Shrubwise Data Challenge

Winter/Spring 2026

NIST UC San Diego SCIL Societal Computing and Innovation Lab

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	Training Data	Model Development	Model Application
Define the target outputs your model will predict (e.g., shrub cover and height) and develop reliable label products to train against.	Prepare and curate your inputs by integrating multiple data sources, cleaning artifacts and noise, and validating data quality and alignment.	Train and refine your ML/AI model using the curated dataset, iterating on features, architecture, and hyperparameters to improve performance on target metrics.	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	Introductory Data	Relevant Data Products	Workflow Development
You will have completed onboarding in FireForge, learned how to launch the JupyterHub service, and practiced working collaboratively in shared workspaces.	You will have explored core datasets and foundational concepts needed to understand the challenge goals.	You will have identified additional relevant data products and learned how to incorporate them into your pipeline.	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

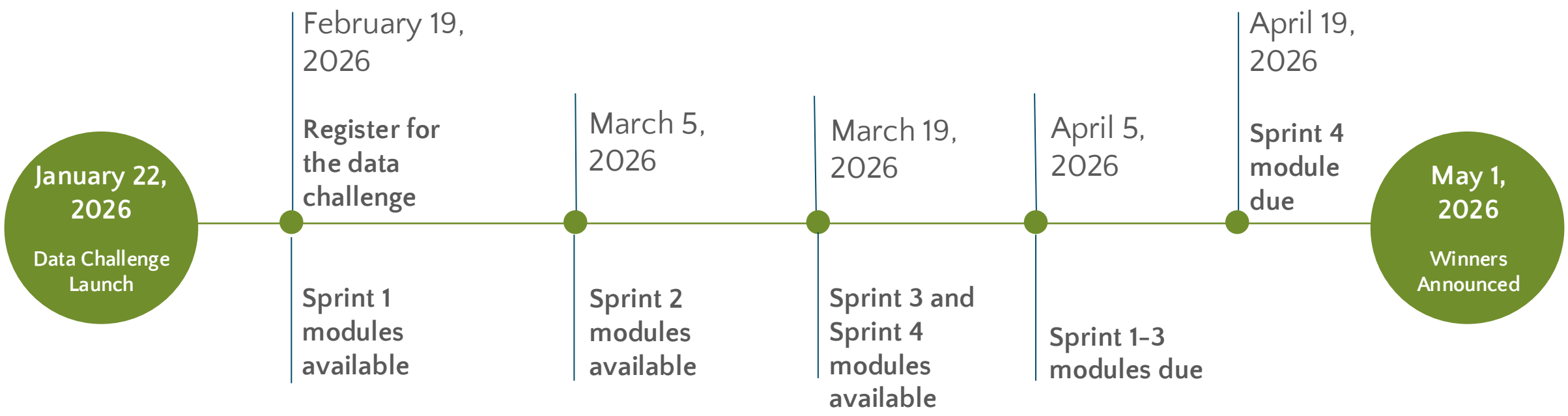
Schedule – Virutal Office Hours

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

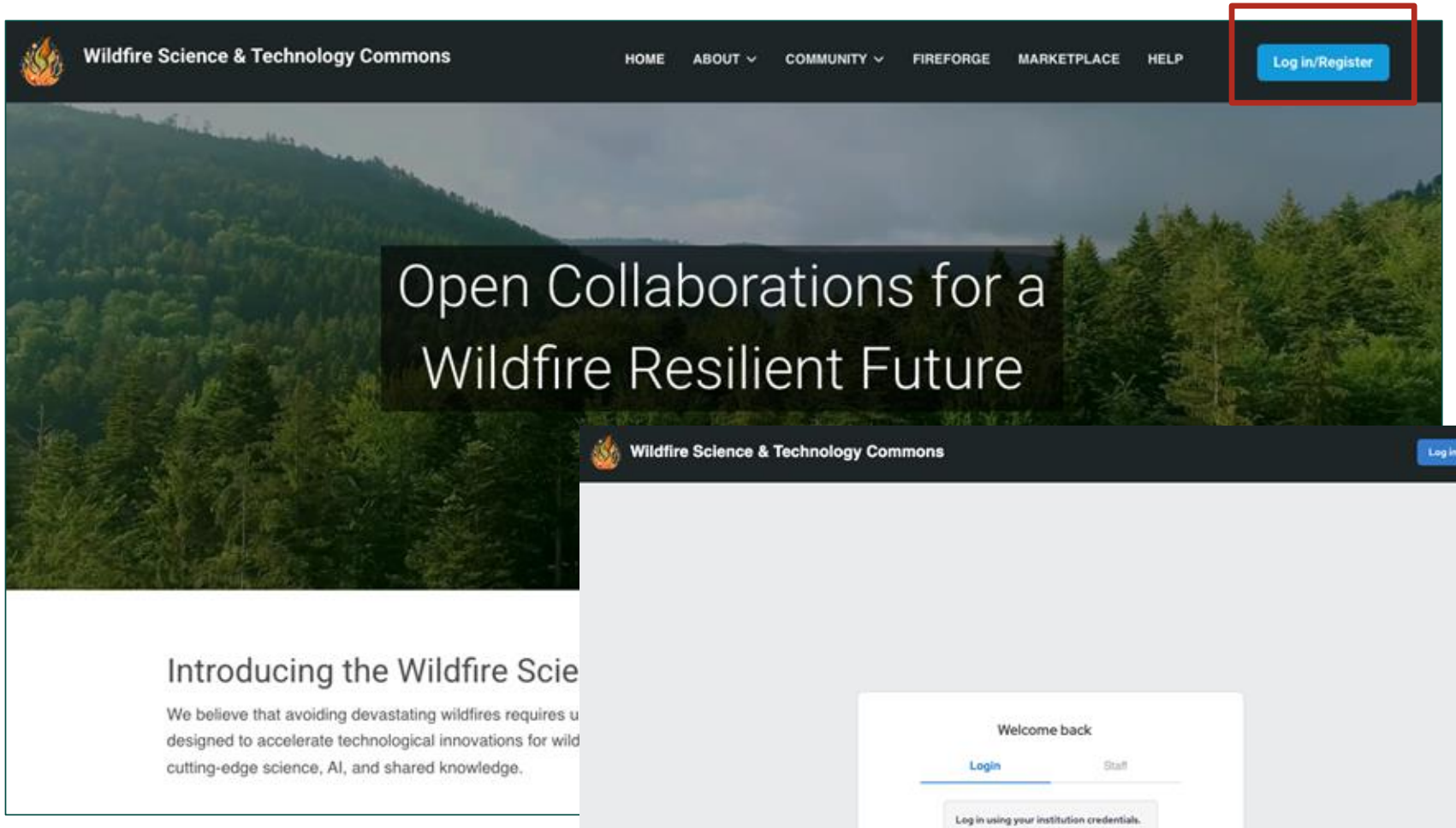
- Wildfire Commons Documentation
- **Onboarding Module**: This onboarding module is designed to familiarize Data Challenge participants with the user experience of working with workspaces.

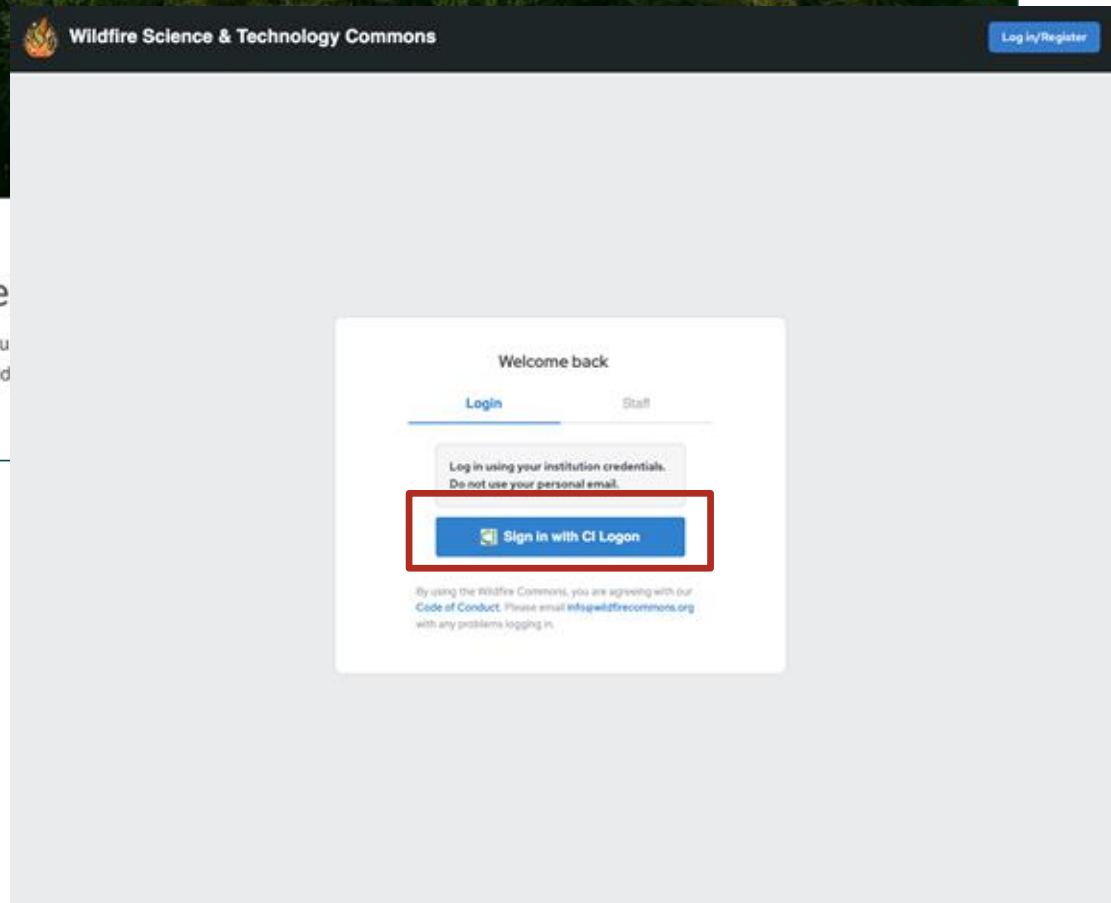
  Participants are encouraged to complete this onboarding module after registering for the Shrubwise Data Challenge.

Onboarding

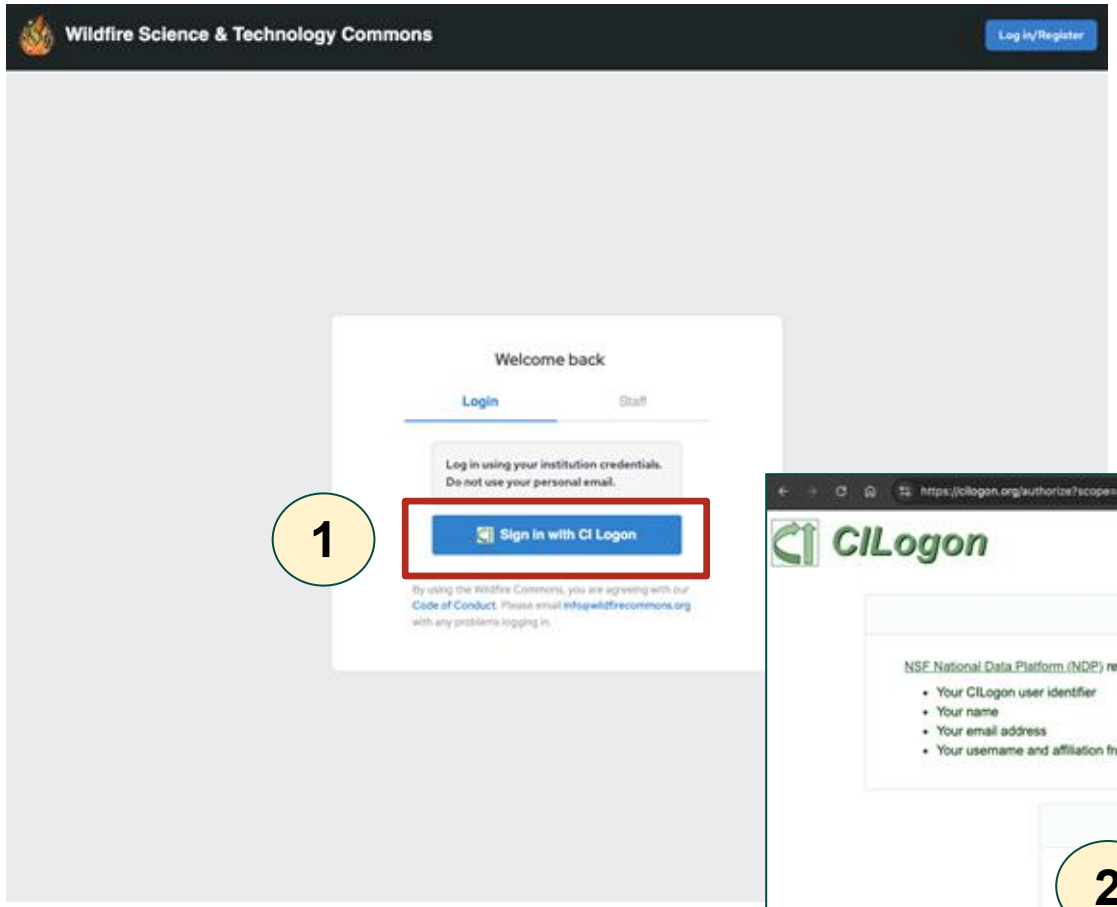
The following slides will show you the onboarding process of FireForge, including how to join and start working on the Onboarding Module.

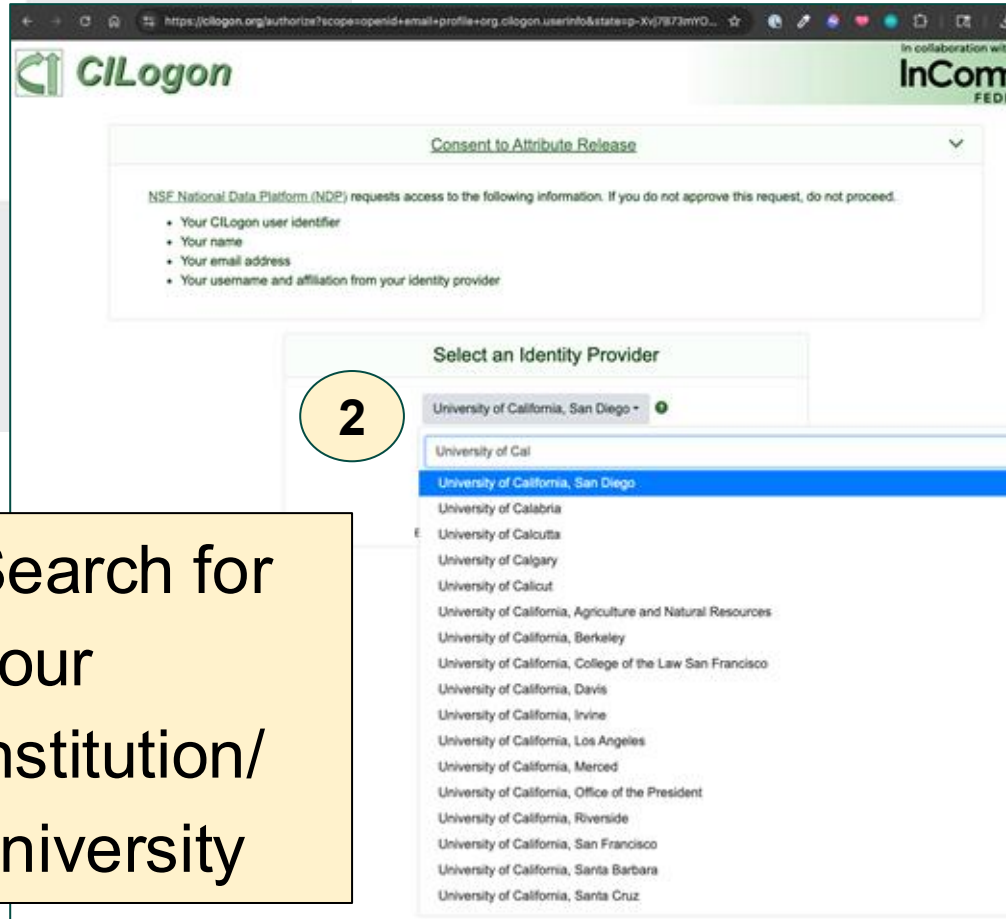
Register with your institution/ university account





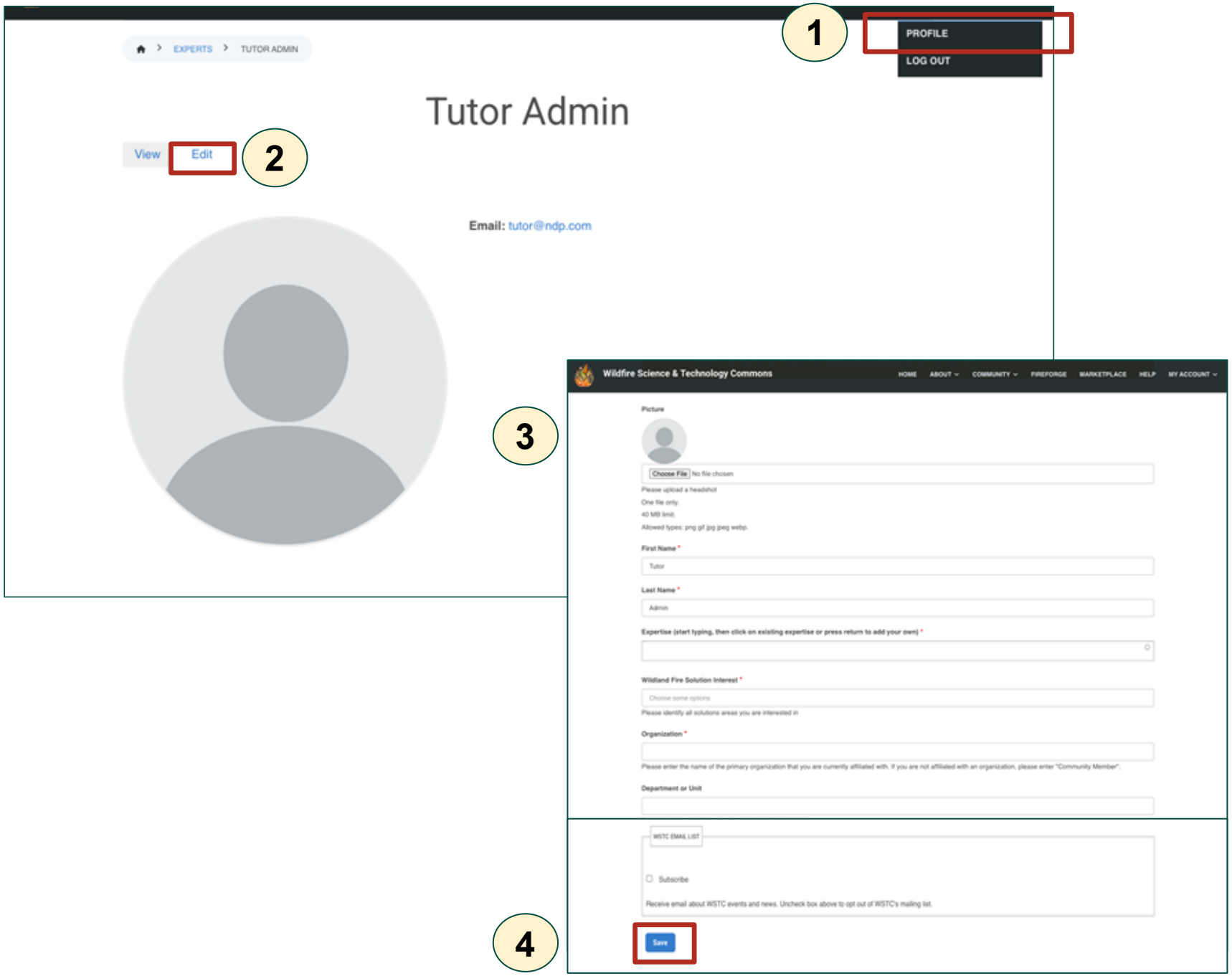
Register with your institution/ university account





Search for your institution/ university

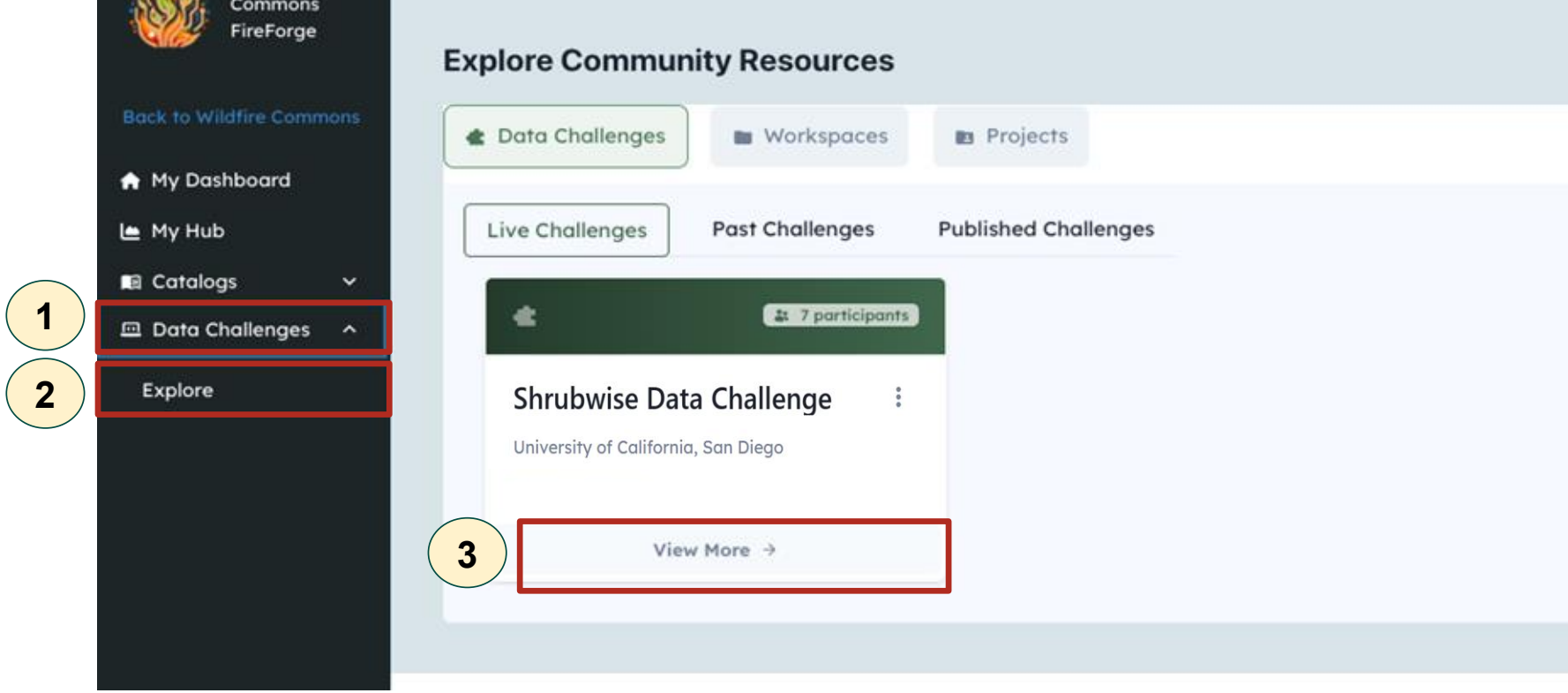
Set-up your profile

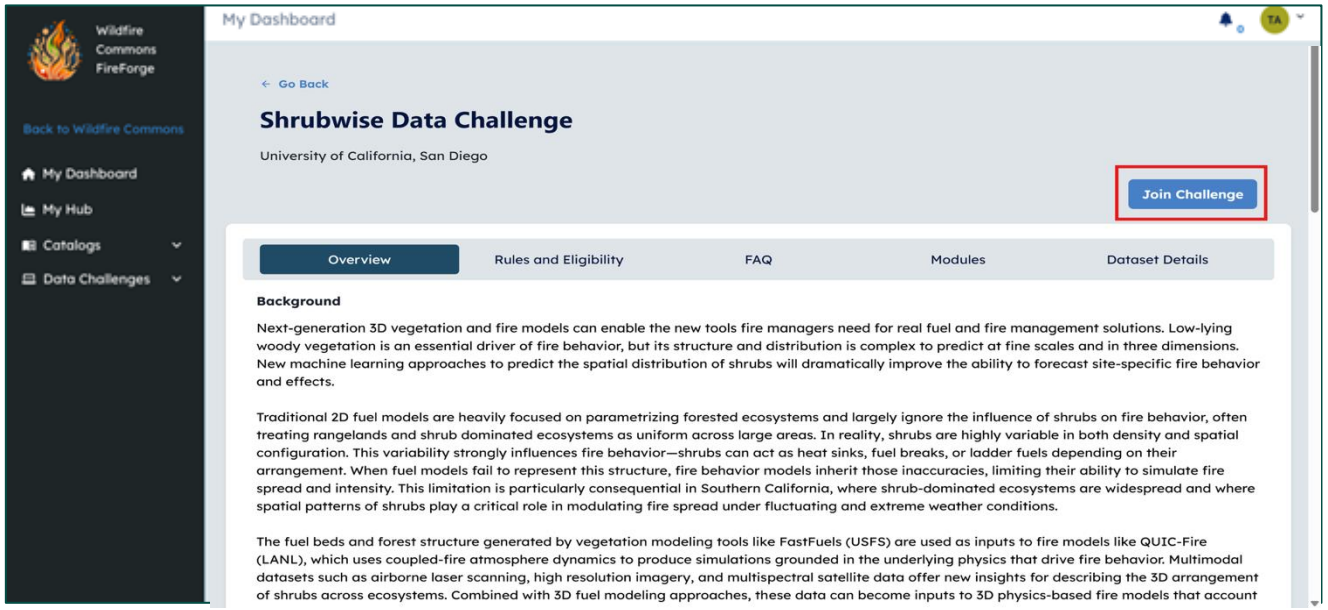


Navigate to data challenge

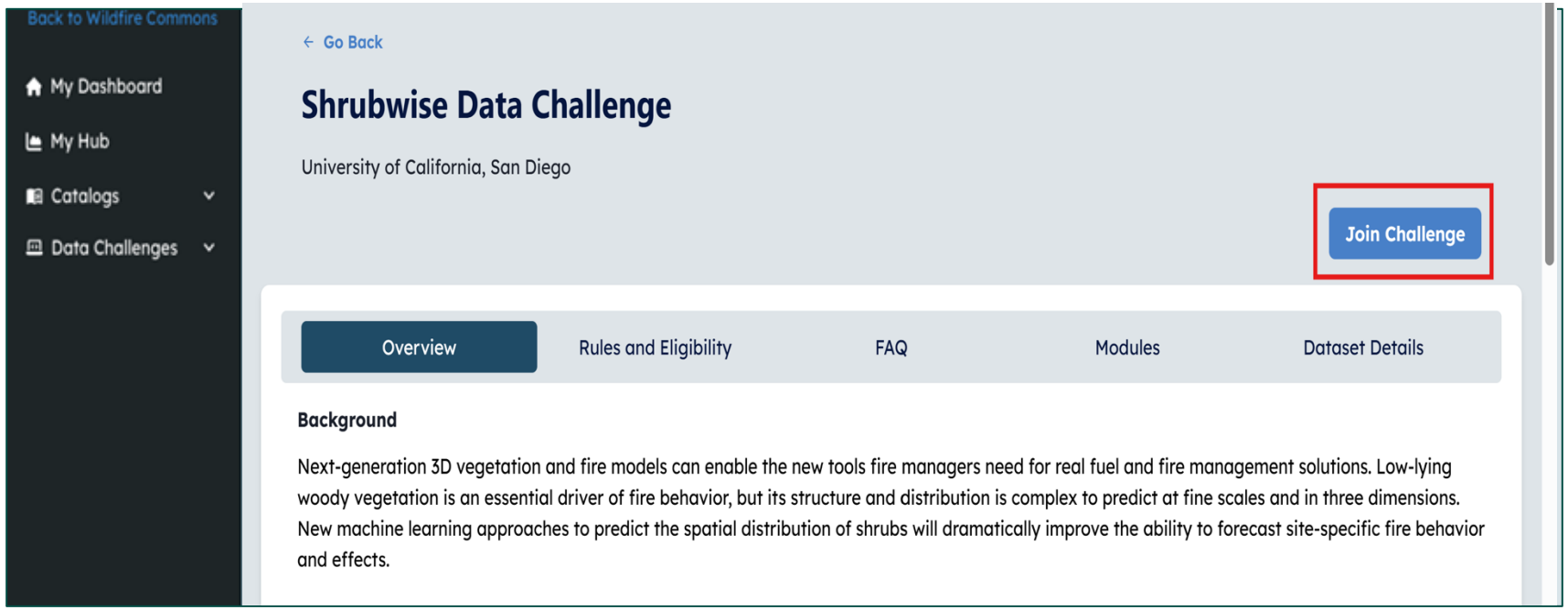
Once your profile is set-up, navigate to fireforge.wildfirecommons.org

Navigate to data challenge and join challenge



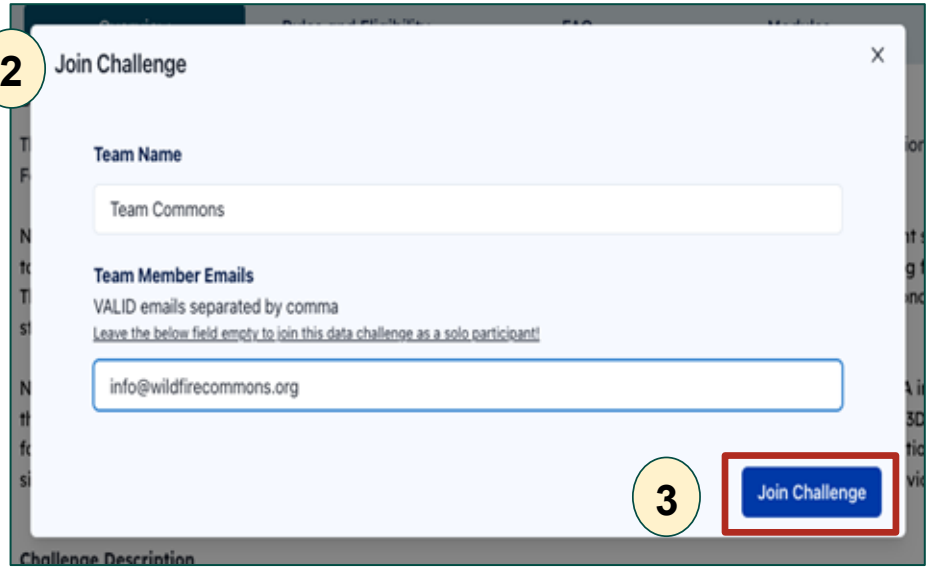


Navigate to data challenge and join challenge



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.



Data challenge modules

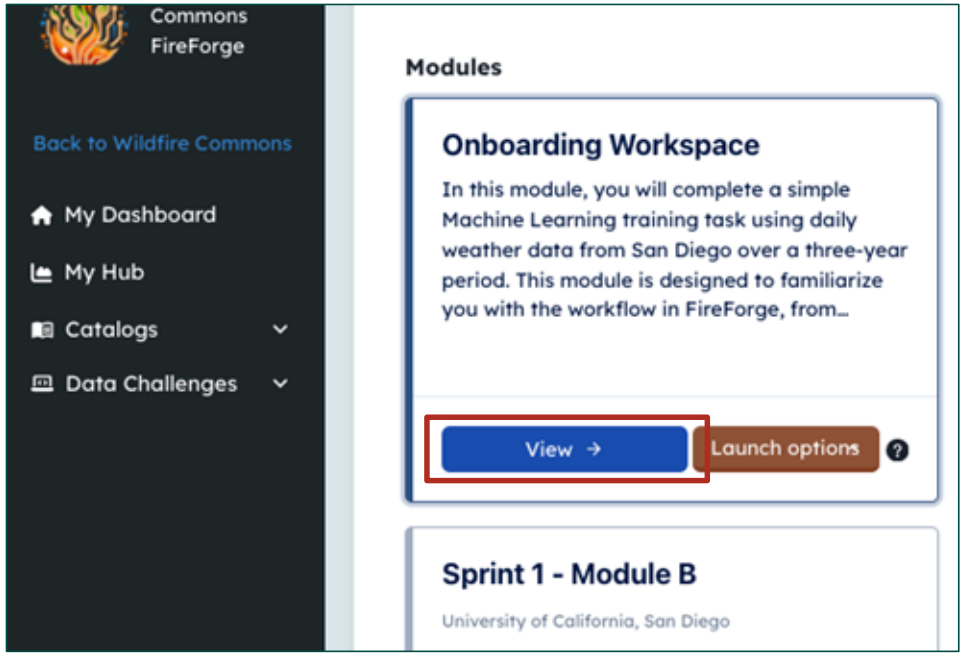
Once you join a Data Challenge, your modules will appear. To view the content of the onboarding module, click on the “View” button.

<u>Guidance: Read modules carefully</u>

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.

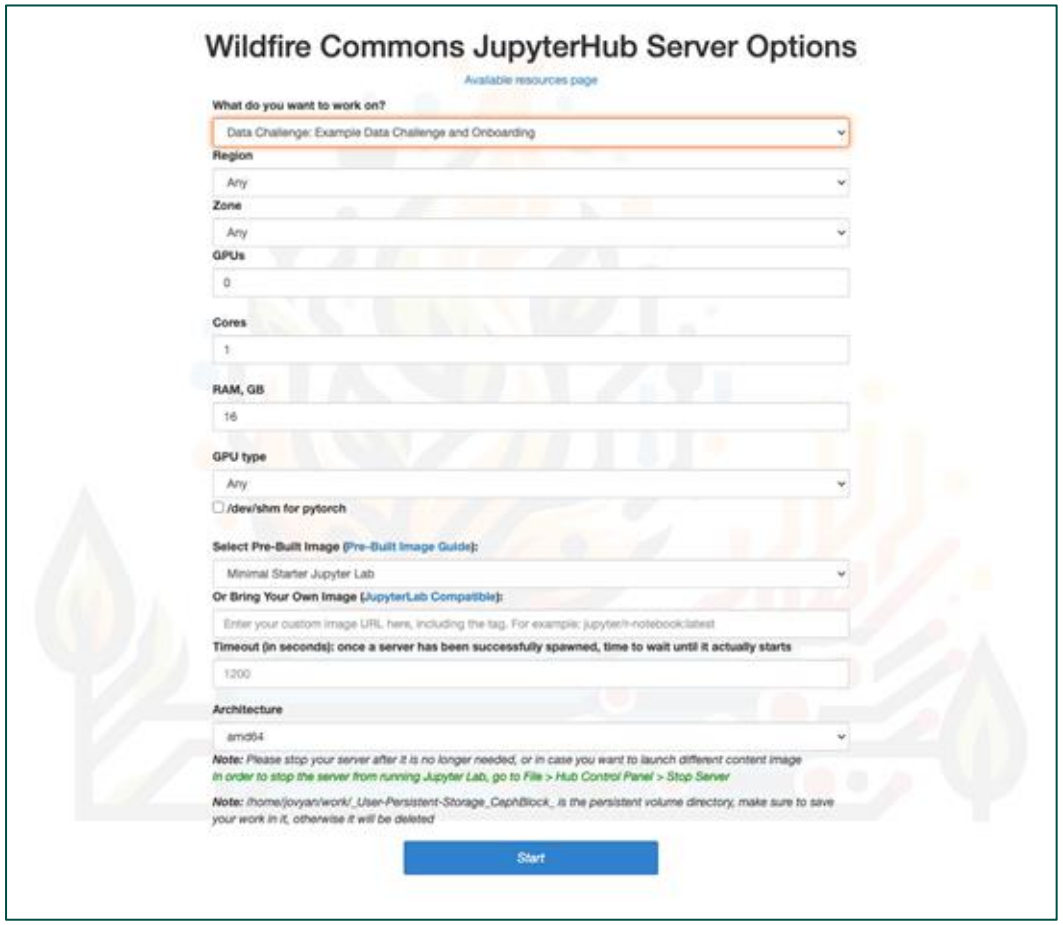
Launching JupyterHub

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

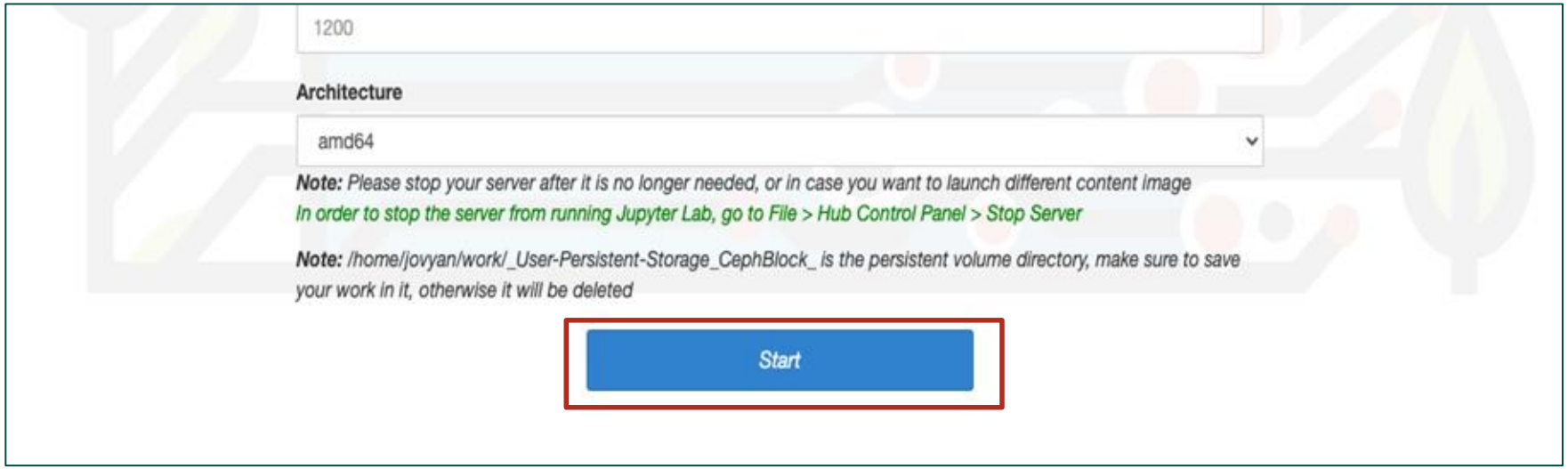
There will be multiple offerings for compute resources.

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



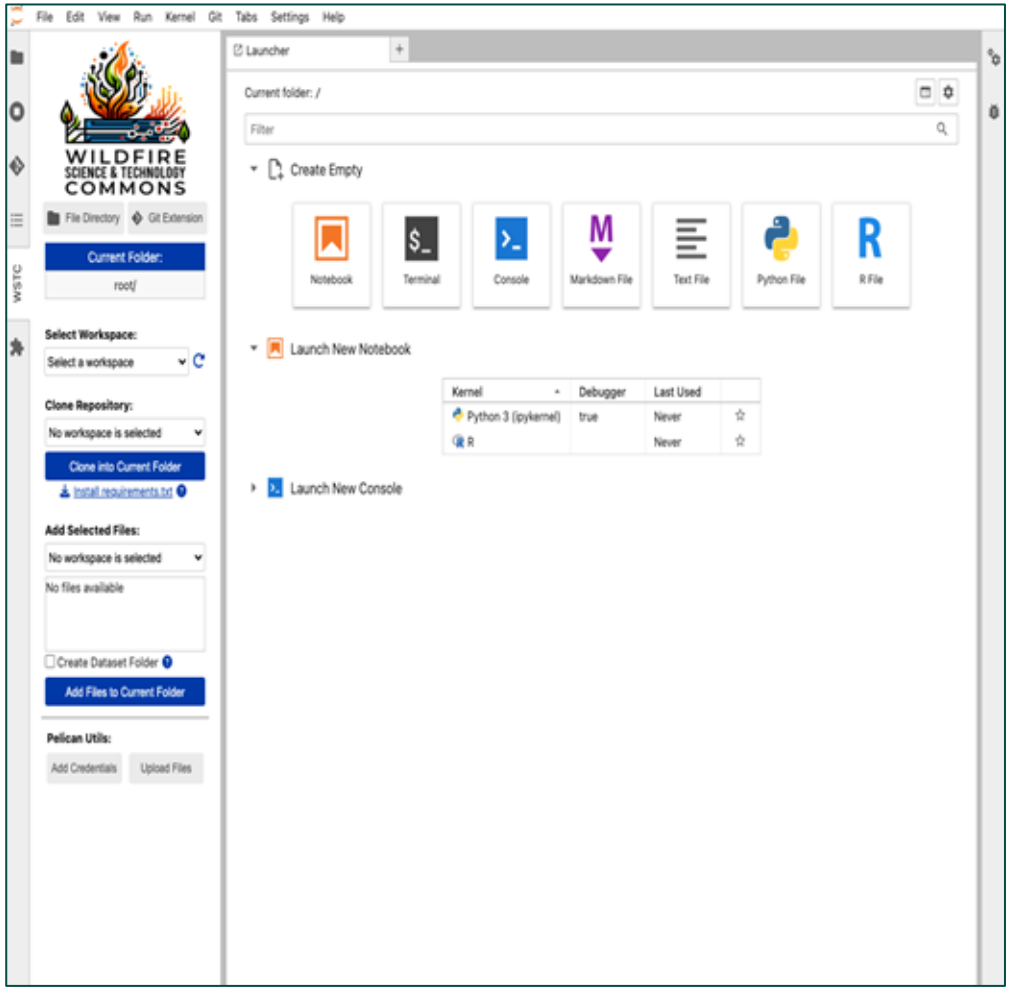
Launching JupyterHub

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



Using JupyterHub

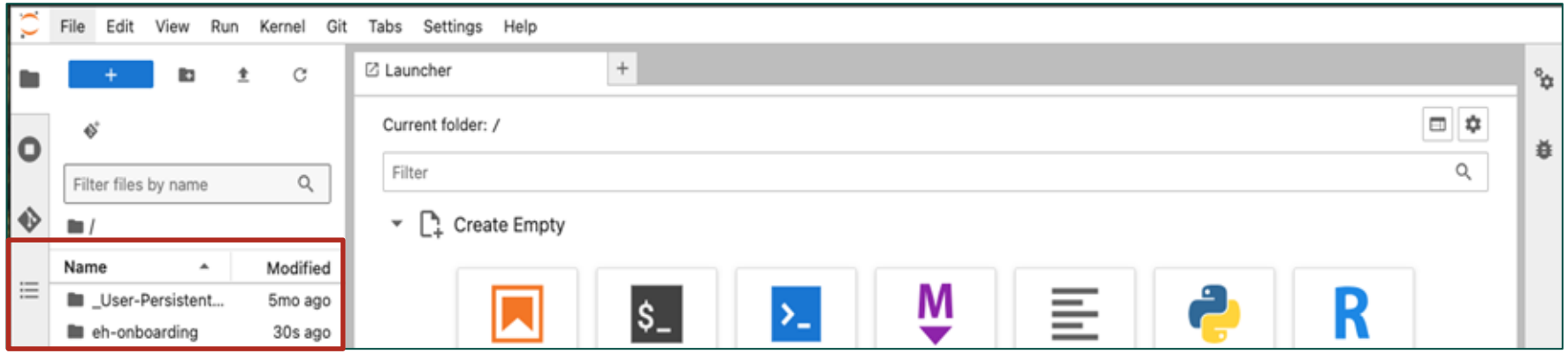
Once your server initiates, it will look like this screen →



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.**

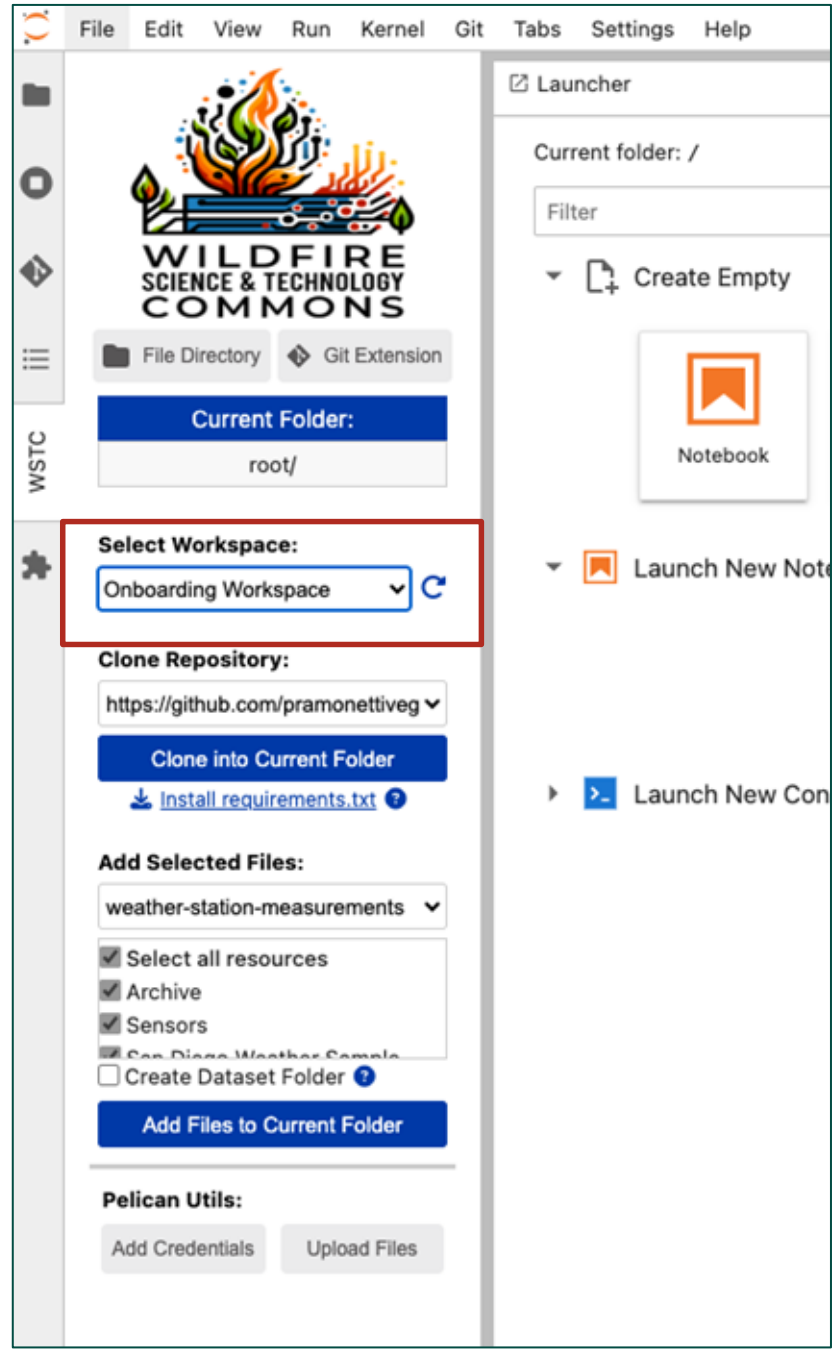
Good practices about shared storage

- 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

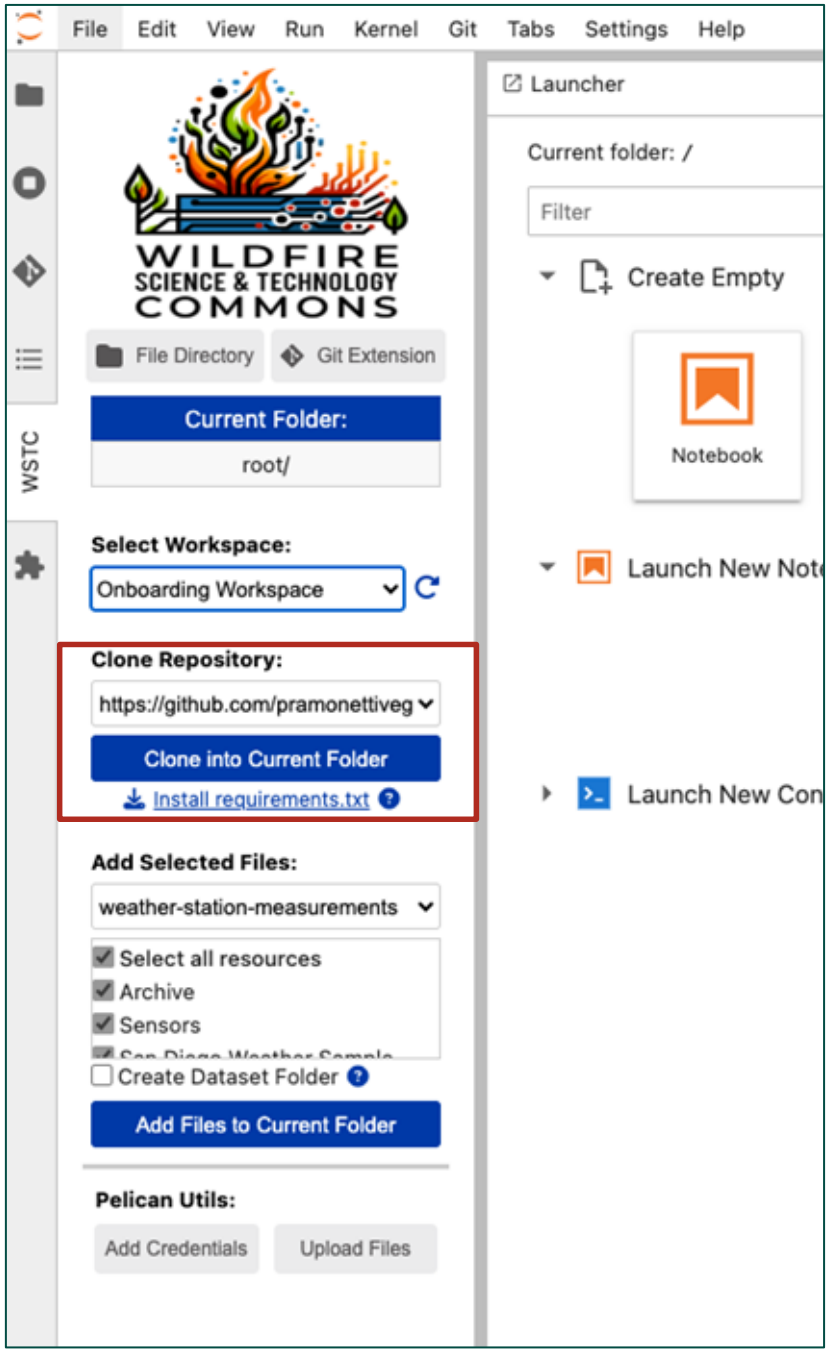
In this dropdown menu →

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



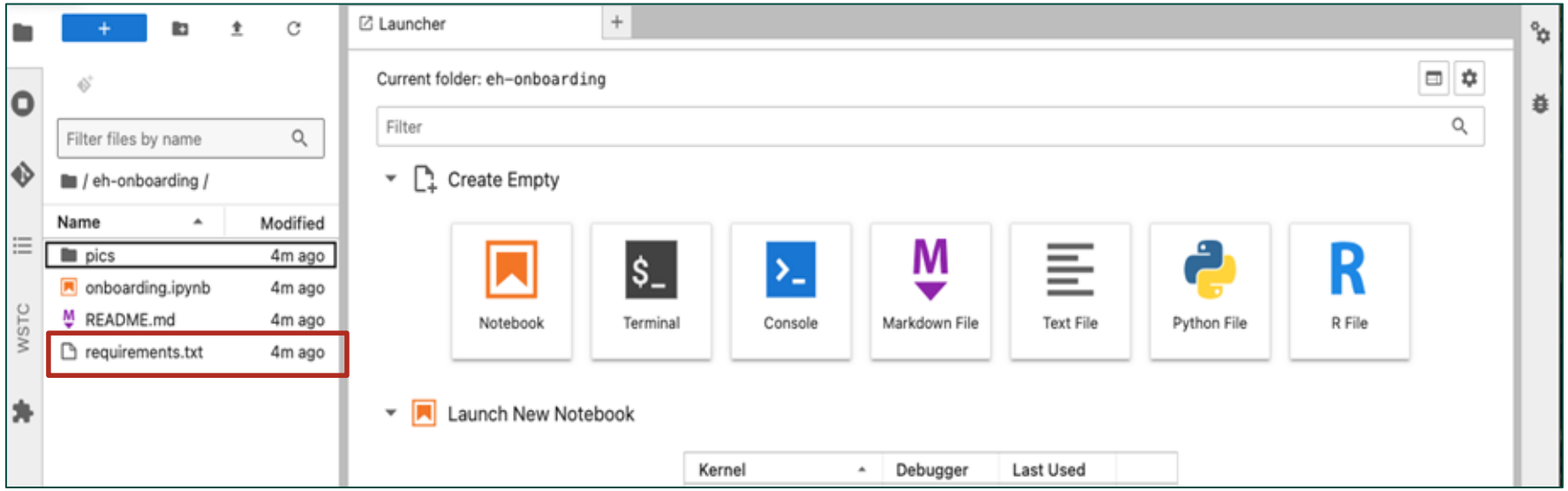
Using JupyterHub

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



Using JupyterHub

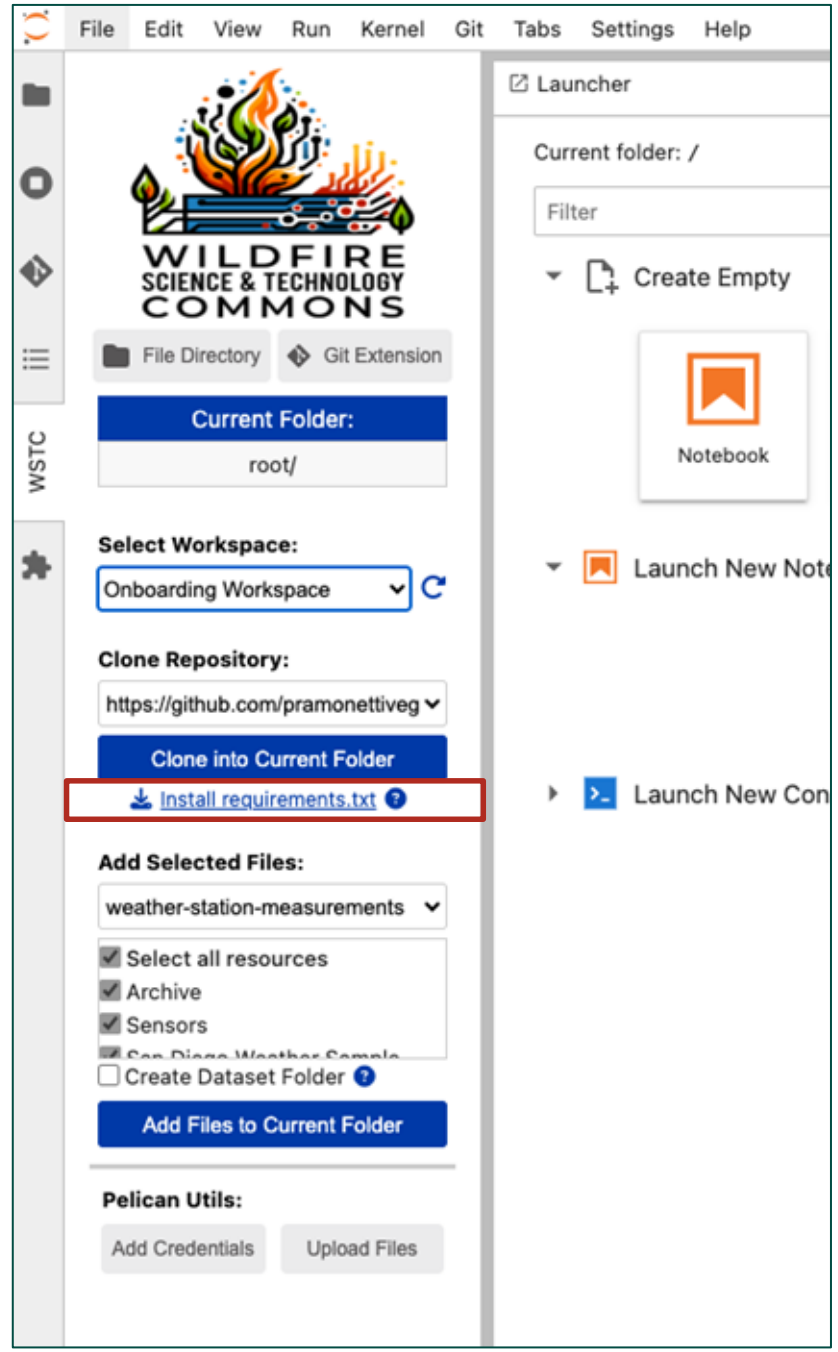
After cloning the repository and moving the cloned folder, the repository may contain a requirements.txt file.



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.

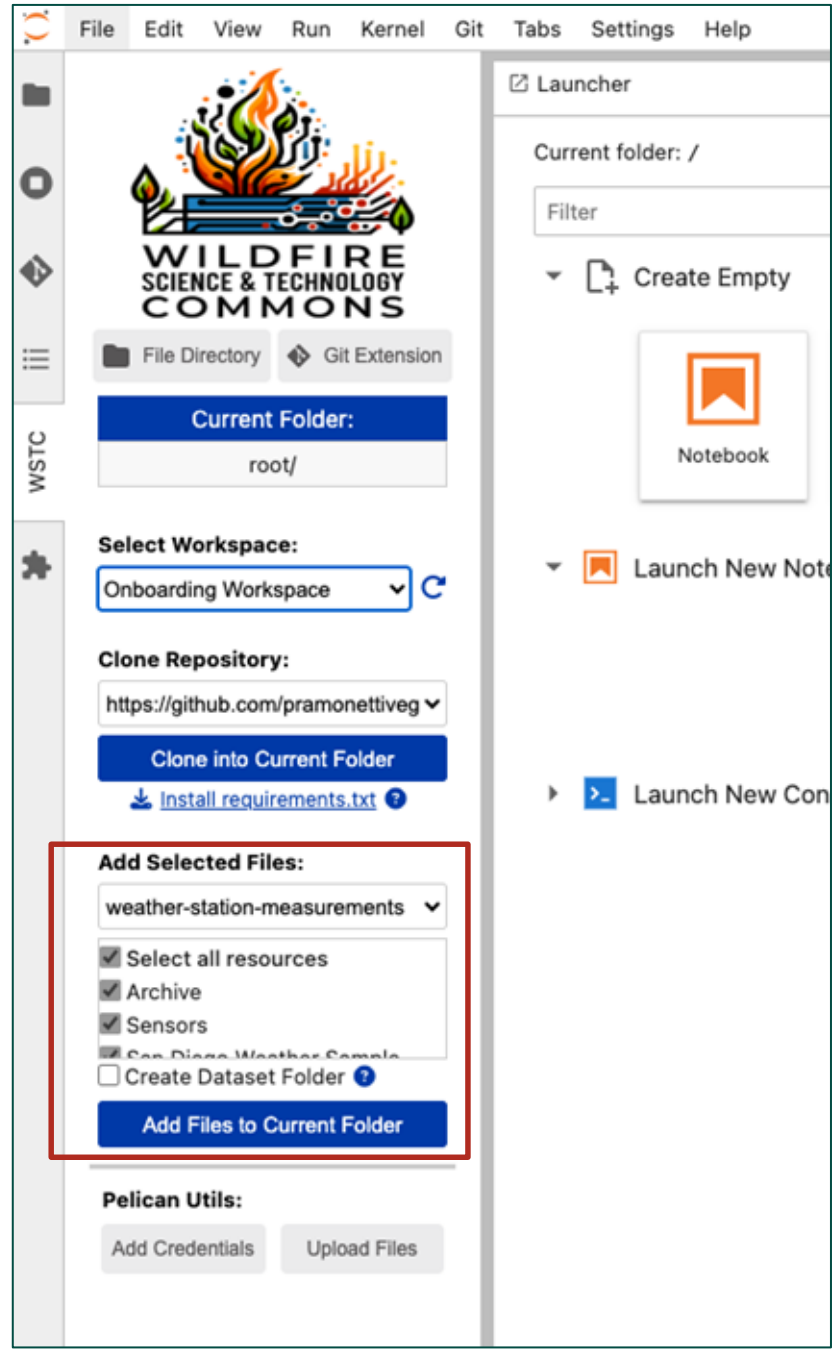


Using JupyterHub

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.

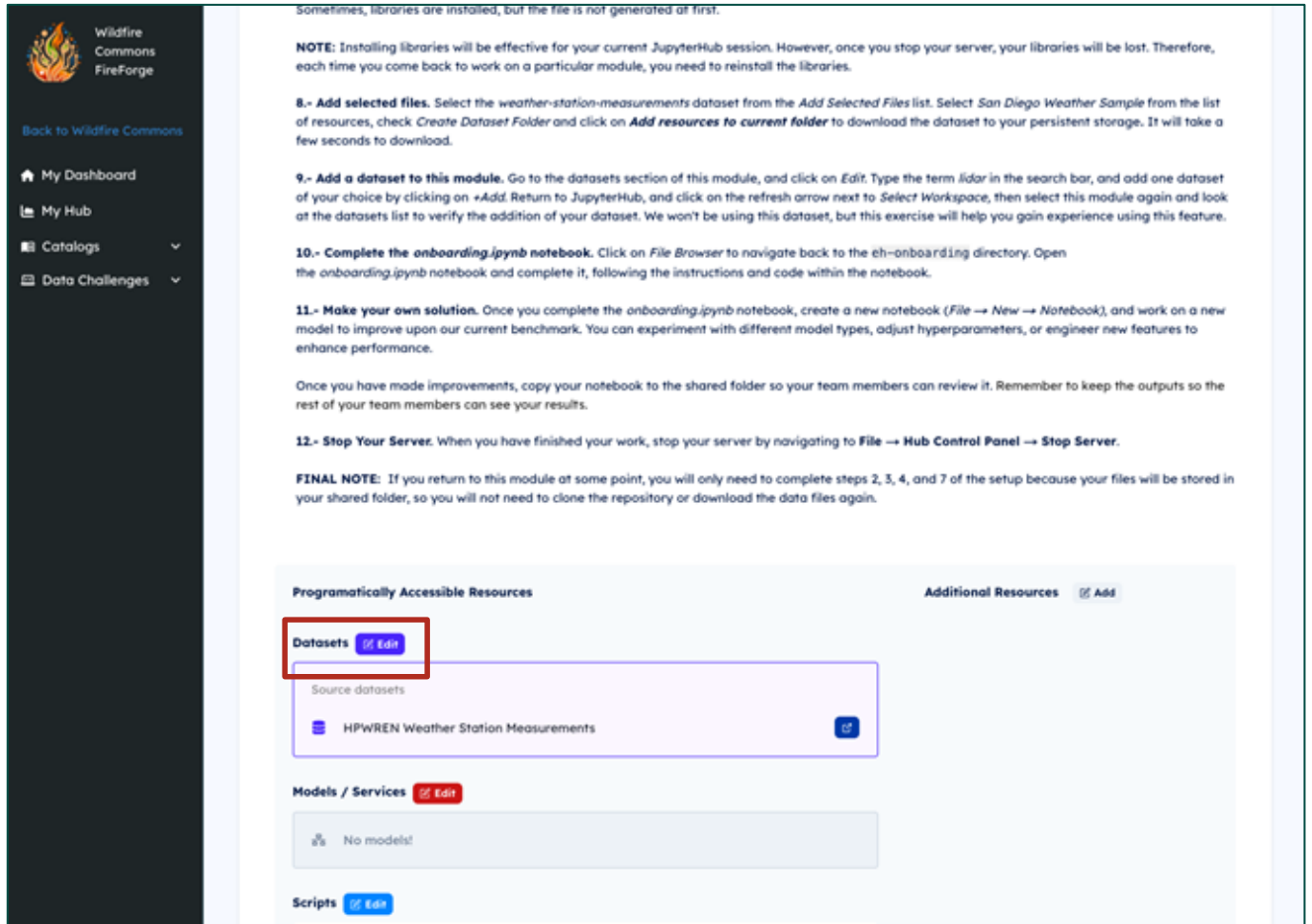
Click “Add files to Current Folder” to download them.



Using JupyterHub

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

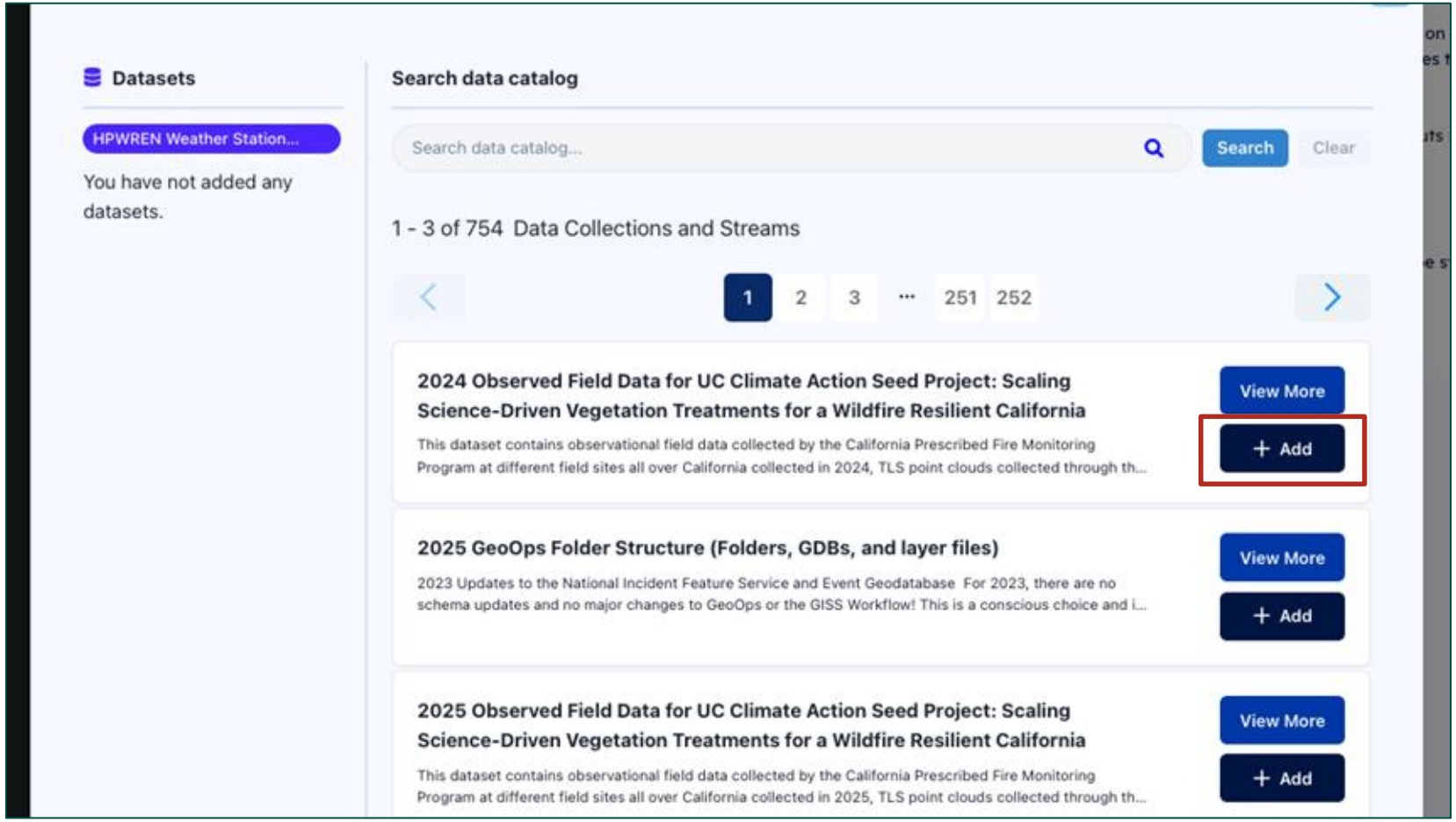
If this is the case, go to the module *(in a separate window)*, and click on the Datasets “Edit” button.



Using JupyterHub

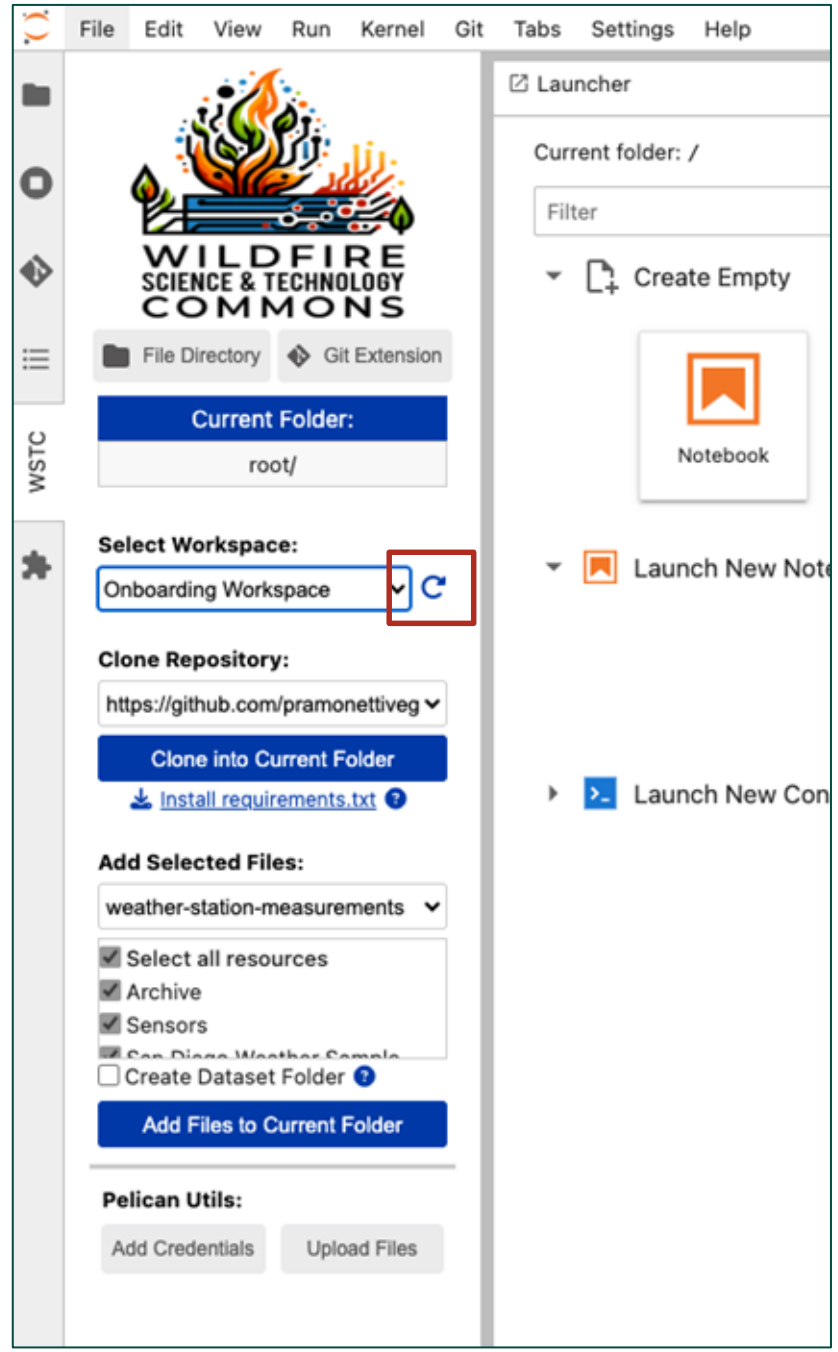
Search for the dataset of your interest.

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



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

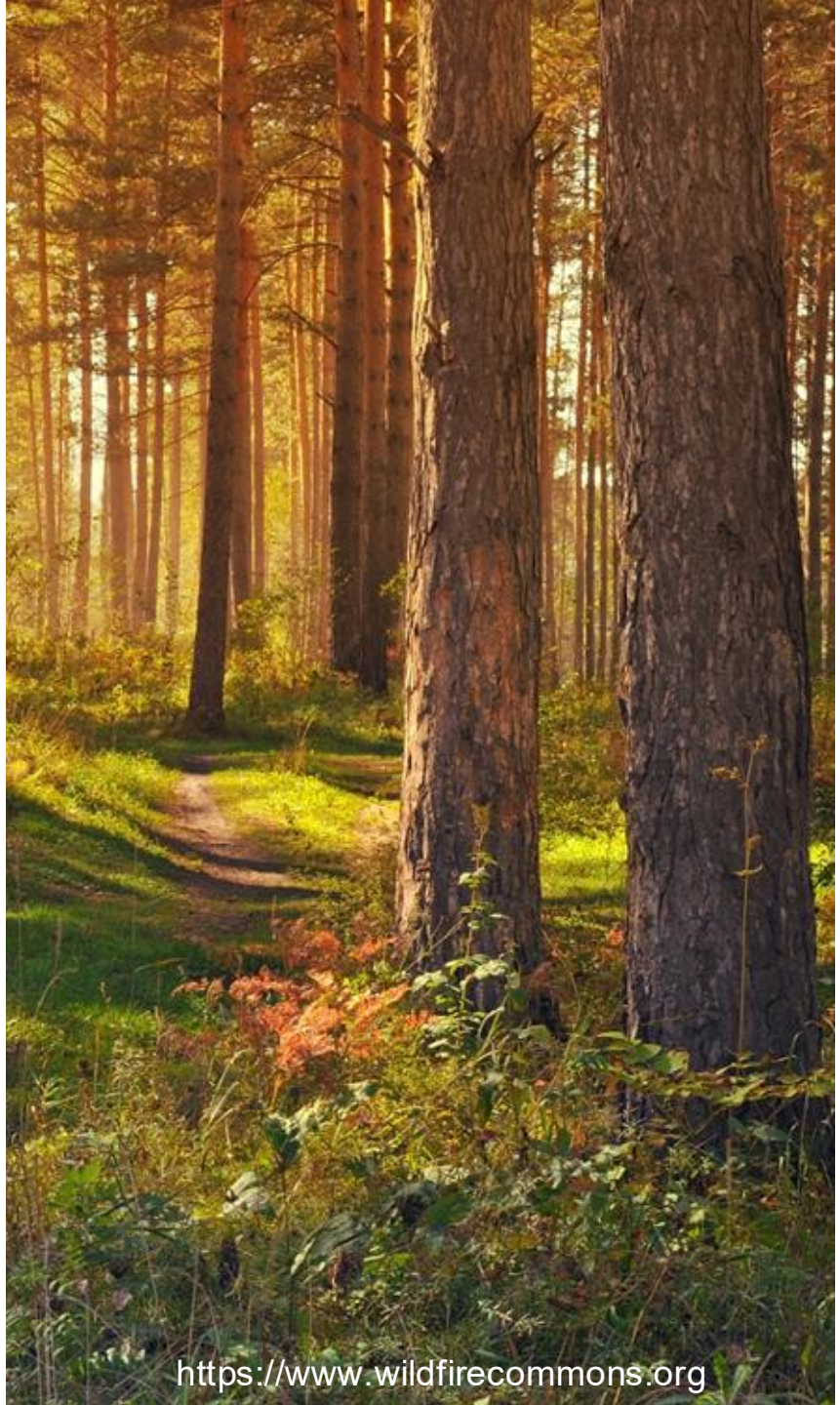
Work on the Module

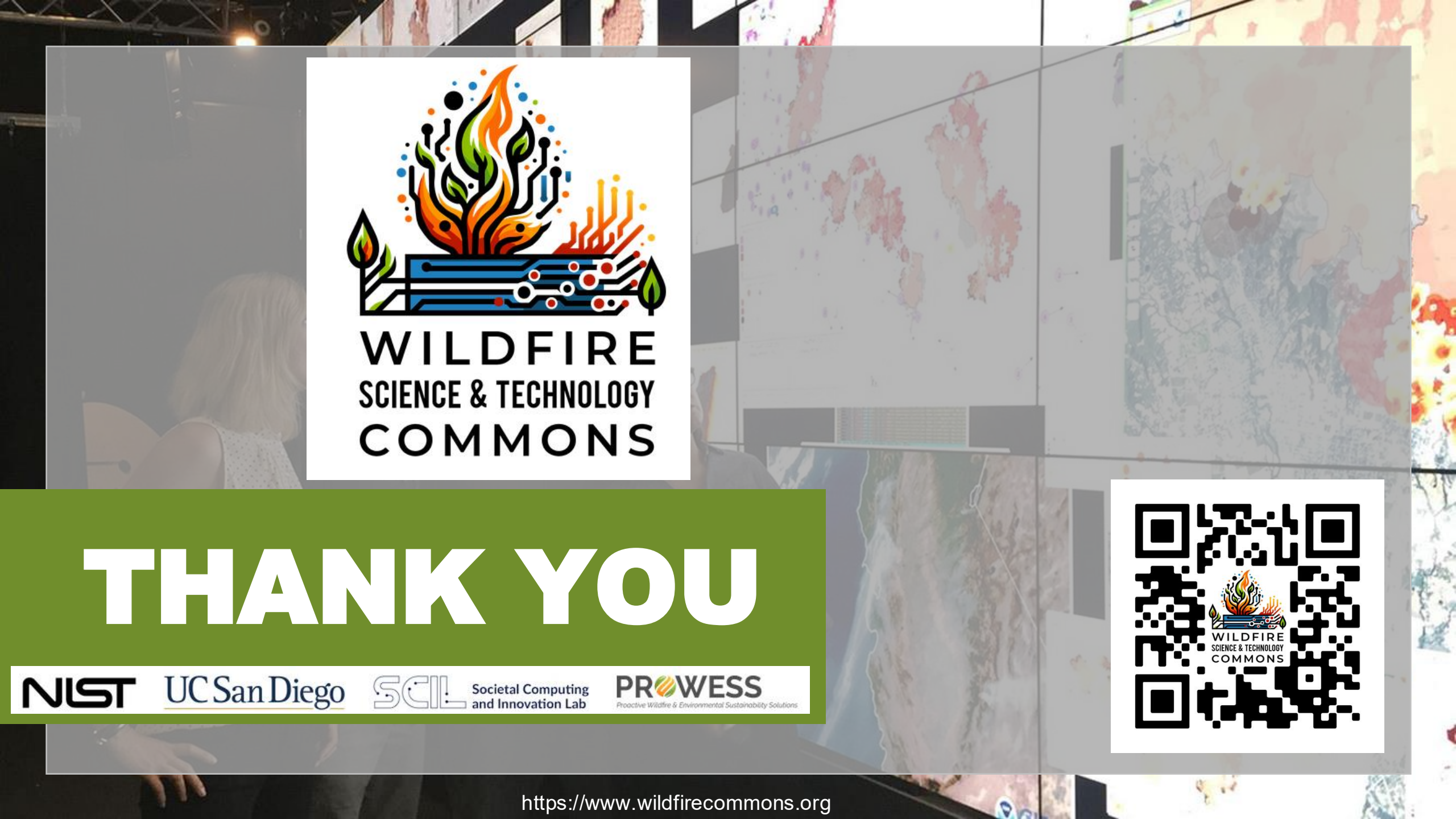
Begin working on the instructional modules as soon as they are released

05

Submit Your All Your Work by April 19

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





THANK YOU

NIST UC San Diego SCIL Societal Computing and Innovation Lab PROWESS Proactive Wildfire & Environmental Sustainability Solutions

https://www.wildfirecommons.org