



## **Enabling a new era of science on the cloud (NASA ROSES 80NSSC21K0564 Award # 20-TWSC20-2-0003)**

### **Annual Progress Report: Year 2 of 3; PI Lowndes, co-lead Robinson**

The following is the second annual progress report for Openscapes. This report covers the period from February 1, 2022 through December 7, 2022 (today's date). This represents most of Year 2 of this grant. UPDATED through February 24, 2023 following submission to NSSC-Grant-Report.

This project supports scientists using data from NASA Earthdata served from the Distributed Active Archive Centers (DAACs) as they migrate workflows to the cloud. Priorities throughout are promoting Open Science through skill development and role-modeling, and diversity, equity, and inclusion as part of daily work.

## **Table of Contents**

<b>Year 2 Summary</b>	<b>2</b>
<b>Maintain and grow DAAC Mentors program</b>	<b>2</b>
ESDSWG engagement	2
Nominations, selection, onboarding	2
Participating mentors	3
Maintain GitHub Organization: NASA-Openscapes	3
Maintain NASA-Openscapes website	4
Maintain JupyterHub 2i2c Cloud infrastructure	4
Opportunistic in-person meetups (community building)	4
Open science & DEI summary	5
<b>Support DAAC Mentors with emerging needs</b>	<b>5</b>
Openscapes Mentorship: Cohort Calls, Clinics and Co-working	5
Maintain & update Earthdata Cloud Cookbook	5
Support DAAC Internal trainings	6
Support External trainings	6
Support earthaccess python library	6
Support Cheatsheets and Guides	7
Support White Paper: JupyterHub as NASA Core Services	8
Support Mentors' presentations and posters	9
Carpentries Instructor Training	10
Open science & DEI summary	10
<b>Lead the Champions Program</b>	<b>10</b>
<b>Community Engagement within NASA and Beyond</b>	<b>11</b>
Talks & Community Calls	11
Preprint: The Openscapes Flywheel	12
Open science & DEI summary	12
<b>What's next</b>	<b>12</b>
<b>Appendix: Testimonials</b>	<b>12</b>
	<b>1</b>



## Project Impact Summary (Years 1+2)

NASA Openscapes is a 3-year effort to grow a cross-DAAC Mentor community supporting Open NASA Earth Science in the Cloud. We approach this work as movement building, and developed the Openscapes Flywheel with NASA Earth science data centers, using the concept where transformations occur from relentlessly pushing a giant, heavy flywheel that builds momentum over time. The Flywheel supports teams across NASA DAACs to grow morale and technical capacity across their organizations by (1) Engage bright spots, through welcoming them and creating space and place; (2) Empower a learning culture through investing in learning and trust and working openly (3) Amplify Open science leaders, through leveraging the common and inspiring the bigger movement ([Robinson & Lowndes 2022](#)). Through this work, some highlights of impact to date: (1 Engage): 7 DAACs participating (NSIDC, PO.DAAC, LP DAAC, GES DISC, ASDC, ASF, ORNL); have a JupyterHub and Notebook-Quarto-GitHub workflow for documentation and publishing; have co-created a consistent set of tutorials and teaching style; co-led the 2021 Cloud Hackathon and 2022 Champions program; have documented our work through the Flywheel pub and Approach Guide; and given many invited talks & keynotes. (2 Empower): Mentors have led 10 Workshops: internal with DAAC staff and external with researchers; developed the [Earthdata Cloud Cookbook](#); Reused tutorials, slides, graphics and facilitation and open practices; were more aware cross-DAAC, less recreating; from user feedback developed [Cheatsheets](#) and the [earthaccess](#) python library; wrote the [Value of Hosted JupyterHubs \(White paper RFI\)](#); Collaborating on Hackweeks, 2i2c cont'd access after workshops. (3 Amplify): Mentors are amplifying across-DAACs and beyond: Career advancement & bringing mindset to new places; Speaking up in other meetings (User Needs TIM, TRAIN, Cloud Playground); Connecting & consulting based on experiences - Pathfinder for 2i2c, comparing w/ SMCE; AWS; Engaging beyond (Pangeo Forge, Ladies of Landsat, pyOpenSci).

From one Mentor, Cassandra Nickles (PO.DAAC):

Openscapes has created a collaborative environment for DAAC staff to collectively support open science initiatives for NASA Earthdata users. It enables us to work more openly with other DAACs toward our common goal of making the Earthdata ecosystem more accessible and inclusive. We've developed awesome material to help Earthdata users such as [workflow cheatsheets](#), a python package ([earthaccess](#)), and data recipes hosted in the cross-DAAC [NASA Earthdata Cloud Cookbook](#). Perhaps just as important as what we've done however, are mindsets we've grown into along the way. It's okay to share imperfect works in progress. The virtual environment can be conducive to laughter and connection. Ideas are not too big or too small to share. We are better at dreaming and implementing the future together.

## Year 2 Summary

The overarching vision of our project is to support scientific researcher teams using NASA EOSDIS data as they migrate their workflows to the cloud. We are doing this working with NASA Distributed Active Archive Centers (DAACs) over three years by:

1. Developing a cross-DAAC Mentor community



2. Empowering science teams through the Champions program
3. Scaling the Champions program with DAAC Mentors

In Year 2 we focused on supporting DAAC Mentors with their existing needs and shifting our planned agenda. That showed up specifically in delaying a 2022 Cloud Hackathon – we had expected this in Fall 2022 but it did not align with existing DAAC priorities. DAAC Mentors were more-than-busy with leading workshops, giving invited presentations, and developing new software and conceptual support from what they learned working directly with users. Several things that emerged as priorities this year that we continue to support:

- Internal DAAC trainings - Openscapes Mentors teaching their NASA colleagues about Cloud in hands-on workshops in our Openscapes 2i2c JupyterHub (poster at AGU Fall Conf)
- earthaccess python library - a Python package to search and access NASA datasets; Fall was a big development push (poster at AGU Fall Conf)
- Cheatsheets - friendly visual workflows and guides to introduce topics, delineate common cloud terminology, and describe general data workflows, access, tools and services. (poster at AGU Fall Conf)
- White paper - “The Value of Hosted Jupyter Hubs in enabling Open, NASA Earth Science in the Cloud” co-written by the Mentors in Fall 2023
- Flywheel paper - “The Openscapes Flywheel: A framework for managers to facilitate and scale inclusive Open science practices” by Robinson & Lowndes (talk at AGU Fall Conf)

## **Maintain and grow DAAC Mentors program**

We are supporting and strengthening the community of DAAC folks that are already creating cloud learning resources, having formalized this as a community in Year 1. We designed the program with active engagement with NASA communities and using open source software for all communication.

### **ESDSWG engagement**

As in Year 1, in Year 2 we also gave a short talk in the opening session of EDSWG about the momentum and accomplishments of NASA Openscapes. Working with Justin Rice, we thanked DAAC Managers for supporting the current Mentors so far and asked to nominate members of their team to give paid time to join the NASA Openscapes DAAC Mentors program.

### **Nominations, selection, onboarding**

Our Year 1 Cohort included Mentors from 5 DAACs: PO.DAAC, NSIDC, ASDC, LP DAAC, GES DISC, as well as IMPACT and JPL. In Year 2, ORNL and ASF pairs and trios joined the Cohort, including additional Mentors from LP DAAC, PO.DAAC, and GES DISC.

We held 1:1 intro calls to meet the nominees and describe more about the program. We accepted all teams, and onboarded the new DAAC Mentors in late Spring 2021, before breaking for the Summer. Cohort Calls, as described more below, created space for Mentors across DAACs to start building relationships and seeing the common parts of their work.



## Participating mentors

We are grateful to the 2021 Mentors from five DAACs continuing on this year, noted with an \* below, and we are excited to see the DAAC mentor community expand across the initial five DAACs to seven DAACs with ORNL and ASF Mentors joining us in Year 2. List as of December 2023

- Makhan Viridi, ASDC \*
- Sargent Shriver, ASF
- Alexander Lewandowski, ASF
- Brandon Bottomley, ASF
- Alexis Hunzinger, GES DISC \*
- Chris Battisto, GES DISC
- Brianna Pagan, GES-DISC
- Aaron Friesz, LP DAAC \*
- Mahsa Jami, LP DAAC
- Brianna Lind, LP DAAC
- Amy Steiker, NSIDC \*
- Luis Lopez, NSIDC \*
- Andy Barrett, NSIDC \*
- Catalina Oaida, PO.DAAC \*
- Jack McNelis, PO.DAAC \*
- Cassandra Nickles, PO.DAAC
- Celia Ou, PO.DAAC
- Jess Welch, ORNL DAAC
- Michele Thornton, ORNL DAAC
- Muthukumaran Ramasubramanian, IMPACT \*
- Shubhankar Gahlot, IMPACT \*
- Iksha Gurung, IMPACT \*

## Maintain GitHub Organization: NASA-Openscapes

We maintain the GitHub Organization where Mentors continue to collaborate. It is a shared neutral space (i.e. not within any specific DAAC) and acts as a sandbox for us all to try things out, test things, and develop together. As described more below, we held GitHub clinics, co-working sessions, and 1:1 screensharing help as needed to help everyone learn and share together so that all Mentors became confident GitHub users and confident collaborating and sharing imperfect work with each other.

GitHub Org: <https://github.com/nasa-openscapes>

## Maintain NASA-Openscapes website

We updated our website - a place to share updates about our project. First built with R (distill) and GitHub, this is also part of role modeling open science and how to communicate about open projects. In 2022 we migrated the site to Quarto, which works natively for Python as well as R users, so that all Mentors are able to



update and contribute. This website includes a list of presentations by the Mentors and makes the Earthdata Cloud Cookbook more prominent from the menu.

Website: <https://nasa-openscapes.github.io/>

Presentations: <https://nasa-openscapes.github.io/about.html#slides>

Blog posts: a selection from 2023:

- The why, what, and how of our NASA Openscapes cloud infrastructure: 2i2c JupyterHub and corn environment
  - <https://www.openscapes.org/blog/2022/11/17/nasa-earthdata-cloud-infrastructure/>
- Quarto Keynote at the RStudio Conference
  - <https://www.openscapes.org/blog/2022/08/10/quarto-keynote/>
- Hello Quarto! A Chat with NASA Openscapes, co-hosted with R-Ladies Santa Barbara
  - <https://www.openscapes.org/blog/2022/06/22/hello-quarto/>
- From downloading data to Cloud access: NASA Openscapes Champions Wrap-up
  - <https://www.openscapes.org/blog/2022/05/12/nasa-2022-champions/>

## Maintain JupyterHub 2i2c Cloud infrastructure

We maintain the Openscapes JupyterHub created in 2021 partnering with [2i2c](#) that is available to DAAC Mentors, Hackathon and AGU Workshop participants. Mentors have said this has been critical to support researchers migration to the cloud - without this hands-on experience it is very difficult to advise. To extend the utility of this “sandbox”, we’ve made it available to other DAAC staff. As of December 2023 there are 567 users in our JupyterHub.

JupyterHub: <https://openscapes.2i2c.cloud/hub/>

## Opportunistic in-person meetups (community building)

Co-leads Robinson and Lowndes have opportunistically met with folks involved in NASA Openscapes as travel has opened up. These were the first visits in person since the pandemic, which has been ongoing since this project started. These were a chance to celebrate together and strengthen relationships within the project.

- Goddard, Summer 2022 - Justin Rice & Julie Lowndes
- Pasadena, California 2022 - Catalina Oaida Tagliatela & Luis Lopez
- Colorado, Summer 2022 - Erin Robinson & Julie Lowndes
- California, Fall 2022 - Julie Lowndes & Amy Steiker
- Colorado, Fall 2022 Amy Steiker, Aaron Friesz, Michele Thorton, Catalina Oaida Tagliatela, & Erin Robinson

## Open science & DEI summary

The design of the NASA Openscapes Framework builds on Openscapes’ investment in increasing diversity, equity, and inclusion in all activities and facilitation style to create spaces that are psychologically safe for folks to learn together and share imperfect work. For new examples in 2022, see:



- [A Journey to Data Science: Tools for Equity and Diversity in STEM](#) - July 28, 2022: Talk by Ileana Fenwick at RStudio Conference ([video](#)) (20 minutes)
- [Hello Quarto: share • collaborate • teach • reimagine](#) — July 28, 2022: Keynote by Julia Lowndes at RStudio Conference ([video](#)) (60 minutes)

There is diversity in the Mentors cohort across gender, race, job title, technical level.

## Support DAAC Mentors with emerging needs

All activities begin with a summary of our Code of Conduct: [openscapes.org/code-of-conduct/](https://openscapes.org/code-of-conduct/).

### Openscapes Mentorship: Cohort Calls, Clinics and Co-working

Supporting the DAAC Mentors means creating space (time) and place (through collaborative software) to build trust and find the common, and welcoming new Mentors to this community. We continue to hold twice-monthly regular meetings and coworking sessions on Zoom, have standing Google Docs and Slack for asynchronous communication and so folks stay connected and up-to-date whether or not they are able to attend meetings. All calls conclude with Efficiency and Inclusion Tips.

Through a GitHub Clinic and co-working sessions, we skillshared and learned how to use GitHub collaboratively with branches, pull requests, and are beginning to formalize using it for code review.

### Maintain & update Earthdata Cloud Cookbook

We continued developing and supporting the a tutorial book architecture using Quarto, which combines Jupyter Notebooks into beautiful online books that can be easily shared, navigated, and used by learners of all skill levels (i.e. it is a friendlier experience than notebooks in a GitHub repository). In 2023 we've revisited the organization and in Fall 2023 made a further push to rework the HowTos section from what Mentors learned in teaching, and added the Cheatsheets section (more detail below).

We partnered with RStudio as their first external testers using Quarto. Quarto builds from what RStudio learned from RMarkdown. It makes collaborating to create technical documentation streamlined because we work in plain text documents that can have executable code (Python, R) and are rendered using Jupyter and Knitr engines.

Earthdata Cloud Cookbook: <https://nasa-openscapes.github.io/earthdata-cloud-cookbook>

We also moved the tutorials page from the main NASA-Openscapes website to the Cookbook itself - all available for reuse and remix. As NASA DAACs transition data to the cloud, tooling and workflows will evolve. We organize tutorials in books, with the idea that each teaching event has its own book that is developed for a concrete audience and is a snapshot in time, and that the Earthdata Cloud Cookbook will be a collection of the most up-to-date tutorials.



## Support DAAC Internal trainings

As Mentors led DAAC internal trainings, for example at GES DISC, they reused materials and workflows co-developed by the Mentors and iterated since the 2021 Cloud Hackathon. Support involved creating and managing the Quarto book, helping establish norms around scope and style (including live-coding) coordinating tutorial reviews and reviewing tutorials, dry runs for each teacher to practice live-coding and get feedback.

## Support External trainings

External trainings for NASA Earthdata users - at least 5 events, including ECOSTRESS, SWOT, and User Working Groups, reused materials and workflows developed from the Hackathon. We develop tutorials for teaching events that each has its own e-book, using Quarto, and linked in the Cookbook. Tutorials are developed to teach open science and Cloud workflows for specific audiences. They are a snapshot in time as workflows with NASA Earthdata Cloud emerge and evolve. Some examples:

### 2023 EMIT Workshop

<https://github.com/nasa/EMIT-Data-Resources>

This workshop was designed to be completed locally, however it was offered to use the Openscapes 2i2c JupyterHub cloud workspace to revisit the EMIT Data Tutorials Workshops hosted February 3, 10, and 17th 2023. The Feb 17th event had 91 active users in large python instances in the 2i2c JupyterHub; Aaron Friesz, Bri Lind (LP DAAC) and Luis Lopez (NSIDC) worked ahead of time with 2i2c engineers to accommodate the increased CPU needs for this workshop.

### 2022 ECOSTRESS Cloud Workshop

<https://nasa-openscapes.github.io/2022-ECOSTRESS-Cloud-Workshop/>

The goal of the workshop is expose ECOSTRESS data users to ECOSTRESS version 2 (v2) data products in the cloud. Learning objectives focus on how to find and access ECOSTRESS v2 data from Earthdata Cloud either by downloading or accessing the data on the cloud. The LP DAAC is the NASA archive for ECOSTRESS data products. ECOSTRESS v2 data products will be hosted in the NASA Earthdata Cloud, hosted in AWS.

### 2022 SWOT Ocean Cloud Workshop

<https://podaac.github.io/2022-SWOT-Ocean-Cloud-Workshop/>

The goal of the workshop is to get ready for Surface Water and Ocean Topography ([SWOT](#)) and enable the (oceanography) science team to be ready for processing and handling the large volumes of SWOT SSH data in the cloud. Learning objectives focus on how to access the [simulated SWOT L2 SSH data](#) from Earthdata Cloud



either by downloading or accessing the data on the cloud. PO.DAAC is the NASA archive for the SWOT mission, and once launched will be making data available via the NASA Earthdata Cloud, hosted in AWS.

## Support earthaccess python library

Luis Lopez (NSIDC) has been developing the earthaccess python library (renamed from earthdata) - a software response to pain points identified by users and Mentors through all of this work. Art is part of our original budget; we worked with artist Allison Horst to design a logo for earthaccess, borrowing the practices from the R community of creating logos for code packages and making them a hex shape so they can be displayed with others.



### 2022 AGU Abstract:

Title: earthdata: a Python package to search and access NASA datasets.

Authors: Luis López, Andy Barret, Amy Steiker, Jessica Scheick, Julie Lowndes, Erin Robinson, The NASA Open Scapes mentors team.

The NASA Earth Observing System Data and Information System (EOSDIS) contains almost 9,000 datasets spanning a wide range of science domains and data processing levels. Across this data archive, many data search and access options are available suiting a variety of needs and use cases. We can use NASA Earthdata Search, we can use tools supported by individual NASA Distributed Active Archive Centers (DAACs), and we could even use data.gov! All of these portals are great, but they are not designed for programmatic access and reproducible workflows. Reproducible workflows are extremely important in the age of cloud data access, cloud computing, and open science. In this context, we are developing earthaccess, a python library that aims to simplify data discovery and access for those using the PyData ecosystem (xarray, dask, numpy). Using this library eliminates the need to know the intricacies of NASA's Application Programming Interfaces (APIs) and cloud data storage systems.

earthaccess is a result of a cross-DAAC collaboration through the NASA Openscapes program, which aims to build a cross-DAAC community to help scientists working with NASA Earth science data migrate their workflows to the cloud. earthaccess aims to simplify data discovery and access to address some of the pain points of programmatic access. It provides a higher abstraction for NASA's Common Metadata Repository (CMR) search API so that searching for the data can be done using a simpler notation instead of low level HTTP queries. With earthaccess, scientists do not need to concern themselves with where data is being distributed from (i.e. cloud-hosted or on-premises at a data center) because the library handles both workflows transparently and consistently. earthaccess provides authenticated sessions that can be used with xarray and other PyData libraries to access NASA EOSDIS datasets directly. earthaccess allows scientists to get to their science in an easier, simpler, and faster way, reducing barriers to cloud-based data analysis.





## Support Cheatsheets and Guides

We developed a workflow where each cheatsheet is its own Google Slide that can be embedded in multiple places, for example the Earthdata Cloud Cookbook and PO.DAAC's website. This way, any updates to the Cheatsheet (slide) will propagate to all these places with no extra maintenance. The Google Slides are organized internally and linked from the Cheatsheets page in the Cookbook.

### 2022 AGU abstract:

Title: Curating information to guide NASA Earthdata users into the cloud

Authors: Cassandra Nickles<sup>1</sup>, Catalina Oaida Taglialatela<sup>1</sup>, NASA Openscapes Team

<sup>1</sup>Jet Propulsion Laboratory, California Institute of Technology

### Abstract:

Having migrated all its datasets to the cloud this past year, NASA's Physical Oceanography Distributed Active Archive Center ([PO.DAAC](#)) has recognized data accessibility is not synonymous with data availability. Accessing and analyzing data in the cloud can be an intimidating task, especially for new users or those comfortable with pre-existing local (e.g. local computer) workflows. When new at something, it can be difficult to even know where to begin. We at PO.DAAC, together with our other DAAC colleagues through NASA Openscapes, have been asking ourselves how to curate our data and information so that end users from a variety of backgrounds have guidance on specific cloud access pathways. Here, we present a series of diagrams we have developed that outline high level introduction topics, delineate common cloud terminology, and describe general data workflows, access, tools and services. We have also created a NASA Earthdata Cloud data Workflow Cheatsheet as a one-stop-shop linking to data recipe tutorials and informational websites for users who are starting to take conceptual pieces and explore their own workflows. These diagrams have been implemented in the [NASA Earthdata Cloud Cookbook](#) developed by the [NASA Openscapes](#) team as a resource across all NASA DAACs. We encourage feedback and discussion on how best to improve such data user resources to help us all as we enter into the cloud data era!

## Support White Paper: JupyterHub as NASA Core Services

Erin Robinson spearheaded a White Paper, following conversations with Aaron Friesz (LP DAAC) with Katie Baynes. The title is The Value of Hosted Jupyter Hubs in enabling Open, NASA Earth Science in the Cloud.

### Draft executive summary:

There is a paradigm shift from traditional geospatial workflows that are changing the typical access and analysis patterns: discover, access, subset, download, and transform data from various collections to analyze on local machines. Cloud-based data collections and workflows allow traditionally disparate data to reside in common locations which are available through similar access and reduction patterns and for which a researcher can bring their analysis to the data. We see substantial growth and popularity in this model in, for example, Google Earth Engine (GEE) and Microsoft Planetary Computer, emphasizing the common locations of data collections and analysis platforms. NASA Earthdata holds authoritative collections of growing amounts of data collections which are estimated to grow to 250 petabytes by 2025. NASA, along with many other federal agencies, is moving the data to the cloud to allow a coupling of for analysis and to be done alongside



the data in the cloud. This change is transformative, but it comes with costs. Scientists new to the cloud – which is most scientists – spend more time on the technicalities of the cloud than on their science. Shared cloud-hosted computational environments such as the 2i2c Openscapes Jupyter Hub and the OpenSARLab Jupyter Hub allow multiple users to access cloud computing resources using tools and interfaces that are familiar to many researchers. These hubs have been invaluable for NASA Openscapes, Mentors, many other staff members across the DAACs, and researchers accessing NASA data in the cloud. Hubs are centrally administered, removing the need for scientists to be cloud-experts, reducing the barrier to entry, and allowing hundreds of users to access and analyze NASA data in the cloud for the first time. In addition to supporting end users, these hubs provide a space for NASA staff to learn and develop resources to support researchers across the DAACs. This white paper describes our experiences working with these hubs: the current hub landscape, the types of users we have supported, the benefits we have seen, and articulates the need for this kind of infrastructure on a permanent basis.

## Support Mentors' presentations and posters

In Year 2, Mentors were increasingly invited to give presentations to share about their work and lessons learned. As Cloud migration has been moving fast and DAAC Mentors and user support has been responsive, presentations have proved to be a powerful way to communicate not only to the target audience, but to reuse and amplify.

We have supported DAAC Mentors in shaping, creating, and reusing content between talks, as well as adding them to the NASA Openscapes website and Cookbook for broader reach.

We're currently preparing for a poster session at 2022 AGU: "IN22C - Environmental Data User Support with Cloud-Based User Services II Poster".

The following talks are included as a full list at <https://nasa-openscapes.github.io/about.html#slides>.

- [Working with NASA Earthdata in the Cloud](#) - January 24, 2023 - Amy Steiker, Bri Lind, Julie Lowndes, Luis López, Michele Thornton, and the NASA Openscapes Mentors. [ESIP Winter Meeting](#) "Enabling Open Science with NASA's Earthdata in the Cloud" Session
- [earthaccess: A Python Library for NASA Earthdata](#) - December 13, 2022 - Luis López, Andrew P. Barrett, Julie Lowndes, Joseph H. Kennedy, Erin Robinson, Amy Steiker, Jessica Scheick, The NASA Openscapes mentors team. AGU Fall Conference Poster ([slide](#))
- [NASA Openscapes: Lessons Learned supporting Cross-DAAC User Services to migrate to the Cloud](#) - December 13, 2022, Aaron Friesz, Alexis Hunzinger, Amy Steiker, Catalina Oaida Taglialatela, Luis López, Cassandra Nickles, Bri Lind, Mahsa Jami, Celia Ou, Julia Stewart Lowndes, Erin Robinson, NASA Openscapes DAAC Mentor Community. AGU Fall Conference Poster ([slide](#))
- [Curating information to guide NASA Earthdata users into the cloud with workflow diagrams and cheatsheets](#) - December 13, 2022 - Cassandra Nickles, Catalina Oaida Taglialatela, Julie Stewart Lowndes, Amy Steiker, Alexis Hunzinger, Aaron Friesz. AGU Fall Conference Poster ([slide](#))
- [Early lessons learned from supporting end user's transition to the cloud](#) - November 16, 2022, Alexis Hunzinger, LAADS DAAC User Working Group 2022
- [NASA Openscapes Cloud Infrastructure](#) - October 13, 2022, Luis Lopez ([video](#))



- [Efforts to support end users in the journey to the cloud](#) - September 27, 2022: [Open Source Science Data Repositories Workshop 2022](#) Steiker, Hunzinger, López, Oaida Taglialatela, Friesz ([video](#))
- [NASA Earthdata Cloud: Myths, Truths, Questions](#) - September 7, 2022. Steiker, Heightley (NSIDC)
- [Efforts to support end users in the journey to the cloud](#) - August 17, 2022: [ESDIS SE TIM 2022](#) Oaida Taglialatela, Hunzinger, Smit, López, Steiker
- [NASA Briefing to Unidata](#) - July 2022, Christine Smit
- [NSIDC DAAC User Working Group](#) - May 20, 2022. Barrett, Steiker, Meier, Roebuck, Beig, Lopez, (NSIDC)
- [NASA Earthdata Cloud & The Cloud Paradigm](#) - April 2022. Friesz (LP DAAC)
- [NASA Openscapes Lessons from Year 1](#) - April 21, 2022: ESDSWG Robinson, Lowndes, López
- [NASA Openscapes Cloud Infrastructure](#) - February 9, 2021: NSIDC by Luis López

**A third 2022 AGU abstract**, in addition to earthdata and cheatsheets abstracts above:

Title: NASA Openscapes: Lessons Learned supporting Cross-DAAC User Services to migrate to the Cloud  
Authors: Aaron Friesz, Alexis Hunzinger, Amy Steiker, Catalina Oaida Taglialatela, Luis López, Cassandra Nickles, Bri Lind, Mahsa Jami, Celia Ou, Julia Stewart Lowndes, Erin Robinson, NASA Openscapes DAAC Mentor Community; Artwork by Allison Horst.

The NASA Openscapes project answers a call from NASA Earthdata to support both the NASA Distributed Active Archive Centers (DAACs) User Services staff and the initial, early adopter scientific researcher teams as they transition their workflows from to the cloud. In order to support this change, the project has three goals: (1) Engage a cross-DAAC Mentor community of collaborative cloud data instructors, that co-create, curate, and use shared resources (“make once, use often”) (2) Empower science teams through the Champions program to migrate their download- intensive data analysis workflows to the cloud and open, kinder science and (3) Scale Open Science leaders - to support more teams transforming their workflows towards open, kinder science and the cloud.

We will highlight the cross-DAAC mentor community NASA Earthdata Cloud Cookbook. The Cookbook content was initially developed for the November 2021 NASA Earthdata Cloud Hackathon hosted by PO.DAAC in collaboration with the NSIDC DAAC and the LP DAAC and supported by the NASA Openscapes Project, ASDC DAAC and GES DISC. From that initial work, the material has been refined and reused across DAACs highlighting the common elements to accessing NASA Earthdata in the cloud. We then used this material in early 2022 with the NASA Openscapes Champions cohort, ten research teams that are early adopter research teams interested in migrating their data access and analysis workflows. Through all of this work with early adopters, we collectively addressed many of the initial challenges to accessing NASA Earthdata on the cloud. We will share those lessons learned. Moving forward we are expanding the cross-DAAC mentor community to new DAACs and continuing to create and share materials. Our hope is that through this work, NASA DAACs will be well-positioned to support more research teams as they migrate to the cloud and help enable better science through both capacity building and the power of the cloud.

## **Carpentries Instructor Training**

Our Year 2 Carpentries Platinum Membership has 15 seats available for Instructor training.



Mentors take Carpentries Instructor Training to learn teaching pedagogy and see and practice live coding when teaching. Live coding means typing live at a pace where your learners can follow along with you, and is a deeper learning experience than only seeing a demo scroll that learners do not experience themselves.

## **Open science & DEI summary**

All events (Mentor Cohort Calls and Hackathons) begin with a Code of Conduct. We work to normalize open, imperfect, reuse through building trust and psychological safety.

## **Lead the Champions Program**

The Openscapes Champions Program is our flagship program and we adapted it to support research teams interested in migrating workflows to the Cloud, with support from the NASA Mentors.

We began recruitment at AGU in December and in February accepted 10 research teams (~36 researchers) to join our Mentors and plan migrating their workflows to the cloud. These research teams are interested in a wide variety of NASA Earthdata and various stages of cloud technology.

The ten research teams who participated were interested in a wide variety of NASA Earthdata and various stages of cloud technology familiarity. You can learn more about their research below. Together as a Champions cohort they discussed what worked and didn't work as they migrated workflows to the cloud, with a focus on collaboration and open science. We met as a cohort five times over two months, on alternating Fridays. Each cohort call included a welcome and code of conduct reminder, two teaching sessions with time for reflection in small groups or silent journaling and group discussion, before closing with suggestions for future team meeting topics ("Seaside Chats"), Efficiency Tips, and Inclusion Tips. Additional hands-on clinics and coworking sessions were scheduled within this period and will extend for the next two months to support these teams as they continue to work on the cloud workflow migration. In addition, the teams were supported by the Openscapes DAAC mentors and staff and Element84 and had access to Openscapes' 2i2c Jupyter Hub, which will continue for the next year.

Blog post describing more details, and information about the 10 teams:

<https://www.openscapes.org/blog/2022/05/12/nasa-2022-champions/>

A recent 3-month check re-connected Champions and Mentors and revamped support for R and Matlab users on the Cloud. Peter Cornillon estimates using the Cloud will take his processing time from 1.5 years to 1 month. Allan Just focused on documentation for their Cloud workflows, and has reused our materials to promote open science and reproducible research ([slides](#); [repo](#), made with quarto!). They are interested in what NASA datasets are on the cloud today (seems like a lot since we started cohort)

Announcement website: <https://nasa-openscapes.github.io/champions>

Cohort website: <https://nasa-openscapes.github.io/2022-nasa-champions>



## Community Engagement within NASA and Beyond

Lowndes & Robinson gave fewer talks to NASA audiences in Year 2 as compared to Year 1, as NASA Openscapes Mentors were increasingly invited to give talks about their work directly. The big exception was a keynote at the RStudio Conference in July 2022, where we shared about NASA Openscapes to a large global tech audience. Below are our direct activities for community engagement; see the previous section for the Mentor talks we supported.

Our community engagement approach focuses on inspiring and attracting folks to join the open science movement, using original artwork, storytelling, and practical tips.

Full list of slides shared from our project are:

- <https://nasa-openscapes.github.io/about.html#slides>
- <https://openscapes.org/media>

## Talks & Community Calls

**Openscapes Community Call - Hello Quarto** - Our 6th Openscapes Community Call co-hosted with [R-Ladies Santa Barbara](#) featured a “celebrity interview” with [NASA Openscapes Mentors](#) Amy Steiker, Catalina Oaida Tagliatela, Aaron Friesz, with J.J. Allaire, lead Quarto developer and CEO of [RStudio](#). The conversation was led by [Sam Csik](#) of R-Ladies Santa Barbara. <https://www.openscapes.org/blog/2022/06/22/hello-quarto/>

**Keynote at RStudio Conference - “Hello Quarto: share • collaborate • teach • reimagine”, Summer 2022** - We shared how we support NASA Earthdata DAACs as part of the global launch of Quarto, an open source publishing framework that we contributed to as first external users and has been a big part of “place” - having a place to collaborate across the DAACs. Blog post: <https://openscapes.org/blog/2022/08/10/quarto-keynote>

**AGU Fall Conference Plenary Talk, “The Future is Open” session, Fall 2022** - We were on an invited panel with NASA TOPS chaired by Dr. Chelle Gentemann to share our work on a big stage: [Supporting open science as a daily practice](#) - December 16, 2022 - Julie Lowndes, Erin Robinson, Openscapes Mentors.

**At the [ESIP Winter meeting](#) in January 2023** we coordinated a panel discussion “Better science for future us: Openscapes stories and approaches for the Year of Open Science”. Mentors from NASA, NOAA Fisheries, CalEPA, and Pathways to Open Science programs shared their stories of open science movement building through Openscapes. Blog post upcoming in March 2023.

## Preprint: The Openscapes Flywheel

We spent time this summer writing a manuscript called **The Openscapes Flywheel: a framework for managers to facilitate and scale inclusive Open science practice**. This is now a preprint in Eartharxiv, and in review at the Data Science Journal: <https://eartharxiv.org/repository/view/4560/>

This work builds from our invited talk at AGU 2021, where we and several DAAC Mentors were invited to give talks at the [Open Science in Action](#) Session, co-organized by Chelle Gentemann, focused on “Open project



design”. This is part of our work to make the implicit explicit, to make more visible the technical and people skills involved with building community around open science for true culture change.

## Open science & DEI summary

With a priority of connecting NASA Openscapes with the global Open Science movement, here are a few ways we centered DEI in our work:

- partnered with RLadies Santa Barbara, a chapter of a global movement to increase gender diversity in the R programming language, for a community call. This conversation also bridged R and Python communities.
- focus on inclusion in Keynote: emphasizing 1) making the implicit explicit; 2) onboarding learners as contributors; 3) psychological safety and growth mindset
- Flywheel publication making the implicit explicit around community building in open science
- developed artwork with Allison Horst, again bridging the R and Python communities by borrowing a community norm from R (creating logos in a hex shape) and prioritizing welcoming visuals

## What's next

2023 will focus on supporting the DAAC Mentor community as they continue leading DAAC internal and external trainings, developing the earthdata library, cheatsheets, the White Paper, and other emerging needs. As one experiment, we will offer DAAC Mentors to receive additional professional development through coaching training as a cohort with Openscapes. We will onboard new Mentors in Fall 2023.

In Winter 2023 we will offer professional development for Mentors: **“Mentoring with a Coach Approach”**.

This is to develop coaching skills to help us as open science mentors, while also building skills and habits to promote inclusion. This is also exciting because it’s a chance to expand our Openscapes Mentor community - our mentor-coach cohort will include 25 other Openscapes mentors from NOAA Fisheries, EPA, and academia.

In Spring 2023 we will lead a second NASA Champions Cohort with science teams, and are exploring combining this with a Hackathon or Hackweek.

## Appendix: Testimonials

**March 23, 2022 - Jinbo Wang**

We just had a coding club kick off meeting at JPL/PODAAC with about 10 oceanographers and hydrologists including @Severine Fournier. @Jack McNelis (Helper) participated too. I would like to send my huge thanks to the openscapes team for the continuous support of the jupyterhub to the podaac scientists. It is an important element that enables our activities especially for our foreign nationals. @Julie Lowndes @Erin Robinson



**Jinbo Wang** 1:10 PM

We just had a coding club kick off meeting at JPL/PODAAC with about 10 oceanographers and hydrologists including @Severine Fournier. @Jack McNelis (Helper) participated too. I would like to send my huge thanks to the openscapes team for the continuous support of the jupyterhub to the podaac scientists. It is an important element that enables our activities especially for our foreign nationals. @Julie Lowndes @Erin Robinson



### Winter 2022 - Catalina Oaida Tagliatela

From Catalina: +1 that within DAACs we (PODAAC) are also taking what we've learned and built last year through NASA Openscapes and applying it to upcoming workshop; both the tools we've discovered and used like quarto books, github workflows, using JupyterHub in aws but also the open science open source way of thinking, how to structure events, what brings most value to science etc;

I still get kudos from those that attended on xarray, hvplot, zarr and even the use of slack for getting help at the hackathon. A cloud hackathon participant shared what she learned and liked about the hackathon with another early adopter group she is in and the organizer is looking to implement or share those practices at their own workshops - yay for spreading open science

from Chelle Gentemann [NASA HQ] she/her  
(Ext) to Everyone: 1:56 PM

I love this & used it - so great to get rid of  
all the auth access code - a great example  
of making data more accessible! Great  
work Luis!!!!

from Julie Lowndes (Ext) to Everyone: 1:56  
PM

!! So great to hear Chelle!

from Luis Lopez (Ext) to Everyone: 1:56 PM

Thanks Chelle! you're one of the authors of  
the library too!

## Appendix: Screenshots





A view of the Cookbook, July 2021:

<https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/issues/44> . This is recently after we made the Cookbook and the left navigation shows the original architecture.

EarthData Cloud Cookbook

Search

Welcome

Getting Started

How to use this book

EarthData Login

Authentication

API Primer

NASA Cloud Data Discovery

CMR and CMR-STAC API

Earthdata Search UI

CMR Virtual Directories

NASA Cloud Data Access

Data access demo

Earthdata Search

OPeNDAP in the Cloud

Direct in-region

COF via Earthdata Harmony

Download to local

NASA Cloud Data Transformations

Harmony API

HarmonyPy Introduction

netCDF to Zarr

L2 Subsetter

Earthdata Search UI

OPeNDAP in the Cloud

End-to-end Examples

Python example with Quarto

RMarkdown example.Rmd

Notebook example

## EarthData Login

Modified From [PODAAC AGU 2020](#) and [LPDAAC HLS Tutorial](#)]

An Earthdata Login account is required to access data, as well as discover restricted data, from the NASA Earthdata system. Thus, to access NASA data, you need EarthData Login. Please visit <https://urs.earthdata.nasa.gov> to register and manage your Earthdata Login account. This account is free to create and only takes a moment to set up.

To avoid being prompted for credentials every time you run and also allow clients such as curl to log in, you can create and/or add the following to a “netrc” file (pronounced “Net RC”). On a Mac, this is a file called `.netrc`, and on Windows it is `_netrc`. (There are no extensions on either file).

Below are ways to create a “netrc” file in your home directory.

## R for Windows or macOS

Create a “netrc” file by running the following code in your R console (or after saving to an R script).

Show the code

## Common questions

### How do I know if I already have a netrc file?

Your netrc file will likely be in your root directory. It is a hidden file that you will not be able to see from your Finder (Mac) or Windows Explorer (Windows), so you'll have to do this from the Command Line. Navigate to your root directory and list all:

### On a Mac:

```
cd ~  
ls -la
```

If you see a `.netrc` file, view what's inside (perhaps with `nano`), and if you'd like to delete the current version to start afresh, type `rm .netrc`.

← How to use this book Authentication →

On this page

R for Windows or macOS

Common questions

Edit this page

Report an issue