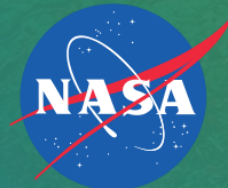


Construction of an Airborne Data Inventory for Improved Data Discoverability and Access

Deborah K. Smith, Stephanie M. Wingo, Carson R. Davis,
Kaylin Bugbee, Rahul Ramachandran, Brian Freitag

ADMG / IMPACT / UAH / NASA MSFC



What is ADMG?

The **Airborne Data Management Group (ADMG)** was officially established in September 2018 as part of IMPACT - the Inter Agency Implementation and Advanced Concepts Team at NASA Marshall Space Flight Center in Huntsville, AL

ADMG's Primary Role is to **support data producers and DAACs** in making sure that NASA airborne science data are discoverable and usable by the broader research community

- Serve as primary **point-of-contact and resource** for airborne information for data producers, DAACs, ESDIS, project teams, applied users, and the research community
- Promote a **consistent data producer experience** for data producers and users needing to interact with multiple DAACs
- Identify and **help solve issues** with NASA airborne data transfer, publication, discovery, and distribution, suggesting best practices for improved data stewardship
- Locate **historical data** to be systematically archived by DAACs
- Construct an agency-wide **Airborne Data Inventory** containing lots of campaign, flight, aircraft, instrument and data product metadata

- ADMG works to address airborne data problems
- Taking full assessment of NASA Airborne Earth Science data
- Adding important metadata for easier and faster data discovery

NASA Earth Science Airborne Data Inventory

Dictionary

inventory



in·ven·to·ry

/ˈɪnvənˌtɔɪəri/

noun

noun: **inventory**; plural noun: **inventories**

a complete list of items such as property, goods in stock, or the contents of a building.

Similar: [list](#) [listing](#) [catalog](#) [directory](#) [record](#) [register](#) [checklist](#)

• NORTH AMERICAN

a quantity of goods held in stock.

"in our warehouse you'll find a large inventory of new and used bicycles"

• (in accounting) the entire stock of a business, including materials, components, work in progress, and finished products.

verb

verb: **inventory**; 3rd person present: **inventories**; past tense: **inventoried**; past participle: **inventoried**; gerund or present participle: **inventorying**

make a complete list of.

• enter in a list.

"about forty possible sites were inventoried"

Similar: [list](#) [catalog](#) [record](#) [register](#) [make a list of](#) [file](#) [log](#)

Origin

LATIN

invenire
come
upon

LATE LATIN

inventarium
a list of what
is found

MEDIEVAL LATIN

inventorium

inventory

late Middle English

late Middle English: from medieval Latin *inventorium*, alteration of late Latin *inventarium*, literally 'a list of what is found', from Latin *invenire* 'come upon'.



Photo by [Julia Joppien](#) on [Unsplash](#)

The ADMG Inventory Process

Inventory Goals:

- To identify and **locate all** NASA airborne Earth science data
- To **identify existing metadata**
- To **enhance metadata** for improved information and data discovery, more complex queries
- To **identify issues** in airborne data archival and discovery
- To make searching for airborne data **more intuitive** and complete no matter where data are archived
- To **improve and speed data access**



Identify all airborne data at DAACs

Data that are discoverable in CMR do not represent all DAAC data



Identify Data at other NASA Repositories

Access any data external to DAACs that require publication



Google Scholar Searches

Publications mention airborne field campaign data that exists somewhere



NASA Ames Airborne Science

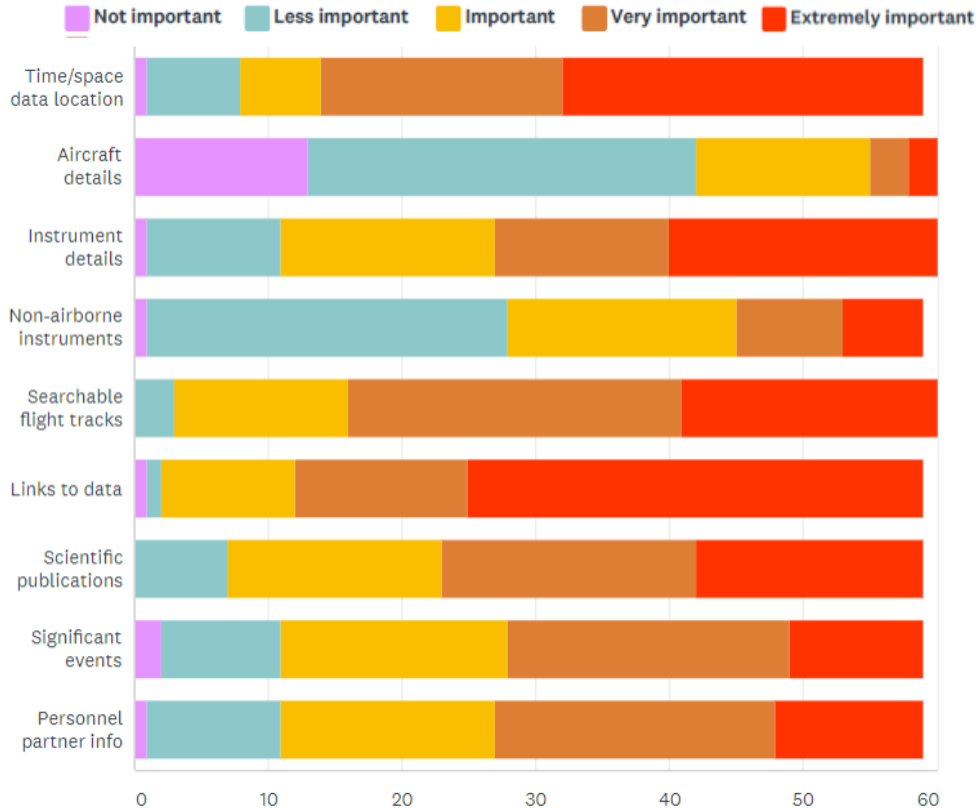
Work with Airborne science program to find what Earth science to publish



NASA LaRC Atmospheric Chemistry

Data from many atmospheric chemistry airborne campaigns are at LaRC

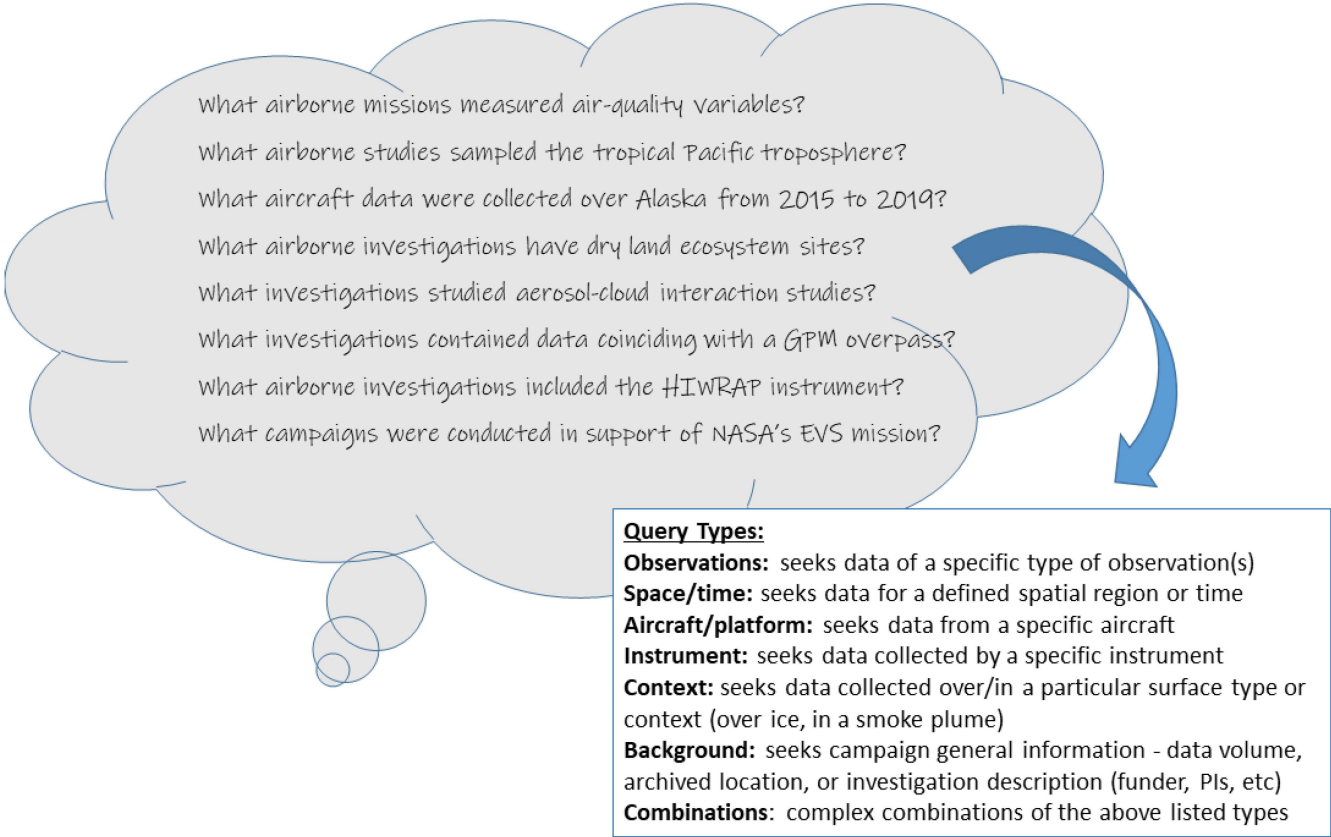
Prospective User Survey to Assess Needs



Features Prioritization:

- 1 - Links to data
- 2 - Time/space data location
- 3 - Searchable flight tracks
- 4 - Scientific publications
- 5 - Instrument details
- 6 - Personnel & partner info
- 7 - Significant events
- 8 - Non-airborne instruments
- 9 - Aircraft details

Prospective User Survey to Assess Needs



What airborne missions measured air-quality variables?
What airborne studies sampled the tropical Pacific troposphere?
What aircraft data were collected over Alaska from 2015 to 2019?
What airborne investigations have dry land ecosystem sites?
What investigations studied aerosol-cloud interaction studies?
What investigations contained data coinciding with a GPM overpass?
What airborne investigations included the HIWRAP instrument?
What campaigns were conducted in support of NASA's EVS mission?

Query Types:

Observations: seeks data of a specific type of observation(s)

Space/time: seeks data for a defined spatial region or time

Aircraft/platform: seeks data from a specific aircraft

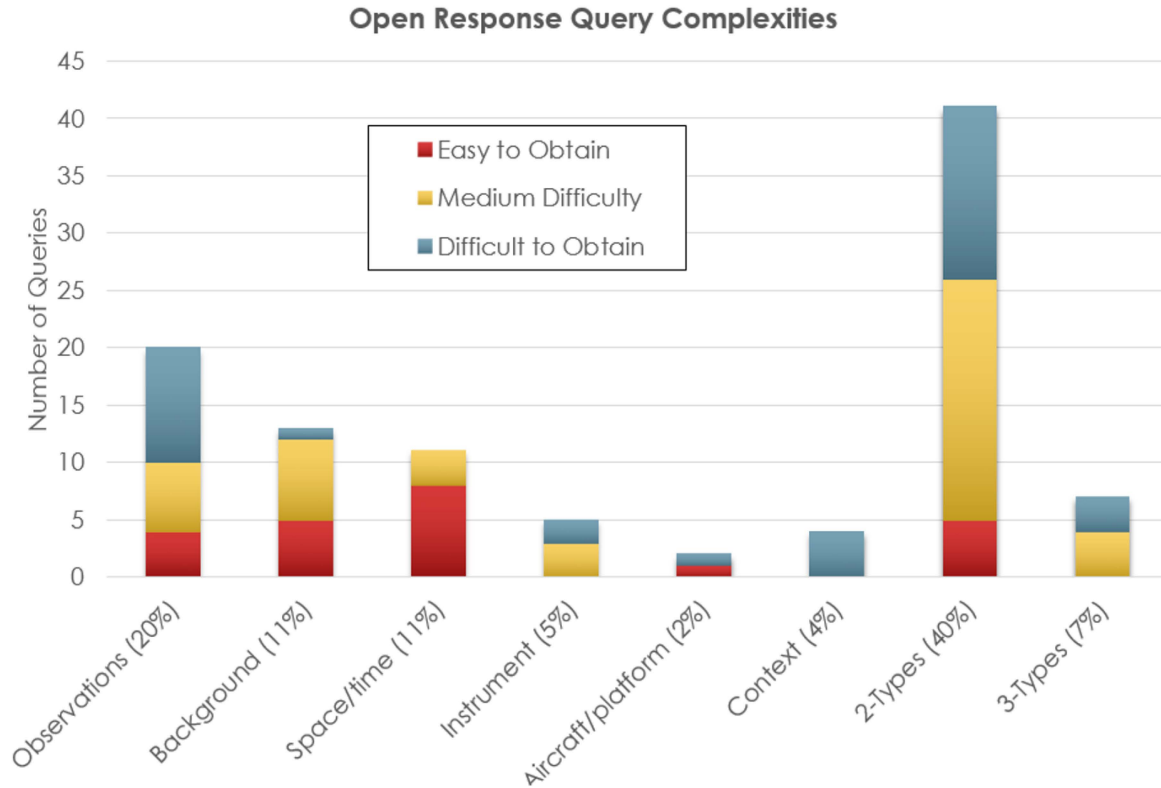
Instrument: seeks data collected by a specific instrument

Context: seeks data collected over/in a particular surface type or context (over ice, in a smoke plume)

Background: seeks campaign general information - data volume, archived location, or investigation description (funder, PIs, etc)

Combinations: complex combinations of the above listed types

Prospective User Survey to Assess Needs



Importance of Meeting User Needs

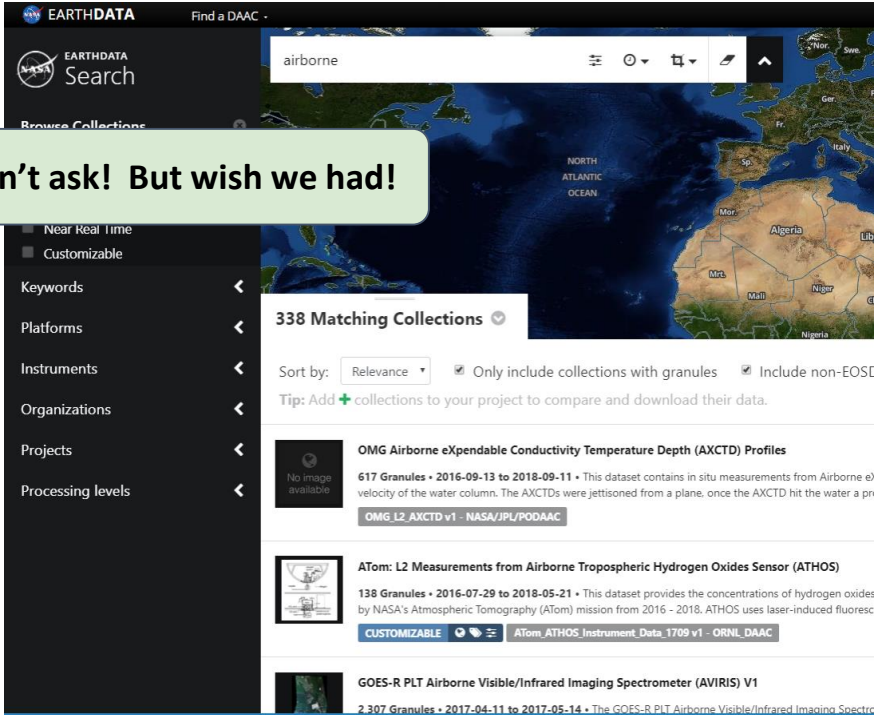
Are the requested queries already possible?

- ~60% of identified user queries are already possible using Earthdata Search

So why is NASA Earthdata Search not used?

- Interface:
 - Scientists interested in airborne data think in queries, want **contextualized responses**
- Vocabulary disconnects:
 - “Data analysis” != data quality;
 - “Data analysis” = scientific process leading to a result
- Adding additional metadata, will allow for more queries, a total of 80% possible
- The remaining 20% of queries would require additional metadata and information that is too costly (in time or effort) to add

We didn't ask! But wish we had!



The screenshot shows the NASA Earthdata Search interface. The search bar contains the word "airborne". The left sidebar lists various filters: Near Real Time, Customizable, Keywords, Platforms, Instruments, Organizations, Projects, and Processing levels. The main content area displays "338 Matching Collections". The first result is "OMG Airborne eXpendable Conductivity Temperature Depth (AXCTD) Profiles" with 617 granules from 2016-09-13 to 2018-09-11. The second result is "ATom: L2 Measurements from Airborne Tropospheric Hydrogen Oxides Sensor (ATHOS)" with 138 granules from 2016-07-29 to 2018-05-21. The third result is "GOES-R PLT Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) V1" with 2,307 granules from 2017-04-11 to 2017-05-14. A map of the North Atlantic Ocean is visible in the background.

Clear Priorities for Inventory Content & Functionality

Major Survey Takeaways:

- Help airborne scientists access information and data quickly - without frustration
- Utilize **user input during inventory interface development** to build what scientists will use
- Invest time in **curating existing metadata** to improve quality and science relevance
- **Add additional metadata** beyond what is currently available to allow for complex questions and searches

Airborne Inventory Complexity

The NASA Earth Science Airborne Inventory

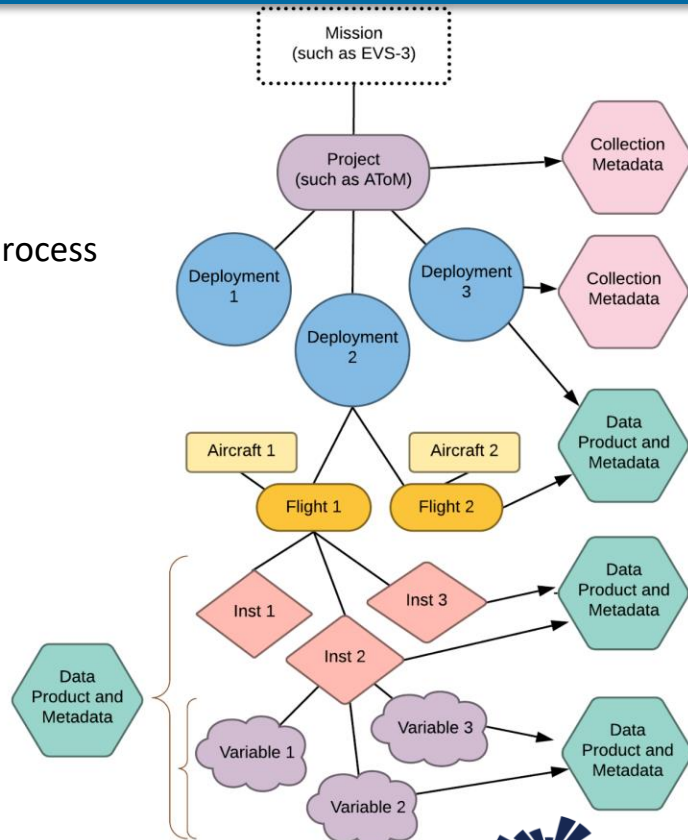
A full list of

- Archived NASA airborne Earth science data
- Existing campaign data needing archival - work the process
- Identification and archival of “Lost Data”

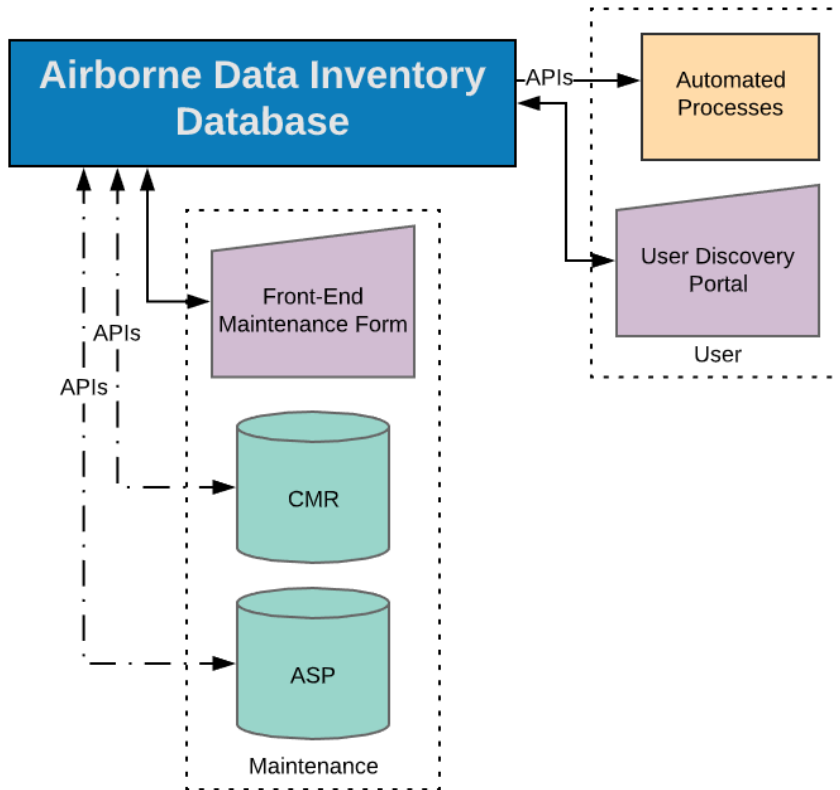
Information organized in a relational database containing

- Additional metadata adding context
- Aircraft data for every campaign
- Flight tracks
- Data product DOIs and links
- Instrument information
- Deployment details
- Significant events
- MORE!

To make NASA airborne data **Findable** and **Accessible**!



Relational Database Needed



The inventory utilizes existing Common Metadata Repository (CMR) metadata and other information from Airborne Science Program (ASP) via APIs.

Maintenance Interface - curation

- adding new information
- performing routine checks for errors or missing data/ broken links
- access inventory user statistics

Data Discovery Portal - public user interface

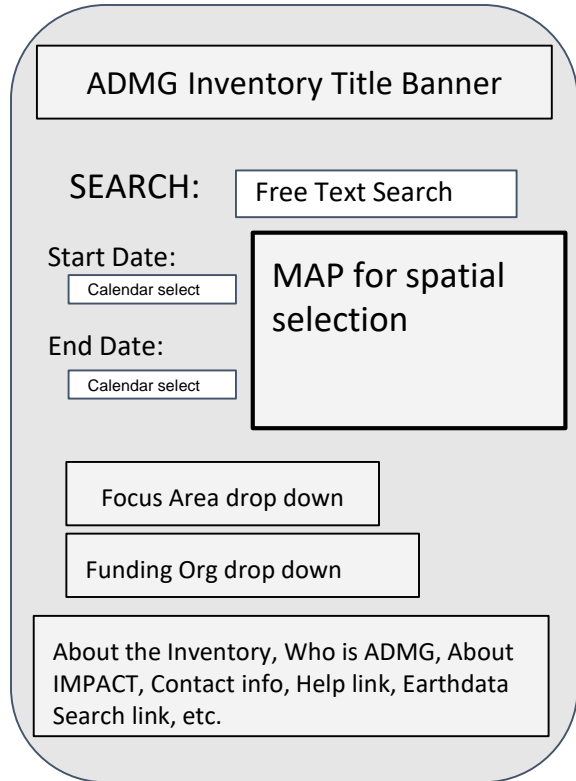
- submitting queries
- obtaining information
- accessing data
- finding campaign details

Setting User Interface Functional Requirements

Users need to:

- Easily **locate information** by campaign, instrument, measurements, data products (DOIs), aircraft, search regions, time frames and other relevant related data
- Access data and campaign/investigation resources in as **few steps** as possible (clicks, programming steps)
- **Access data** via Earthdata Search URLs no matter where the data are stored
- **View images** such as deployment flight tracks, campaign study regions, instrument ground sites
- **Provide feedback** on inventory contents, noting inaccuracies or asking ADMG for more details

Data Discovery Portal - User Interface



Data Discovery Portal - Inventory Homepage

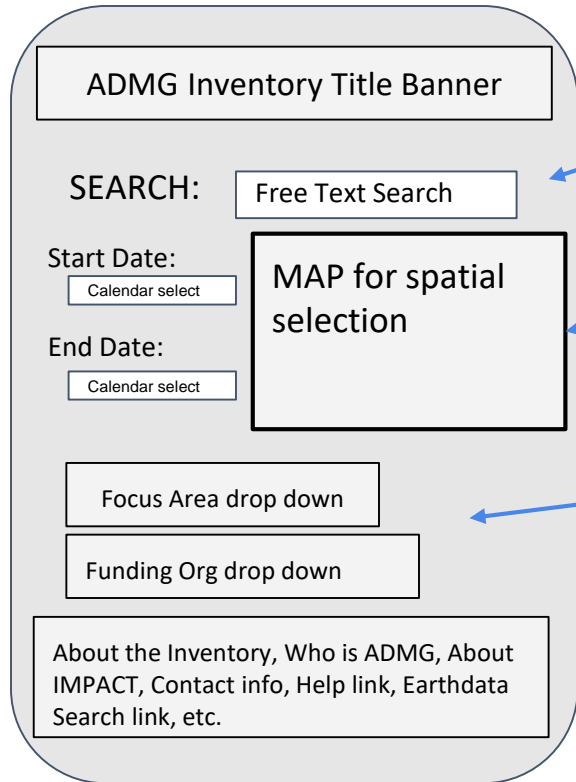
- Overview information
- Temporal and Spatial Search
- Footer links:
 - About ADMG, IMPACT, ESDIS
 - Link to Earthdata
 - **Contact / Feedback**
- Lists of useful **airborne data tools**
 - Descriptive list with links

No login credential required

- **open to public** for reading

Data Discovery Portal Homepage - simplified wireframe concept of potential interface

Data Discovery Portal - User Interface



- Free text search:
 - User-provided string
 - Applied to metadata elements at each information level
- Date search
- Map search:
 - User-drawn area
- Closed list selection:
 - Campaigns
 - NASA Earth Science Focus Areas
 - Season
 - Aircraft
 - Instrument Types
 - Funding Organization
 - Measurement Types
 - Vertical Measurement Region
 - Ground type

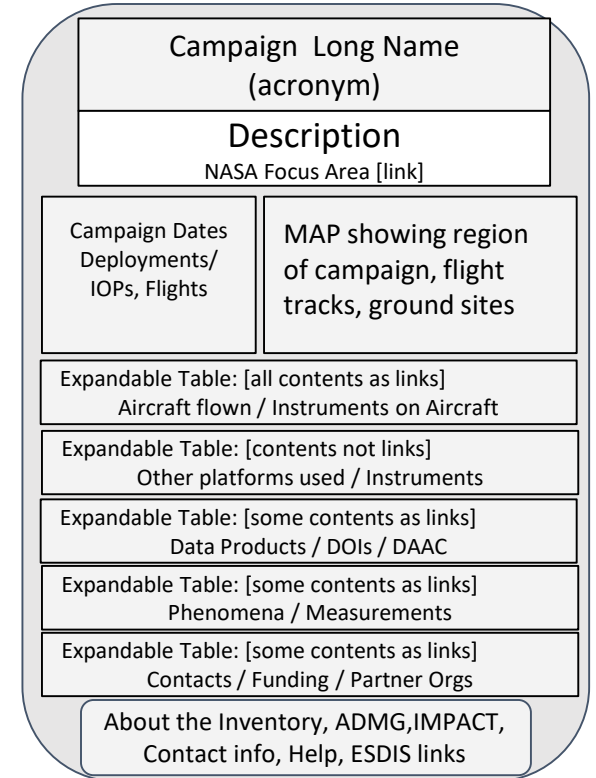
Data Discovery Portal Homepage - simplified wireframe concept of potential interface

Data Discovery Portal - User Interface

User enters

- Campaign name → campaign landing page
- Instrument name → instrument landing page
- Aircraft name → aircraft landing page
- Measurement → measurement landing page

- Collapsible frames for each information category
- Highly linked information
- Figures for study region, flight tracks, ground sites
- Campaign description



Data Discovery Portal - Campaign Landing Page (simplified wireframe concept)

Development Progress

- ADMG is now collecting and curating metadata for over 100 previous NASA airborne campaigns
- Development Phases:
 - Phase 1: Relational database, API development - *in process*
 - Phase 2: Maintenance Interface, Limited public interface: fixed selection, map area search
 - Phase 3: Full Data Portal interface, broad public release: Free text search
- Deliver user documentation, introductory webinars and videos
- We are ***looking for early interface testers...*** ***Please let us know if you would like to participate!***

Summary

- Full accounting of NASA Airborne Earth Science data is in progress
- Additional metadata and information is curated in order to improve data discovery and access
- Limited inventory available in late 2020

- Looking for interested scientists to test User Interface
- Contact: Stephanie.M.Wingo@nasa.gov