



# The Openscapes Approach: Building Open Science Communities

Aaron Friesz

KBR Inc. contractor to the USGS EROS Center

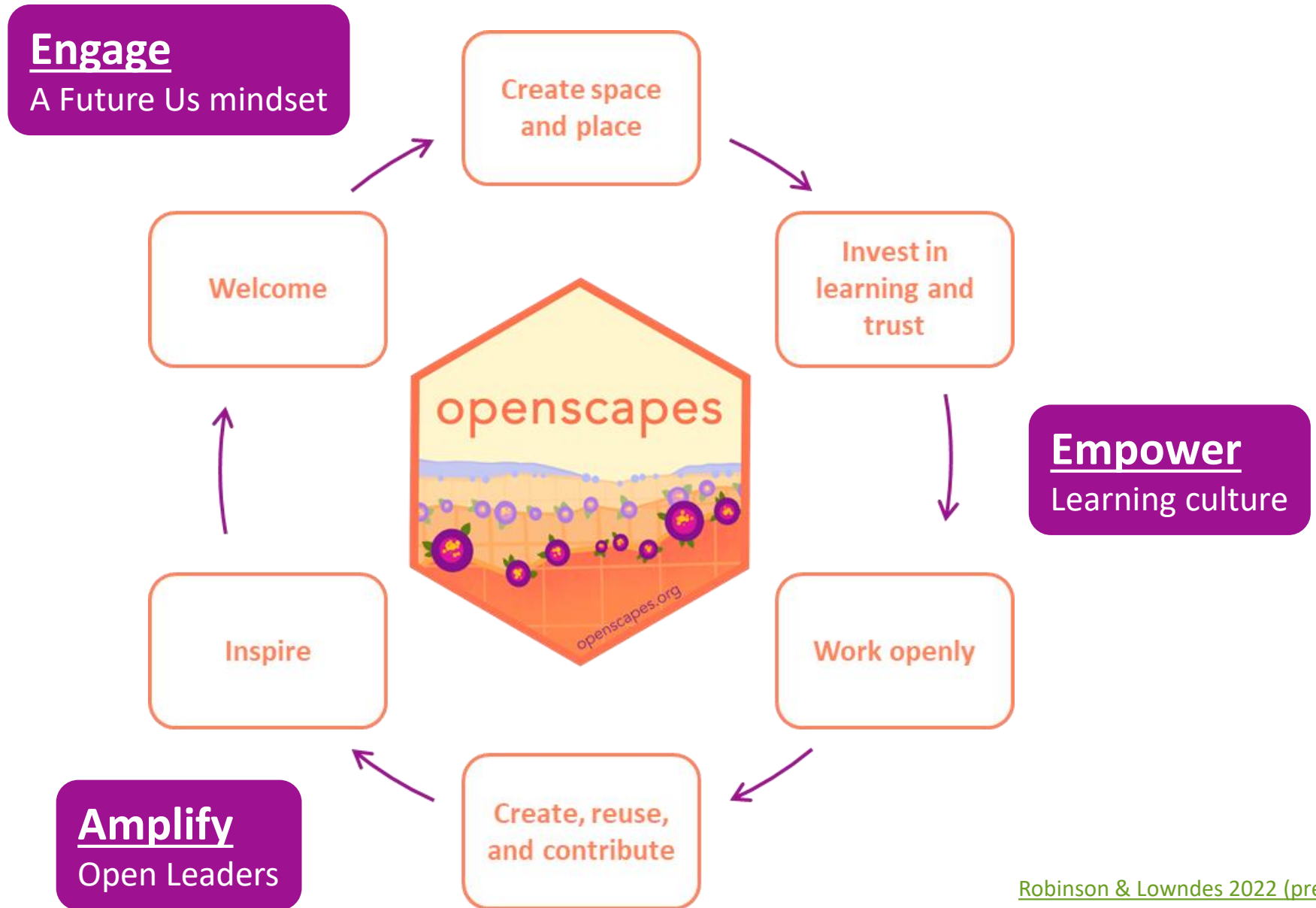
AGU 2023

**Openscapes** helps researchers move from lonely science to open and collaborative science as they explore and navigate the open science landscape safely and learn new technical skills.

**Culture shift = technical + human**

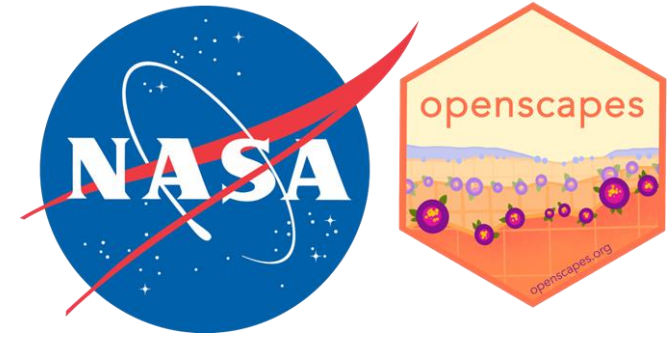


**The Openscapes Flywheel**  
A framework for managers to facilitate and scale inclusive Open science practices





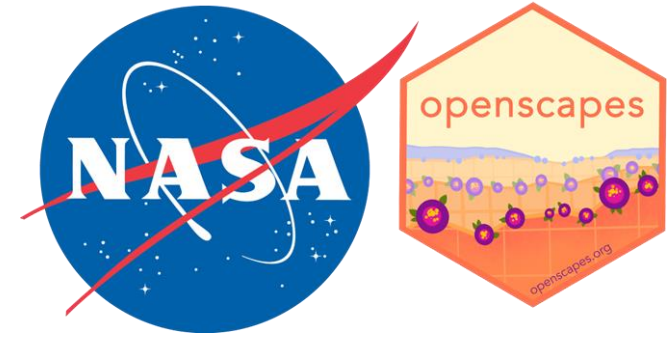
# NASA Openscapes



We are a mentor community across multiple NASA Distributed Active Archive Centers (DAACs)



# NASA Openscapes Framework



**The overarching vision** is to support scientific research teams using NASA EOSDIS data as they migrate their workflows to the cloud. We are doing this work with NASA DAACs by:

1. Developing a cross-DAAC Mentor community
2. Providing access to a cloud playground environment
3. Empowering science teams through workshops, hackathons, and the Champions program
4. Scaling the Openscapes Champions program with DAAC Mentors

# NASA Openscapes – Migration to Earthdata Cloud



## NASA Openscapes Mentor Community

### DAAC Staff

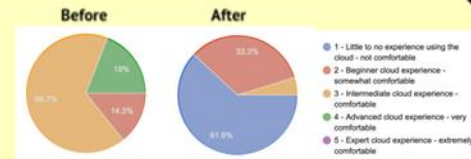
- Lay a foundation with **cloud terminology and concepts**
- Provide resources that are **easy to revisit**
- **Continued support** and education are critical
- Significant **learning curve and time investment** required for cloud adoption



Shared 2i2c cloud environment featuring JupyterHub

### End-Users

- Improved **conceptual understanding** of why and when to use, or not use, the cloud
- Inconsistent data and service availability leads to **difficulties reusing** a given workflow
- Lack of common and robust resources
- Earthdata Cloud ecosystem is **complex and overwhelming**



Sentiments from cloud workshop

### Open Science Community

- Recognizing **easy cloud access as a core service**
- Continuing to close the loop between the users we work with and our engineers to **build solutions together**



**Cheatsheets** are a one-stop shop for cloud data access vocabulary & roadmaps (see poster IN22C-0320 for all cheatsheets)

### NASA Earthdata Cookbook

is a central resource for common tutorials, use cases, and self-guided learning

### earthaccess

**Python library** is an open-source library to simplify Earthdata Cloud search and access





# 2021 Cloud Hackathon

## Preparation

- 9 co-created tutorials for data access
- User-friendly book with Quarto
- Notebook review, teaching dry runs
- Shared facilitation & teaching practices

## The event

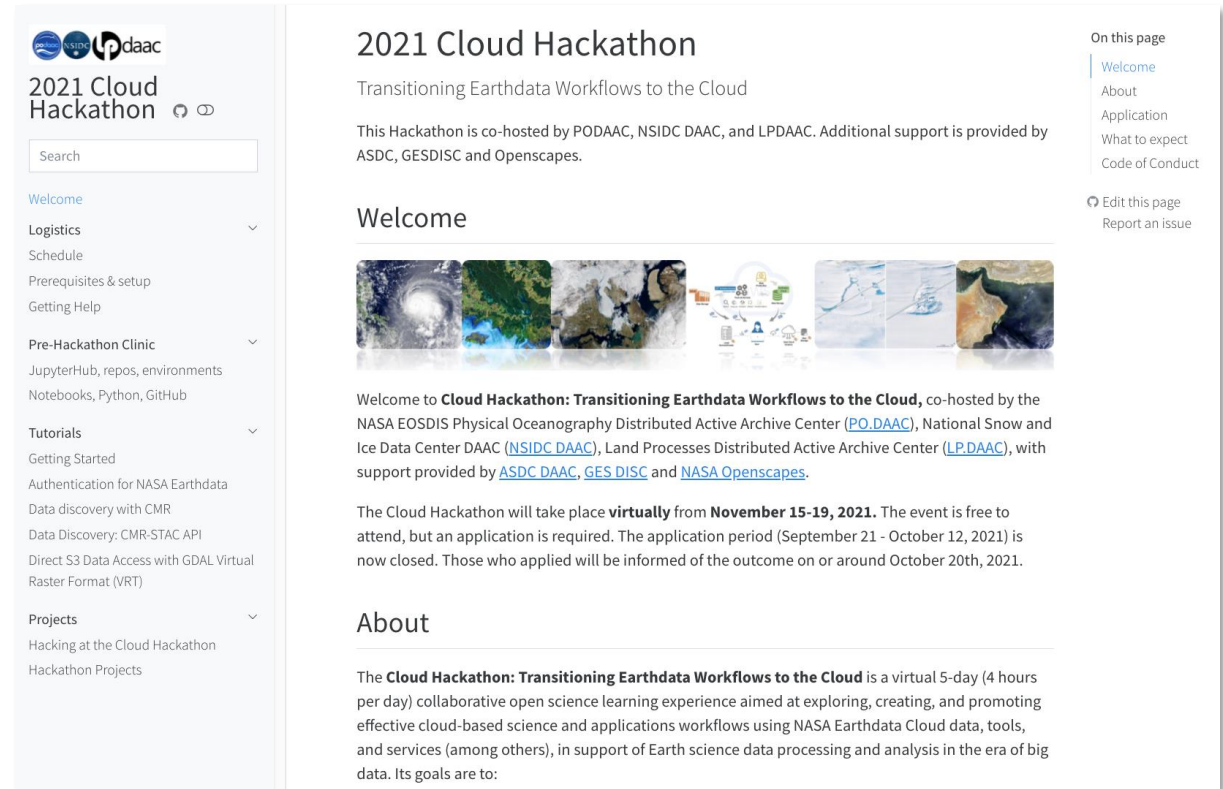
- 65 2i2c JupyterHub AWS instances
- 50 forks of the GitHub repo
- 8 hack-team projects presented on Day 5

*"It was a really great week. The tutorials were AMAZING. Everyone did a great job, and everyone was very nice. I really appreciated welcoming environment. I don't have a strong python background. But i was supported in learning all around"*

## Blog summaries

[earthdata.nasa.gov/learn/articles/2021-cloud-hackathon](https://earthdata.nasa.gov/learn/articles/2021-cloud-hackathon)

[podaac.jpl.nasa.gov/announcements/2021-12-15-The-2021-Cloud-Hackathon](https://podaac.jpl.nasa.gov/announcements/2021-12-15-The-2021-Cloud-Hackathon)



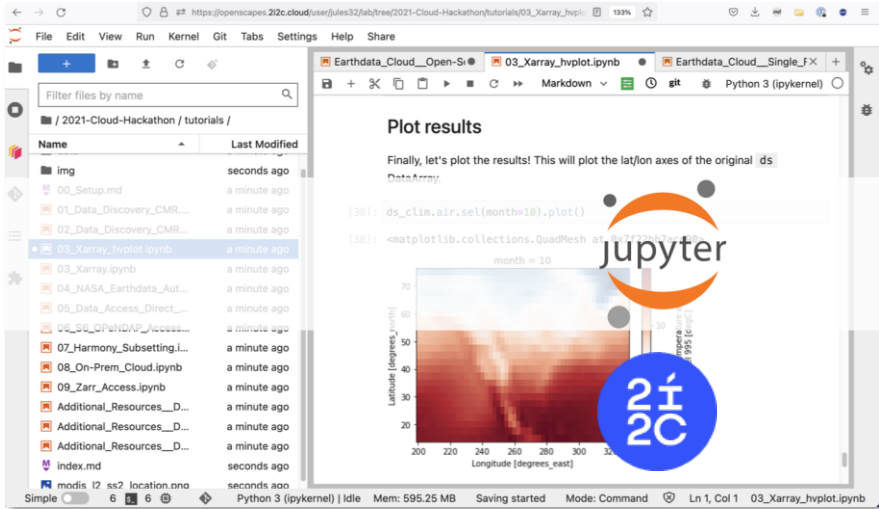
The screenshot shows the website for the 2021 Cloud Hackathon. The header includes logos for PODAAC, NSIDC, and LPDAAC. The main title is "2021 Cloud Hackathon" with a subtitle "Transitioning Earthdata Workflows to the Cloud". A search bar is present. A navigation menu on the left lists sections: Welcome, Logistics, Pre-Hackathon Clinic, Tutorials, and Projects. The main content area has a "Welcome" section with a banner image of Earth data and a paragraph introducing the event, co-hosted by NASA EODIS Physical Oceanography Distributed Active Archive Center (PO.DAAC), National Snow and Ice Data Center DAAC (NSIDC DAAC), and Land Processes Distributed Active Archive Center (LP.DAAC). It also mentions support from ASDC DAAC, GES DISC, and NASA Openscapes. Below this is an "About" section describing the event as a virtual 5-day collaborative open science learning experience.

On Day 1, Mentors stepping in to teach due to an emergency: trust + teamwork + familiarity with the material

<https://nasa-openscapes.github.io/2021-Cloud-Hackathon/>

# Identifying & Responding to User Needs

15+ workshops & talks led by Mentors since: reusing & extending tutorials with software & conceptual solutions



**2i2c JupyterHub:**  
Python, R, Matlab,  
**corn** base image:  
built on Pangeo  
stack

Python  
library



**Cookbook:**  
current  
tutorials &  
onboarding

NASA Earthdata Cloud Cookbook  
Supporting NASA Earth science research teams' migration to the cloud

AUTHOR: NASA OpenScapes Team  
PUBLISHED: May 1, 2022

Welcome

Welcome to the NASA OpenScapes EarthData Cloud Cookbook!

This Cookbook is learning-oriented to support scientific researchers using NASA Earthdata from Distributed Active Archive Centers (DAACs) as they migrate their workflows to the cloud. It has resources to learn how to work with NASA Earthdata on the cloud, as well as documentation for how to contribute to these resources. It is under active, open development.

The Cloud Paradigm

NASA Distributed Active Archive Centers (DAACs) are in the process of moving their data holdings to the cloud. In the new paradigm, data storage (orange in the illustration) will migrate to the cloud (green) and DAAC-provided tools and services built on top of the data are co-located in the Earthdata Cloud.

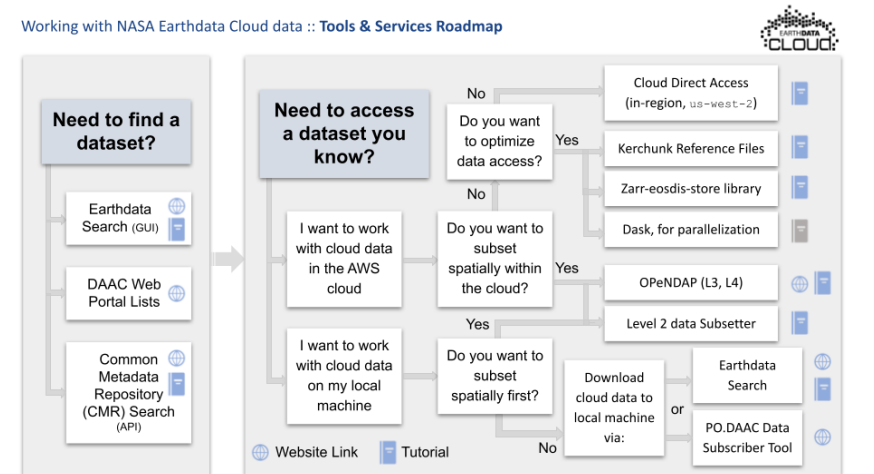
## Value of Hosted JupyterHubs White paper / RFI

The Value of Hosted JupyterHubs in enabling Open NASA Earth Science in the Cloud

Response to Aspect 1, the question on user needs and use cases for scientific data and computing in support of Open Science at SMD

Relevant NASA SMD scientific Division: *Earth Science*

**Cheatsheets  
& guides**





# NASA Openscapes Champions

NASA Openscapes Champions is a mentorship and professional development opportunity for research teams using data from NASA DAACs and interested in open science and migrating their analytical workflows to the cloud.

To date we have supported 17 teams in migrating their workflows to the cloud.

## Benefits to the Champions:

- Access to the 2i2c Hub for 1-year
- More direct support from the DAACs
- Pathway toward cloud migration

## Benefits to NASA DAACs

- Clear identification of user needs
- Success stories to showcase



Cohort Call Topics	Open science resources	Guest Teachers
1. Openscapes mindset, Better science in less time	<a href="#">mindset, better science in less time</a>	<a href="#">Jinbo Wang</a> , Caltech/JPL; <a href="#">Allan Just</a> , Mount Sinai
2. Team culture and data strategies for future us	<a href="#">team culture, data strategies for cloud</a>	<a href="#">Andy Barrett</a> , NSIDC
3. Open communities and coding strategies for future us	<a href="#">open communities</a> , coding strategies for cloud	<a href="#">Amy Steiker</a> , <a href="#">Luis Lopez</a> , NSIDC
4. NASA Earthdata Cloud Clinic, hands-on lesson from NASA Mentors	<a href="#">NASA Earthdata Cloud Clinic</a>	<a href="#">Amy Steiker</a> , NSIDC
5. Pathways share	<a href="#">Earthdata Cloud Cookbook</a>	<a href="#">Cassie Nickles</a> , PO.DAAC



# Scaling Openscapes – Building Open Science Communities

## The Openscapes Approach

### Diverse Perspectives

Organizations may have similar goals but rarely pursued them together. Without cooperation, the scope and quality of efforts is more likely to be more difficult and lacking impact in comparison to what is possible with diverse minds.

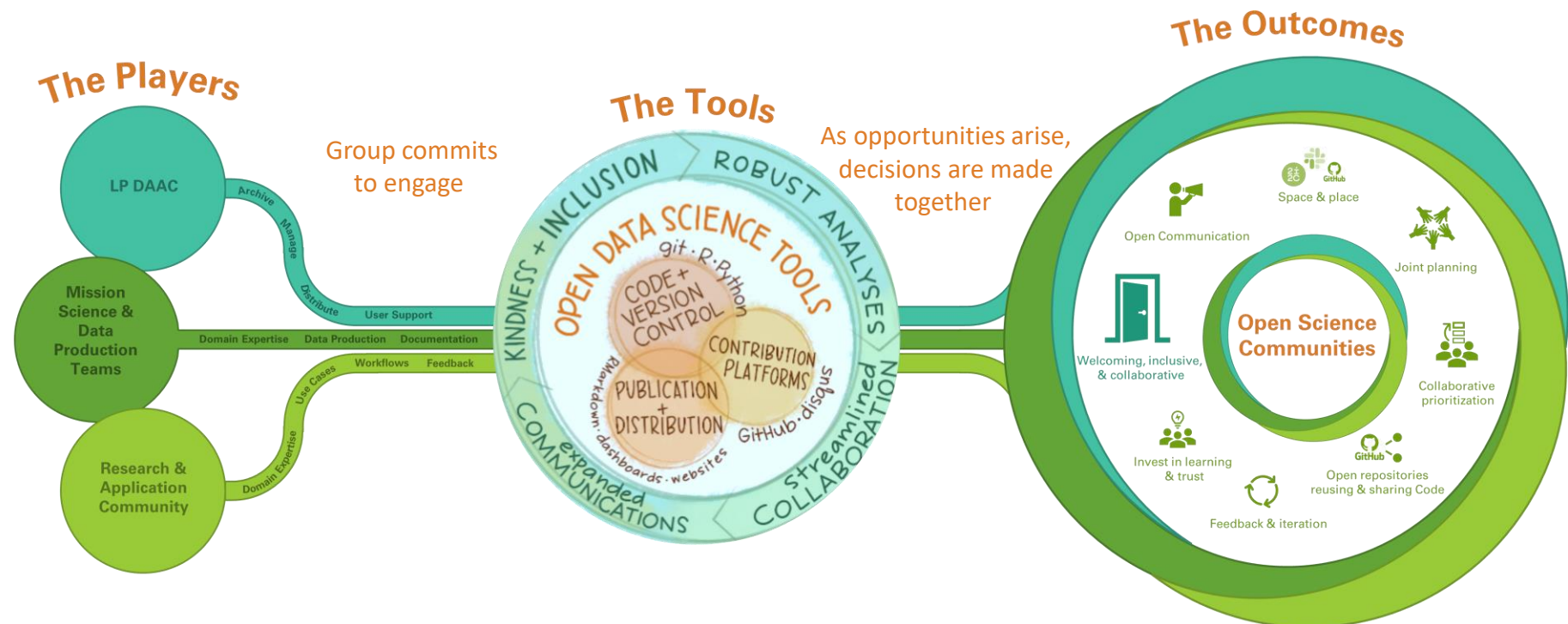
### Structural Consistency

People and organizations have different lived experiences and norms. Providing extremely consistent, accessible, and friendly spaces to experience the fullness of cooperation lets everyone engage at their own pace. Consistency converts skeptics and there are specific plans to co-develop resources.

### Better for everyone

The accumulation of individual and group successes facilitates a large-scale change in organizational and individual mindsets which has beneficial effects for everyone involved. Positive environments facilitate the production of timely reproducible resources as well as attracts more people, ideas, and materials that sustainably grown into better outcomes for more people.

Open Science Communities are groups of people openly creating, sharing, teaching, and collaborating around shared interests, with a culture of shared & continued learning; prioritizing diversity, equity, and belonging.



# Successes

daac  
VITALS

If running this notebook locally, you will find instructions to set up a compatible environment in the setup folder. If running on the OpenScapes 2I2c Cloud Instance for a Workshop, no additional setup is required.

## 1 Finding Concurrent ECOSTRESS and EMIT Data

Summary

Both the Ecosystem Spaceborne Thermal Radiometer Experiment on Space Station and the Earth surface Mineral dust source Investigation (EMIT) instruments are located on the International Space Station (ISS). Their overlapping fields of view provide an unprecedented opportunity to demonstrate the compounded benefits of working with both datasets. This notebook will show how to utilize the [earthaccess Python library](#) to find concurrent ECOSTRESS and EMIT data.

Search

### NASA's Earth Surface Mineral Dust Source Investigation (EMIT...)

NASA Earthdata  
4 videos • 428 views • Last updated on Jul 7, 2023

Play all Shuffle

This playlist includes information about NASA's Earth Surface Mineral Dust Source Investigation (EMIT) mission, measurements, and applications and also provides workshop

- EMIT Data Tutorial Series Workshops Week 1: Intro to EMIT Mission and Data**  
NASA Earthdata • 1.2K views • 7 months ago
- EMIT Data Tutorial Series Workshops Week 2: Working with EMIT Data - The Basics**  
NASA Earthdata • 711 views • 7 months ago
- EMIT Data Tutorial Series Workshops Week 3: Working with EMIT Data - Advanced**  
NASA Earthdata • 409 views • 6 months ago
- Meet EMIT, the Newest Imaging Spectrometer on the International Space Station**  
NASA Earthdata • 1.1K views • 4 months ago

## EMIT Webinar Series



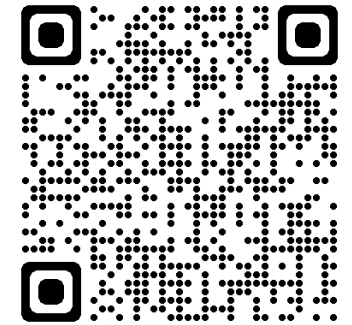


# Successes

The image is a collage of four main components:

- Top Left:** A YouTube playlist titled "NASA's Earth Surface Mineral Dust Source Investigation [EMIT...]" with a "Play all" button.
- Top Center:** A screenshot of a Jupyter notebook page titled "1 Finding Concurrent ECOSTRESS and EMIT Data". The page includes a "Summary" section and a map of the Earth's surface.
- Top Right:** A Zoom meeting window titled "UP-EMIT-ECOSTRESS joint workshop log up" showing several participants in a grid view.
- Bottom Center:** A screenshot of the GitHub repository page for "nasa / VITALS". The repository is public and contains 9 branches and 1 tag. The file list includes folders like "data", "img", "modules", "python", "setup", "user\_contributed" and files like ".gitignore" and "CHANGE\_LOG.md".

**VSWIR Imaging and Thermal Applications, Learning, and Science Repository**



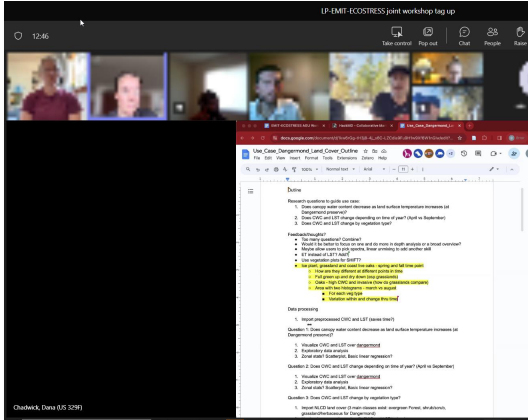
# Successes

The collage illustrates various aspects of the project's success:

- GitHub:** A screenshot of the 'VITALS' repository, showing a file tree with folders like 'data', 'img', 'modules', 'python', 'setup', 'user\_contributed', and 'gitignore'.
- YouTube:** A video player showing a series of workshops titled 'EMIT Data Tutorial Series Workshops Week 1: Intro to EMIT Mission and Data' through 'Week 3: Working with EMIT Data - Advanced'.
- Jupyter Notebook:** A notebook titled '1 Finding Concurrent ECOSTRESS and EMIT Data' with a table of contents and a map of the Los Angeles area.
- Zoom Meeting:** A screenshot of a Zoom meeting titled 'LP-EMIT-ECOSTRESS joint workshop tag up' with several participants visible in a grid view.
- Google Docs:** A document titled 'Use Case\_Dangermond\_Land\_Cover\_Outline' containing research questions, feedback thoughts, and data processing steps.

Co-working / co-developing for the AGU Workshop

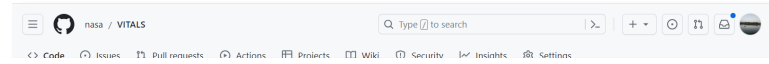
# Successes



VITALS



- Welcome >
- 2023 AGU Workshop
- Repository Description
- Workshop Schedule
- Background
- Setup Instructions >
- Prerequisites
- Cloud Workspace Setup
- Local Setup
- Python Notebooks >
- 1 Finding Concurrent Data
- 2 EMIT Reflectance and ECOSTRESS LST
- 3 Canopy Water Content
- Contributing >
- Contributing to this Repository
- Contributor Code of Conduct

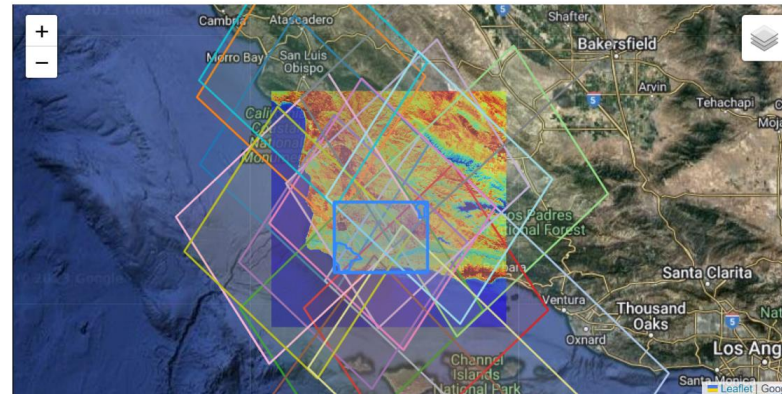


If running this notebook locally, you will find instructions to set up a compatible environment in the [setup](#) folder. If running on the Openscapes 2i2c Cloud Instance for a Workshop, no additional setup is required.

## 1 Finding Concurrent ECOSTRESS and EMIT Data

### Summary

Both the Ecosystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) and the Earth surface Mineral dust source Investigation (EMIT) instruments are located on the International Space Station (ISS). Their overlapping fields of view provide an unprecedented opportunity to demonstrate the compounded benefits of working with both datasets. In this notebook we will show how to utilize the [earthaccess](#) [Python library](#) to find concurrent ECOSTRESS and EMIT data.



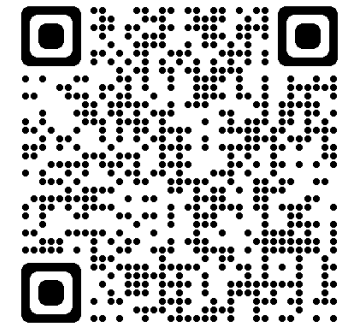
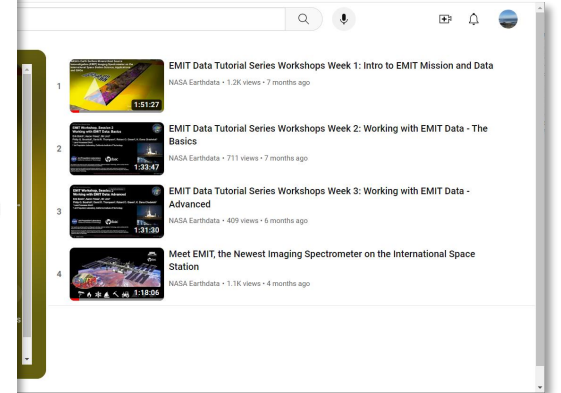
On this page

### 1 Finding Concurrent ECOSTRESS and EMIT Data

1. Setup
2. Search for ECOSTRESS and EMIT Data
3. Organizing and Filtering Results
4. Visualizing Intersecting Coverage
5. Generating a list of URLs and downloading data

Contact Info:

Appendices



## VITALS Webbook for the AGU Workshop

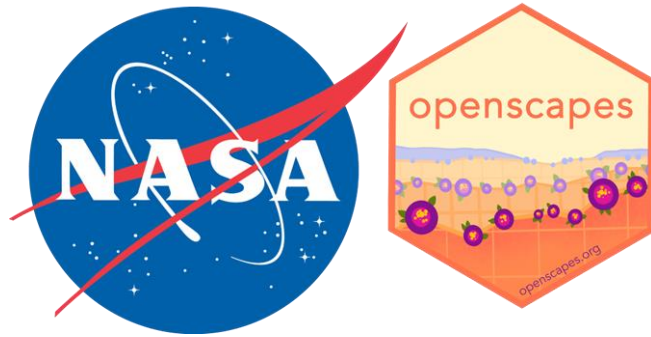


# Challenges



- Learning curve for cloud, Git, and GitHub is steep
- Continued availability to address questions and problems
- Breaking down complex topics (tech/terminology and science)
- Organizing resources / reducing duplication
- Making diverse resources (Python, R, etc.)
- Managing resource updates, issues, and pull requests (PRs)
- Confidence in sharing code
- Mission oriented funding for outreach and community participation





Thank You



Openscapes artwork by Allison Horst