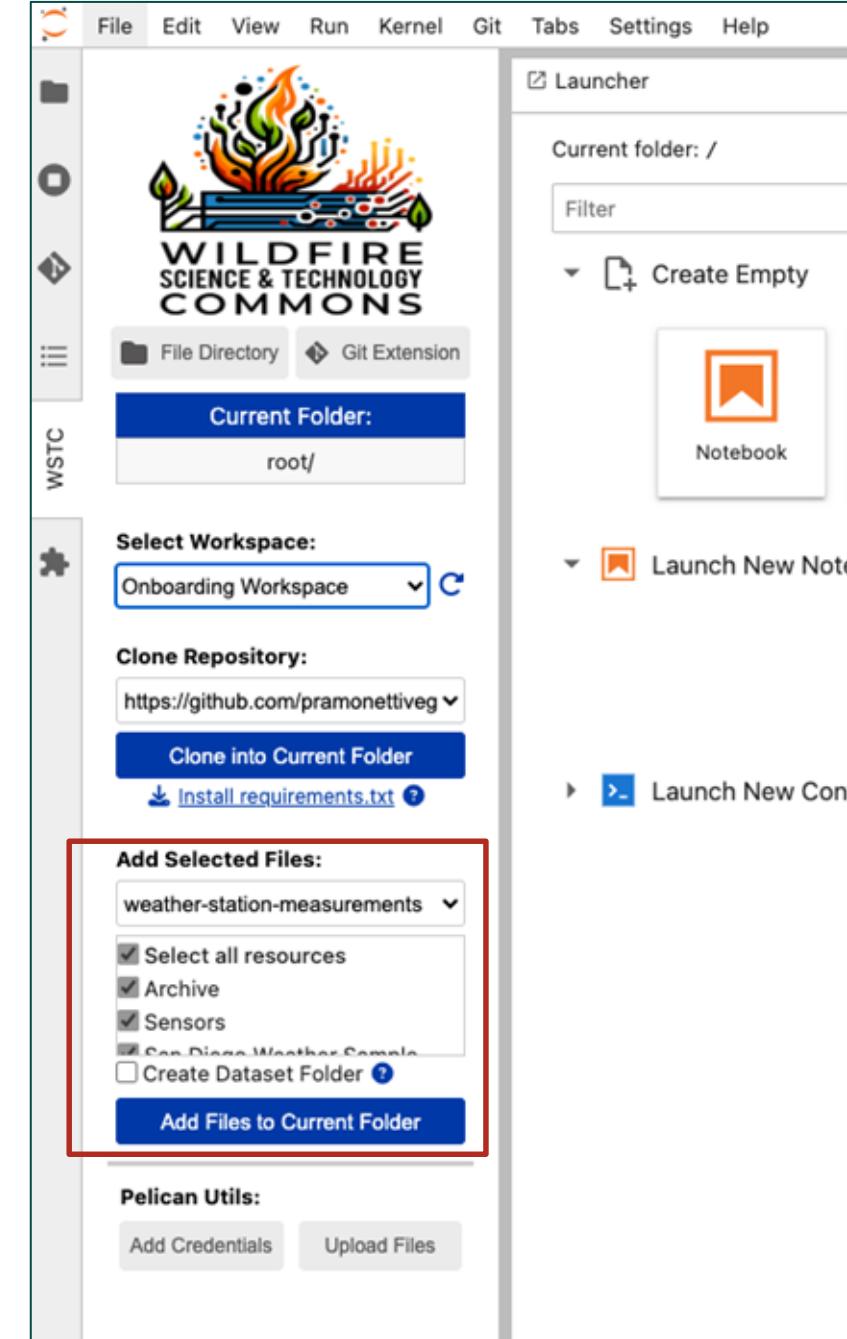


Using JupyterHub

The datasets attached to the module appear in this list.

Each record can contain several files. You can select the files you want to download.

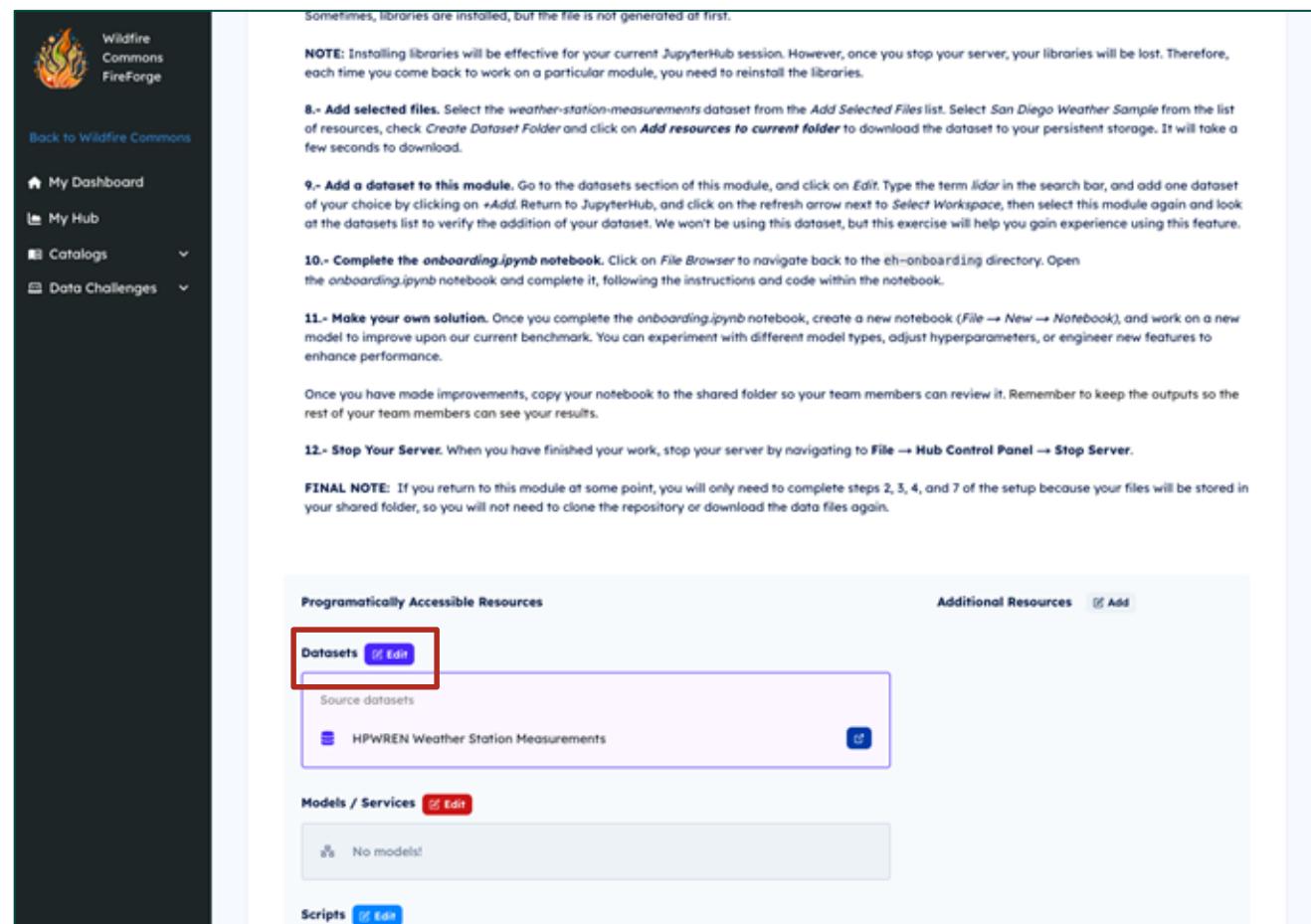
Click “Add files to Current Folder” to download them.



Using JupyterHub

In some cases, a module might require you to add additional data from the catalog.

If this is the case, go to the module (*in a separate window*), and click on the Datasets “Edit” button.



Sometimes, libraries are installed, but the file is not generated at first.

NOTE: Installing libraries will be effective for your current JupyterHub session. However, once you stop your server, your libraries will be lost. Therefore, each time you come back to work on a particular module, you need to reinstall the libraries.

8.- Add selected files. Select the weather-station-measurements dataset from the **Add Selected Files** list. Select San Diego Weather Sample from the list of resources, check **Create Dataset Folder** and click on **Add resources to current folder** to download the dataset to your persistent storage. It will take a few seconds to download.

9.- Add a dataset to this module. Go to the datasets section of this module, and click on **Edit**. Type the term `lidar` in the search bar, and add one dataset of your choice by clicking on **+Add**. Return to JupyterHub, and click on the refresh arrow next to **Select Workspace**, then select this module again and look at the datasets list to verify the addition of your dataset. We won't be using this dataset, but this exercise will help you gain experience using this feature.

10.- Complete the onboarding.ipynb notebook. Click on **File Browser** to navigate back to the `eh-onboarding` directory. Open the `onboarding.ipynb` notebook and complete it, following the instructions and code within the notebook.

11.- Make your own solution. Once you complete the `onboarding.ipynb` notebook, create a new notebook (`File` → `New` → `Notebook`), and work on a new model to improve upon our current benchmark. You can experiment with different model types, adjust hyperparameters, or engineer new features to enhance performance.

Once you have made improvements, copy your notebook to the shared folder so your team members can review it. Remember to keep the outputs so the rest of your team members can see your results.

12.- Stop Your Server. When you have finished your work, stop your server by navigating to `File` → `Hub Control Panel` → `Stop Server`.

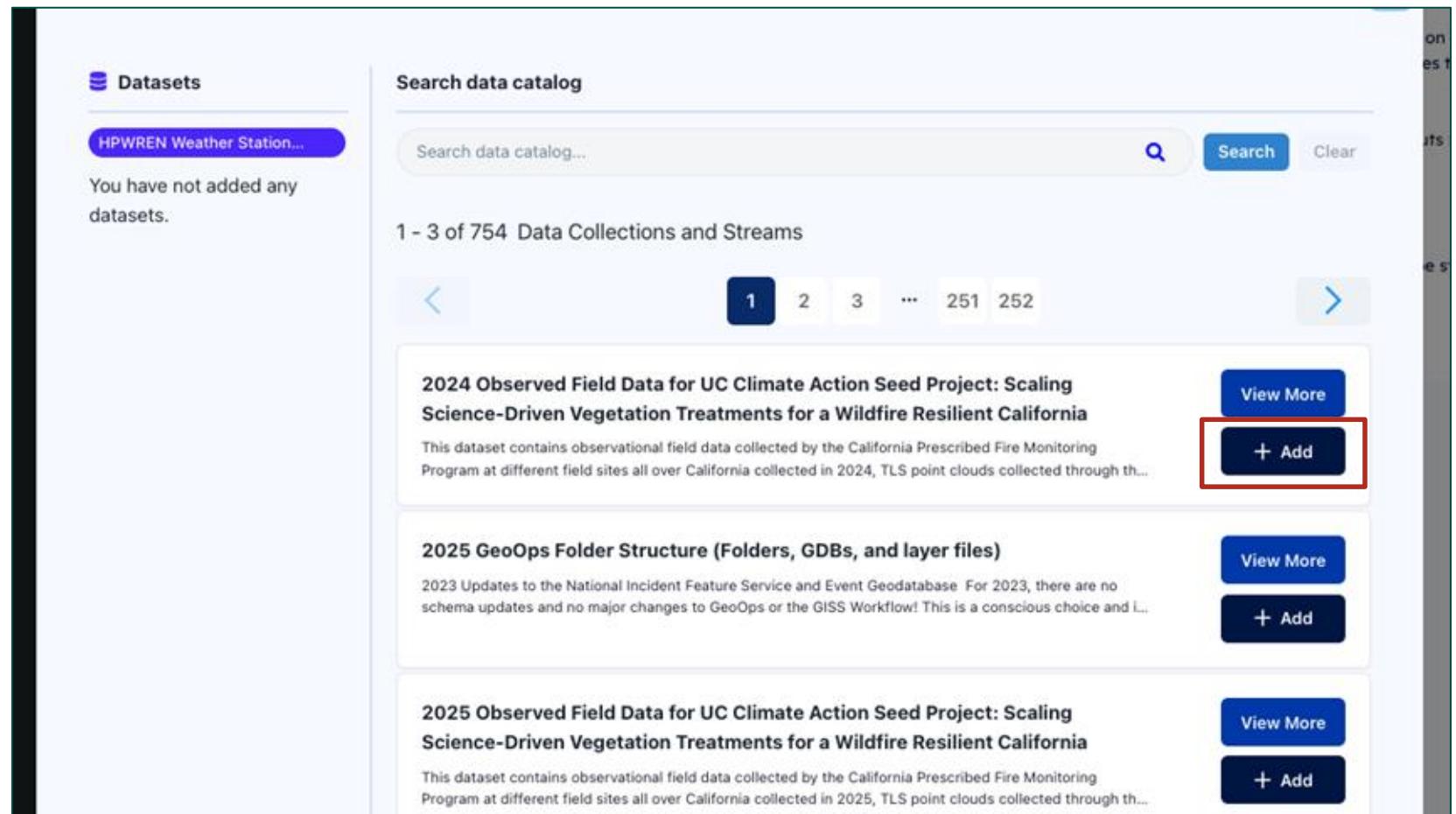
FINAL NOTE: If you return to this module at some point, you will only need to complete steps 2, 3, 4, and 7 of the setup because your files will be stored in your shared folder, so you will not need to clone the repository or download the data files again.

Programmatically Accessible Resources	
Datasets	Edit
Source datasets	
HPWREN Weather Station Measurements	Edit
Models / Services	
No models!	
Scripts	

Using JupyterHub

Search for the dataset of your interest.

Click “Add” to add the dataset to your module.



Datasets

HPWREN Weather Station...

You have not added any datasets.

Search data catalog

Search data catalog...

1 - 3 of 754 Data Collections and Streams

1 2 3 ... 251 252

2024 Observed Field Data for UC Climate Action Seed Project: Scaling Science-Driven Vegetation Treatments for a Wildfire Resilient California

This dataset contains observational field data collected by the California Prescribed Fire Monitoring Program at different field sites all over California collected in 2024, TLS point clouds collected through th...

2025 GeoOps Folder Structure (Folders, GDBs, and layer files)

2023 Updates to the National Incident Feature Service and Event Geodatabase. For 2023, there are no schema updates and no major changes to GeoOps or the GISS Workflow! This is a conscious choice and i...

2025 Observed Field Data for UC Climate Action Seed Project: Scaling Science-Driven Vegetation Treatments for a Wildfire Resilient California

This dataset contains observational field data collected by the California Prescribed Fire Monitoring Program at different field sites all over California collected in 2025, TLS point clouds collected through th...

View More

+ Add

View More

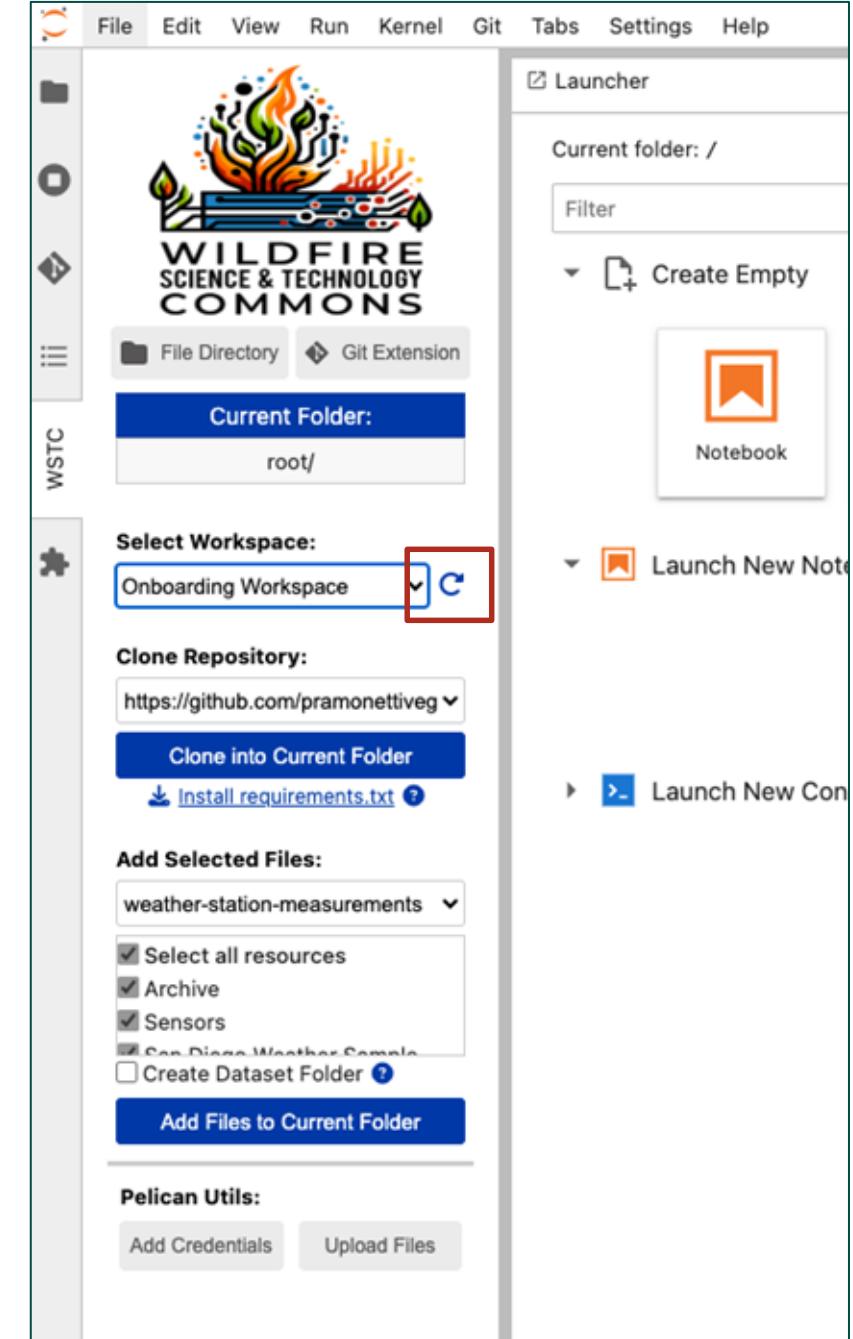
+ Add

View More

+ Add

Using JupyterHub

Click the Refresh button to reload your updated modules with your newly added dataset.



Next steps

01

Register

Set-up your profile on the Wildfire Commons using your academic email / institutional email

02

Form Your Team

Select your team members and register as a team

03

Complete Onboarding Module

Get familiar with FireForge, JupyterHub and working with your team

04

Work on the Module

Begin working on the instructional modules as soon as they are released

05

Submit Your All Your Work by April 19

Complete all your modules with your team and submit by April 19, 2026





THANK YOU



UC San Diego



Societal Computing
and Innovation Lab



Proactive Wildfire & Environmental Sustainability Solutions

<https://www.wildfirecommons.org>

