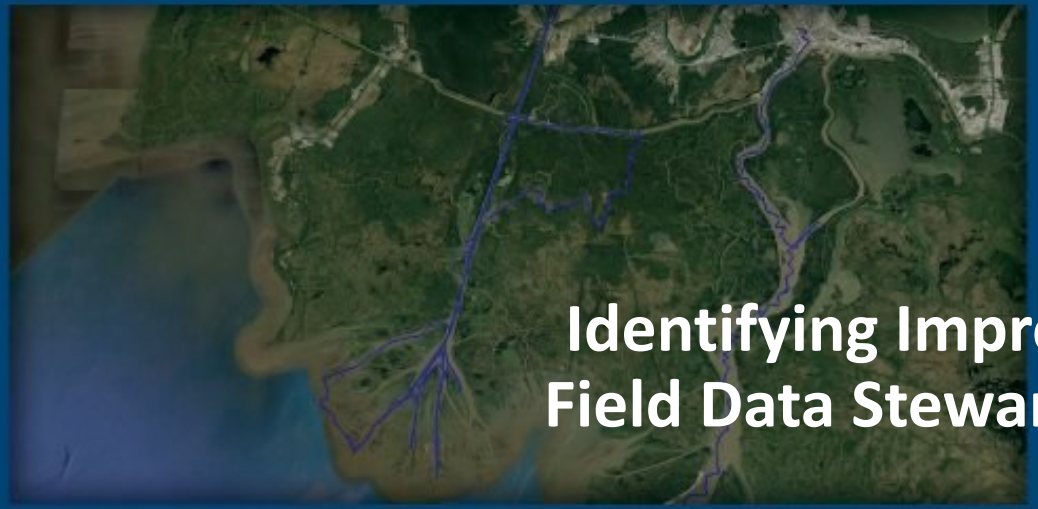


Identifying Improvements to Airborne and Field Data Stewardship from Data User and Producer Needs

Findings from the March 2022 Airborne and Field Data Workshop

Geoffrey Stano (GHRC) - Presenter

Deborah Smith (ADMG), Sara Lubkin (ESDIS), Megan Buzanowicz(ASDC), Leigh Sinclair(GHRC)

Workshop Purpose and Goals

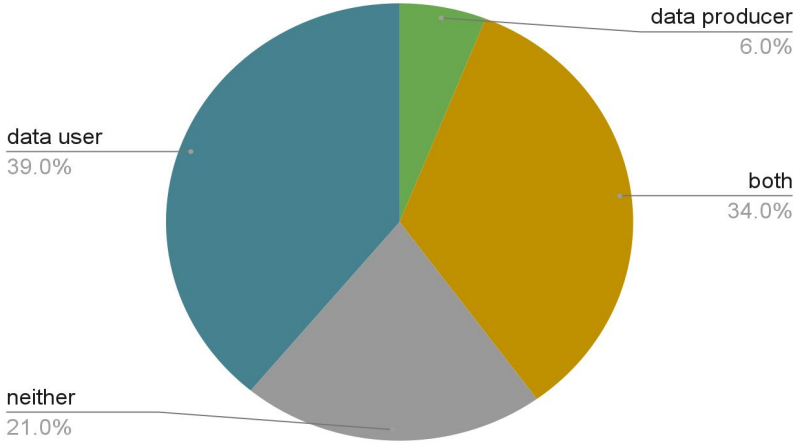
Workshop Purpose: to gather valuable feedback from data users *and* data producers in order to improve the acquisition and use of NASA's airborne and field data.

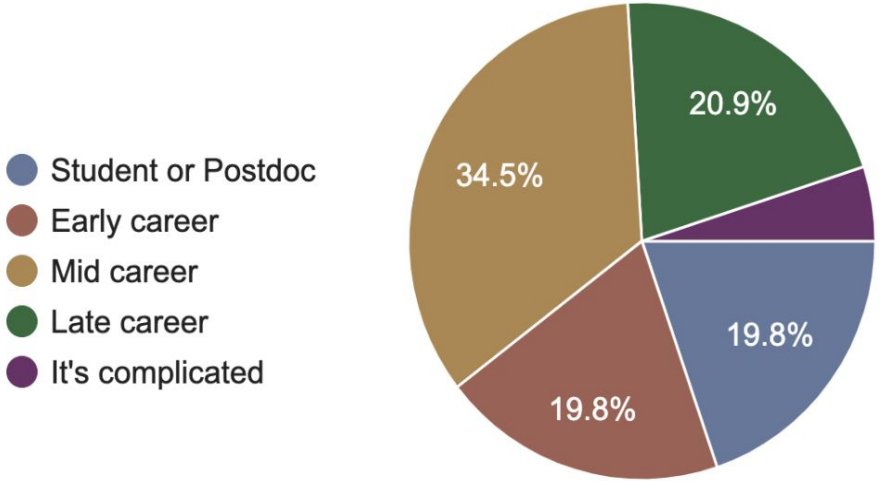
Workshop Goals:

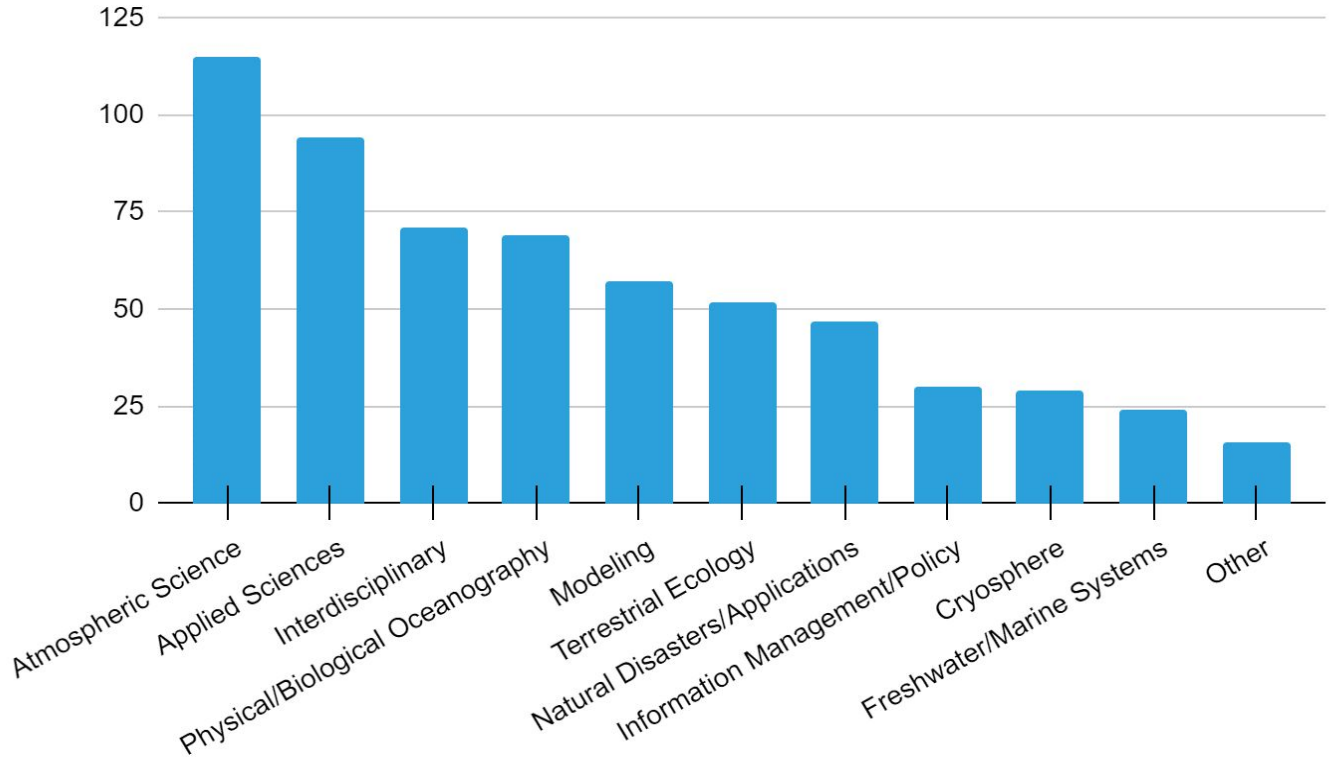
- Identify new **ways to improve the discovery, access, and reuse** of NASA's airborne and field data
- Intentionally include the airborne community in future planning to **ensure the community needs are met**
- Understand how NASA can better help researchers **realize the full value of airborne and field data**
- **Enable Open Source Science** and determine what this means for airborne and field data communities

Workshop Attendees

- 272 people registered for the workshop
- Between 50 and 100 attendees at any given time
- Maximum attendance occurred on Day 1



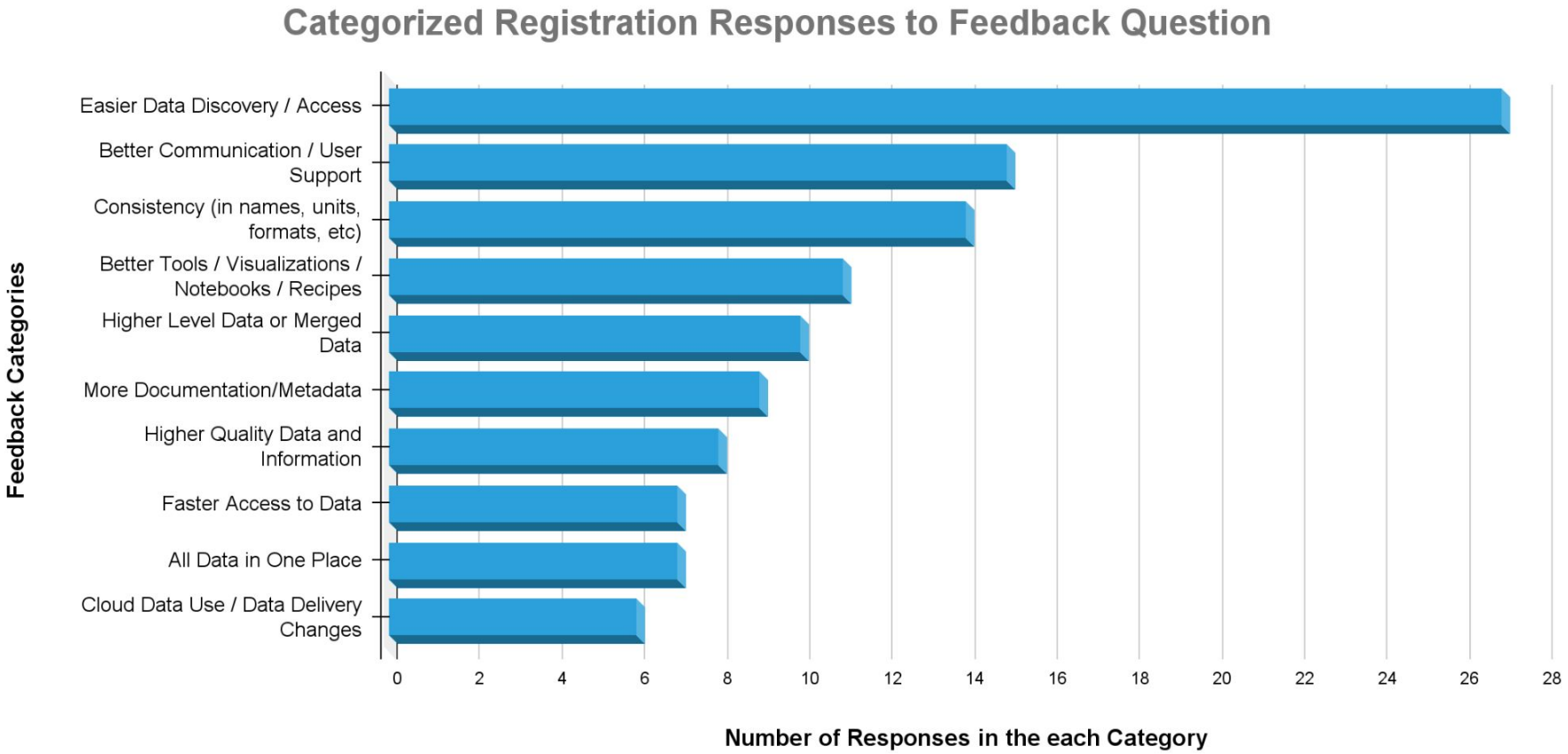




Self-Reported Primary Area of Interest

*Collected during workshop Registration

“What is your most important feedback about NASA airborne and field data discovery, access, and/or use?”



Day 1: Data User Focus

Session 1: **The airborne *data user* community shares their experience** (Moderator: Bruce Wilson)

- **Invited Data User Talks:** (*Phil Townsend, Qing Liang, Mark Tschudi, Timothy Lang, Kenton Ross*)
- **Moderated User Panel:** (Panel Members: *Tim Bailey, Owen Cooper , Rebecca Hornbrook, Liz Hoy, Fred Huemmrich, Joe Munchak, Sean Serbin, Alexey Shiklomanov,* Moderator: Michele Thornton)
- **Breakout Groups Discussion**: *What is your data user experience?*

Session 2: **Resources for airborne data users** (Moderator: Bruce Wilson)

- Airborne Data in the Cloud Presentation (*Bruce Wilson*)
- NASA Tools and Resources for Users of Airborne/Field Data (Moderator: Deborah Smith)
  - CASEI - *Stephanie Wingo, ADMG*
  - SOOT - *Jennifer Tindell, ASDC*
  - ADV and SDAT - *Michele Thornton, ORNL DAAC*
  - FCX - *Geoffrey Stano, GHRC*
- **Audience Feedback Activity**: *What tools do you need to do your work?*

Day 2: Data Producer (Scientist) Focus

Summary of Day 1: Key Takeaways

Breakout Discussions - *What did we learn from Day 1: Data Users?*

Session 3: **The airborne *data producer* community shares their experience**

Flash Talks:

- Emma Yates (AJAX): The Unusual Case of AJAX
- Daniel Jensen (Delta-X/AVIRIS): AVIRIS-NG Data in the Delta-X Campaign
- Fred Bingham (S-MODE):
- Stacy Brodzick (IMPACTS):
- Gao Chen (Atmospheric Chemistry): Suborbital Science Data for Atmospheric Composition

Fishbowl Panel: Moderator: Kasey Phillips

Session 4: **Working together to make data accessible**

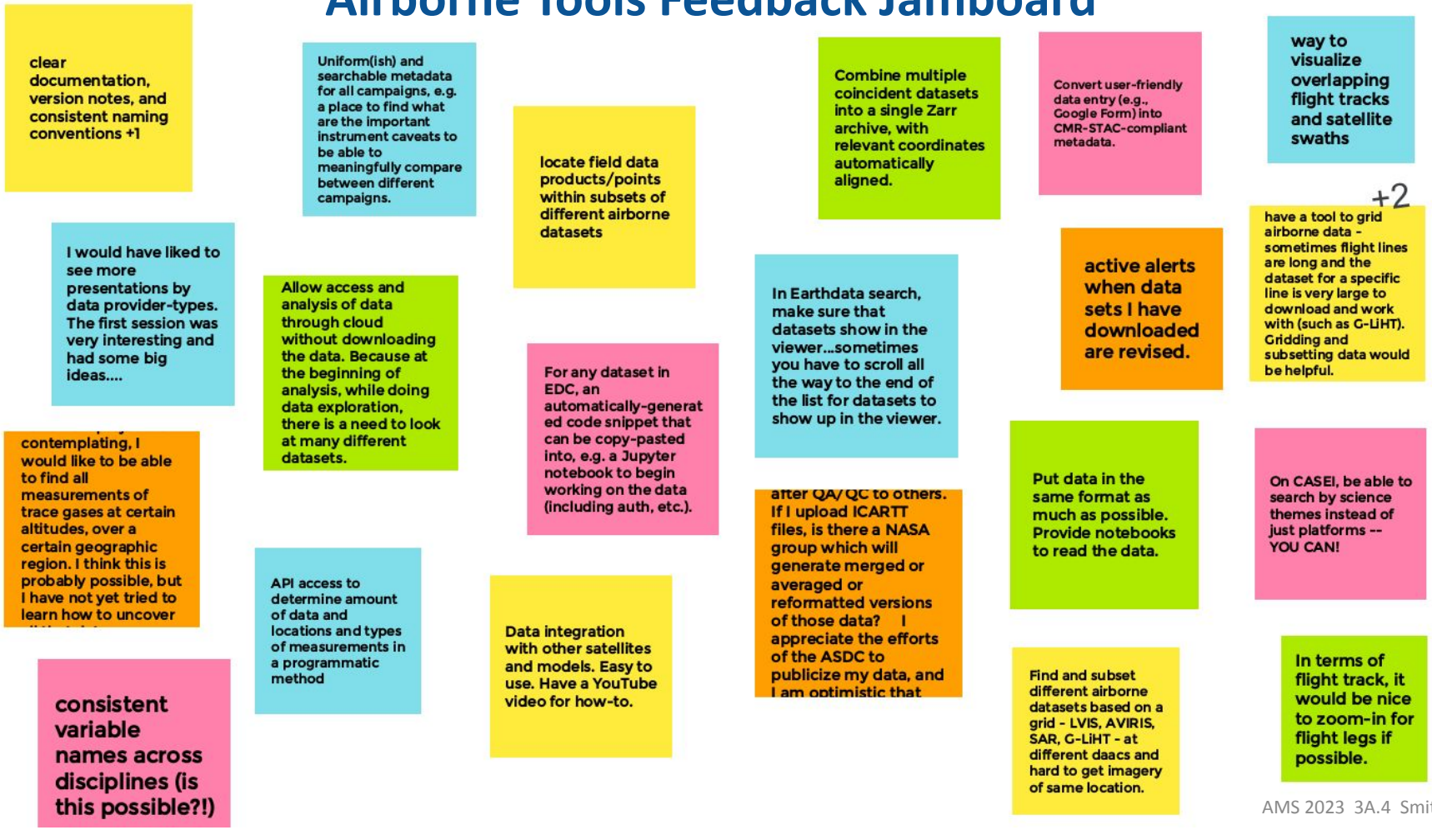
- Enhancing Communication Throughout the Project Lifecycle - Amanda Leon and Deborah Smith
- DAAC Processes and Responsibilities / Data Producer Responsibilities - Bruce Wilson
- Breakout Room Discussions
- Role of Standards in Airborne Data - Siri Jodha Khalsa and Joseph Koch
- NASA's Open Source Science Initiative - Elena Steponaitis
- Open Discussion (Moderator: Ge Peng)

Closeout: Looking to the future - Deborah Smith and Sara Lubkin

Collecting Workshop Feedback

- Registration Questions
- Breakout Room Discussions
- Data User and Data Producer presentations
- Meeting chats
- Panel Discussions
- Slido Polls
- Jamboard Activities

Airborne Tools Feedback Jamboard



Feedback Categories

Data and information access	Communication	Consistency	Merged/higher level data
Metadata	OSS	Cloud	Programmatic change
Quality	Search and discovery	Support	Tools

Next: **Developed User Stories**

- Shared user stories with Summer 2022 ESIP Session attendees
  - Asked to prioritize user stories [must have, need, want, etc]
  - Estimated effort and cost needed to implement potential solution ideas for all must have stories
- Organizing committee continued this work in subsequent monthly meetings until all 60 stories were categorized, prioritized, and solutions generated and estimated for cost/effort

Next: **Evaluated status of potential solutions and identified report recommendations**

Add one issue or concern per line written as a user story					high priority item to address		start here
Priority (must have, need, want, not imp)	**Effort (low, high)**	**Cost (low, high)**	**Category Short Title**	**User Story (as a user I want to....)**	**Data User / Data Producer / DAAC / Developer / All**	**Status**	**Ideas for How to Improve**
Must	medium	medium	Access	As a power data user, I want to easily download bulk data.	Data User	parts exist	"one-stop shopping", unify bulk download soln, cloud solution?, instrument scientist; does every DAAC have API access, pieces already in existence
Must	medium	medium	Access	As a data user, I want to be able to access data grouped by campaign or geographic area or some other variable by accessing virtual collections	Data User	in progress	virtual collections, CMR effort with variables
Must	medium	medium	earch / Discove	As a data user who is used to Earthdata search, I want more airborne data in EDS because I don't know which DAAC to look at for the data products I need.	Data User	in progress	Tag as airborne - as filter option
Must	medium	medium	Access	As a data user I want to get all the instrument, platform, or campaign data accessible from one place.	Data User	parts exist	Everything in Earthdata - filter by instrument, place, campaign. Can do in CASEI; comprehensive landing pages; collection of collections DOIs; instrument microarticles - need consistency among DAACs because each do things in their own way
Need	high	high	Access	As a data user I want to combine multiple coincident datasets into a single Zarr archive, with relevant coordinates automatically aligned.	Data User	not exist	
Need	high	high	Access	As a data user I want to to get all the instrument, platform or campaign data accessible from one place	Data User	in progress	CASEI is built to solve this for discovery/search, but not download
Need	low	low	Access	As a data user I want to have a Jupyter notebook to help me use some particular airborne or field data	Data User	in progress	already available for some tutorials at some DAACs
Need	medium	low	Support	As a data producer, I want to understand and have guidance on providing standardized data formats that meets the protocol for incorporation into community tool resources.	Data Producer	parts exist	Include airborne in DPDG; some DAACs better than others at working with DPs; early involvement of DAACs will help.
Need	high	low	Support	As a tool developer, I want to understand the common data formats/standards in which airborne data are provided in order to create software products for the user community.	Developer	parts exist	DPDG is a starting point. What in DPDG is useful for developers - could there be a quick start guide for developers; ADMG knowledge center - pointers to other existing information
Need	low	low	Metadata	As a data user, I want to identify and work with all data that is part of a campaign, which spans multiple airborne platforms over discontiguous space and time/instrument used in multiple campaigns. I need metadata about these different flights/campaigns that will allow me to search by "virtual collection" or "group"	Data User	in progress	Virtual collections in progress in CMR; DOI policy done
Need	medium	medium	Metadata	As a data user who wants to search by instrument/campaign, I need metadata that is more consistent, uniform(ish) and searchable for all campaigns and instruments	Data User	in progress	Earthdata Pub/DPDG may help for new data; more effort for really old data - need to weigh value of updating old metdata; low effort going forward; high effort going back; ARC - assessments across DAACs; learning resources - "walk in user shoes"
Need	medium	medium	Consistency	As a data user or data producer, I want consistency in file names, units, file formats and variable names	All	parts exist	DPDG resource; easier to do going forward that going back; different approach across DAACs - can it be unified? Would best practice resources be helpful?
Need	low	low	Consistency	As a data user who obtains data from more than one DAAC and Earthdata, I want more consistency between archives and all Earthdata sites (e.g. download procedures for bulk download also vary between DAACs)	Data User	in progress	Earthdata cloud through Earthdata Search will help - also bulk download; currently good bulk downloaders DAAC specific
Need	high	high	Tools	As a data user, I want to easily find and subset different airborne instrument data products and have the data from these products available on the same grid to make it easier to get imagery from the same Earth location.	Data User	parts exist	Parts exist at different DAACs, but need to coordinate across DAACs - role for Harmony/GIBS

Immediate changes
(high return, high need, low cost)

- Early DAAC assignment
- Airborne and Field Resource Center on Earthdata
- Improved DAAC Campaign or Instrument Information
- More consistency in units, file formats and variable names, terms used in communications

Recommendations

New Tools and Services

- Ensure airborne and field data can be found via Earthdata Search
- Provide ability to search for multiple variables
  - In situ air samples (i.e targeted compounds)
  - Air column (i.e. total column)
  - Remote sensing of surface - (location, bands, altitude)
  - Ocean surface or ocean subsurface
- Utilize campaign catalog/flight catalogs across all DAACs and gather contextual information for campaigns
- Provide ability to subset data based on location, date, altitude, instrument, platform, flight
- Expand/improve existing tools such as Field Campaign Explorer (FCX)
- Enable data subscriptions to update users when new data available

Recommendations

Program Changes
(high return, low cost, culture change)

- Develop uniform and consistent practices across DAACs for documentation and user guides
- Plan another workshop in 2 years to assess changes
- Provide a means to share Science Team developed code

Recommendations

Thank you!

Workshop Information

https://www.earthdata.nasa.gov/esds/impact/admg/nasa-airborne-and-field-data-workshop

We are happy to share details of the findings from the workshop.

Final workshop report to be released in February.

Contact for more information:

Sara Lubkin sara.h.lubkin@nasa.gov (ESDIS)
Deborah Smith deborah.smith@uah.edu (ADMG)