

# DIARieS: Developing a Vision for Maturing the Heliophysics Infrastructure towards Open Science (AGU 2022)

Rebecca Ringuette<sup>1</sup>, Alec Engell<sup>2</sup>, Oliver Gerland<sup>3</sup>, Ryan M Mcgranaghan<sup>4</sup>, and Barbara Thompson<sup>4</sup>

<sup>1</sup>Affiliation not available

<sup>2</sup>NextGen Federal Systems

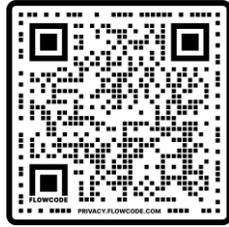
<sup>3</sup>Ensemble Government Services

<sup>4</sup>NASA/GSFC

December 27, 2022



LIKED publication DOI:  
10.1016/j.asr.2022.10.051



DIARieS publication DOI:  
10.1016/j.asr.2022.05.012

# DIARieS

---

DEVELOPING A VISION FOR MATURING THE  
HELIOPHYSICS INFRASTRUCTURE TOWARDS  
OPEN SCIENCE

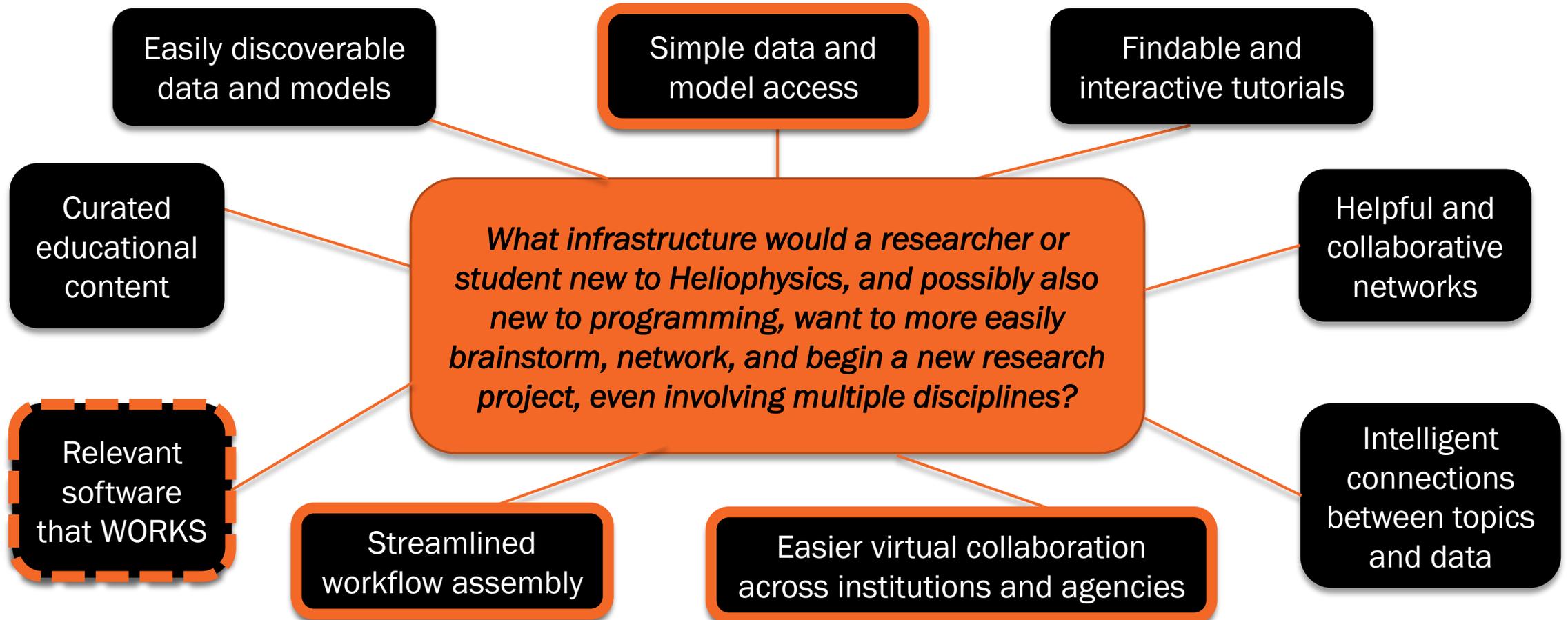
*R. Ringuette, A. Engell, O. Gerland, R. M. McGranaghan, and  
B. Thompson.*

AGU 2022 Fall Meeting



# Motivation

---



## **DIARieS**

An ecosystem to simplify  
**D**iscovery, **I**mplementation,  
**A**nalysis, **R**eproducibility,  
and **S**haring of scientific  
results and environments.

A  
**D**iscovery online resource for  
discovering and  
implementing knowledge,  
data, and infrastructure  
resources.

# Proposed Solution

# The DIARieS Ecosystem

## Discover

*Search for data, catalogs, models, and software using uniform parameters:*

- Discipline
- Domain
- Phenomenon
- Time Range
- Standardized Variable Name

## Implement

Add the selected data, catalogs, models, and software to your ecosystem to customize your research environment **by clicking a button.**

## Analyze

Perform your specialized analysis in an *automatically versioned notebook* with *built-in interactive graphics* and *high performance computing* capabilities.

## Reproduce

**Preserve your analysis** workflow in an *automatically containerized ecosystem*, complete with all research components included.

## Share

Create *customized narrated tutorials* to present your research or dashboard to the community or with colleagues with **more transparency.**

# Welcome to the DIARieS Ecosystem

File Add Templates Packages Variables Citations Recording Share Help

## Menu options (in above drop-down menus)

### File:

- New
- Open
- Save/Save as
- Close

### Add:

- **Data/Model Outputs/Catalogs**
- Software
- Widget
- Notebook
- Computing Resource

### Templates:

- Template search
- Import template
- Create new template

### Packages:

- None added

### Variables:

- None added

### Citations:

- (Pop up window with citation list and export options)

### Recording:

- Create recording
- Edit recording
- View recording

### Share:

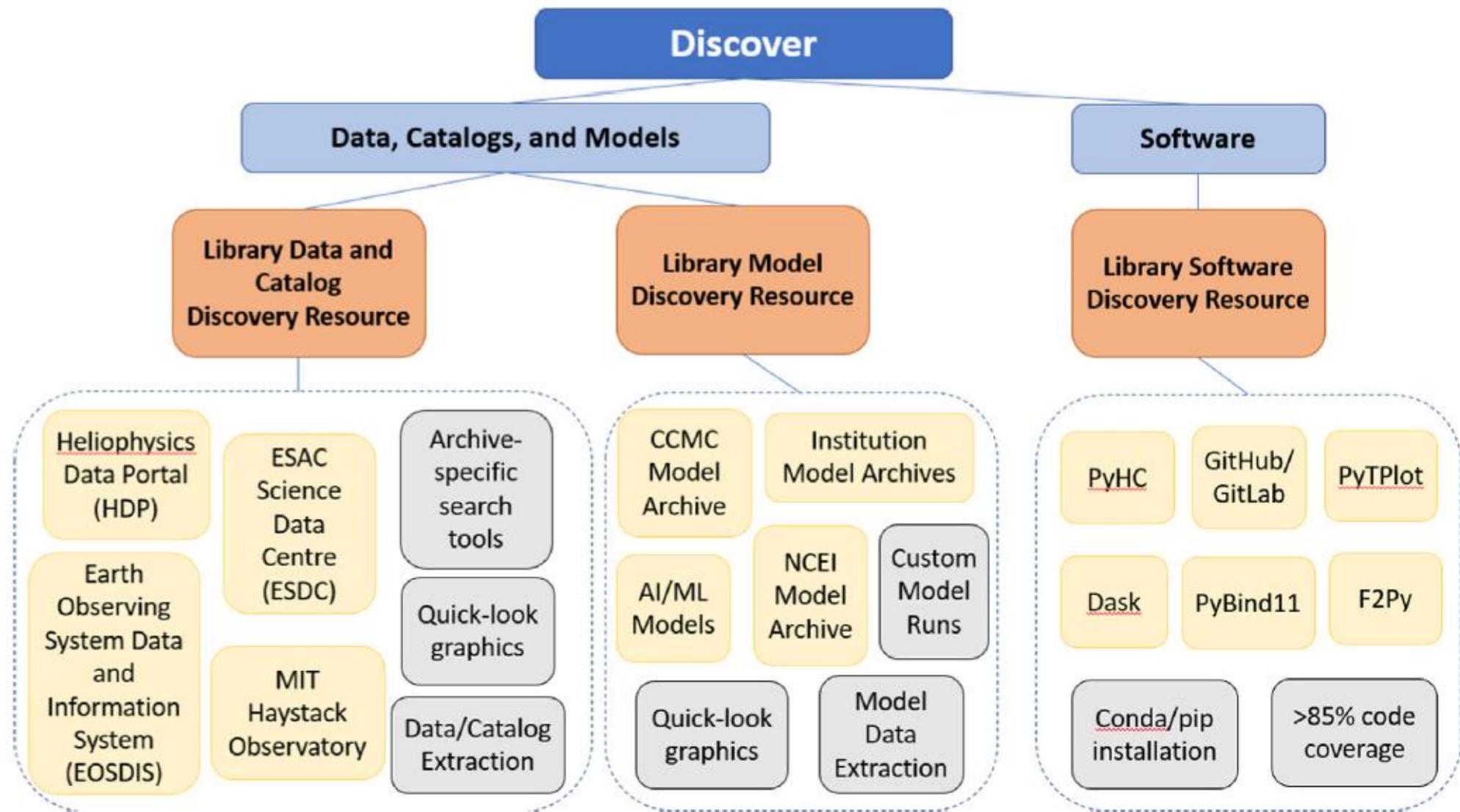
- Add/Remove authors/viewers
- Change permissions
- Create demo/user mode
- Publish

### Help

Documentation/Help  
Metadata/Properties  
Remove/Update

Documentation/Help  
Metadata/Properties  
Remove/Change Name

Create or open a DIARieS ecosystem to get started  
or click here for a narrated tutorial



- Built on LIKED – an intelligent search interface built using data science technology.
- Two sections based on type of expected mechanisms: Data and Software

# DIARieS container: Data Discovery and Implementation Interface

cdaweb.gsfc.nasa.gov/cgi-bin/eval2.cgi

NASA GODDARD SPACE FLIGHT CENTER  
Space Physics Data Facility

+ Goddard Home  
+ NASA Home

+ SPDF HOME + MISSION DATA + MODELS at CCMC + SCIENCE ENABLED + AND MORE

+ CDAWeb Home  
**CDAWeb**  
+ FEEDBACK

*Coordinated Data Analysis Web*

## CDAWeb Data Explorer

Select start and stop times from which to GET or PLOT data:

Start time (YYYY/MM/DD HH:MM:SS.mmm): 2015/11/25 04:00:00.000

Stop time (YYYY/MM/DD HH:MM:SS.mmm): 2015/11/26 04:00:00.000

Compute uniformly spaced binned data for scalar/vector/spectrogram data (not available with noise filtering)

Use spike removal to filter data without binning (not available with noise filtering)(Warning: Experimental !!).

Select an activity:

Data Availability Chart : Generate a chart showing when data is available for the selected data set(s) and time range (Select > 1day). **NEW**

Plot Data : select one or more variables from list below and press submit.

List Data (ASCII/CSV): select one or more variables from list below and press submit. (Works best for < 31 days)

Download original files : press submit button to retrieve list of files. (Max. 200 days - use [HTTPS site](#) for larger requests)

Create V3.8 CDFs for download or Autoplot demonstration: select one or more variables from the list below and press submit.

Create Version 2.7.2 compatible CDFs (Default is Version 3.8)

Create audio files based on data from selected variables. [More information about audification.](#)

Note: [CDF patch](#) required for reading Version 3.8 CDFs in IDL or MATLAB.  
Get [CDFX](#) - IDL GUI plotting/listing toolkit software. To be used with either the daily or "created" CDF files available above.

Pressing the "Submit" button will spawn a new window/tab in order to support the new "Previous" and "Next" functions.

Submit Reset

- Users first use the LIKED interface (accessible inside and outside of DIARieS) to find the desired archive (e.g. SPDF as shown here),
- Then use the search interface and mechanisms of the archive to find desired data.
- DIARieS simplifies how to add that data to the analysis environment.

# DIARieS container: Data Discovery and Implementation Interface

cdaweb.gsfc.nasa.gov/cgi-bin/eval3.cgi

**GODDARD SPACE FLIGHT CENTER**  
Space Physics Data Facility

+ Goddard Home  
+ NASA Home

+ SPDF HOME   + MISSION DATA   + MODELS at CCMC   + SCIENCE ENABLED   + AND MORE

+ CDAWeb Home  
**CDAWeb**  
+ FEEDBACK

**Coordinated Data Analysis Web**

## CNOFS\_CINDIS\_IVM\_500MS

Download your new file (REQUIRES latest IDL/CDF\_patch file from CDF): [cnofs\\_cindis\\_ivm\\_500ms\\_20151125040000\\_20151126034441.cdf](#)

OR

[Interactively plot this file \(using non-NASA software provided by Autoplot.org\)](#)

NOTE: On a Mac OS X the dynamic jnlp could open as an ascii file. To fix that please follow these [instructions](#).

<< Previous time range   Next time range >> **NEW**

>> Zoom IN time range <<   << Zoom OUT time range >> **NEW**

< Pan left   Pan right > **NEW**

Return to: CDAWeb Data Explorer **NEW**

[notes and caveats](#)

**NASA Official:** Robert M. Candey  
(301)286-6707, [Robert.M.Candey@nasa.gov](mailto:Robert.M.Candey@nasa.gov)  
Curator: Tami Kovalick  
Last Modified: 5 Dec 2022

**Contact SPDF:** NASA-SPDF-  
[Support@nasa.onmicrosoft.com](mailto:Support@nasa.onmicrosoft.com)  
+ Privacy Policy and Important Notices

- Open link in new tab
- Open link in new window
- Open link in incognito window
- Add to container
- Link to container
- Save link as...
- Copy link address
- Add link to list
- Inspect

Saves data to DIARieS container

Links data container to DIARieS container

# DIARieS container: Data Discovery and Implementation Interface

cdaweb.gsfc.nasa.gov/cgi-bin/eval3.cgi

GODDARD SPACE FLIGHT CENTER | + Goddard Home

Save	Description	Units	Default Name	Enter custom name
<input type="checkbox"/>	RPA data quality flag: 0=high 9=low ( <a href="https://spdf.gsfc.nasa.gov/pub/data/cnofs/cindi/QualityFlagDesc.doc">https://spdf.gsfc.nasa.gov/pub/data/cnofs/cindi/QualityFlagDesc.doc</a> )		RPAflag	
<input checked="" type="checkbox"/>	Ion density in cm-3	cm-3	ionDensity	rho_ion
<input checked="" type="checkbox"/>	Ion temperature in Kelvin	K	ionTemperature	T_ion
<input checked="" type="checkbox"/>	Ion 1 (normally O+) -----> Atomic Mass Number	AMU	atomicMass_ion1	Oplus
<input checked="" type="checkbox"/>	Ion 2 (normally H+) -----> Atomic Mass Number	AMU	atomicMass_ion2	Hplus
<input type="checkbox"/>	Ion3 (normally He+) -----> Atomic Mass Number	AMU	atomicMass_ion3	
<input type="checkbox"/>	Ion4 -----> Atomic Mass Number	AMU	atomicMass_ion4	
<input type="checkbox"/>	Ion 5 -----> Atomic Mass Number	AMU	atomicMass_ion5	
<input checked="" type="checkbox"/>	UT in seconds	s	time	
<input checked="" type="checkbox"/>	Geographic latitude in degrees	deg	glat	
<input checked="" type="checkbox"/>	Geographic longitude in degrees	deg	glon	
<input checked="" type="checkbox"/>	Altitude in km	km	altitude	

Reset Form      Add all with defaults      Save      Cancel

# DIARieS container: Data Discovery and Implementation Interface

← → ↻ cdaweb.gsfc.nasa.gov/cgi-bin/eval3.cgi

**GODDARD SPACE FLIGHT CENTER**  
Space Physics Data Facility

+ Goddard Home  
+ NASA Home

+ SPDF HOME + MISSION DATA + MODELS at CCMC + SCIENCE ENABLED + AND MORE

+ CDAWeb Home  
**CDAWeb**  
+ FEEDBACK

*Coordinated Data Analysis Web*

## CNOFS\_CINDIS\_IVM\_500MS

Download your new file ([REQUIRES latest IDL/CDF patch file from CDF](#)): [cnofs\\_cindis\\_ivm\\_500ms\\_20151125040000\\_20151126034441.cdf](#)

OR

[Interactively plot this file \(using non-NASA software\)](#)

NOTE: On a Mac OS X the dynamic jnlp could open [instructions](#).

<< Previous time range Next time range >>

>> Zoom IN time range << << Zoom OUT time range >> **NEW**

< Pan left Pan right > **NEW**

Return to: CDAWeb Data Explorer **NEW**

[notes and caveats](#)

NASA Official: Robert M. Candey  
(301)286-6707, [Robert.M.Candey@nasa.gov](mailto:Robert.M.Candey@nasa.gov)  
Curator: Tami Kovalick  
Last Modified: 6 Dec 2022

Contact SPDF: NASA-SPDF-  
[Support@nasa.onmicrosoft.com](mailto:Support@nasa.onmicrosoft.com)  
+ Privacy Policy and Important Notices

When data is added to the container...

- A link for each variable is added to the variable list,
- And the proper citation is added to the container's citation list.
- The link in the variable list links to three items:
  - Documentation/Help
  - Metadata/Properties
  - Remove/Change Name Options

# DIARieS container: Software Discovery and Implementation Interface

The screenshot shows the 'Projects' page on heliopython.org. The page has a search bar and a 'Show Keyword Filters' button. Below the search bar is a table of projects. The first project is 'HAPI Client' by Bob Weigel, and the second is 'Kamodo' by Rebecca Ringuette. A blue button labeled 'Add to container' is overlaid on the bottom of the table.

**Projects**

To add a project to this page, please refer yourself to the [project addition instructions](#).

### Core packages

These packages each offer a wide range of functionality in their area, and conform to the PyHC com

Table **Cards**

Search:  Show Keyword Filters

Name	Description	Code	Docs	Site	Contact
<b>HAPI Client</b>	Access time series data access from many sources				Bob Weigel
<b>Kamodo</b>	CCMC tool for access, interpolation,				Rebecca Ringuette becca.ringuette@nasa.gov

**Add to container**

The functionality is similar for the software discovery interface:

- The interface will look for the package on conda-forge, then pip, then on GitHub and similar websites.
- If found, the package will be installed with the first option found.
- If the package installation script/command cannot be found, then the user will be presented with an option to install it manually.
- Any package conflicts will be handled automatically by currently existing tools (e.g. Kubernetes clusters).

# DIARieS container: Software Discovery and Implementation Interface

The screenshot shows the 'Projects' page on heliopython.org. The page has a dark header with navigation icons and the URL 'heliopython.org/projects/'. Below the header is a 'Projects' section with a sub-header 'Core packages'. A text block says 'To add a project to this page, please refer yourself to the [project addition instructions](#).' Below this are tabs for 'Table' and 'Cards', a search box, and a 'Show Keyword Filters' button. A table lists projects with columns for Name, Description, Code, Docs, Site, and Contact. The first row is for 'HAPI Client' by Bob Weigel, and the second is for 'Kamodo' by Rebecca Ringuette. A blue 'Package Added' button is at the bottom center. A yellow callout box on the right contains text and a bulleted list.

Projects

To add a project to this page, please refer yourself to the [project addition instructions](#).

## Core packages

These packages each offer a wide range of functionality in their area, and conform to the PyHC comm

Table Cards

Search:  Show Keyword Filters

Name	Description	Code	Docs	Site	Contact
HAPI Client	Access time series data access from many sources				Bob Weigel
Kamodo	CCMC tool for access, interpolation,				Rebecca Ringuette rebecca.ringuette@nasa.gov

Package Added

When software is added to the container...

- A link for each package, named with the package name, is added to the package list,
- And the proper citation is added to the container's citation list.
- The link in the package list links to three items:
  - Documentation/Help
  - Metadata/Properties
  - Remove/Update Options

# Welcome to the DIARieS Ecosystem

File Add Templates Packages Variables Citations Recording Share Help

## Menu options (in above drop-down menus)

### File:

- New
- Open
- Save/Save as
- Close

### Add:

- Data/Model Outputs/Catalogs
- Software
- Widget
- Notebook
- Computing Resource

### Templates:

- Template search
- Import template
- Create new template

### Packages:

- cdflib
- Xarray
- Kamodo
- ...

### Variables:

- rho\_ion
- T\_ion
- Oplus
- time
- ...

### Citations:

- (Pop up window with citation list and export options)

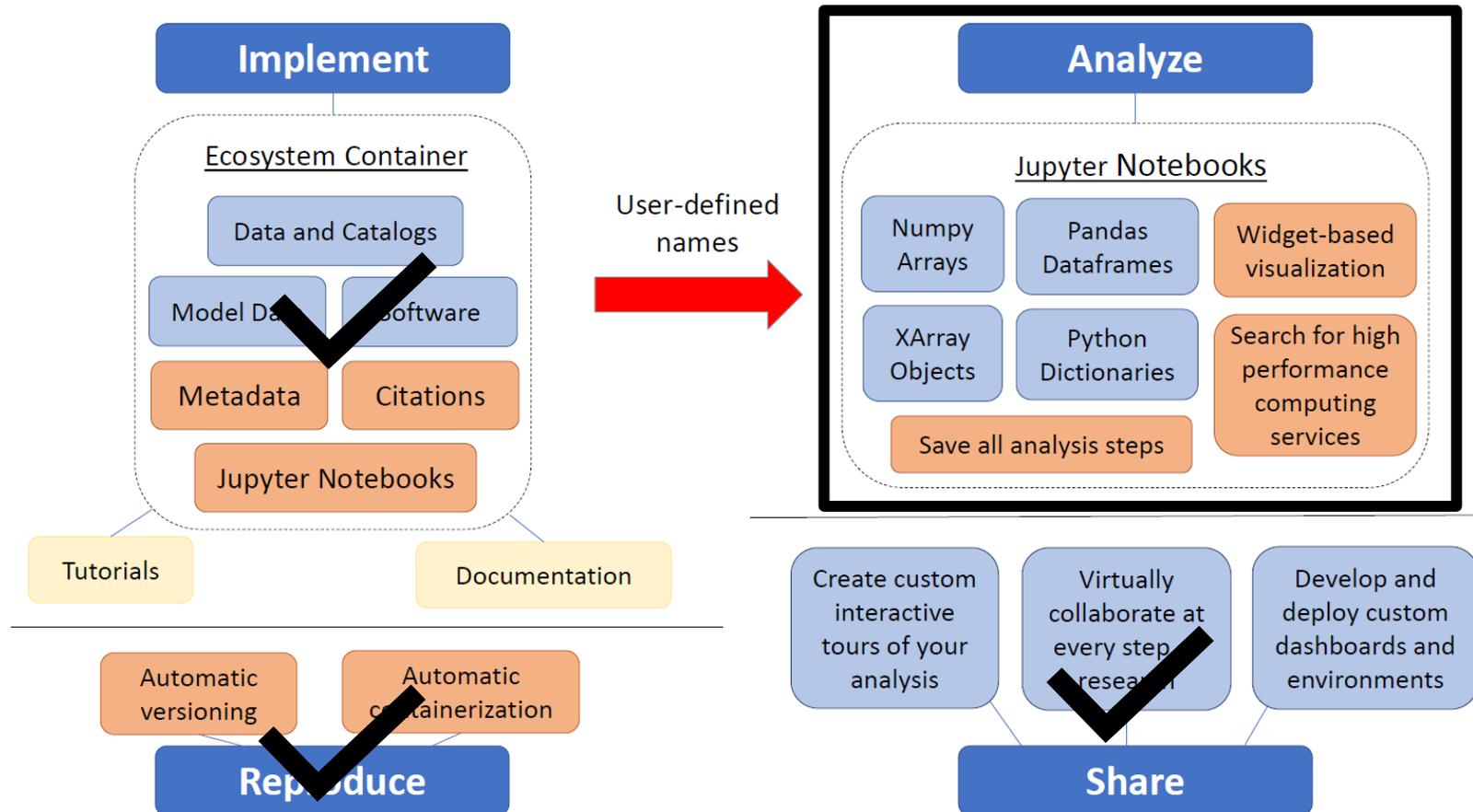
### Recording:

- Create recording
- Edit recording
- View recording

Documentation/Help  
Metadata/Properties  
Remove/Update

Documentation/Help  
Metadata/Properties  
Remove/Change Name

# DIARieS: Simplifying AIR



(AIR = Accessibility, Interoperability, and Reproducibility)

# Welcome to the DIARieS Ecosystem

File Add Templates Packages Variables Citations Recording Share Help

## Menu options (in above drop-down menus)

### File:

- New
- Open
- Save/Save as
- Close

### Add:

- Data/Model Outputs/Catalogs
- Software
- Widget
- Notebook
- Computing Resource

### Templates:

- Template search
- Import template
- Create new template

### Packages:

- cdflib
- Xarray
- Kamodo
- ...

### Variables:

- rho\_ion
- T\_ion
- Oplus
- time
- ...

### Citations:

- (Pop up window with citation list and export options)

### Recording:

- Create recording
- Edit recording
- View recording

Documentation/Help  
Metadata/Properties  
Remove/Update

Documentation/Help  
Metadata/Properties  
Remove/Change Name

# DIARieS container: Analysis Interface

## Widget Library

Search: \_\_\_\_\_

- Kamodo visualization
- CDF visualization
- PlasmaPy function library
- pysat data retrieval
- Kamodo model data interface
- SpacePy SWMF tools
- SunPy data filtering
- SciPy

## Loaded Widgets

Search: \_\_\_\_\_

### PyHC Coordinate Converter

Choose input variables  Select coordinate system

time	Select variable	Select units	Select coordinate type
X or longitude	Select variable	Select units	Select coordinate type
Y or latitude	Select variable	Select units	Select coordinate type
Z, radius, or altitude	Select variable	Select units	Select coordinate type

Choose output variables  Select coordinate system

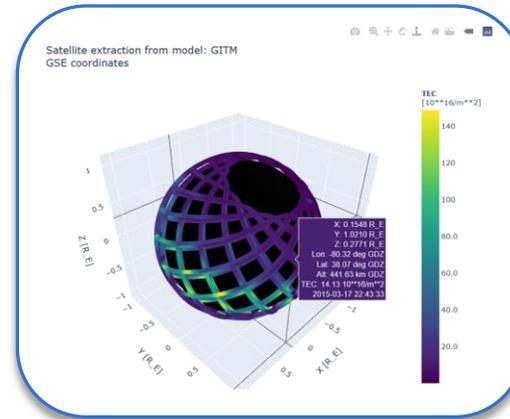
time	Type new name	Select units	Select coordinate type
X or longitude	Type new name	Select units	Select coordinate type
Y or latitude	Type new name	Select units	Select coordinate type
Z, radius, or altitude	Type new name	Select units	Select coordinate type

### AstroPy Unit Converter

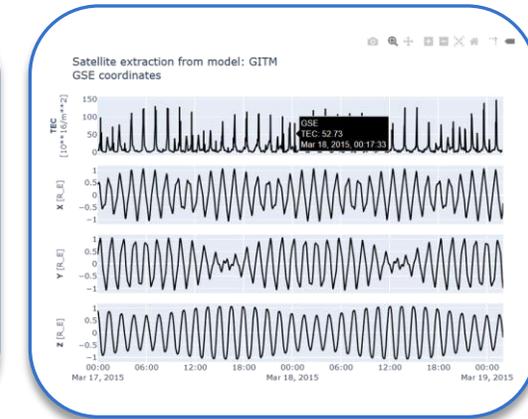
### Kamodo Flythrough

## Active Widgets

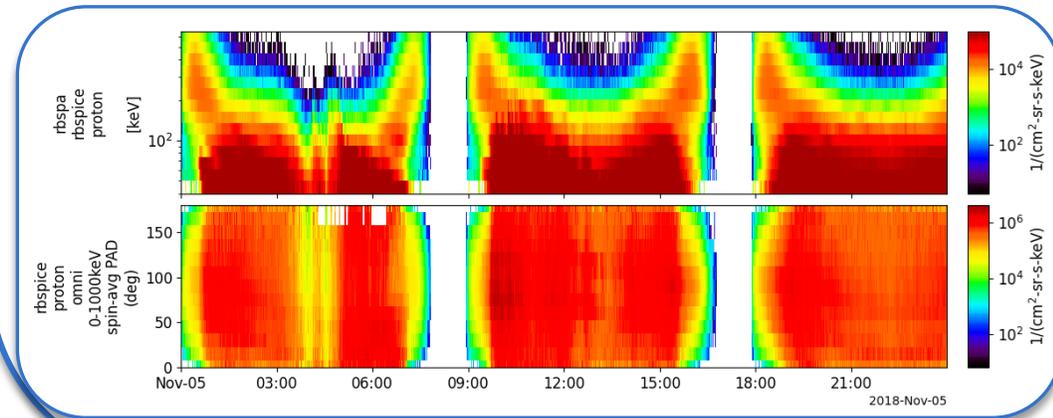
Search: \_\_\_\_\_



Kamodo 3D visualization



Kamodo 1D visualization



pySPEDAS visualization

# DIARieS container: Analysis Interface

## Variable Library

- rho\_n
- T\_ion
- X\_GEO
- Y\_GEO
- Z\_GEO
- time.UTC
- TEC\_GITM
- glat

## Loaded Widgets

### PyHC Coordinate Converter

Choose input variables Select coordinate system

time	Select variable	Select units	Select coordinate type
X or longitude	Select variable	Select units	Select coordinate type
Y or latitude	Select variable	Select units	Select coordinate type
Z, radius, or altitude	Select variable	Select units	Select coordinate type

Choose output variables Select coordinate system

time	Type new name	Select units	Select coordinate type
X or longitude	Type new name	Select units	Select coordinate type
Y or latitude	Type new name	Select units	Select coordinate type
Z, radius, or altitude	Type new name	Select units	Select coordinate type

## Interactive Notebook

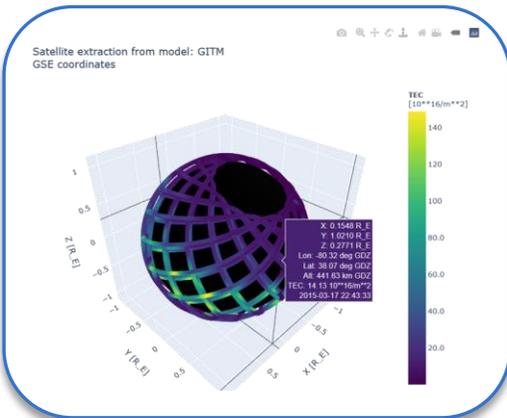
### 2.3.1 Perform SpacePy Calculations with MMS Data

```
def spacecraft_magnetopause_calculations(mms_mec_vars):  
    """Returns epoch,  
    pos,  
    distance between spacecraft and magnetopause  
    magnetopause's distance from Earth,  
    spacecraft's distance from Earth,  
    and solar zenith angle.  
    """  
    data = pyplot.get_data(mms_mec_vars[0])  
    pos_gsm = data.y  
    ticks = spacepy.time.Ticktock(data.times, dtype='UNIX')  
    epoch = ticks.UTC  
    c = spacepy.coordinates.Coords(pos_gsm, 'GSM', 'car', u  
    pos = c.convert('GSE', 'car').data  
  
    # Get the Shue coefficients  
    alpha = [] # Shue flaring angle  
    standoff = spacepy.empiricals.getMPstandoff(ticks, alpha  
    alpha = np.array(alpha)  
  
    # Solar zenith angle of s/c position (angle with GSE +x  
    sza = np.arctan2(pos[:, 0], np.sqrt((pos[:, :2] ** 2).s  
    # Radial distance to MP along Earth-SC line (applicatio  
    mp_dist = standoff * (2. / (1 + np.cos(sza))) ** alpha  
    # Radial distance to SC  
    sc_dist = np.sqrt((pos ** 2).sum(axis=1)) / 6378  
    # How far is SC outside of MP?  
    sc_to_mp = sc_dist - mp_dist  
  
    return epoch, pos, sc_to_mp, mp_dist, sc_dist, sza
```

### 2.3.2 Detect Magnetopause Crossings

Since we can calculate the distance between the spacecraft and the magnetopause, we can find crossings by collecting the indices where that distance changes 0 (i.e., changes sign):

## Active Widgets



## Widget Help

Select a widget to see the tutorials and documentation.

## Presentation Tools

Add a  
narration

Record a  
tutorial

Presentation  
Mode



Link to recorded talk on YouTube:  
<https://youtu.be/aP7VAHMO6M4>

# Summary



LIKED publication DOI:  
10.1016/j.asr.2022.10.051



DIARieS publication DOI:  
10.1016/j.asr.2022.05.012

Heliophysics currently needs a coordinated push to make our work FAIR and efficient.

*Connect*  
our archives based on a 'system of systems' approach.

*Build*  
on the successes of current efforts and technology.

*Assemble*  
our various advances into a streamlined workflow

*We call on the community to work together on this goal to launch our field into the future.*