

**2021 ESIP Summer Meeting (July 19th-23rd, 2021)**

<https://2021esipsummermeeting.sched.com/>

[Meeting Takeaways Doc](#)

Community Participation Guidelines & Reporting: Click [HERE](#)

---

## **Foraging for Dataset-Usage Relationships**

**Thursday, July 22 • 11:00am - 12:30pm**

**More info & slides on Sched Session Page:** <https://sched.co/jMOT>

---

### **Attendance & Check in**

*Add your Name / affiliation / pronouns / location / social media to a bullet*

- Jonathan Blythe, BOEM
- Sara Lafia / ESIP Fellow (University of Michigan)
- Christopher Lynnes, NASA\
- Myranda Uselton | NOAA NCEI | she/they | Nashville, TN | [myranda.uselton@noaa.gov](mailto:myranda.uselton@noaa.gov)
- Vanessa Chatman / NASA ESDIS / She, Her / Riverside, CA
- Cindy Lin / Cornell / she/her / NYC
- Tyler Christensen / NOAA / she/her / Silver Spring MD
- Stephanie Schmidt / Forest Service / Corvallis, OR
- Vincent Inverso / NASA (ESDIS) / he/him / Philadelphia
- Karl Benedict / University of New Mexico/ Albuquerque, NM / he/him.
- Ryan McGranaghan / he-him / ASTRA LLC & NASA GSFC / Mount Rainier, MD / @AeroSciengineer / podcast: [Origins](#)
- Ed Armstrong / NASA JPL PO.DAAC / Pasadena, CA
- Chung-Lin Shie / UMBC/JCET, NASA/GSFC / he, him / Greenbelt, MD
- Chris Erdmann / AGU / he, him / Raleigh, NC / @libcce
- Margaret O'Brien, UCSB, Environmental Data Initiative (EDI)
- Erin Antognoli / USDA National Agricultural Library / she/her / Beltsville, MD / <https://www.linkedin.com/in/erinantognoli/>
- Brian Wee / Massive Connections / he him his / Washington, DC.
- Angela Li/NASA GESDISC/She, her/Greenbelt MD
- Nate James/NASA Greenbelt, MD
- Lori Hager / NOAA NCEI / Asheville, NC
- Bruce Wilson / ORNL DAAC / he,him / Oak Ridge, TN
- Debbie Roth /MBLWHOI Library, WHOI / Woods Hole, MA
- Matt Mayernik / NCAR / Boulder, CO
- Ted Habermann / Metadata Game Changers (<https://ror.org/05bp8ka05>)
- Susan Shingledecker / ESIP / She/her / Severna Park, MD / @sshingledecker
- Sarah O'Connor / NOAA-NCEI / Silver Spring MD
- Xiaohua Pan/NASA GSFC GES DISC/She/Greenbelt, MD
- Doug Newman / NASA EED / [douglas.j.newman@nasa.gov](mailto:douglas.j.newman@nasa.gov)
- Jeffrey Campbell / USDA / Beltsville Maryland

## 2021 ESIP Summer Meeting (July 19th-23rd, 2021)

<https://2021esipsummermeeting.sched.com/>

[Meeting Takeaways Doc](#)

Community Participation Guidelines & Reporting: Click [HERE](#)

- Andrea Pörsch / Helmholtz Metadata Collaboration / GFZ German Research Centre for Geosciences, Potsdam
- Irina Gerasimov / NASA GES DISC / Greenbelt, MD
- Madison Langseth / USGS / she/her / Denver, CO
- Allan Doyle / NASA ESO / he/him / Boston, MA
- Dave Meyer/NASA/GES DISC
- H. K. "Rama" Ramapriyan/SSAI, NASA GSFC (ESDIS)/Camarillo, CA
- Lisa Kaser, NSIDC, Boulder, CO
- Robert R. Downs / NASA SEDAC
- Katy Luquire / NOAA NCEI / Asheville, NC
- John Porter/ VCR-LTER UVA / he/him / Charlottesville, VA
- Andrew Cramer/ NASA ESDIS / he/him / Houston, Tx
- Dean Carter/NOAA/NCEI/Asheville, NC
- Teddy Gelabert / NASA (ESDS/SSAI) / GSFC / teddy.gelabert@nasa.gov
- Daniel Fuka/Virginia Tech/Albuquerque, NM/
- Corinna Gries/Environmental Data Initiative
- Rohan Dayal, NASA, He/him, NY, rohan.dayal@columbia.edu
- Danie Kinkade (she/her)/ BCO-DMO / Woods Hole
- Amrutha Elamparuthy / USGCRP/ Washington DC
- Robert Casey / IRIS Data Services / he/him / Seattle, WA
- Nathaniel Crosby / NASA GES DISC / he/him / MA
- Mathew Biddle / NOAA-IOOS / he/him / Annapolis, Maryland
- David Moroni / JPL/PO.DAAC / Pasadena, CA
- Kristina Stoyanova (she/her) / NASA GES DISC / Boston, MA
- Amanda Dean / NOAA NCEI / Asheville, NC
- Elena Steponaitis / NASA HQ
- Binita KC, NASA GSFC/Adnet Systems, Greenbelt MD

### Agenda

**11:00-11:20: Welcome**

**11:20-12:15: Forage, forage, forage...(Breakout rooms)**

**12:15-12:30: What the heck just happened?**

At session start:

- **Welcome**
  1. Orient attendees to tech (Zoom and QiqoChat)
  2. Invite attendees to share their name/affiliation/location in the Zoom chat
- **Community Participation Guidelines Reminder**

## 2021 ESIP Summer Meeting (July 19th-23rd, 2021)

<https://2021esipsummermeeting.sched.com/>

[Meeting Takeaways Doc](#)

Community Participation Guidelines & Reporting: Click [HERE](#)

- By participating in this session, you agree to adhere to the ESIP [Community Participation Guidelines](#)
- Report an Issue: <https://www.integritycounts.ca/org/esip>
- **Goals of this Session:** Provide *human-curated* dataset-article relationships for Usage-Based Discovery's Graph Database, and article classification to enable machine-learning automation in the future. We call this Foraging for Dataset-Usage Relationships.
  - [What are we doing today? Foraging!](#)
- **Forage, forage, forage...**
  1. **Find** research articles that **use** Earth science datasets
  2. **Classify** the article according to topic
  3. **Identify** the **datasets** used in the article
  4. **Submit** the **article DOI** and **dataset DOI** via Google Form:
    - a. [Team Flood](#) (Coaches: Irina Gerasimov, Vinny Inverso)
    - b. [Team Fire](#) (Coaches: Sara Lafia, Vanessa Chatman)
    - c. [Team Life](#) (Coaches: Jonathan Blythe, Brian Wee)
    - d. [Team Sea Level](#) (Coaches: Ed Armstrong, Andrew Cramer)
    - e. [Cru Motley](#) (Coaches: Chris Lynnes and Doug Newman)

Once verified, these will then show up in the [Usage-Based Discovery tool](#). (If time allows, team coaches will enter some of them into the actual database during ESIP.)

For a short demo video<sup>1</sup> of the tool, see here: [Usage Based Discovery Demo](#).

### Resources for future use

- **DataCite Commons** (<https://commons.datacite.org/>)
- **PID Graph** (<https://www.project-freya.eu/en/pid-graph/the-pid-graph>)
- **XDD support for ML phase**
- **Coleridge Initiative “Rich Context” project using NLP find hidden data citations in journal articles** (<https://coleridgeinitiative.org/data-products/>) -- contact [tyler.christensen@noaa.gov](mailto:tyler.christensen@noaa.gov)
- 

### Foraging strategies

- **Search by dataset “short name” (requires insider knowledge)**

---

<sup>1</sup> Thanks to Vincent Inverso, NASA, for creating the video.

## Resources for this session

### Quick Start

- In [Google Scholar](#), combine a research topic term with a dataset DOI in the search blank in order to get dataset-article relationships and article classification at the same time.

Example:

flood 10.5067/TRMM/TMPA/DAY/7

- **See All Results:** If it shows just one result with the text “Showing the best result for this search”, click on “[See all results](#)” to see the rest of the articles..
- **Article DOIs:** You may need to click through to the article to find the article DOI.
- **Dataset ShortNames:** Once you finish with DOI searches, you might try searching by dataset ShortName (e.g., TRMM\_3B42, or “TRMM 3B42”) instead...
- **More Google Scholar search tips:**
  - Quote multi-word phrases to avoid false positives from isolated word occurrences

“sea level change”
  - Use a “-” sign before a word or phrase to filter *out* results with that word or phrase (useful in avoiding duplicates from two searches of different keywords)

“-sea level pressure”
  - Use “OR” to search for two search terms at the same time:

“sea level change” OR “sea level rise”
- **How To Classify Articles**
  - **By Reading:** Read the first part., starting with the title and then the abstract. Note that many papers can be classified under more than one topic.<sup>2</sup>
  - **By Searching:** Add the topic to your dataset search, letting Google Scholar do the “work”.
- [Starter Pack of Search Terms and Dataset DOIs](#)
- [More about Foraging and Classification](#)

---

<sup>2</sup> In most cases, you don’t need to dive into the body of the paper; major topics are *usually* mentioned in title or abstract.


## 2021 ESIP Summer Meeting (July 19th-23rd, 2021)

<https://2021esipsummermeeting.sched.com/>

[Meeting Takeaways Doc](#)

Community Participation Guidelines & Reporting: Click [HERE](#)

- **Pro Tips**

- In the [DataONE search portal](#), look for the “quotes” icon  to get a quick idea of which datasets have been cited in publications. Then click on the dataset to see and click through to the actual articles. (Brian Wee)
- DataOne is not just about Bio stuff. There are other datasets in there too, e.g., [Programme for monitoring of the Greenland ice sheet \(PROMICE\): Moderate Resolution Imaging Spectroradiometer \(MODIS\) Greenland albedo, 2000-2019](#). But try to focus on ones related to Earth Observations (vs., say, genomes).
- In [Google Datasets Search](#), look for this in the results:

”6 scholarly articles cite this dataset ([View in Google Scholar](#))”,

Then click on “View in Google Scholar”, add your topic search term and search. (Sara Lafia).

- (Your Tip Here)

### What the Heck Just Happened??

- Please answer our [Foraging After-Action Survey!](#)
- Capture 3 Takeaways & Shared insights

- **Closing circle**

- Share any reminders
  - Come to our next Cluster Call on **Thursday, August 19 (3:00-4:00 ET)** - Discovery cluster meets on the 3rd Thursday of the Month
  - Join our cluster’s email list serve (<http://lists.esipfed.org/mailman/listinfo/esip-discovery>)
  - Check out our wiki ([https://wiki.esipfed.org/Discovery\\_Cluster](https://wiki.esipfed.org/Discovery_Cluster))
- Shout-outs & thanks (see also Slack #kudos channel)

### 5 ways to stay involved with ESIP

- Join the conversations on our [ESIPFed Slack Channel](#) and say hello on #general.
- Keep up with the latest from in and around ESIP by signing up for the [ESIP Monday Update](#) or the mailing list of this group.
- Share your feedback, questions, and suggestions with our Twitter community [@esipfed](#) and [#esipfed](#).

## 2021 ESIP Summer Meeting (July 19th-23rd, 2021)

<https://2021esipsummermeeting.sched.com/>

[Meeting Takeaways Doc](#)

Community Participation Guidelines & Reporting: Click [HERE](#)

- Hop on an [ESIP Telecon](#) that interests you - no need to RSVP!
- At this meeting, check out the Research Showcase Poster and Demo Gallery.

### Feedback

Please take a moment to share your thoughts on this session in Sched by going to the Sched session page linked at the top of this document.

### Paste in Zoom Chat notes that you would like to save here

*Notes taken by ESIP Community Fellow - Cindy Lin*

Chris Lynnes:

1. Current discovery interfaces present dataset characteristics
2. Need 3 pieces of information to drive the tool: Research Topic, Research Article, and Datasets used
3. Foraging for Dataset-Usage Relationships
  - a. Find research articles
  - b. Classify articles
  - c. Identify the datasets used in article
  - d. Submit the article DOI and dataset DOI via Google

*10:54:04 From cindy lin to Everyone:*

*Hi Session Organizers, I am your community to help with the note taking, muting mic, etc.*

*10:54:12 From cindy lin to Everyone:*

*Please do make me co-host as well! Thank you!*

*10:56:02 From MyPhone to Everyone:*

*Hi everyone!! I'll be mainly in listen only mode for a while.*

*10:56:24 From Ed Armstrong to Everyone:*

*Hello*

*10:56:35 From Karl Benedict to Everyone:*

*Good morning*

*11:03:12 From Sara Lafia (ESIP Fellow) to Everyone:*

*Hello everyone! Welcome, please sign in with your name, affiliation, etc.*

*<https://docs.google.com/document/d/13TdHZICCA0YN5geWGG7N2-vqX3LAHHwIAddCuf0JiJ0/edit?usp=sharing>*

*11:05:18 From tedhabermann to Everyone:*

*Ted Habermann, Metadata Game Changers (<https://ror.org/05bp8ka05>), [ted@metadatagamechangers.com](mailto:ted@metadatagamechangers.com)*

*11:19:48 From Tyler Christensen - NOAA to Everyone:*

**2021 ESIP Summer Meeting (July 19th-23rd, 2021)**

<https://2021esipsummermeeting.sched.com/>

[Meeting Takeaways Doc](#)

Community Participation Guidelines & Reporting: Click [HERE](#)

*Question: are you looking specifically for NASA datasets? Or any earth observation data, from any source?*

11:20:05 From Ed Armstrong to Everyone:

*Any source @tyler*

11:20:07 From Doug Newman to Everyone:

*If it has a DOI...*

11:20:19 From Tyler Christensen - NOAA to Everyone:

*Got it-- thanks!*

11:20:25 From Doug Newman to Everyone:

*Don't we have 'all of them' Tyler? If not, get working on that mate!*

11:25:22 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*Hi All - in case it helps, here is an example on how to search NOAA OneStop's 84,962 collections for a known DOI: Go to [data.noaa.gov/onestop](https://data.noaa.gov/onestop) and put in the search field:*

11:25:23 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*doi:="doi:10.7289/V5BR8Q7N"*

11:25:37 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*The quotes are needed.*

11:25:51 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*Clickable link example:*

*<https://data.noaa.gov/onestop/collections?q=doi%3A%3D%22doi%3A10.7289%2FV5BR8Q7N%22>*

11:25:54 From Ed Armstrong to Everyone:

*Thanks Ken. Good tip*

11:26:37 From Mathew Biddle (NOAA-IOOS) to Everyone:

*+1 @Ken, didn't know you could do that*

11:27:10 From Doug Newman to Everyone:

*You can search CMR/EDSC using DOIs too*

11:27:39 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*And here is a search that will show all the datasets using the 10.7289 DOI "shoulder":*

11:27:45 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*doi:10.7289\**

11:27:56 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*or clickable: [https://data.noaa.gov/onestop/collections?q=doi%3A10.7289\\*](https://data.noaa.gov/onestop/collections?q=doi%3A10.7289*)*

11:28:23 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*I think NOAA has another shoulder or two. Any NOAA participants know offhand?*

11:31:09 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*If you do simply doi:10\* you get lots more hits*

11:32:58 From Sara Lafia (ESIP Fellow) to Everyone:

*Perhaps the topic tags and comments are adding value too?*

11:33:04 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*(some of those are datasets that came to NOAA with a DOI already attached)*



**2021 ESIP Summer Meeting (July 19th-23rd, 2021)**

<https://2021esipsummermeeting.sched.com/>

[Meeting Takeaways Doc](#)

Community Participation Guidelines & Reporting: Click [HERE](#)

11:33:35 From Ed Armstrong to Everyone:

*In the starter pack in the spreadsheet there is a "search terms" free text suggestions. Its a separate tab.*

11:33:49 From Madison Langseth to Everyone:

*Is there an API for the Usage Based Discovery Tool or will there be?*

11:34:17 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*@Ed Armstrong and others planning on going to the Sea Level Crew. Here is a OneStop search that returns 30 datasets with DOIs AND the phrase "sea level":*

*[https://data.noaa.gov/onestop/collections?q=doi%3A10\\*%20AND%20%22sea%20level%22](https://data.noaa.gov/onestop/collections?q=doi%3A10*%20AND%20%22sea%20level%22)*

11:34:42 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*Oops, sorry, copy and paste error I think.*

11:34:54 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*Try this:*

*[https://data.noaa.gov/onestop/collections?q=doi%3A10\\*%20AND%20%22sea%20level%22](https://data.noaa.gov/onestop/collections?q=doi%3A10*%20AND%20%22sea%20level%22)*

11:35:25 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*I can't figure out why but that is the right link but it doesn't seem to survive the Zoom translation.*

11:35:28 From Sara Lafia (ESIP Fellow) to Everyone:

*The only stuff that shows up in PID graph uses DOIs to cite data explicitly (i.e. in a Reference list) so we can emphasize discovering data citations that are informal (where data are cited in the main body of the text)*

11:36:17 From Sarah to Everyone:

*10.25921 is another one of the NOAA doi shoulders*

11:37:09 From Ed Armstrong to Everyone:

*Very nice Ken. Excellent starting point for sea level*

11:37:10 From Sara Lafia (ESIP Fellow) to Everyone:

*+1 @bruce yes, I agree the coverage is patchy so we will be adding value!*

11:37:10 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

*Thanks Sarah! Note I think I am doing something wrong with my clickable links.*

11:37:31 From Bruce Wilson (ORNL DAAC; he/him) to Everyone:

*Thanks, Irina*

11:40:50 From Sara Lafia (ESIP Fellow) to Everyone:

*<https://forms.gle/BPyKzrnUU2ztXPTW6>*

11:41:47 From Sara Lafia (ESIP Fellow) to Everyone:

*Starter search terms and DOIs*

*[https://docs.google.com/spreadsheets/d/1\\_vjkPBZK7dALCXBvegTNJkFEBP7kl6WEgFHrryxarPQ/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1_vjkPBZK7dALCXBvegTNJkFEBP7kl6WEgFHrryxarPQ/edit?usp=sharing)*

11:41:59 From Cindy Lin - ESIP Fellow to Everyone:

*Support is much appreciated! : )*

11:49:45 From Vanessa Chatman to Everyone:

*<https://scholar.google.com>*



**2021 ESIP Summer Meeting (July 19th-23rd, 2021)**

<https://2021esipsummermeeting.sched.com/>

[Meeting Takeaways Doc](#)

Community Participation Guidelines & Reporting: Click [HERE](#)

11:49:59 From Vanessa Chatman to Everyone:

<https://ui.adsabs.harvard.edu>

11:50:12 From Sara Lafia (ESIP Fellow) to Everyone:

*I like to use Google Dataset Search too! <https://datasetsearch.research.google.com/>*

11:50:12 From Vanessa Chatman to Everyone:

<https://www.dataone.org>

11:50:21 From Vanessa Chatman to Everyone:

<https://datasetsearch.research.google.com/search?query=10.3334%20wildfire&docid=L2cvMTFqY2tnMmJ2NA%3D%3D>

11:50:29 From cindy lin - ESIP Fellow to Everyone:

*Thank you!*

11:51:48 From Sara Lafia (ESIP Fellow) to Everyone:

*Once you find an article that uses your dataset, please enter it here using this form*

<https://forms.gle/uAzbLxZg3bVTsb998>

11:51:55 From Sara Lafia (ESIP Fellow) to Everyone:

*We have 1 so far :)*

11:54:15 From Vanessa Chatman to Everyone:

*I'm starting off with: Forecasting carbon monoxide on a global scale for the ATom-1 aircraft mission: insights from airborne and satellite observations and modeling*

11:54:29 From cindy lin - ESIP Fellow to Everyone:

*Wow super specific*

11:54:57 From cindy lin - ESIP Fellow to Everyone:

*Mine is just peatland fires viirs for now - mainly because national agencies in indonesia are transiting from aqua and terra to using viirs wholly for fire detection*

11:59:19 From Sara Lafia (ESIP Fellow) to Everyone:

*My interest is a lot broader; I'm looking for datasets used in modeling applications. Finding some cool examples of wildfire modeling case studies that use VIIRS active fire products*

11:59:50 From Sara Lafia (ESIP Fellow) to Everyone:

*Go team! We're on FIRE*

12:06:03 From Sara Lafia (ESIP Fellow) to Everyone:

*DOI resolver in case you want to check validity <https://dx.doi.org/>*

12:07:08 From Vanessa Chatman to Everyone:

*Can satellite-based data substitute for surveyed data to predict the spatial probability of forest fire? A geostatistical approach to forest fire in the Republic of Korea — very interested in this topic!*

12:18:22 From Sara Lafia (ESIP Fellow) to Everyone:

*The fire team did great! We had over 20 contributions, woohoo!*

12:19:04 From Ed Armstrong to Everyone:

*How do we find each teams contribution number ?*

**2021 ESIP Summer Meeting (July 19th-23rd, 2021)**

<https://2021esipsummermeeting.sched.com/>

[Meeting Takeaways Doc](#)

Community Participation Guidelines & Reporting: Click [HERE](#)

12:20:45 From Vincent Inverso to Everyone:

18 for team flood

12:20:54 From Sara Lafia (ESIP Fellow) to Everyone:

@ed if you go into edit the Google survey form your team used, you can see the questions and responses

12:21:23 From Ed Armstrong to Everyone:

I forgot you can do that !

12:21:58 From Ed Armstrong to Everyone:

29 responses for team sea level

12:22:05 From Doug Newman to Everyone:

This ties for 'most fun session' with Graph yesterday

12:22:08 From Sara Lafia (ESIP Fellow) to Everyone:

You also get a view of contributions by ORCID, so we can find out who submitted the most entries :-)

12:22:21 From Sara Lafia (ESIP Fellow) to Everyone:

Wow, go team Sea Level! Prolific

12:23:32 From cindy lin - ESIP Fellow to Everyone:

I realized most of the articles about fires in Indonesia and Brazil are rarely linked to datasets that have DOI. These datasets are usually provincial, national, or local datasets regarding burn maps, geospatial maps etc. ORCID does open up to a social network of authors who have worked in the region though.

12:24:27 From Doug Newman to Everyone:

Space Weather effects on climate change articles will be very careful about citing the 'space weather' datasets but not the 'climate change' datasets. Where did this rainfall data come from?

12:29:00 From Ken Casey - NOAA/NCEI and ESIP Pres to Everyone:

NOAA now using DataCite with separate prefixes for archived data (10.25921), publications (10.25923), and GMD data (10.25925)

12:29:36 From Robert Downs to Everyone:

SEDAC uses 10.7927

12:29:52 From Madison Langseth to Everyone:

USGS uses 10.5066

12:30:16 From Chris Lynnes to Everyone:

<https://forms.gle/VEWZofZWfY7xwno6A>

12:30:20 From Nate James to Everyone:

NASA ESDIS 10.5067

12:30:37 From Doug Newman to Everyone:

I feel like I need to add some more 'offline'

Ken Casey - NOAA/NCEI and ESIP Pres to Everyone (12:30 PM)

Older NOAA DOI prefix was 10.7289

**2021 ESIP Summer Meeting** (July 19th-23rd, 2021)

<https://2021esipsummermeeting.sched.com/>

[Meeting Takeaways Doc](#)

Community Participation Guidelines & Reporting: Click [HERE](#)

*Ed Armstrong to Everyone (12:31 PM)*

*PO.DAAC doi is 10.5067*