Science Platforms Coordina

IHDEA Working Group Update

GOAL

To develop international standard software computing environments for Heliophysics

MEMBERS

Jan Reerink
Shawn Polson

Anne Spalding
Arnaud Masson
Brian Thomas
James Parr
Jeffrey Bradford
Jim Colliander
Julie Barnum
Rebecca Ringuette
Sandy Antunes
Sarah Rourke

INTRODUCTION

01
INSPIRATION

02
WHAT WE DONE

03
SURVEY RESULTS

04
FUTURE WORK

01
INSPIRATION

IHDEA Cloud Workshop July:

"*We want a shared Heliophysics software environment for cloud computing*"

<> Code ⊙ Issues 9 ⅂⅃ Pull requests ⊡ Discussions ⊙ Actions ⊞ Projects ⊞ Wiki ⛨ Security ⬓ Insights ⚙ Settings

🐍 **pyhc-docker-environment** Public

Edit Pins ▾ ⊙ Unwatch 4 ▾ ⅂⅃ Fork 0 ▾ ☆ Star 1 ▾

⅂ main ▾ ⅂⅃ 2 Branches ⬭ 0 Tags

🔍 Go to file t Add file ▾ <> Code ▾

About

Pipeline to generate and dockerize Python environments containing all published PyHC packages

⬚ github-actions[bot] Update Docker image requirements, spreadsheet, README 50f5e55 · 2 weeks ago ⟳ 166 Commits

⬓ .github/workflows	Update to Python 3.10	4 months ago
⬓ docker	Update Docker image requirements, spreadsheet, README	2 weeks ago
⬓ spreadsheets	Update Docker image requirements, spreadsheet, README	2 weeks ago
⬓ utils	Temporarily remove OMMBV	3 weeks ago
⬚ .gitattributes	Manage executable paper with Git LFS	8 months ago
⬚ .gitignore	Initial commit	8 months ago
⬚ README.md	Update Docker image requirements, spreadsheet, README	2 weeks ago
⬚ __init__.py	Initial commit	8 months ago
⬚ pipeline.py	Temporarily specify numpy==1.26.4	3 months ago
⬚ pipeline_requirements.txt	Setuptools<60.0 didn't fix anything	3 weeks ago

📖 Readme
⟋⟍ Activity
▤ Custom properties
☆ 1 star
⊙ 4 watching
⅂⅃ 0 forks

Report repository

Releases

No releases published
Create a new release

Packages
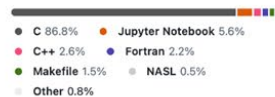
No packages published
Publish your first package

📖 README ✎ ☰

Contributors 2

👤 **sapols** Shawn Polson

⬓ github-actions[bot]

PyHC Environment Pipeline

Overview

The PyHC Environment Pipeline automates the creation of Docker images with Python environments that come pre-loaded with the latest versions of all published PyHC (Python in Heliophysics Community) packages. The GitHub Actions workflow rebuilds and pushes the Docker images to Docker Hub each time a PyHC package releases a new update. It also keeps the source files in this repo in-sync with what's in Docker Hub.

Features

Languages

● C 86.8% ● Jupyter Notebook 5.6%
● C++ 2.6% ● Fortran 2.2%
● Makefile 1.5% ● NASL 0.5%
● Other 0.8%

Explore / spolson/pyhc-environment

spolson/pyhc-environment ☆0

Manage Repository

By spolson · Updated 13 days ago

A JupyterLab session using an environment containing every published PyHC package.

↓ Pulls 191

IMAGE

Overview   Tags

This `pyhc-environment` image runs a JupyterLab session with an environment containing every published PyHC package.

Inspiration

The environment used by this JupyterLab session is backed by the pyhc-docker-environment 🔗 GitHub repo. That repo maintains up-to-date package version requirements for a unified Python virtual environment holding the latest versions of each PyHC package, automated by GitHub Actions.

Caveats

It is difficult with so many dependency requirements to create a valid Python virtual environment capable of even importing all these PyHC packages. There is one `import-test.ipynb` notebook added to demonstrate that all the imports do work, but little further compatibility testing has been done.

Docker Pull Command

```
docker pull spolson/pyhc-environment
```
Copy

"*Sounds like a new working group*"

02
WHAT WE'VE DONE

WHAT WE'VE DONE

OCT '23
First kick-off meeting

DEC '23
Working group starts meeting monthly

FEB '24
Helio software env survey created
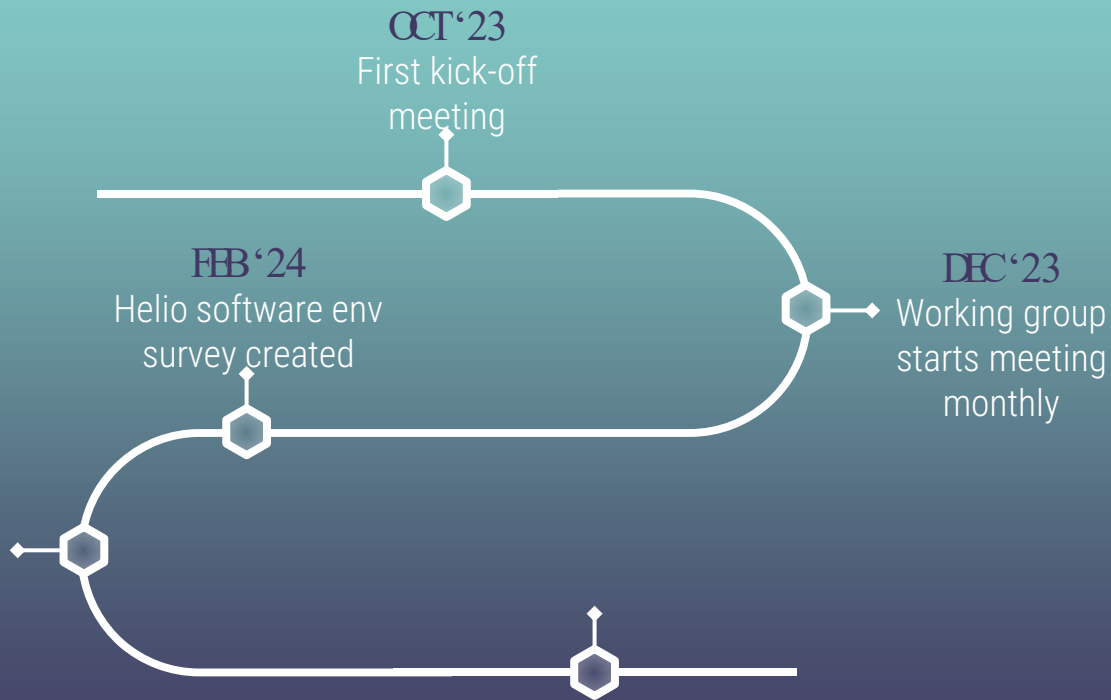
SURVEY: Software Environments for Heliophysics

"We are developing multiple Python-based software environments tailored to current research in our fields, so we are interested in which Python packages you currently use in your research. By developing these software environments based on community input, we intend to increase the reproducibility of your research for others, which research shows is expected to increase citations to your work and increase collaboration across our communities. So please open up your favorite software environment and answer the brief questions below"

Software Environments for Heliophysics

Hi there!

We are an international working group that is working to create sharable science platforms for the Heliophysics, Space Physics, and related communities. We are developing multiple Python-based software environments tailored to current research in our fields, so we are interested in which Python packages you currently use in your research. By developing these software environments based on community input, we intend to increase the reproducibility of your research for others, which research shows is expected to increase citations to your work and increase collaboration across our communities. So please open up your favorite software environment and answer the brief questions below. If you have more than one environment, please enter your answers for those environments on separate forms. We intend to give reports on our progress at the DASH workshop and at AGU. Thanks for your input!
IHDEA Science Platforms Coordination Working Group
Co-chairs: Shawn Polson and Jan Reerink
https://ihdea.net/

shpo9723@colorado.edu Switch account

The name and photo associated with your Google account will be recorded when you upload files and submit this form. Your email is not part of your response.

Any files that are uploaded will be shared outside of the organization they belong to.

* Indicates required question

If you would like us to contact you about this information, please enter your email address.

Your answer

What broad science category(ies) does the work you perform in this software environment fall into? *

☐ Solar

☐ Heliophysics

☐ Geospace

☐ Space Physics

☐ Magnetosphere

☐ Ionosphere/Thermosphere

WHAT WE'VE DONE

OCT '23
First kick-off meeting

DEC '23
Working group starts meeting monthly

FEB '24
Helio software env survey created

AUG '24
HelioCloud-based env prototype built

NOW
Deploying and iterating
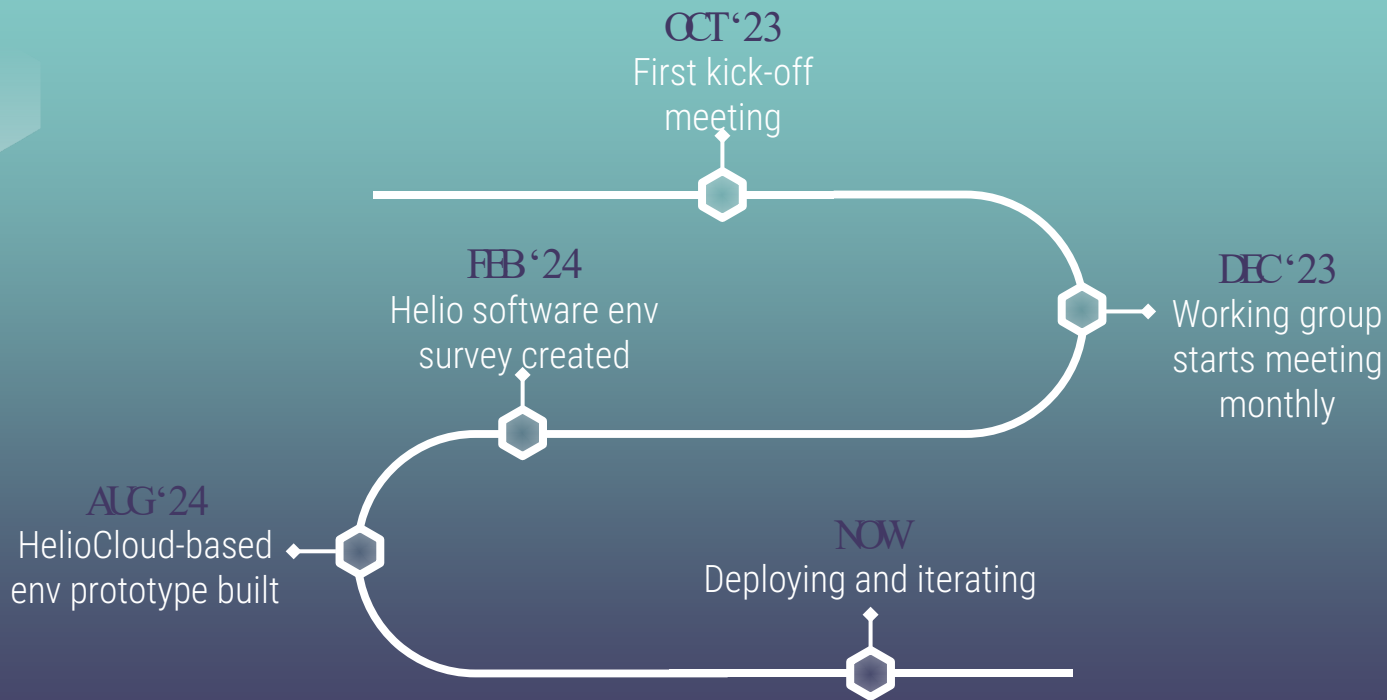
DEPLOYMENT
PARTNERS

2i2c

The International Interactive Computing Collaboration—offer scalable and access-controlled Binder/Jupyter resources

ESADatalabs

A "public moderated beta" of browser-based tools for scientists with a focus on bringing questions and code to data

2i2c.org

A global network of community hubs for interactive learning and discovery

Our interactive computing platform gives research and education communities a digital home to create and share knowledge with a global network of communities to learn from.

Learn about our interactive computing platform   Join our network ›
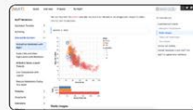
We help communities build their own interactive computing hub in the cloud with open infrastructure

Resource and user management   **Interactive interfaces**   **Computational knowledge bases**

2i2c's community hub platform and consultancy services ensure your community makes the best use of open infrastructure for interactive computing in the cloud.

We serve **over 90 communities across the globe** with **over 7000 active users** dedicated to creating and sharing knowledge. See our community impact stories for inspiration.

NASA   Smithsonian   hhmi | Howard Hughes Medical Institute   National Center for Atmospheric Science   PANGEO   UNIVERSITY OF TORONTO   COLUMBIA UNIVERSITY

Join our network of community hubs

datalabs.esa.int

DEPLOYMENTS

2i2c Binder

BinderHub where NASA foots the bill (small & big)

2i2c JupyterHub

Authenticated hub with many images, resource options

ESA DATALABS

JupyterLab platform connected to ESA infrastructure

❽ binder

Turn a Git repo into a collection of interactive notebooks

Have a repository full of Jupyter notebooks? With Binder, open those notebooks in an executable environment, making your code immediately reproducible by anyone, anywhere.

New to Binder? Get started with a Zero-to-Binder tutorial in Julia, Python, or R.

Build and launch a repository

GitHub repository name or URL

| GitHub ▾ | https://github.com/heliophysicsPy/science-platforms-coordination/ |

Git ref (branch, tag, or commit)

heliocloud-base

Path to a notebook file (optional)

Path to a notebook file (optional) | File ▾ | launch

Copy the URL below and share your Binder with others:

https://binder.opensci.2i2c.cloud/v2/gh/heliophysicsPy/science-platforms-coordination/heliocloud-base 📋

Expand to see the text below, paste it into your README to show a binder badge: ❽ launch binder ▶

How it works

① **Enter your repository information**

Provide in the above form a URL or a GitHub repository that contains

<> Code   ⊙ Issues   ⇄ Pull requests   ⊙ Actions   ⊞ Projects   ▥ Wiki   ⊙ Security   ⊵ Insights   ⚙ Settings

🐍 science-platforms-coordination   Public

⚲ Edit Pins ▾   ⊙ Unwatch 2 ▾   ⤙ Fork 0 ▾   ☆ Star 0 ▾

⑂ heliocloud-base ▾   ⑂ 4 Branches   ⊙ 0 Tags   🔍 Go to file   Add file ▾   <> Code ▾

About

This branch is 13 commits ahead of main .   ⇄ Contribute ▾

No description, website, or topics provided.

👤 sapols Update README.md   1e558b7 · 1 hour ago   ⊙ 18 Commits

▢ Readme
▥ MIT license

🗋 Dockerfile	Fix typo	last month
🗋 LICENSE	Initial commit	2 months ago
🗋 README.md	Update README.md	1 hour ago
🗋 apt.txt	First implementation with HelioCloud/Pangeo base	last month
🗋 conda-lock.yml	Upload conda-lock.yml	2 weeks ago
🗋 environment.yml	First implementation with HelioCloud/Pangeo base	last month
🗋 install_cdflib.sh	First implementation with HelioCloud/Pangeo base	last month
🗋 jupyter_notebook_config.py	Add jupyter_notebook_config.py	last month
🗋 start	First implementation with HelioCloud/Pangeo base	last month

∿ Activity
⊞ Custom properties
☆ 0 stars
👁 2 watching
⑂ 0 forks

Report repository

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

● Dockerfile 100.0%

Suggested workflows
Based on your tech stack

🐳 Docker image   [Configure]
Build a Docker image to deploy, run, or push to a registry.

▢ README   ⚖ License

Science Platforms Coordination Base Image

(For Binder/JupyterHub)

A Docker image inspired by HelioCloud's base image to be used by the Science Platforms Coordination IHDEA working group.

⊗ binder

Turn a Git repo into a collection of interactive notebooks

Have a repository full of Jupyter notebooks? With Binder, open those notebooks in an
executable environment, making your code immediately reproducible by anyone,
anywhere.

New to Binder? Get started with a Zero-to-Binder tutorial in Julia, Python, or R.

Build and launch a repository

GitHub repository name or URL

| GitHub ▾ | https://github.com/heliophysicsPy/science-platforms-coordination/ |

Git ref (branch, tag, or commit)	**Path to a notebook file (optional)**	
heliocloud-base	Path to a notebook file (optional)	File ▾

launch

Copy the URL below and share your Binder with others:

https://binder.opensci.2i2c.cloud/v2/gh/heliophysicsPy/science-platforms-coordination/heliocloud-base 📋

Expand to see the text below, paste it into your README to show a binder badge: ⊗ launch | binder ▸

How it works

① **Enter your repository information**

Provide in the above form a URL or a GitHub repository that contains

Untitled.ipynb — JupyterLab

File   Edit   View   Run   Kernel   Git   Tabs   Settings   Help

Untitled.ipynb

Code   Python 3 (ipykernel)

git

```
[2]: !pip list
```

Package	Version
accessible-pygments	0.0.5
adal	1.2.7
adlfs	2024.7.0
aiapy	0.7.4
aiobotocore	2.13.2
aiofiles	24.1.0
aioftp	0.22.3
aiohappyeyeballs	2.3.5
aiohttp	3.10.3
aioitertools	0.11.0
aiosignal	1.3.1
alabaster	0.7.16
alembic	1.13.2
annotated-types	0.7.0
anyio	3.7.1
appdirs	1.4.4
argon2-cffi	23.1.0
argon2-cffi-bindings	21.2.0
asciitree	0.3.3
asdf	3.4.0
asdf-astropy	0.6.1
asdf_coordinates_schemas	0.3.0
asdf_standard	1.1.1
asdf_transform_schemas	0.5.0
asteval	1.0.2
astropy	6.1.2
astropy_healpix	1.0.3
astropy-iers-data	0.2024.8.5.0.32.23
asttokens	2.4.1
async_generator	1.10
async-timeout	4.0.3
attrs	24.2.0
awscli	2.17.30
awscrt	0.21.2
azure-core	1.30.2
azure-datalake-store	0.0.51
azure-identity	1.17.1
azure-storage-blob	12.22.0
Babel	2.14.0
backcall	0.2.0
backports.tarfile	1.0.0
beautifulsoup4	4.12.3
black	24.8.0
bleach	6.1.0
blinker	1.8.2
blosc	1.11.2
bokeh	2.4.3

Simple   0   1   Python 3 (ipykernel) | Idle   Mem: 280.70 / 4096.00 MB   Mode: Command   Ln 1, Col 1   Untitled.ipynb

Repositories   Account

sapols / helio-science-platform ☆

Repository Activity

JULY   AUGUST   SEPTEMBER
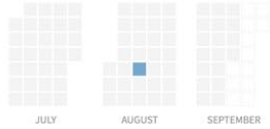
Recent Repo Builds

No builds have been run for this repository.

Click on the ▦ Builds tab to start a new build.

Description

Click to set repository description
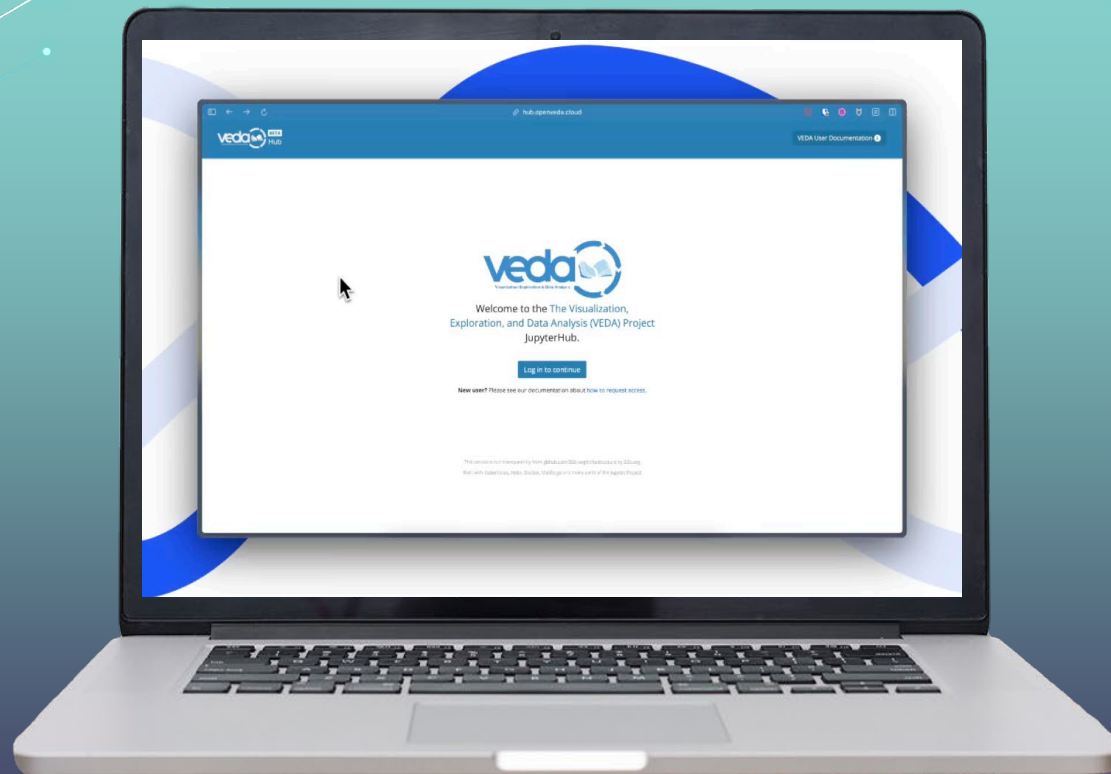
Pull this container with the following Podman command:

`podman pull quay.io/sapols/helio-science-platform`

Pull this container with the following Docker command:

`docker pull quay.io/sapols/helio-science-platform`

Imagining multiple flavors:
HelioCloud, PyHC, Survey results

03

SURVEY RESULTS

META ANALYSIS

(#) = number of respondents

DATA PROCESSING & ANALYSIS TOOLS

Specific science questions (10)
In Situ data analysis (7)
Data & image processing (5)
Event labeling & ML (3)

MODELING SIMULATIONS

(6) Specific science questions
(5) Comparing models/theory
(2) Performing model/sim runs
(2) Event detection & forecasts

DATA FUSION & VISUALIZATION

Data, images, & datacubes (5)
Data wrangling (3)
Making spacecraft data interoperable (3)
Time-series data (2)
Publication-quality data (2)
Plotting telemetry (1)
FAIR data (1)

SOFTWARE DEVELOPMENT & DATA PRODUCTION
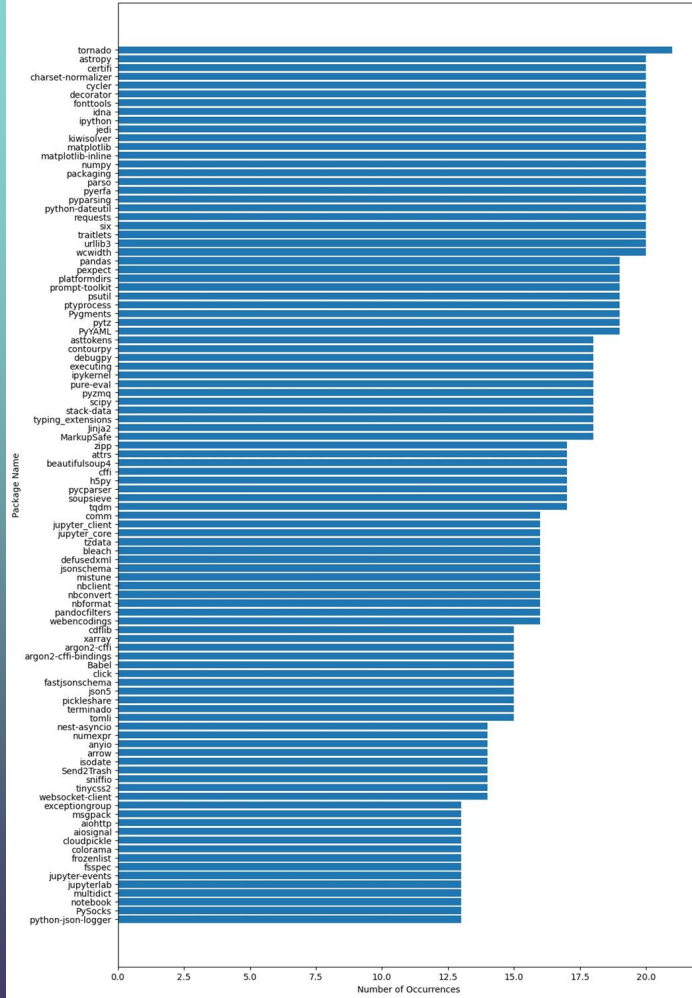
(2) Improve specific packages
(2) Create data from telemetry
(1) Operationalize research s/w
(1) Quality check data

Top 100 Packages by Occurrence

04

FUTURE WORK

THANKS

Any questions?

shawn.polson@lasp.colorado.edu

Heliophysics
env survey:

HelioCloud base
image in Binder:

2i2c
website:

ESA Datalabs
website: