

Earthdata Cloud Cookbook - NASA Openscapes

This is our planning and development; and we'll be increasingly using GitHub Issues and Projects

Repo: <https://github.com/NASA-Openscapes/earthdata-cloud-cookbook>

Book: <https://nasa-openscapes.github.io/earthdata-cloud-cookbook/>

[This doc](#) (public)

[Inventory spreadsheet](#) (started by Catalina Feb 3)

May 10 Hackday #4

Roll-call

Name / share something going on in May for you

- Julie / heading to Monterey to work and see friends
- Stef / started [Dragon Boating](#) Monday
- Cassie / swing dancing thursdays all the time
- Amy / gardening - springtime gardening is my fav 😊
- Alexis / Ski to Sea relay race at the end of month - I'm sea kayaking! > so cool!
- Chris / Excited to go wedding outfit shopping for my friend at the end of the month
- Owen / finally getting back out into the Colorado mountains for the year!
 - Architect on Transformation Train when Patrick Quinn left
-
-
-
-

Catchups, Goals & Planning (~30 mins)

Welcome

- Quick preamble - Julie
 - Hackdays to help move Cookbook forward. Erin/Julie/Stef help shape w/ skills; your expertise + contributions drives content. We can be opinionated together.
 - Communications plan: Slack/[GitHub Discussions](#)
 - please [say hi in Erin's GitHub Discussion](#) thread
- Welcome Owen Littlejohns!! 🙌

Updates

- Cloud playground - Cassie + Amy
 - Moving along and fast (for NASA). With VEDA. Collab'ing with 2i2c; using Openscapes 2i2c as basis. Will be "friends & fam", not open to public

- Would we like to help build tuts? Yes yes yes, since basis already in Cookbook. Theirs will be advertised well.
- What are next steps?
 - how do we have VEDA hub mimic our envt
 -
-
- Alexis: these tuts need to come from a nasa.gov website
 - will talk to Justin about this
 - we'll keep working here for now; whole idea is to get this integrate in NASA
 - “operational hardening”
 -
 -
-

Review past work

- [Summary from Hackday 3](#) (link to bookmark below)
- **Get-started update** - Jess updated 🎉
 - <https://nasa-openscapes.github.io/earthdata-cloud-cookbook/our-cookbook.html>
 - She reviewed files in the repo but not published in the book, kept/improved what was relevant, deleted old cruft
- **OpenDAP** - Cassie, Amy,
 - Amy: @Cassie Nickles I was just thinking about the conversation we had with the OPeNDAP team last week and was wondering if there's anything we could work on during our Hackday in that context? Like are there other OPeNDAP tutorials that we should be incorporating into the cookbook, or adding links back to the OPeNDAP github.io page? Just wondering if there are any other dots to connect.
 - Cassie: In the how do I subset data section, we could probably link to their opendap starter page <https://opendap.earthdata.nasa.gov/> that links to their access tutorials <https://opendap.github.io/documentation/tutorials/DataAccessTutorials.html>. I'll join in Wednesday and we can discuss
- “Help our tutorials be more user-friendly” - Hyrax, server who supports. (also an animal)



-



- wikipedia: "...small, thickset, ...well-furred, rotund"
- They have a list of access tutorials; we can provide feedback and add more. Timeline: getting tutorials by early July, give feedback by early August.
- Cassie: would love having a link in the Cookbook to the access tutorials
- https://nasa-openscapes.github.io/earthdata-cloud-cookbook/how-tos/access_data.html
-
- (Stef moved notes [below](#))
-

Pitch ideas for today:

(add your ideas)

- **Tutorials review** [Inventory spreadsheet](#) (started by Catalina Feb 3)
 - Idea for today: We review by running in 2i2c and seeing if relevant
 - Cloud Hackathon tutorials, then publish in cookbook (they're already in the repo)
- **Harmony?** Amy previously pushed to branch, draft PR work in progress
 - <https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/pull/187>
- **Examples folder** - call to review & add to published book (Julie pitch)
 - We review by running in 2i2c and seeing if relevant

earthdata-cloud-cookbook / examples /

Andy Barrett add examples stub eda1332 · 3 months ago History

Name	Last commit message	Last commit date
..		
GESDISC	added GESDISC kerchunk to quarto book, updated kerchunk notebook titles	last year
LPDAAC	added point buffer search and access example	last year
NSIDC	add USGS Landsat example to quarto, tidy sidebar titles	last year
ORNL	added kerchunk example for podaac ECCO SSH and gesdisc MERRA2; ornl's...	last year
PODAAC	added GESDISC kerchunk to quarto book, updated kerchunk notebook titles	last year
USGS_Landsat	add USGS Landsat example to quarto, tidy sidebar titles	last year
img	added img and updated markdown	last year
Transform-workflow.ipynb	add USGS Landsat example to quarto, tidy sidebar titles	last year
index.qmd	add examples stub	3 months ago

-
- (earthaccess intro) - Jess & Luis have made progress but neither are here today
-

Breakouts (45 mins)

- **OPeNDAP: Cassie & Chris**
 - OPeNDAP: Cassie & Chris added OPeNDAP access to how-do-i > access data > if I'm local
 - Merged branch into cookbook main!
<https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/pull/202#pullrequestreview-1421457010>
- **harmoniypy: Amy, Owen, Julie**
 - harmony-py: Amy and Owen added Harmony to how-do-i > subset data
<https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/pull/187/commits> . Nearly ready to merge but zoom screensharing stopped
-

```

[25]: f = s3_fs.open(url, mode='rb')
      ds = xr.open_dataset(f)
      ds

AttributeError                                Traceback (most recent call last)
Cell In[25], line 1
----> 1 f = s3_fs.open(url, mode='rb')
      2 ds = xr.open_dataset(f)
      3 ds

File /srv/conda/envs/notebook/lib/python3.9/site-packages/fsspec/spec.py:1135, in AbstractFileSystem.open(self, path, mode, block_size, cache_options, compression, **kwargs)
    1133 else:
    1134     ac = kwargs.pop("autocommit", not self._intrans)
-> 1135     f = self._open(
    1136         path,
    1137         mode=mode,
    1138         block_size=block_size,
    1139         autocommit=ac,
    1140         cache_options=cache_options,
    1141         **kwargs,
    1142     )
    1143     if compression is not None:
    1144         from fsspec.compression import compr

File /srv/conda/envs/notebook/lib/python3.9/site-packages/s3fs/core.py:649, in S3FileSystem._open(self, path, mode, block_size, acl, version_id, fill_cache, cache_type, autocommit, requester_pays, cache_options, **kwargs)
    646 if cache_type is None:
    647     cache_type = self.default_cache_type
-> 649 return S3File(
    650     self,
    651     path,
    652     mode,
    653     block_size=block_size,
    654     acl=acl,
    655     version_id=version_id,
    656     fill_cache=fill_cache,
    657     s3_additional_kwargs=kw,
    658     cache_type=cache_type,
    659     autocommit=autocommit,
    660 )

```

```

File /srv/conda/envs/notebook/lib/python3.9/site-packages/fsspec/spec.py:1135, in AbstractFileSystem.open(self, path, mode, block_size, cache_options, compression, **kwargs)
    1133 else:
    1134     ac = kwargs.pop("autocommit", not self._intrans)
-> 1135     f = self._open(
    1136         path,
    1137         mode=mode,
    1138         block_size=block_size,
    1139         autocommit=ac,
    1140         cache_options=cache_options,
    1141         **kwargs,
    1142     )
    1143     if compression is not None:
    1144         from fsspec.compression import compr

File /srv/conda/envs/notebook/lib/python3.9/site-packages/s3fs/core.py:649, in S3FileSystem._open(self, path, mode, block_size, acl, version_id, fill_cache, cache_type, autocommit, requester_pays, cache_options, **kwargs)
    646 if cache_type is None:
    647     cache_type = self.default_cache_type
-> 649 return S3File(
    650     self,
    651     path,
    652     mode,
    653     block_size=block_size,
    654     acl=acl,
    655     version_id=version_id,
    656     fill_cache=fill_cache,
    657     s3_additional_kwargs=kw,
    658     cache_type=cache_type,
    659     autocommit=autocommit,
    660     requester_pays=requester_pays,
    661     cache_options=cache_options,
    662 )

File /srv/conda/envs/notebook/lib/python3.9/site-packages/s3fs/core.py:1995, in S3File.__init__(self, s3, path, mode, block_size, acl, version_id, fill_cache, s3_additional_kwargs, autocommit, cache_options)
    1990 def __init__(
    1991     self,
    1992     s3,
    1993     (...)
    1994 ):
    1995     bucket, key, path_version_id = s3.split_path(path)
    1996     if not key:
    1997         raise ValueError("Attempt to open non key-like path: '%s' % path)

File /srv/conda/envs/notebook/lib/python3.9/site-packages/s3fs/core.py:428, in S3FileSystem.split_path(self, path)
    411 """
    412 Normalise S3 path string into bucket and key.
    413 """
    414 (...)
    425 ['mybucket', 'path/to/versioned_file', 'some_version_id']
    426 """
    427 path = self.strip_protocol(path)
-> 428 path = path.lstrip('/')
    429 if '/' not in path:
    430     return path, '', None

AttributeError: 'list' object has no attribute 'lstrip'

```

Regroup & iterate

- Cassie+Chris - OPeNDAP
 - https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/blob/177-opendap-tutorials/how-tos/working-with-data-in-cloud/Earthdata_Cloud_Data_Access_OPeNDAP_Example.ipynb
- Amy+Owen - harmonypy!
 - Amy screenshared with Owen+Julie, made a [second commit](#) before leaving; then Owen finished.
 - Question about workflows with .qmds in JupyterHub. Hard to execute.
 - Amy made a “copy” .ipynb to test that she didn’t commit
 - Owen ran `quarto render subset.qmd` from the terminal but that means you have to render the whole notebook each time
-

Breakouts #2 (45 mins)

- OPeNDAP: Cassie & Chris
- harmonypy: Owen, Julie

Wrap up

- Cassie & Chris merged pull request [added opendap access #202](#). If Cassie has time another day, she’ll see what other places in Cookbook she can add opendap to
-
-
-

Mar 30 Hackday #3

Post-hackday 3 Summary

Brief planning check-in and then mostly worked in breakouts the whole time -

- OPeNDAP tutorial (Chris) - goal to not use pyDAP, instead earthaccess/Xarray and Zarr.
 - Progress: Added auth steps and CMR query. [Notebook](#); [issue #177](#)
- How-Tos: Harmony-py subset example with PO.DAAC data (Amy)
 - Progress: Pushed work-in-progress to branch, [draft PR](#)
- How-Tos: Populated “access data”; created “read data” for python. (Cassie & Andy)
 - Progress: [merged PR](#)

- Next steps: Tagging R people fill in their wisdom - please go through and populate some of the R parts
- Tutorials: Import AppEEARS tutorial (Mahsa & Julie)
 - Progress: successful import and [draft PR](#)
 - Question: not re-import all notebooks each time (see draft PR)
- Tutorials: earthaccess walk-through (Jess). Shared workflow with Luis (conversation below). Focus: how to use earthaccess outside the JupyterHub
- Get-Started section (Jess) - reviewed [old chapters commented out in quarto .yml](#) to see how to incorporate/delete.
 - Progress: notes/suggestions below
 - Next steps: review & do!
- Environments chapter (Erin) started, notes here: [#186](#)
- DOI, Zenodo, Citation, Documentation (Stef) - Deposit in Zenodo & get DOI for cookbook [#178](#). Added DOI badge to README: [commit](#)
 - woohoo <https://zenodo.org/record/7786711>
 - Next steps: review pull request for citation text: [#191](#)

Roll call (5 min)

Name / how-you're-doing-today emoji

- Jess / 🐱
- Chris / 🌧️ (excited for tomorrow's storms here, at long last)
- Julie / 🏔️ view from my house! We got snow on the mountains last night
- Cassie / ☀️
- Stef / 🔧
- Celia / 🛠️ my work today is not too related to openscapes - fixing/setting up for dataset. Nearly forgot, I recently answered a user question who followed a tutorial on Openscapes Cookbook, but they did not realize it had to run in the cloud. -> a call to boldly label tutorials in-the-cloud
- Erin/ 👍
- Andy / ☕
-
-

Post-hackday 2 Summary: (on cookbook [README](#))

At our 2nd hackday we had a brief overview of cookbook progress since we'd last met before working in breakout groups together. In this session:

- Updated the importer function by adding .md capability and rethinking storing copies of notebooks bc will get out of sync ([commit](#)) (Cassie+Luis)
- added CC-BY [license](#) and added [license text to page footer](#) (Erin+Stef)
- Planned How-Tos restructure through organization & level of granularity on left and right nav ([draft PR](#)) (Andy, Catalina, Mahsa, Alexis, Julie)
- Update and add Cloud OPeNDAP tutorials ([issue](#)) (Chris, Michele)

- Reviewing earthaccess tutorial in cookbook to parse and point to the docs; also [how to sync a forked repo](#) via cmd line (Jess)

TODAY

- Chris - troubleshooting with OPeNDAP, will try to not do it with pyDAP, wants to try it with earthaccess/Xarray and Zarr. Michele on vacation this week
 - <https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/issues/177>
 - Added auth steps and CMR query: https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/blob/177-opendap-tutorials/how-tos/working-with-data-in-cloud/Earthdata_Cloud_Data_Access_OPeNDAP_Example.ipynb
 -
- Julie, Amy, Cassie - How-Tos we'll focus on filling in content - the plan. What we did below:
 -
- Amy worked on Harmony, pushed to branch, draft PR work in progress
 - <https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/pull/187>
- Mahsa & Julie - import appears notebook and draft PR
 - After the tutorials are there we'll break out to how tos
 -
- Cassie & Andy - populated access data more and created the read data section for python. Tagging R people fill in their wisdom -please go through and populate some of the R parts
 - Prev idea was to go through 2021 hackathon tutorials
 - <https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/pull/190>
- Jess - earthaccess tutorial. Shared workflow with Luis (conversation below). A user would have stopped at step 8 but I got to 13. Focus: how to use earthaccess outside the JupyterHub
- Jess shared in chat:
 - Something interesting I read about git branches/commits: A branch in Git is simply a lightweight movable pointer to one of these commits. The default branch name in Git is master. As you start making commits, you're given a master branch that points to the last commit you made. Every time you commit, the master branch pointer moves forward automatically. [Source](#)
- TODO: add corn chapter - this is what it does, how to
- Jess: Get Started section (unpublished) - reviewed [old chapters commented out in quarto.yml](#) to see how to incorporate/delete
 - SEE: get-started/index.md; get-started/earthdata-login.qmd ## review, paste in appendix, delete; get-started/api-primer.qmd ## move to how-tos/use_apis?
 - Jess's \$0.02
 - Proposed Sources of Information
 - Understanding the Cloud
 - Selected content from "Cheatsheets, Guides, & Slides" page

- API primer
- Authentication
 - Earthdata Login
 - earthaccess EDL programmatic login
 - .netrc creation? (what contexts will this be necessary?)
- How to Code
 - Selected content from "Our Cookbook" page
 - Python installation/environment setup
 - R installation/environment setup
- OR some of the "background" info listed above can go into a new section and Get Started is pertinent only to technical aspect of code/cloud environments. Either way, I think this would replace portions of the "How to..." pages.
- Mahsa - APPEARS
 - How do I
- Erin - will start <https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/pull/186>
- Stef:
 - Add instructions on getting DOI for published materials to Approach Guide [#49](#)
 - Pull request for review <https://github.com/Openscapes/approach-guide/pull/54>
 - Deposit in Zenodo & get DOI for cookbook [#178](#).
 - woohoo <https://zenodo.org/record/7786711>
 - Added DOI badge to README: [commit](#)
 - pull request for citation text: [#191](#)
 - TODOs
 - Create an "Openscapes Community" on Zenodo that will point to all Openscapes "publications" we have in Zenodo. Examples [NASA TOPS](#), [Center for Scientific Collaboration and Community Engagement \(CSCCE\)](#)
 - ~~Q: where in [Approach Guide](#) does "how to get DOI" go, and what do we call it?~~
 -

Jess's Conversation with Luis about *earthaccess* testing

Jess

Hi, Luis! I already got earthaccess "to work" in our 2i2c JupyterHub, so I decided to do user testing on my own machine. I've hit several hiccups, which I will describe next.

I am working on a Windows OS (ZBook) that is managed by ORNL. I started "from scratch" as I assumed a novice python user might.

1. In Visual Studio Code, I tried to run `conda install -c conda-forge earthaccess` but conda was not recognized
 2. I tried `pip install earthaccess` but pip needed to be updated
 3. I updated pip, and tried `pip install earthaccess` again, but the python version was too old to support earthaccess
 4. I opened the VSCode python extension to view my environments, and I opened a terminal for my conda (miniconda3) python environment
 - 5 I tried to run `conda install -c conda-forge earthaccess` but it got stuck in "solving environment"
 6. I removed and added back the conda-forge channel then updated conda to v23.1.0
 7. I enabled libmamba with `conda install -n base conda-libmamba-solver` then set it to be my default solver with `conda config --set solver libmamba`
 8. I tried to run `conda install -c conda-forge earthaccess` again and got back:
CODE BLOCK
 9. I investigated further to find that there was a Python 3.9 installation hiding out and removed it
 10. I set the default python interpreter in VSCode to python v3.10.9 (conda managed)
 11. I deleted all the conda cache from the AppData directory for my user
 12. I tried installing earthaccess again and got back: CODE BLOCK
 13. Since there was a problem with python-magic I tried installing that directly and got back:
CODE BLOCK
 14. Since that didn't help, I set my solver back to "classic" and ran the original code again to get back: CODE BLOCK
- CONCLUSION: It would probably be difficult for even a novice user to get started... at least it is for me. I am pretty good at googling my way through trouble, but this has shown to be difficult given the limited documentation I can find. I imagine another user would have given up back around step 8. What's your advice?

Luis

Jess, can you test in windows if this works: `pip install git+https://github.com/nsidc/earthaccess.git@auth-improvements`

Jess

Hi! Just a heads up... I've been running the command you provided. It's working... still giving me messages... but I started it at about 3 hours ago.

Luis

what? wait... the pip install is taking 3 hours+??

that's not right, I definitely need to create the wheel for Windows, in a couple days with the next version

Mar 2 Hackday #2

Post-hackday 2 Summary: (on cookbook [README](#))

At our 2nd hackday we had a brief overview of cookbook progress since we'd last met before working in breakout groups together. In this session:

- Updated the importer function by adding .md capability and rethinking storing copies of notebooks bc will get out of sync ([commit](#)) (Cassie+Luis)

- added CC-BY [license](#) and added [license text to page footer](#) (Erin+Stef)
- Planned How-Tos restructure through organization & level of granularity on left and right nav ([draft PR](#)) (Andy, Catalina, Mahsa, Alexis, Julie)
- Update and add Cloud OPeNDAP tutorials ([issue](#)) (Chris, Michele)
- Reviewing earthaccess tutorial in cookbook to parse and point to the docs; also [how to sync a forked repo](#) via cmd line (Jess)

Progress since Hackday 1 included:

-  [_CookbookTutorialsInventory_Feb_2023.xlsx](#) that Catalina made to identify Tutorials in the Cookbook and track decisions and progress for updating; and [Imported Cloud Hackathon Tutorials](#) so this content is available to modernize in the cookbook
- [Reorg How-Tos structure as "How Do I..."](#) - the work that Andy led to restructure the How-Tos structure to be more goal-oriented
- [Cookbook GitHub Issues](#) - began organizing ideas in Issues

Roll call (5 min)

Name / some shared joy of any shape or size / how-you're-doing-today emoji

- Julie - sharing our work at [Women In Data Science](#) conf next week! On a panel with NASA Openscapes Champion Nikki Tulley wooo 🌈
- Jess - 😞😞
- Stefanie Butland - yes dogs! had a short riverside walk with my Ruby this morning ☀️❄️
- Michele Thornton - My dogs are my current joy! Always happy and wiggly.
- Cassie - canceled plans yesterday to rest actually was so joyful! 🤗
- Mahsa - I got my Ancestry results back and I am excited to find out about my matches 😊🌈
- Chris - Enjoyed a warm IL day yesterday, preparing for possibly a foot of snow tomorrow!
- Luis - twins at home, we're watching [Mira Detective](#)
- Andy - chilly morning dog walks in the sunshine, followed by coffee.
- Alexis / putting together a beer league softball team for the first time in years / 🤔
- Catalina - seeing snow capped mountains in my "backyard" 🏔️ / massive headache 🤕

Goals & Planning (30 min)

- **Progress since last time (share brief updates/screenshares)**
 -  [_CookbookTutorialsInventory_Feb_2023.xlsx](#) - Catalina
 - Next steps:
 - [Reorg How-Tos structure as "How Do I..."](#) - Andy
 - Next steps: create issues for each how to so we can assign/deal with
 - [Imported Cloud Hackathon Tutorials into Cookbook](#) - Julie

- Next steps: decide which to include in `_quarto.yml`; update with earthaccess
 - [Cookbook GitHub Issues](#)
 - Spotlight: Jess Welch pull request! [#172](#)
 - Other:
- **Ideas for breakouts** - let's assign people to GitHub Issues before breakouts
 - update importer function to include .md files ([#174](#))
 - Luis, Cassie
 - Put false
 - license and citation
 - Stef, Erin
 - added CC-BY [license](#) and added [license text to page footer](#) (need to fix html of cc-by logo)
 - update README, including citation [#167](#)
 -
 - Add Pull Request Review to Contributing section [#173](#)
 - “good first issue”
 - Add mindset and stories chapter [#169](#)
 -
 - Update and add Cloud OPeNDAP tutorials [#177](#)
 - Chris, Michele, other volunteers
 - Make a plan to add Appears to Cookbook, make issue
 - Mahsa
 - How-tos
 - Andy, Julie, Mahsa, Andy
 - Mahsa question: how to import other notebooks like my appears
 - <https://nasa-openscapes.github.io/earthdata-cloud-cookbook/contributing/workflow.html#importing-remote-notebooks>



NASA Earthdata Cloud Cookbook



Welcome

Our Cookbook

Cheatsheets & Slides

How do I... >

find data >

access data >

subset data >

reformat data files >

reproject and regrid >

aggregate data >

plot data >

use APIs >

setup an AWS
instance? >

How-To Guides Parking
Lot

Tutorials >

- [Download Data](#)
 - [How do I subset data granules?](#)
 - [How do I reformat data files?](#)
 - [How do I reproject and resample a data file?](#)
 - [How do I aggregate data?](#)
 - [How do I plot data?](#)
 - [How do I setup an AWS instance?](#)
 - [How-To Guides Parking Lot](#)

-

Breakouts 2 (30 min)

- Chris & Michele:
 - Beginning to add overview and.opendap review to first markdown section

-

-

Importer function screenshare & Discuss

- Luis and Cassie
- Next Steps:
 - Licensing?
 - Add bot to auto-update the notebooks on a weekly basis to ensure they stay up to date (given any updates in its original source)

Regroup & iterate

- How Tos room issue and draft pull request
- <https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/issues/181>
- <https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/pull/182>
-
-
-

Check-out summary before you leave

-

-

Feb 14 ❤️ Pull request review

Amy Julie, Pull request review

To add these notes to the cookbook > [pull request](#) Jul 10

short video of our process as we reviewed a pull request together for the cookbook; it's in our Mentors Cohort Folder and here is the [direct video link](#) if you'd like to watch/share with colleagues.

From the PR page (<https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/pull/171>):

- Start with Conversation tab:
 - We can see all commits and comments on what he has worked on
- Commit Tab:
 - More details on the commits that we saw under Conversation
 - We can see line by line what has changed under that commit (green lines are added, red lines have been removed)
- Files Changed Tab:
 - Sometimes easier to see all the files that changed across the commits
 - In Nav bar: Orange box signifies modified; Green box means somethings' been added
 - We can check the "Viewed" box as we go through them in the Hub
- Checks tab (usually we skip over it!)
 - Has to do with Github Action
 - This means two checks have passed
 - We can know whether this PR has the ability to be built

Now, switching to the Hub:

- Terminal:
 - Switch over to the Main branch and pull so that we have the most recent changes from remote
 - Git checkout to the branch that has the PR (reorg_how_to_guides)
 - Quarto preview - this will build the book with Andy's suggested edits
 - *** The Browse URL actually points to the production Hub! ***
 - We need this to point to staging since that's where we're working, so we updated the URL to staging.openscapes.2i2c.cloud...

We can now view the changes in the built Quarto book and concurrently look at the files changed in GitHub

- Now we see a minor issue with capitalization in the how-tos/subset/index.md file:
 - We can add a suggested by clicking the "plus" button below the line in question
 - We can click "Start a review" button so that Andy gets a single email when we're done reviewing.
 - Click "Viewed" box once we're done with that file

We will review each individual .md file and can come back to the main quarto.yml if we see an issue with the main navigation

- We ran into an issue with the how-tos/find-data/programmatic.qmd title: Even though Andy had changed the title, it was still rendering as the old "Programmatic Search" title. We were troubleshooting the reasons for that but still working on a root cause...

Feb 2 Hackday #1

Roll call

Name / something you've learned/read/seen recently

- Julie / [Birdnote Daily podcast](#) (1:45 min episodes)
- Chris / Groundhogs day was originally celebrated with a badger (German tradition)
- Sargent / Alex from OSL gave the Earthdata webinar for January last week!
 - Set a record for number of participants! And 800 views so far on YouTube (after just two days online)
 - Poll -
 - 270 learn more about OSL
 - 230 help migrating to the cloud
 - How to respond?!
- Cassie / I learned there's a group of scientists working with PO.DAAC data that are accessing it with Julia & want to share some workflow tutorials in this other language!!
- Stef / yes to bird podcasts! [Once Upon a Checklist by Jer Thorp](#) (30-min each) - stories behind the data. e.g. iNaturalist
- Alex / I learned that the need for atmospheric correction of InSAR popped up after an accidental volcanic event warning went out in Hawaii
- Mahsa/ Completed the Carpentries training
- Amy / ~~nothing is coming to mind! That's not a good sign -laughing-~~ I keep thinking about the fabulous ESIP plenary talk from Rebecca Heiss but not sure if the video is available yet
- Luis Lopez / <https://mothership.sg/2023/01/marie-kondo-home-messy/>
- Jess / The Last of Us (on HBO) is awesome... watch it!
- Bri / ebird is pretty cool! <https://ebird.org/home>
- Erin / 1619 docuseries on Hulu & option-command-m + highlighting text in google docs allows you to start a comment.
- Andy / Inside Man on Netflix with David Tennet and Stanley "The Tuch" Tucci.
- Catalina / Star Wars - Rebels series is pretty great

Goals & Planning (30 mins)

- **Proposed plan for today** - Julie
 - Planning, breakouts, regroup, iterate from there
 - This is a public doc; and shift to GitHub Issues more to track dev&progress
 - Public #cloud-cookbook Slack channel
- **Today's goals (proposed)** - screenshare cookbook
 - **Tutorial Inventory** - what are the categories?
 - don't just have a list of them but import/modernize specific ones
 - "I can't find a thing I know should be in the cookbook" - CMR example

- In Development - move into the cookbook proper
- **How-Tos inventory** - tackle these
 - Populate categories we already have
 - Map from In Development and Tutorials, identify
 - Then have mini-hacks for different topics
- **Opinionated** - update with `earthaccess`
 - save other approaches to Appendix
- **Ongoing goals**
 - Add Mindset & Stories - following ESIP convo
 - Contributing - GitHub + Quarto update docs & teach
 - Cheatsheets - update and archive
 - Archive because really important to share the progression and progress
 - Importing notebooks - Luis

Discussion

- What do we want in the cookbook
- With our existing stuff, what does it need / any tweaking, updating, or language to be consistent with cookbook
- “This is how we currently do it, it’s a moment in time. When you revisit this in 6 months it will suggest you visit the Cookbook for most current”
- Workshops - discrete digital entities - give it a DOI
 - with a DOI
-

Breakouts (30 mins)

- Stef Main room - create #cloud-cookbook channel
 - everyone in this mtg has been invited to #cloud-cookbook channel
- Julie & Luis - moving tutorials
- Main room: How-Tos - Amy & Cassie
 - high level inventory
 - Find Data
 - Working with data in the cloud
 - Direct S3 Access
 - Stream HTTPS
 - How to access data with a End User License Agreement
 - How do I convert to zarr to work with COFs?
 - https://www.earthdata.nasa.gov/s3fs-public/2022-05/CSDA_Program_USG_EULA-11-09-20_Rev3.pdf
 - E.g. <https://disc.gsfc.nasa.gov/earthdata-eula>
 - Working with cloud data on your local machine
 - Download HTTPS
 - OPeNDAP (?)
 - Transform

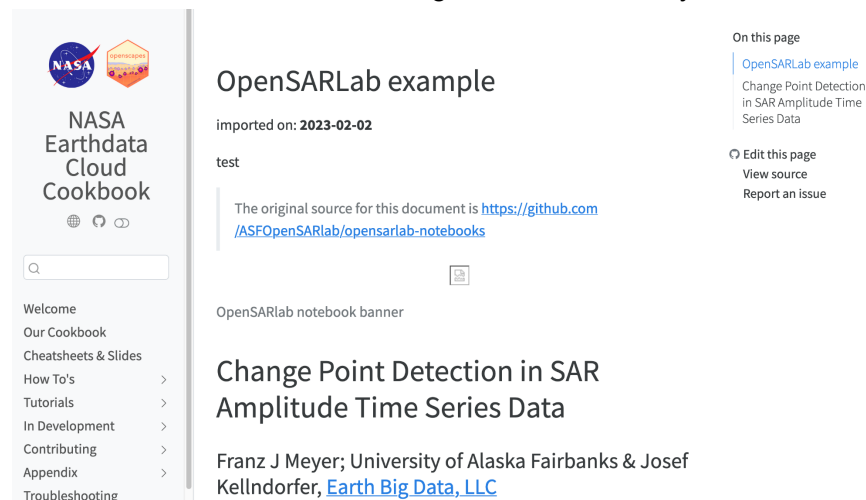
- Subset
 - “How do I download a subset of NetCDF4?”
 - Plot
 - Tech Setup (AWS)
- map in-development
- Good to consolidate the Access section
- Still challenging for a new user to know what a “COG” is - maybe we need to back up a level
- Identify common steps first:
 - Provide a roadmap for how to access different file formats and different access methods (HTTPS, S3)
 - What are the common tools? I.e. Rioxarray for COG
 - Need to keep these short - most of the content should be above the fold.
 - Maybe a separate how-to for set-up
 - How-to’s should be small - “10 or less lines of code”
- We should leverage <https://nasa-openscapes.github.io/earthdata-cloud-cookbook/how-tos/access/> to provide an intro, describe different data formats, access options, etc.
- Where do we stop (when is “access” achieved)?
 - Earthaccess deals with the different modes of access and formats, so we could strip a lot of the guidance out from the older how-tos (that came from 2021 hackathon)
- <https://contributing.ropensci.org/> to help us frame: “I want to...”
- Catalina & Bri- thinking about categorization of tutorials, but also, *what* goes in Tutorials
 - TODO: create [inventory](#) of existing “Tutorials” from
 - *Cloud Hackathon*
 - *AGU workshop*
 - *In Development, EDC Cookbook section*
 - Flag which we may want to work off of and update to put in the Cookbook (do not link from existing Workshops - those are snapshots in time, but use as starting point to build the “most up to date” tutorials in the EDC Cookbook)
 - Start thinking about how to categorize them: do we reflect the “How do I ...” from the How To section?
 - Tutorials from under **In Development**:
 - Search & Download NSIDC Cloud
 - DAAC example
 - Access NSIDC Cloud w/ CMR
 - DAAC example, CMR
 - Direct Access & Harmony Workflow
 - in-cloud direct access, How To?
 - Search & Access USGS Landsat Cloud
 - DAAC example
 - OPeNDAP (Chris Battisto)

- tool/service example, How To?
 - In *Access* version: open and download
 - In *Transform*: subsetting
- Direct S3 Data Access: NetCDF - Daymet v4 Daily TMAX Example
 - DAAC example
- COF Zarr Access via Reformat
- Zarr Example
- Kerchunk Examples
- Matlab & AWS
- CMR STAC in R
- earthdata: Python-R Handoff
- Group discussion:
- Teach people how to do things on their own, not just make it easier by abstracting too much; if something breaks, they might need further help; need a balance
- Troubleshooting info, as helpers to How To's (side bar? Links to more examples? cheatsheets?)
 - Teach people how to fish instead of giving them your fish; but don't need to teach them how to forge a fishing hook
 - In notebooks, we could add functions in hidden code cells. So some things are obscured but easy to display if something breaks.
- LEGO analogy
 - Basic blocks, specialized blocks, kits with suggestion what to build
 - 1. How To's ... complex abstracted blocks; i'm just missing a specific piece of code
 - 2. Tutorials Longer form description of what that step is
 - 3. "Big data/complex examples" ... not meant to be modular; e.g. process MODIS to look at greening; "Skyscraper" examples: combining field data with remote sensing data / a mode, Skyscrapers can copy published research
 - 4. Beyond Python Do we want another section for the R and Matlab examples? Where do those go? (added after hack session)
 - > We'll have one narrative and we use tabsets to show code in different languages, like
<https://nasa-openscapes.github.io/earthdata-cloud-cookbook/how-tos/find-data/programmatic.html>
- make the implicit explicit - tell people what they will find in how-to vs tutorial
- Jess - I'm in a room alone, but Michele will likely join. I'm looking at all the materials so I can think about a contribution. Taking notes here for posterity.
 - Repo README should be fleshed out. An important part is to give a full reference and example of how to cite the cookbook. We, at the ORNL DAAC, will use this reference to point users to the materials within our own tutorials. Will also want to explain to new users how to get to the "gui" for the cookbook if they hit the repo first. Should part of our cookbook include a little demo of GitHub basics?

- <https://nasa-openscapes.github.io/earthdata-cloud-cookbook/>
 - This Cookbook is learning-oriented to support scientific researchers using NASA Earthdata from Distributed Active Archive Centers (DAACs) as ~~they~~ the NASA migrates data and workflows to the cloud.
 - I think the description of the figure can be improved by explaining the workflow step-by-step. Also, there isn't an explanation of AWS, although the name is "in" the cloud graphic.
- <https://nasa-openscapes.github.io/earthdata-cloud-cookbook/our-cookbook.html>
 - Distinguish between "How To" and "Tutorial."
 - I see some small edits to grammar/sentence structure/punctuation that can be made.
- I see a lot of pages in development from contributions from lots of folks. Near the end of "finalizing" this first "version," someone can look across all the pages and edit them to look/feel consistent.

Regroup & iterate

- Import room Report out -
 - Progress:
 - Imported openSARLab test file
 - Quarto error because of `---`, diagnosed and fixed by hand



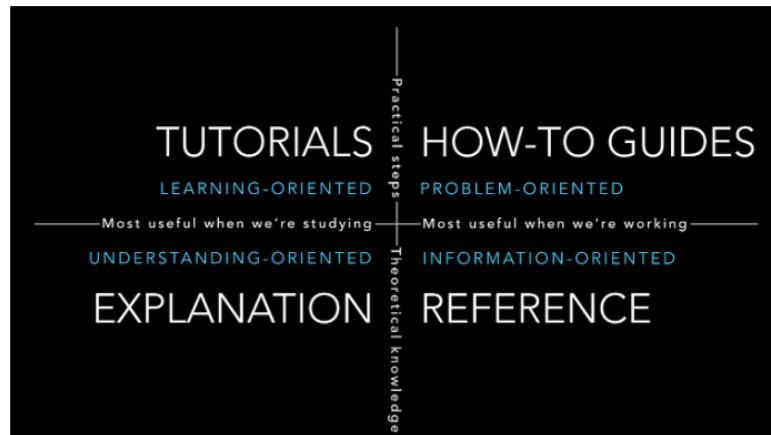
The screenshot displays a Quarto notebook page titled "OpenSARLab example". On the left is a sidebar for the "NASA Earthdata Cloud Cookbook" with a search bar and a navigation menu including: Welcome, Our Cookbook, Cheatsheets & Slides, How To's, Tutorials, In Development, Contributing, Appendix, and Troubleshooting. The main content area shows the notebook title, the import date "imported on: 2023-02-02", the test name "test", and the source URL "https://github.com/ASFOpenSARlab/opensarlab-notebooks". Below this is a banner for "Change Point Detection in SAR Amplitude Time Series Data" by Franz J Meyer and Josef Kelldorfer. On the right, a "On this page" section lists "OpenSARLab example" and "Change Point Detection in SAR Amplitude Time Series Data", with links to "Edit this page", "View source", and "Report an issue".

- Next steps:
 - Media urls
 - Test quarto error in staging hub and if persists Julie talk to quarto
 - Move all 2021 Cloud Hackathon tutorials over
 -
- Main room -
 - (Can all data hosted in the cloud be accessed using https and s3)

- Take a step back - make sure our language is what users would expect; eg “COG”
- See above
-
-
- Tutorial planning room
 - See above
 -
 -

Discussion

- Bri - Big examples for science: one modis file to all of them! And how to do that as a daily thing
- Andy - lego blocks and then kits. Pictures of ideas
 - Lego: more complex abstracted blocks: like a door. I need a door.
 - Bri’s idea are the big skyscraper lego examples.
 - In this Cookbook there is a place for all of these. DOn’t need to be totally hierarchical.
 - Skyscrapers can copy published research
 - yeah or even the skyscraper is combining field data with remotesensing data / a model :)
 -
 - Divio reminder
 - <https://documentation.divio.com/>



- re: “Have a grey block of “Note: if you hit this error” sort of box”, in the Openscapes Approach Guide we use green “tip” boxes and blue “historical aside” boxes.

💡 Tip: Click URL to add to your Google Calendar



Google Meet: Breakout room functionality in Google Meet is limited

💡 Tip: Click URL to add to your Google Calendar



Sometimes we ask people add an event to their own Google Calendar without being added explicitly as a Guest. To do this, in a Calendar Event > click the 3 dots > Publish event > choose Copy Link to event and share link in Slack or email saying "Click URL to add this to your Google Calendar". We have used this to let people in a Cohort know of an optional impromptu session like drop-in GitHub help, or a screenshare-and-tell for example.

- [Source](#) of tip in Quarto
 - `::: {.callout-tip collapse="true"}`
 - `## Tip: Click URL to add to your Google Calendar`
 -
 - Sometimes we ask people add an event to their own Google Calendar without being added explicitly as a Guest. To do this, in a Calendar Event \> click the 3 dots \> Publish event \> choose Copy Link to event and share link in Slack or email saying "Click URL to add this to your Google Calendar". We have used this to let people in a Cohort know of an optional impromptu session like drop-in GitHub help, or a screenshare-and-tell for example.
 - `:::`
- [Source](#) of historical aside, use
- `::: {.callout-note collapse="true"}`
- `### Historical aside`
- text
- `:::`

Breakouts (30-45 mins)

Import

Tutorial Inventory

Cookbook Dev Ideas

Amy Steiker:

Andy Barrett: - re ESIP session:

So many great thoughts. One of the, many, things that stood out, was the survey presented by [@Corey Clatterbuck](#).

I've been asked to co-lead a working group to develop an Open Science goal and tactics for the NSIDC strategic plan. We're a mix of devs, data support specialists, ops, and researchers. Not everyone codes or analyses data. So there are many facets to creating an open science mindset. It really hit home that there is a need to emphasize the human component of creating this mindset. The technological side of OS seems more tangible to those of us familiar with code. However, I suspect it is a barrier to some who are less involved in the tech side.

I wonder if we should include stories about the human side of creating an open science mindset in the cookbook, as well as the technical side of things. I, for one, would be really keen to hear more of these stories.

Amy:

I love the idea of us sharing those stories within the cookbook itself and/or reframing or re-focusing the collaboration tools “step” in the cloud roadmap more on the human side.

Chris Battisto:

For our hackday this week, I'd like to clean up and add to the OPeNDAP page under development:

https://nasa-openscapes.github.io/earthdata-cloud-cookbook/how-tos/access/Earthdata_Cloud__Data_Access_OPeNDAP_Example.html

Currently, my duties for this PI at GES DISC is to create some high-level documentation to prepare users for our Cloud OPeNDAP granules, and the resulting differences between DAP2/DAP4 constraint expressions. I'd like to add some of these with brief descriptions to the page, and I'd also like to see how the earthaccess library handles these DAP4 links. I'll also possibly include Michele and Cassie on this effort-they've created some great notebooks at their DAACs that demonstrate Cloud OPeNDAP access with Python that could serve as useful references.

Let me know if this sounds good! This will also be my first foray into Quarto territory



Sept 29 2022 inventory (historical record)

Cookbook Inventory list! (see discussion just above as well)

- **Find Data** - currently Earthdata Search portal from Hackathon
 - TODO: offer two options: Julie can update structure
 - Earthdata Search - exists
 - Programatically find data
 - https://nasa-openscapes.github.io/2021-Cloud-Hackathon/tutorials/01_Data_Discovery_CMV.html exists, need a full tutorial to understand API.
 - Instead, use earthaccess (as a wrapper for CMV) - does this code exist?
 - Update Oct 3:
<https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/pull/146>
 - Notes: how to think about this for R, wget, curl, etc
 - Bri in progress:
https://raw.githubusercontent.com/NASA-Openscapes/HLSBulkQuery_STAC/main/HLS_bulkdownload.html
 - Repo:
https://github.com/NASA-Openscapes/HLSBulkQuery_STAC
 -
- **Authentication** - currently links to Hackathon and luis's demo
 - Notes: Andy does this make sense to pull this out as a separate section?
 - Search for data
 - Download data - this step includes auth and credentials
 - How to call this out in the cheatsheet. But not necessary as its own section in HowTos because with earthaccess it's really 1 line of code
 - Cassie -
 - TODO
 - Remove Authentication as its own section
 - Build in the 3 auth line in the Access HowTo from
<https://nbviewer.org/gist/betolink/b8a37c6e58da16ba7183c21e6fc731b1>
 - STATUS: removed Auth and created reminder in Access:
<https://github.com/NASA-Openscapes/earthdata-cloud-cookbook/pull/147>
- **Transform**
 - Notes:
 - Should Transform be the next step after Find Data
 - Amy: Optional, and not supported for all data sets
 - Also it's future work: maybe earthaccess can ping all these places and tell you what transform is available for the dataset of your choice

- So if we can't transform as part of the earthaccess, where should we put the Transform piece. (we have a separate harmony python client - there is overlap because you can authenticate with that library)
 - Andy better for earthaccess as a swiss army knife or more modular
 - Constellation of smaller packages -
 - Julie: tidyverse approach <https://tidyverse.org/packages> - can install them each separately but they work together and once they were developed far enough they were packaged into a "metapackage" that installs them all
 - Catalina - Abstract for the user
- Amy: do we still want to think about a basic harmonypy example, will think about how that interacts with earthaccess. Amy - we can def do this if we think it's worth doing
- "This is how you would use this tool" - Catalina. Yes add as a how-to or tutorial. A longer-form tutorial that links the modular packages together: earthaccess and harmonypy
-
- Also, OpenDAP. This offers some functionality and fits with existing workflows
- TODO:
- Instead of Transform, rename to: "Subset, Reformat, Reproject"
 - Subset
 - Harmony
 - Python (Harmony-py), curl (Harmony API), etc.
 - OpenDap
 - Python example, Curl example, etc.
 - Reformat
 - Harmony
 - Python (Harmony-py), curl (Harmony API), etc.
 - Reproject
 - Aggregate
 -
 -
- Notes: curl and wget as options here. Don't have to do it all at once all up front.
- For finding and accessing, use earthaccess
- For access, use opendap. "How to use opendap" - they'll click this if they're looking for it and
 - Catalina: a person at PODAAC will dev these tutorials next PI
 - Michele has some tutorials here too, GESDIS
- You're trying to transform, use harmony
-

- **Access - Cloud Workflow** - current state: dropped AGU tutorials (less like howtos. Lots of overlap with auth etc)
 - Amy - how much of these examples could be consolidated with earthaccess. Is it the same workflow with S3 vs Htpps, and single vs multifile. Maybe not needed . Maybe instead of 4, could have 2
 - Catalina - maybe COG multiple files?
 - Amy - maybe question is “where do we end” with access. Maybe it’s one how to with earthaccess.
 - If we go to plotting etc that depends on format (rioxarray etc)
 - The beauty of earthaccess is that the workflow is the same whether you’re in cloud or not
 - New section: opening files? Diff for COG NetCDF
 - You pull in your python object. Now what: working with data
 - Catalina - is “open” a next step or is it part of it
 - Julie - tidyverse workflow and then constellation of packages fits within (see [Mindset slides 12-13](#))
 - Amy: earthaccess get function
 - Mahsa - how to also advocate how to work in the cloud. Maybe a whole section before download about transform, manipulate, visualize in the cloud. And then the download option
- TODO:
 - Condense these existing tutorials
 - Future conversations about how to have access and handling
- **STATUS (10/21)**: Created [new branch](#) to consolidate to two focused How-To’s using earthaccess for the python guidance, using the tabular format we started under Find Data.
 - Access Cloud-Optimized GeoTIFF
 - Access NetCDF-4/HDF5
- **Download**
- **Tech Setup**
 - Python environment setup - Mahsa’s suggested this Nov 2 . this could be a chapter about what is corn doing and how to install corn locally
 - FAQs here too, and link to OpenSARLab troubleshooting <https://opensarlab-docs.asf.alaska.edu/>
 - Link to Earthdata Forum - help channel that
-