The Openscapes Approach: Building Open Science Communities

Aaron Friesz

KBR Inc. contractor to the USGS EROS Center

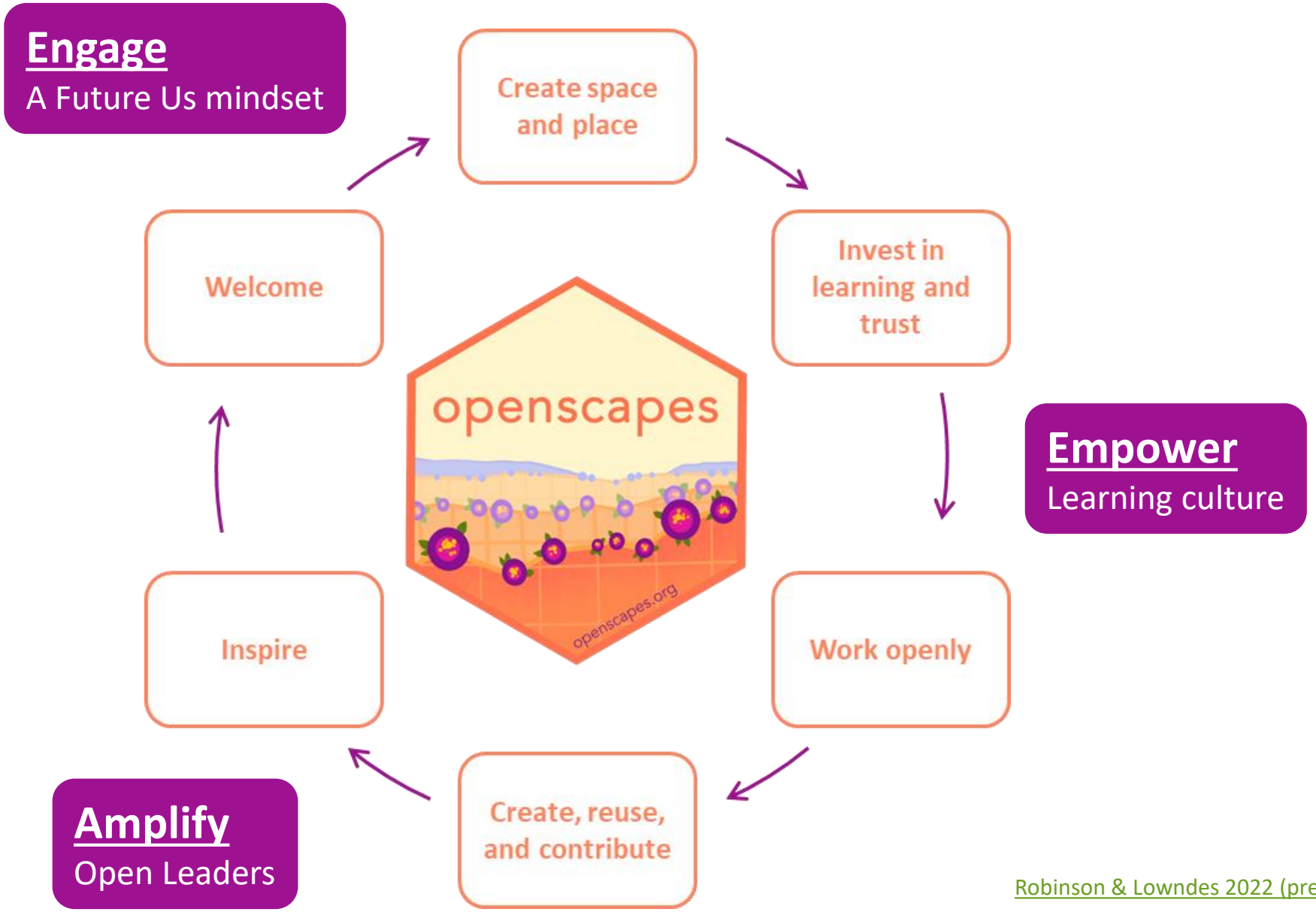
AGU 2023

U.S. Department of the Interior
U.S. Geological Survey

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



Robinson & Lowndes 2022 (preprint)

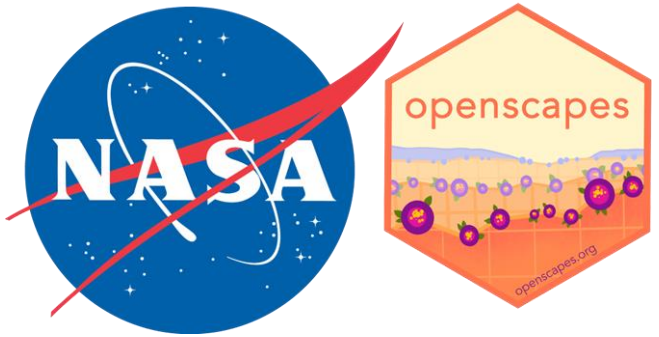
U.S. Geological Survey NASA

NASA Openscapes

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

U.S. Geological Survey

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

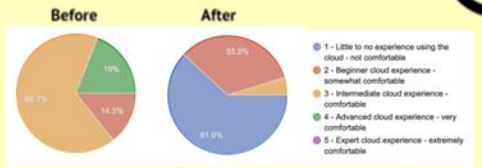
DAAC Staff

- 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

U.S. Geological Survey

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

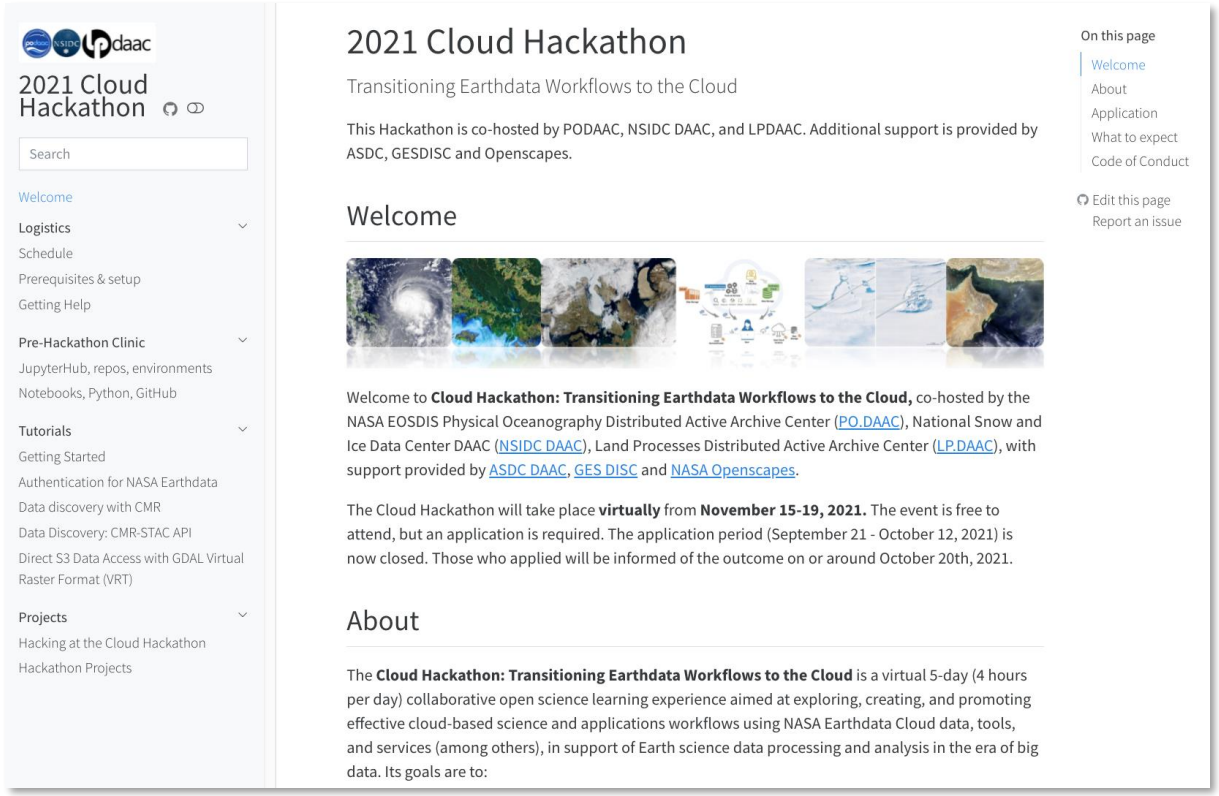
The event

- 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
podaac.jpl.nasa.gov/announcements/2021-12-15-The-2021-Cloud-Hackathon



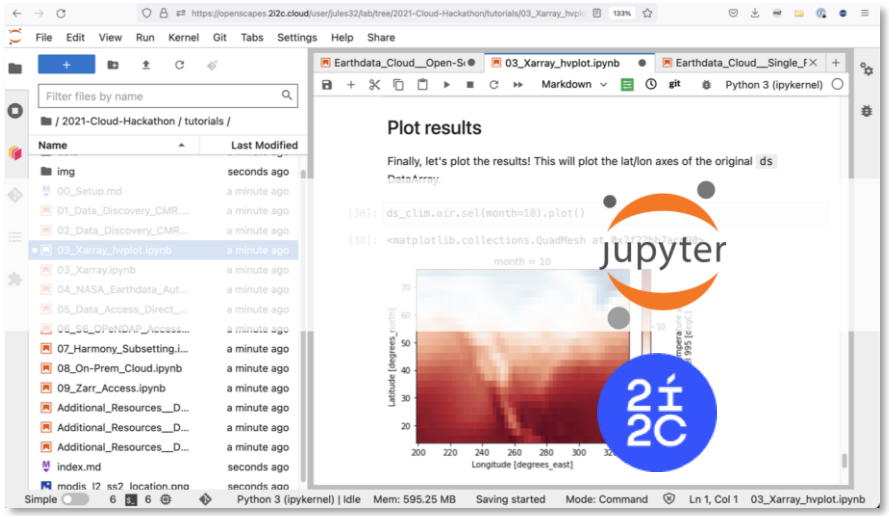
On Day 1, Mentors stepping in to teach due to an emergency: trust + teamwork + familiarity with the material

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

U.S. Geological Survey NASA

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

Cookbook: current tutorials & onboarding

Python library

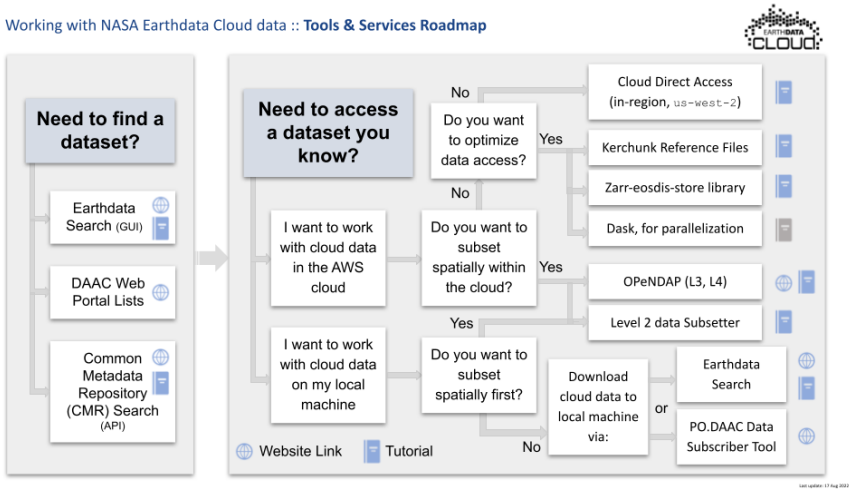
Value of Hosted JupyterHubs White paper / RFI

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.

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

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

U.S. Geological Survey NASA

Scaling Openscapes – Building Open Science Communities

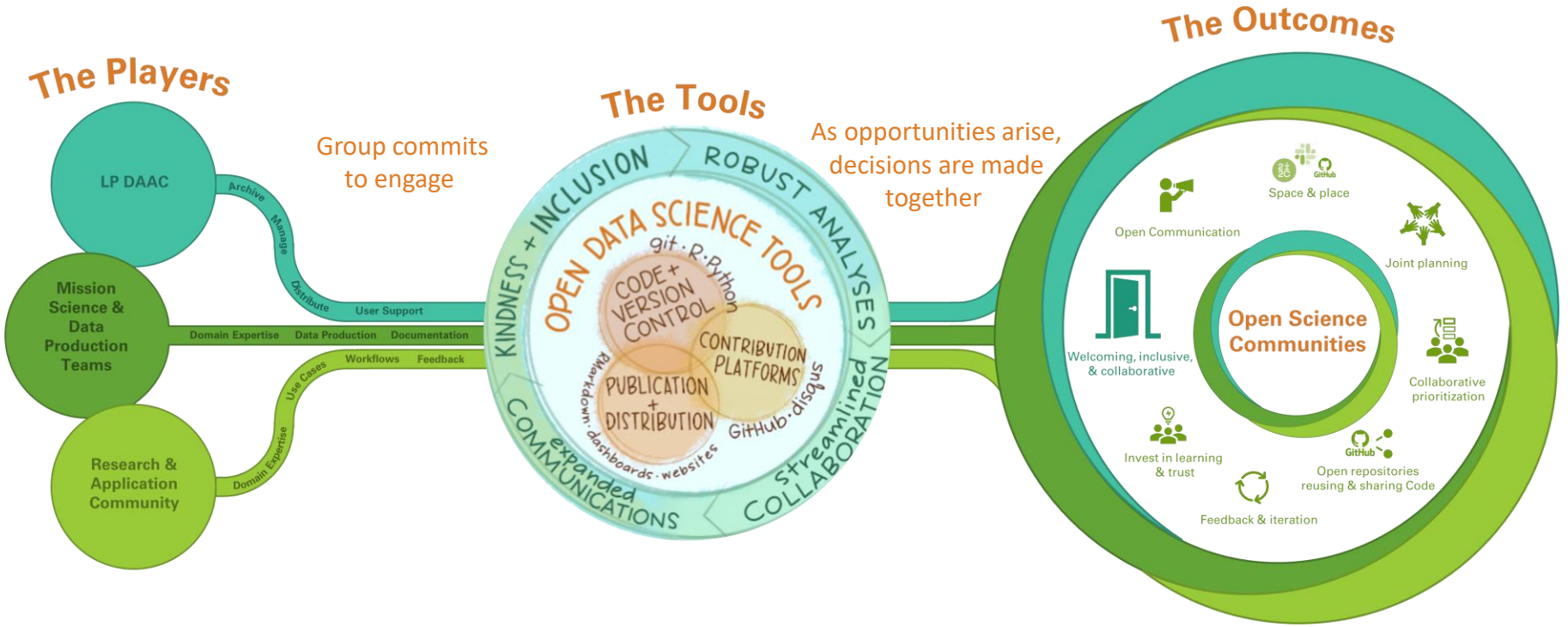
The Openscapes Approach

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.

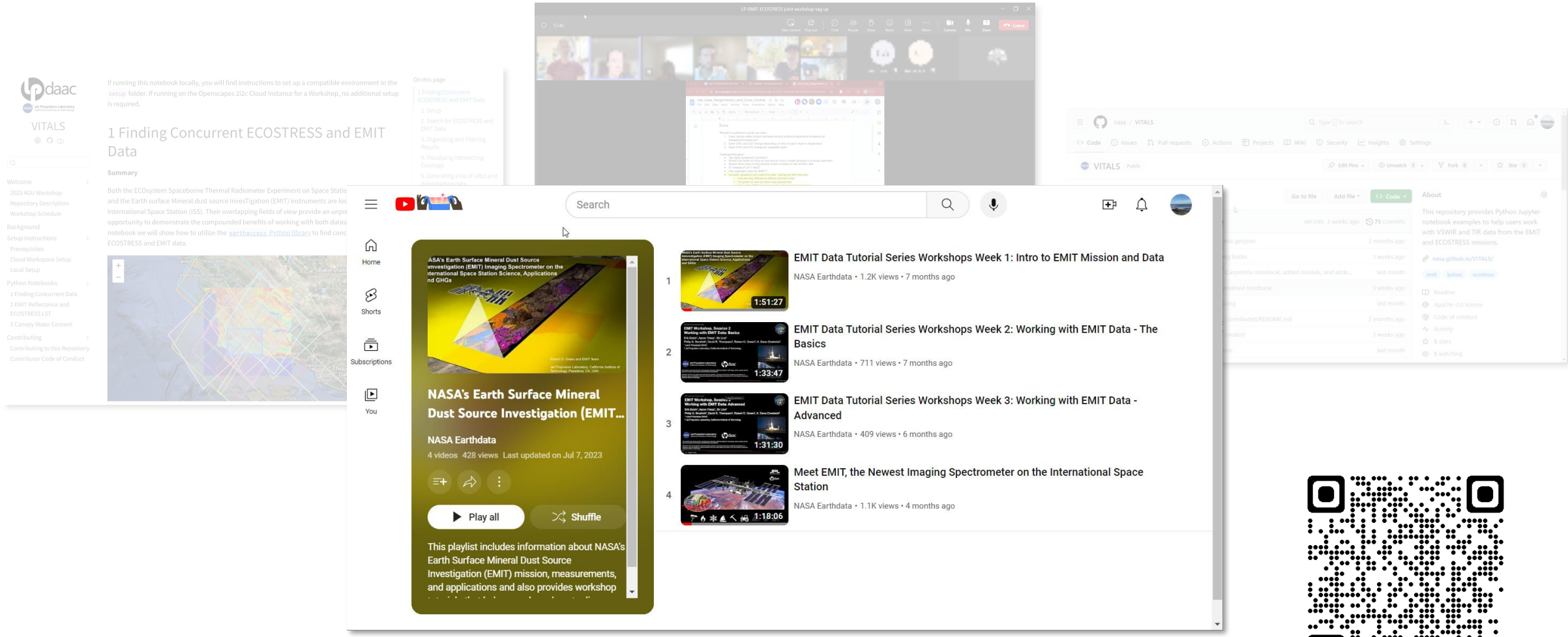
Better for everyone
The accumulation of individual and group successes facilitates a large- scale change in organizational and individual mindsets which has beneficial effects for everyone involved. Positive environments facilitate the production of timely reproducible resources as well as attracts more 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.



U.S. Geological Survey NASA

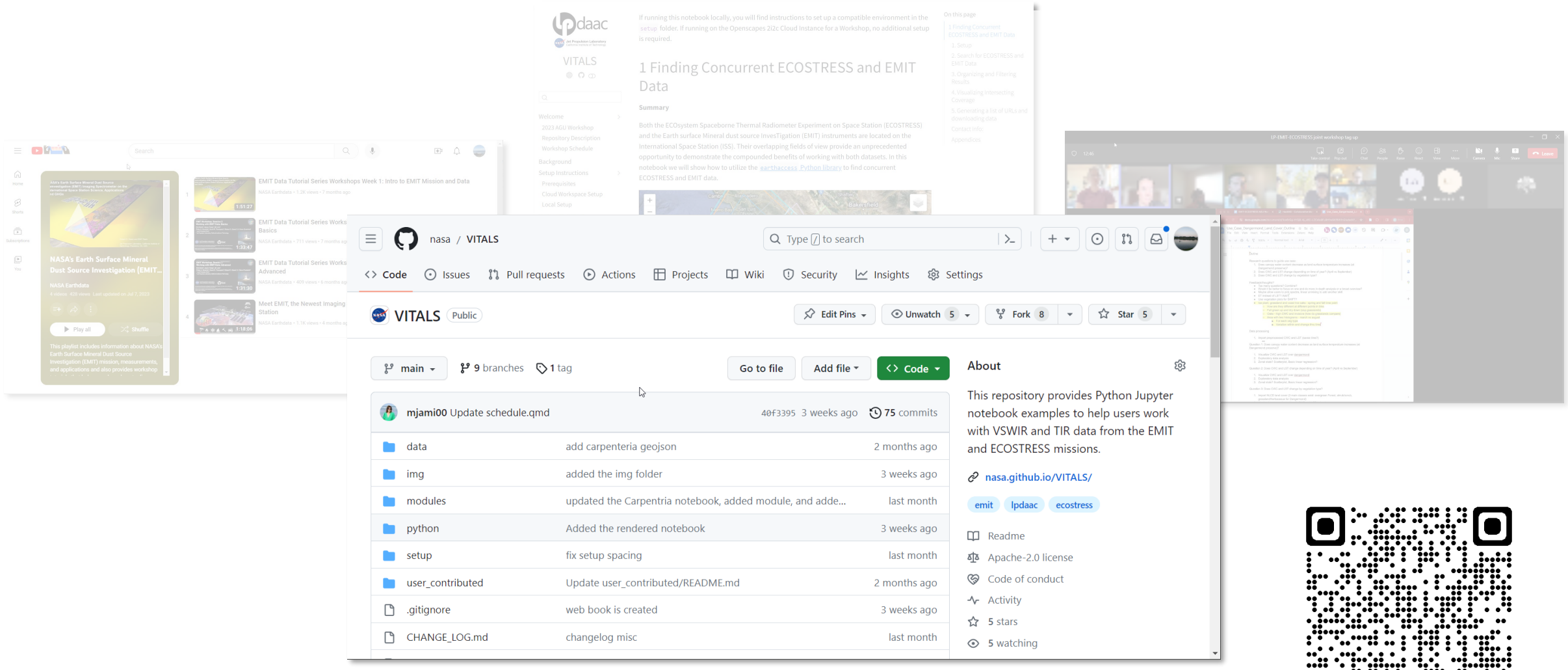
Successes



EMIT Webinar Series

U.S. Geological Survey

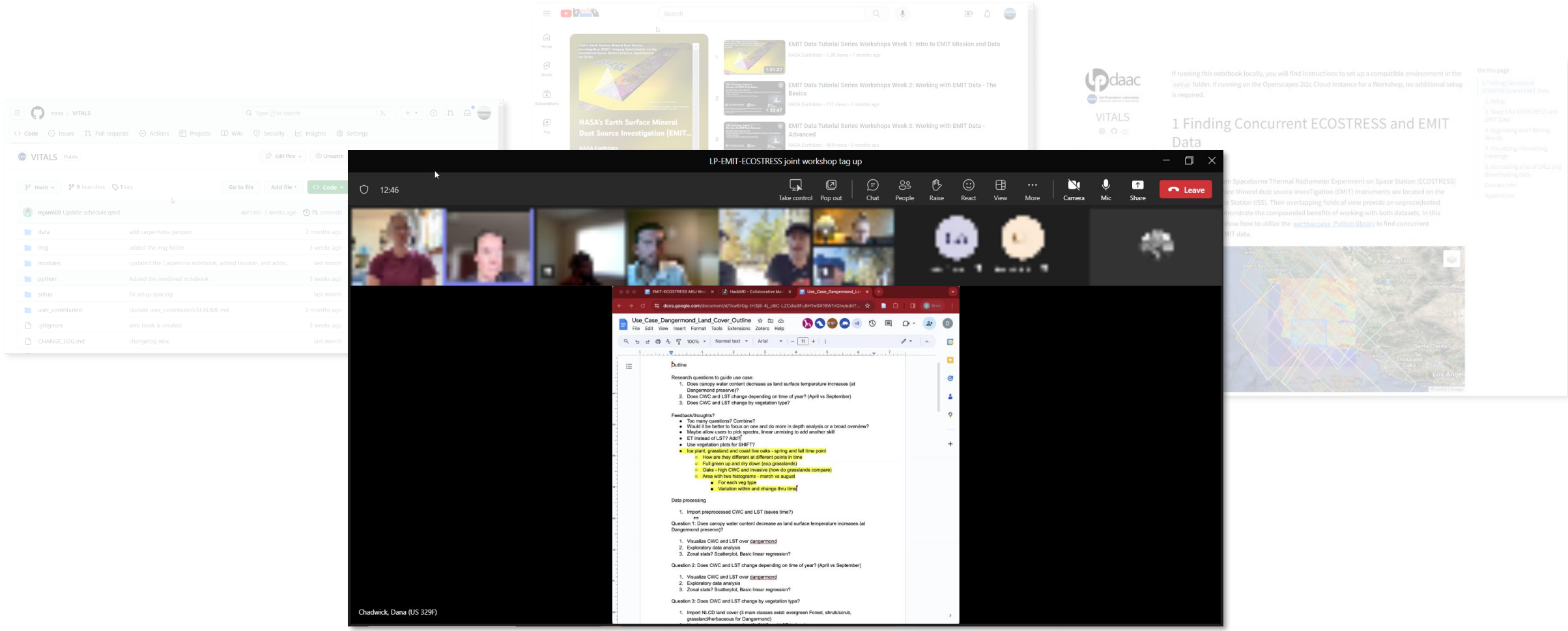
Successes



VSWIR Imaging and Thermal Applications, Learning, and Science Repository

U.S. Geological Survey NASA

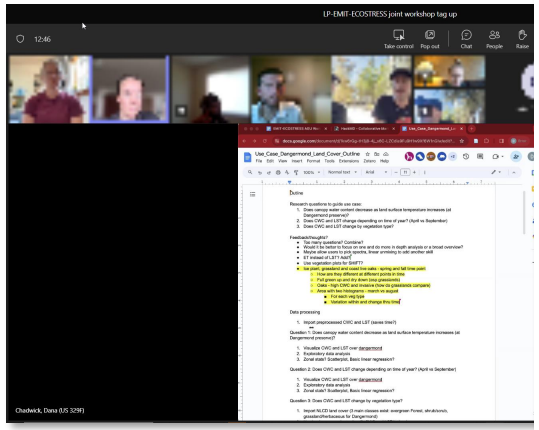
Successes

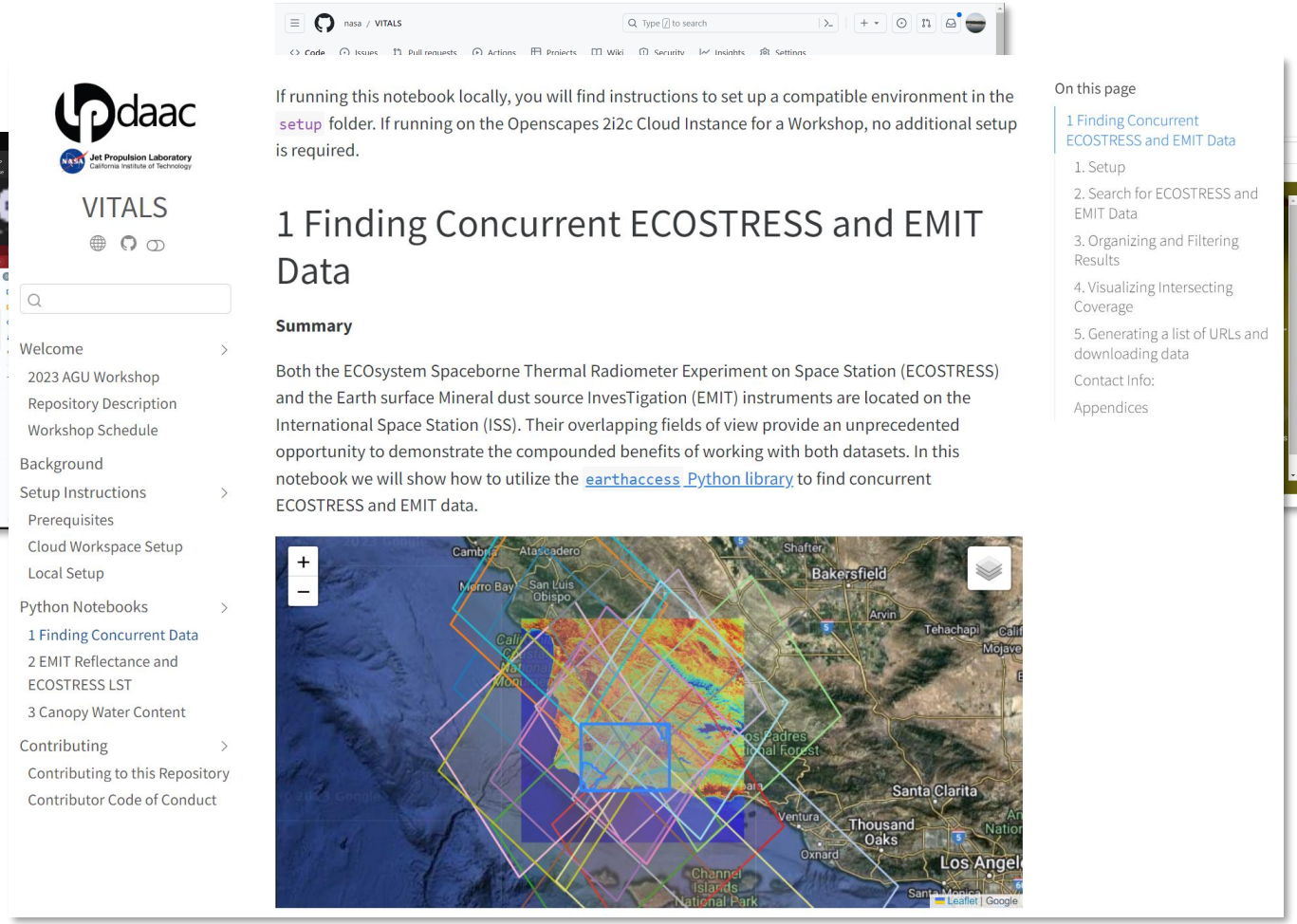


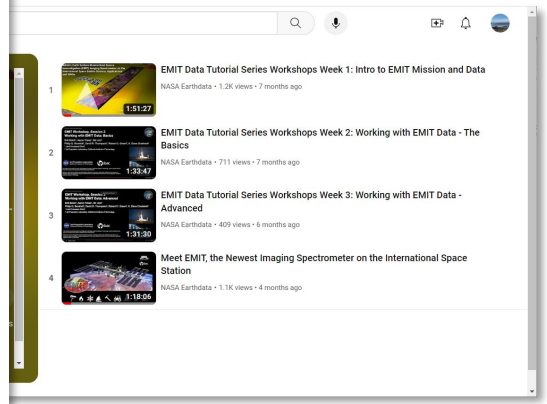
Co-working / co-developing for the AGU Workshop

U.S. Geological Survey

Successes







VITALS Webbook for the AGU Workshop

U.S. Geological Survey

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

Better Science for Future Us

Better Science = More open, reproducible, efficient, diverse, equitable, inclusive, and kind

Future Us = Ourselves, teams, and communities…now, next week, and into the future

Thank You

Openscapes artwork by Allison Horst