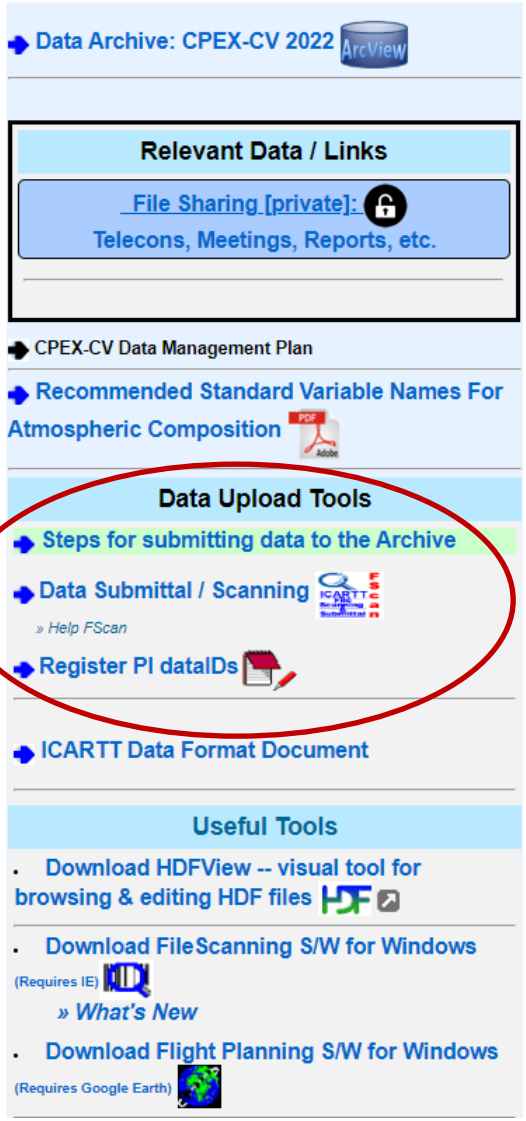
Data Submission Process

- DataID Registration:
  - PI will need to first register dataID(s) before submitting data
  - dataID is the first part of the filename and is used to organize PI files on the data repository
  - Instrument acronym and a brief top level data description required for each dataID
  - The repository website is now open for dataID registration
- Data Submission:
  - File submission is through a scanning tool for checking filenames and ICARTT file integrity:
    - File header: structure and keywords
    - Data Flags for missing data, LOD codes
    - Time Stamps: monotonically increase with no overlaps
  - Publication quality NetCDF and HDF files checked offline for structure as well as global and variable attributes
  - Support zipped multi-file upload and Script-based batch upload available
- Username/password: contact science team lead



DataID Registration Page (Example)

PI Last Name : YANG

PI First Name : MELISSA

Platform (LocationID) : NASA P3-B Aircraft (P3B)

	dataID: (max 45 chars) *Prefix with "camp2ex-" e.g., camp2ex-CO2*	**Data Description:** (max 380 chars) *Describe your measurments; e.g.,* Carbon Dioxide Mixing Ratio	**Instrument(s):** (max 190 chars) *List Instruments; e.g.,* LI-COR 6252
Reset	camp2ex-hsking	Aircraft Navigation Aircraft Attitude	GPS INS

+ Add Another dataID - Remove Last Entry

OVERWRITE my previous record (i.e., ALL previously registered dataIDs for this mission will be removed).

To link instruments with data products, separated by newline

Link to PI webpage, instrument, or experiment description document:

http:// *Optional: to display on LaRC Archive webpage*

Text describing PI experiment or measurements (e.g., NASA LaRC DIAL - Troposphere O3, Aerosols, and Clouds Profiles):

Aircraft Navigational and Attitude Parameters, atmospheric met parameters *Optional: to display on LaRC Archive webpage*

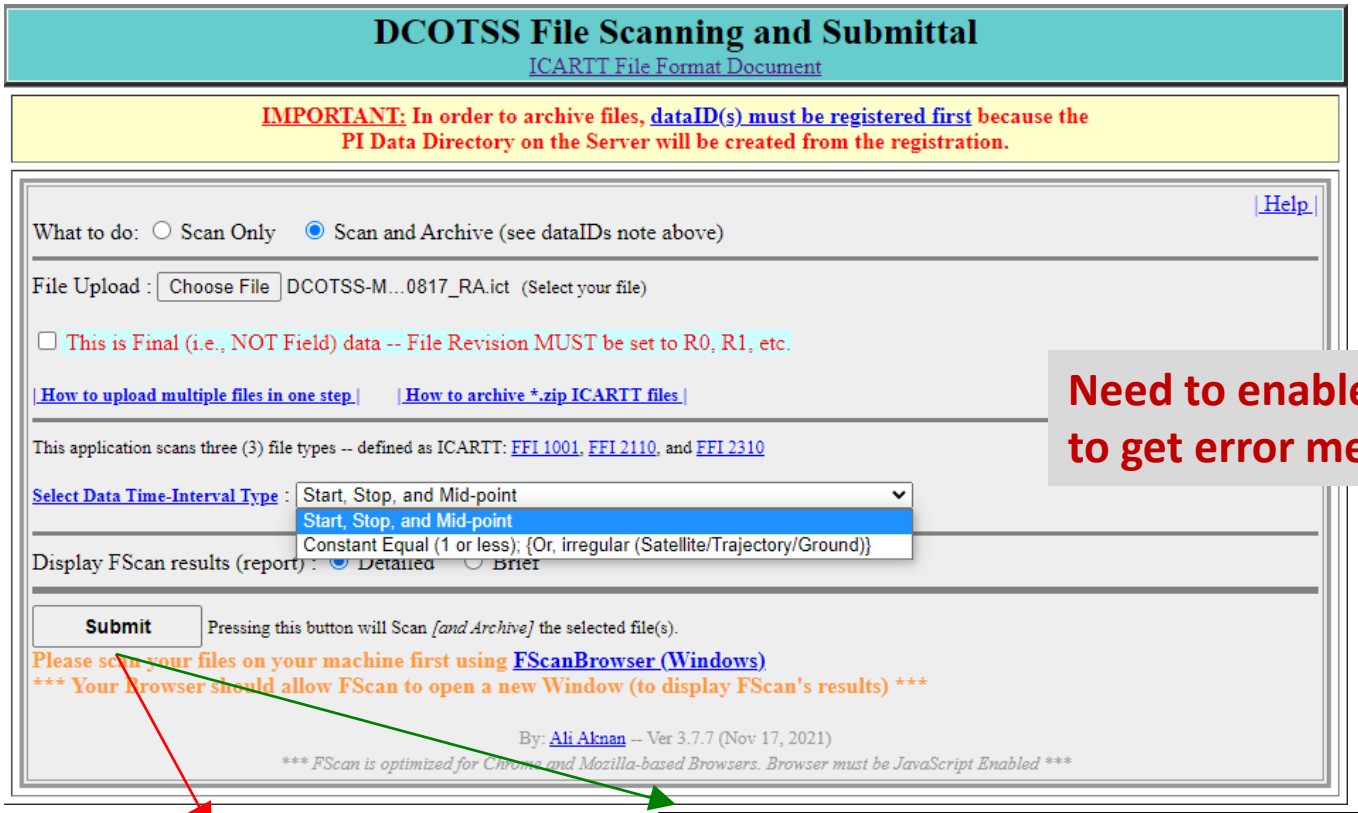
Submit

Current Registered dataIDs on the Server for CAMP2EX 2019

PI Name: Last.First	LocationID	Registered dataIDs
Edit AKNAN.ALI (PI Link)	P3B	camp2ex-test1;camp2ex-TEST3 + Show Description
Edit YANG.MELISSA	P3B	camp2ex-hsking + Show Description
Edit AKNAN.ALI	MODEL	camp2ex-test1;camp2ex-TEST3;camp2ex-TEST4 + Show Description

Click on "Refresh" to retrieve the latest list.

FScan Page (Example)



Need to enable Pop-Ups to get error messages

*** Scanning Failed. See Error Messages Below ***

Please scroll down to the end of this page to view the results.

NOTE: error messages are displayed in RED; warnings and other info in ORANGE

Function Selected: Scan and Archive
FScan Results Report: Detailed
Time-interval Type Selected: Start, Stop, and Mid-point
Filename Submitted: DCOTSS-MMS-1HZ_ER2_20210817_RA.ict
File Size: 3066538 bytes
File Content: application/octet-stream

Receiving file was successful - now onto scanning ...

✓ ALL DONE.

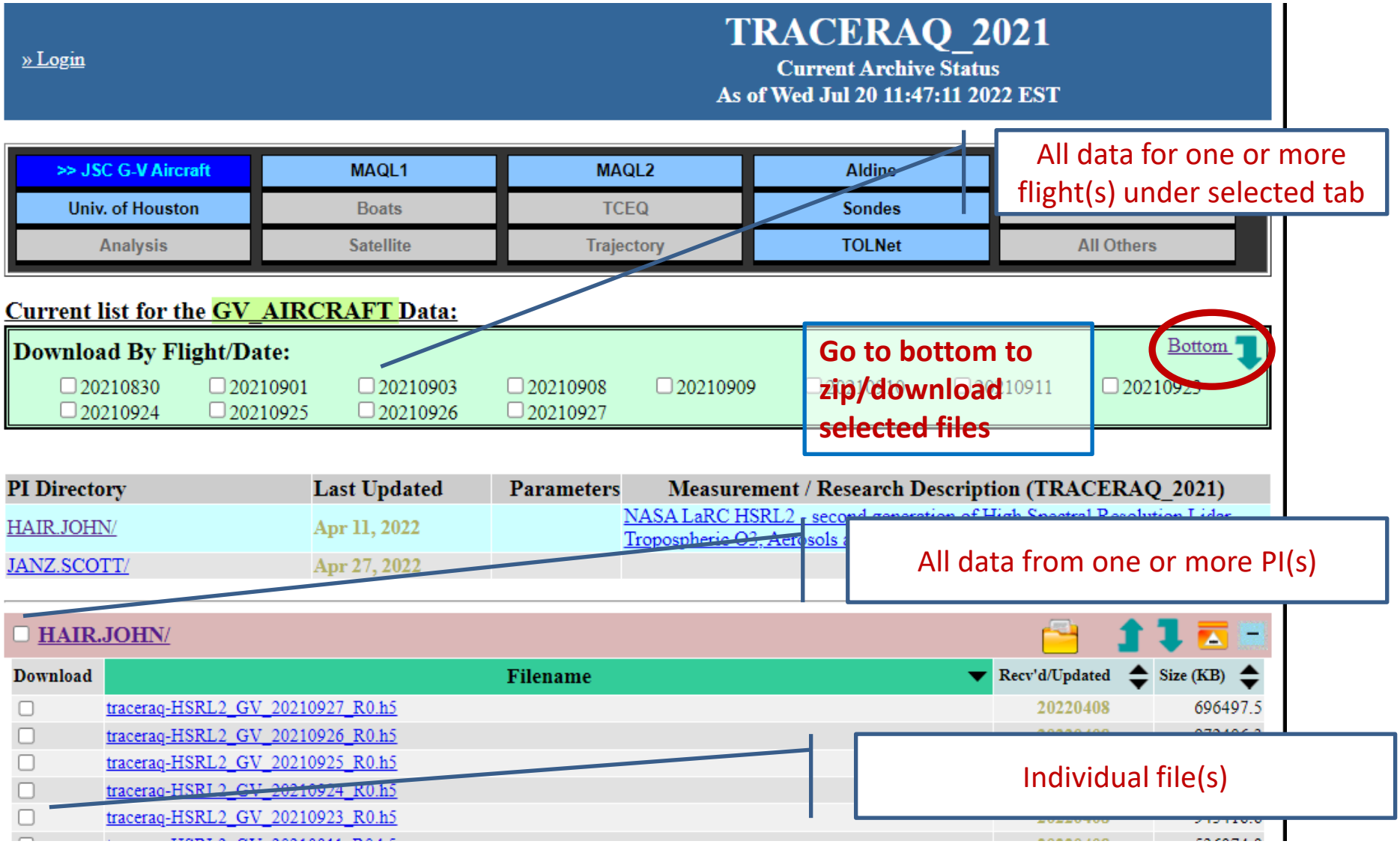
Please scroll down to the end of this page to view the results.

NOTE: error messages are displayed in RED; warnings and other info in ORANGE.

Function Selected: Scan and Archive
FScan Results Report: Detailed
Time-interval Type Selected: Start, Stop, and Mid-point
Filename Submitted: ACTIVATE-LARGE-SMPS_HU25_20220111_RA_L2.ict
File Size: 55987 bytes
File Content: application/octet-stream

Receiving file was successful - now onto scanning ...

Data Download (Example)



- **Login required for preliminary/field data**
- **Web download size limit: 2 GB; batch download script available upon request**

Naming Convention for ALL CPEX-CV Data Files

DataID_LocationID_YYYYMMDD_R#_Description.extension

- DataID: a short description of measured parameter/species, instrument, or model prefixed by “CPEXCV-”
- LocationID: an identifier of measurement platform provided in a drop-down box: “DC8, MERGE, MODEL, SATELLITE, SONDE, ANALYSIS, TRAJECTORY, and OTHER”
- YYYYMMDD: UTC date of takeoff for flight data or the beginning of the measurement for ground sites. Note: CVT = UTC – 1
- R#: Revision identifier. Typically RA, RB, RC, … for field data and R0, R1, R2, … for the publication quality data. Note: archived files cannot be overwritten, *only replaced with subsequent revisions*
- Description: optional additional description of the file if necessary
- Extension: “ict” for ICARTT files, and “h5” for HDF 5 files, “nc” for netCDF files, etc.
- The underscore, “_”, is used ONLY to separate the different fields of the filename
- Examples: the filename for CPEX-CV LARGE CN concentration measurement made on August 29, 2022 flight may be:
  - CPEXCV-LARGE-CPC_DC8_20220829_RA.ict (for field data)
  - CPEXCV-LARGE-CPC_WB57_20220829_R0.ict (for publication quality data)
  - CPEXCV-LARGE-CPC_WB57_20220829_R0_thru20220902-README.pdf (for readme file)

Making CPEX-CV Data F.A.I.R. (I)
ICARTT files

- Keep the same number of variables and variable names for same dataID
  - Needed to facilitate online merge tool as the online merge tool UI is based on the latest submission
  - Required for publication quality files
- Use fixed variable name(s) for Time Stamps, i.e., Time_Start, Time_Stop, and Time_Mid
- No space or special characters in variable names
- Indicate if the data is synched with the sampling time standard determined by the science team
- Align beginning and end of 1 second files to the nav file
- Variable standard names are required for publication quality data .ict files
- Use standardized units: http://codes.wmo.int/wmdr/unit
- Provide a readme file to enhance data usability

Making CPEX-CV Data F.A.I.R. (II)
HDF and NetCDF files

- Use CF compliant coordinate system
  - Time: use “time” for short name and make “time” a dimension scale
    time: long_name = “mid (or start, stop) of the interval”
    time: units = “seconds since YYYY-MM-DD 00:00:00”
    Use “time_bnds” is start and stop times are needed to define the interval and add “bounds” attribute
  - Latitude: use “lat” for short name and attached to “time”
    lat: long_name =
    lat: units = “degrees_north”
  - Longitude: use “lon” for short name and attached to “time”
    lon: long_name =
    Lon: units = “degrees_east”
- Data Product variables
  - 1-D datasets: x(time = n)
    x: long_name =
    x: units = (http://codes.wmo.int/wmdr/unit)
    x: missing_FillValues =
    x: MACIE_standard _name =
    x: uncertainty = uncertainty or uncertainty variable name

Making CPEX-CV Data F.A.I.R. (III)
HDF and NetCDF files (Cont.)

  - 2-D datasets: x(time = n, z = m), e.g., z is scale of vertical profile
    x: long_name =
    x: units = (http://codes.wmo.int/wmdr/unit)
    x: missing_FillValues =
    x: MACIE_standard _name =
    x: uncertainty = uncertainty or uncertainty variable name
- All data variables should be under root, intermediate/ancillary data should be under subgroups
- Global attributes
  - File information
  - PI and data submitter information
  - Project and platform information
  - Version control information
  - Information helpful to others to find your data
  - Instrument/measurement specific information
- CPEX-CV publication quality data will be checked using an offline scanner
- Will work with each PI individually to resolve issues
- Provide a readme file to enhance usability

Making CPEX-CV Data F.A.I.R. (IV)

Standard Names

- Standard name is designed as a "tag" to enhance data discoverability, usability, and ingest processing
  - Measurement category and CoreName are for discoverability across missions and archival process
  - Attributes: information for data use and more detailed search
- Please use the standard names on the list:

https://www-air.larc.nasa.gov/missions/etc/AtmosphericCompositionVariableStandardNames.pdf

- New standard names can be added by contacting Morgan, Gao, or Michael
- Morgan will check the files on the repository and provide feedback to PI as necessary
- Best practices:
  - Use "none" for non-data product variables, e.g., total temperature
  - Use the same standard name for main variable and ancillary variable, e.g., O3 and O3_unc
  - Additional information should be provided in long name/description

Points of Contact

- **Field Repository (www-air)**
  - Michael Shook, NASA Langley Research Center, michael.a.shook@nasa.gov, 757-864-5793
  - Gao Chen, NASA Langley Research Center, gao.chen@nasa.gov, 757-759-5642 (cell)
  - Ali Aknan, SSAI/NASA Langley Research Center, ali.a.aknan@nasa.gov, 757-951-1609
  - Morgan Silverman, SSAI/NASA Langley Research Center, morgan.l.silverman@nasa.gov (standard name issues)
  - Username/password: contact science team lead
- **GHRC**
  - Geffrey Stano, UAH/NASA Marshall Space Flight Center, gts0007@uah.edu
  - Leigh Sinclair, UAH/NASA Marshall Space Flight Center, slb0012@uah.edu
- **ASDC**
  - Megan Buzanowicz, SSAI/NASA Langley Research Center, megan.e.buzanowicz@nasa.gov
  - Sean Leavor, SSAI/NASA Langley Research Center, sean.leavor@nasa.gov