

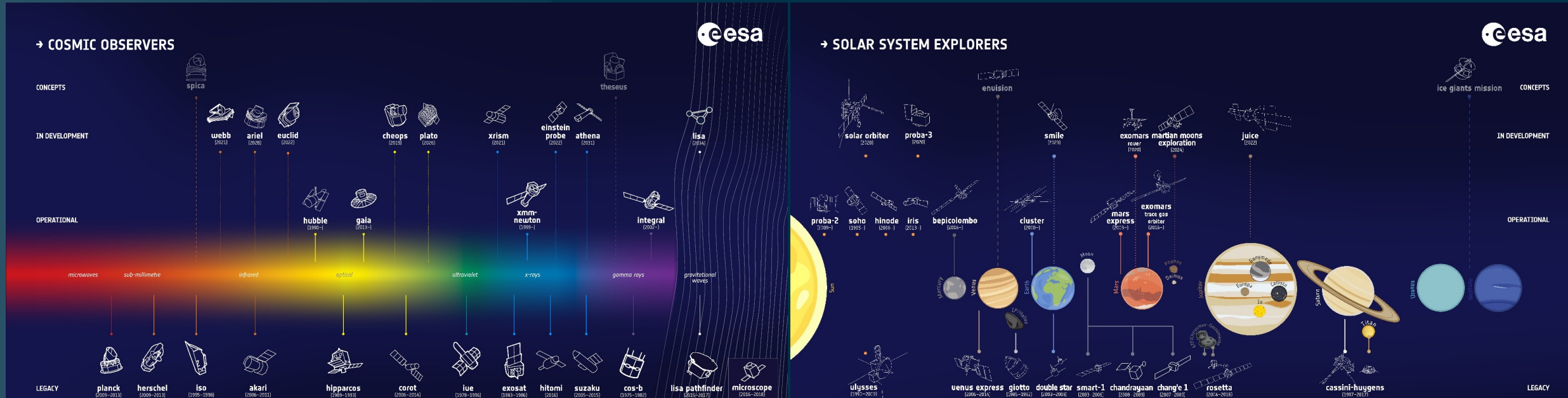
# ESA Datalabs

# Digital Innovation in Space Science

ESA\_Labs – Vicente Navarro

9.05.2023, Venice





ESA Datalabs is available as "Public Moderated Beta"  
If you wish to apply for access, please [submit your request here](#).

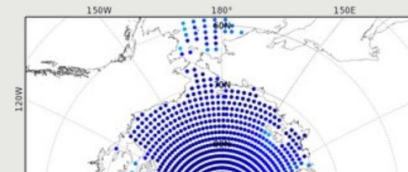


«YOU CAN EITHER MOVE YOUR QUESTIONS  
OR THE DATA. [...] OFTEN IT TURNS OUT  
TO BE MORE EFFICIENT TO MOVE THE  
QUESTIONS THAN TO MOVE THE DATA.»

Jim Gray, eScience: A Transformed Scientific Method


## BRING YOUR QUESTIONS TO THE DATA

There is a new paradigm, opening completely new opportunities for discovery – a data-intensive approach to science. In many domains, we have entered what could be called the golden age of surveys, with several large-scale projects, spanning decades, between finished, ongoing, and planned activities. ESA is responsible, or is a major partner, in several of these initiatives.



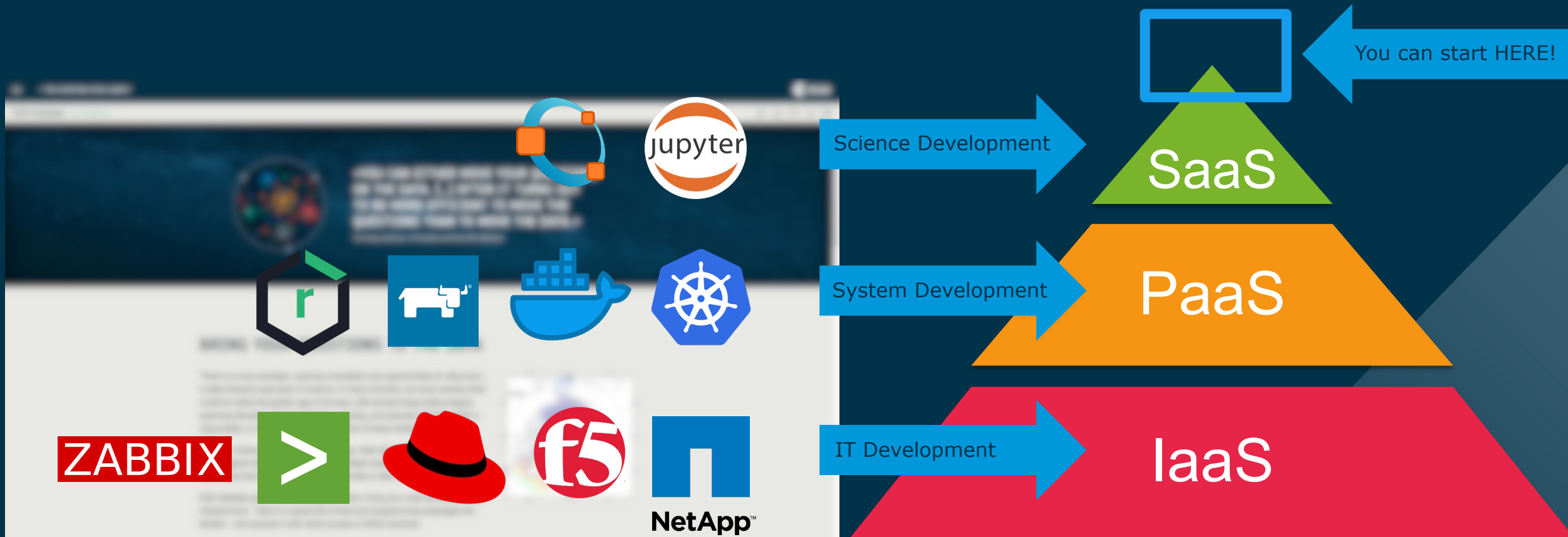
# What does ChatGPT know about ESA Datalabs?

VI Do you know what is "ESA Datalabs"?

 Yes, I'm familiar with ESA Datalabs. ESA Datalabs is a data analytics platform developed by the European Space Agency (ESA) that provides access to a wide range of satellite data and tools for data analysis and processing. The platform is designed to support researchers, scientists, and developers in their work, by providing them with access to data and tools that can help them analyze and make sense of the data they are working with. With its focus on enabling data-driven decision-making, ESA Datalabs aims to provide users with the tools they need to perform advanced data analysis and processing tasks, such as data visualization, data mining, and machine learning. The goal of ESA

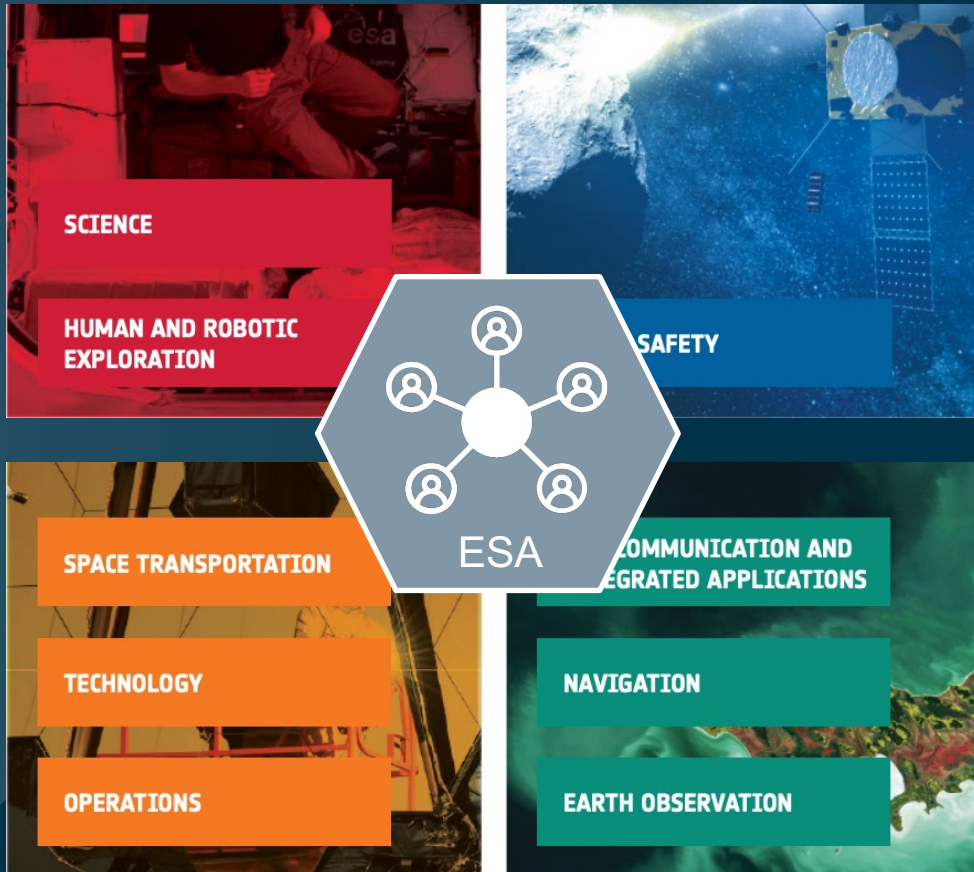
Datalabs is to help researchers, scientists, and developers unlock the full potential of the data they are working with and to support the development of new and innovative applications based on satellite data.

# A Platform Designed to Boost Research Productivity



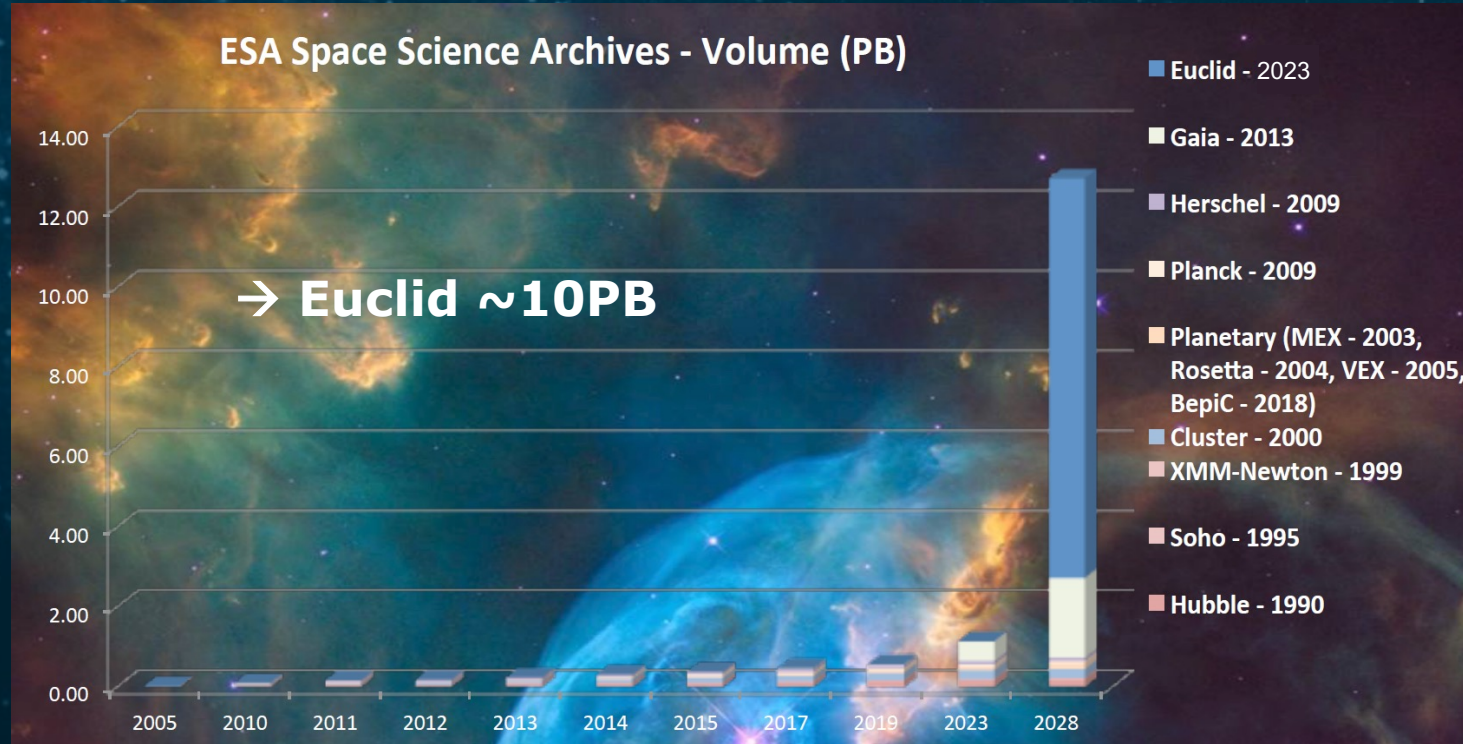
“The tools we use have a profound (and devious) influence on our thinking habits, and, therefore, on our thinking abilities”  
Edsger W. Dijkstra

# SCI-Driven Multi-Domain & Multi-Mission Digital Platform



The background is a complex, abstract digital landscape. It features a series of glowing, wavy green lines that create a sense of depth and movement. Scattered throughout the scene are numerous small, glowing green and white numbers (0-9), some appearing to float or be part of the data flow. The overall color palette is dominated by various shades of green, from light lime to deep forest green, set against a dark, almost black background. The lighting is soft and ethereal, with a central glow that fades towards the edges.

# Data Analysis Innovation

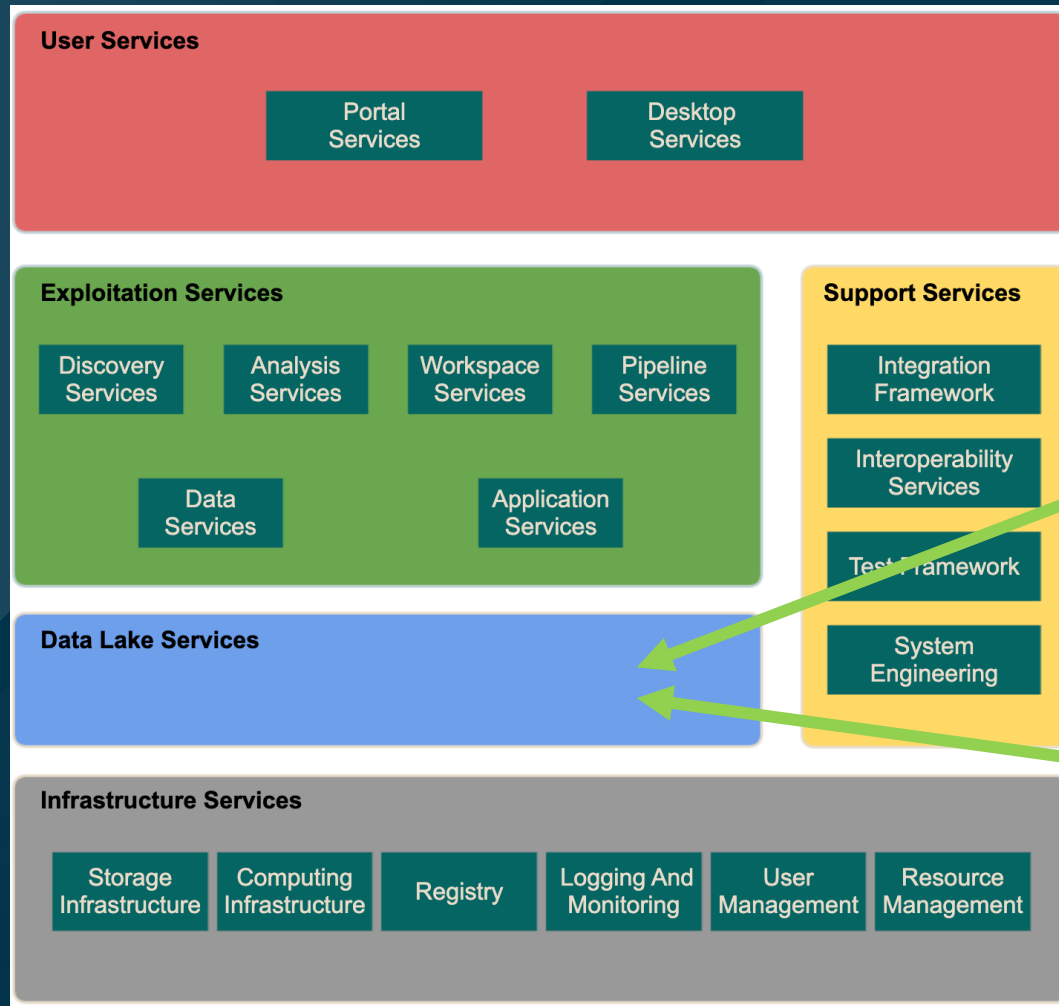


*From bring the data to the user*

*To bring the user to the data*



# Leveraging on existing ESA Data Archives



[datalabs.esa.int](https://datalabs.esa.int)

ESDC
















NAV, HRE, EOP ...

→ THE EUROPEAN SPACE AGENCY esa

ESA Datalabs [0.3.0/BETA] 🔍 📄 📁 🔔 👤

## Create Datalab

Find a datalab in ESA datalabs catalog

 <b>aladin</b> Aladin is an interactive sky atlas allowing the user to visualize digitized astronomical images or full surveys, superimpose entries from astronomical catalogues or databases, and interactively access related data and information from the <i>Simbad database</i> , the <i>VizieR</i> service and other archives for all known astronomical objects in the field	 <b>filezilla</b> FileZilla	 <b>fv</b> FV - An image display and visualization tool for astronomical data
 <b>jl-esdc</b> Jupyterlab ESDC	 <b>jl-euclid-dps</b> Euclid DPS JupyterLab	 <b>jl-herschel</b> Herschel JupyterLab
 <b>jl-juice</b> JupyterLab with JUICE moon coverage tool (0.8.0).	 <b>jl-pangaia</b> PanGaia JupyterLab	 <b>jupyterlab</b> Plain JupyterLab for demonstration of basic functionality.
 <b>jwst</b> Jupyterlab JWST	 <b>jwst-miricle</b> Jupyterlab JWST Miricle	 <b>jwst-nips</b> Jupyterlab JWST NIPS
 <b>jwst-nsrt</b> Jupyterlab JWST NSRT	 <b>qfitsview</b> QFitsView - An image display and visualization tool for astronomical data	 <b>theia-python</b> Theia Python Editor

THE EUROPEAN SPACE AGENCY

ESA Datalabs [0.3.0/BETA]

File Edit View Run Kernel Git Tabs Settings Help

Filter files by name

- / notebooks /
- juice\_kernels
- Moon-coverage\_0.8.0-JUICE\_e... 3 months ago

### Moon-coverage 0.8.0 - JUICE examples

The `moon-coverage` python package is a toolbox to perform surface coverage analysis based on orbital trajectory configuration. Its main intent is to provide an easy way to compute observation opportunities of specific region of interest above the Galilean satellites for the ESA-JUICE mission but could be extended in the future to other space mission.

It is actively developed by the [Laboratory of Planetology and Geodynamics \(CNRS-UMR 6112\)](#) at the University of Nantes (France), under ESA-JUICE founding support.

This notebook provides a list of examples on how-to-use the `moon-coverage` tool version `0.8.0` in the context of the JUICE mission. The complete documentation of this module can be found on [ReadTheDocs](#) for additional details.

```
[1]: from moon_coverage import __version__

if __version__.startswith('0.8.'):
    print(f'Your current version of the moon-coverage tool is {__version__}')
else:
    raise ImportError(
        f'Your version of the moon-coverage tool is {__version__}. '
        'The examples below require the version 0.8.0 and may not work as expected.'
```



THE EUROPEAN SPACE AGENCY

ESA Datalabs [0.3.0/BETA]

File Edit View Run Kernel Git Tabs Settings Help

Output View

J2000 - 21 18 44.613 +00 11 15.28 Fov: 24' X 22'

Search...

Example Query Region + Calibration level (-1) + Filter name (F1130W)

```
[48]: j2a = Jwst.cone_search(coord, radius, ca_level=-1, filter_name='F1130W')
table2a = j2a.get_results()
print(table2a.columns)
print(table2a['observationid', 'dataprodtype', 'calibrationlevel', 'public'])
```

observationid	dataprodtype	calibrationlevel	public
jw01865001001_xx107_00012_miri	image	-1	False
jw01865001001_xx105_00001_miri	image	-1	False
jw01865001001_xx105_00002_miri	image	-1	False
jw01865001001_xx105_00003_miri	image	-1	False
jw01865001001_xx105_00004_miri	image	-1	False
jw01865001001_xx105_00005_miri	image	-1	False
jw01865001001_xx106_00006_miri	image	-1	False
jw01865001001_xx107_00011_miri	image	-1	False
jw01865001001_xx108_00019_miri	image	-1	False
jw01865001001_xx108_00020_miri	image	-1	False

Name	RA (J2000)	Dec (J2000)	Observation ID
J211803.61+000948.7	21h 18m 03.612s	+00° 09' 48.71"	254810010
J211813.90+000806.2	21h 18m 13.902s	+00° 08' 06.28"	254810025



# Desktop Datalabs

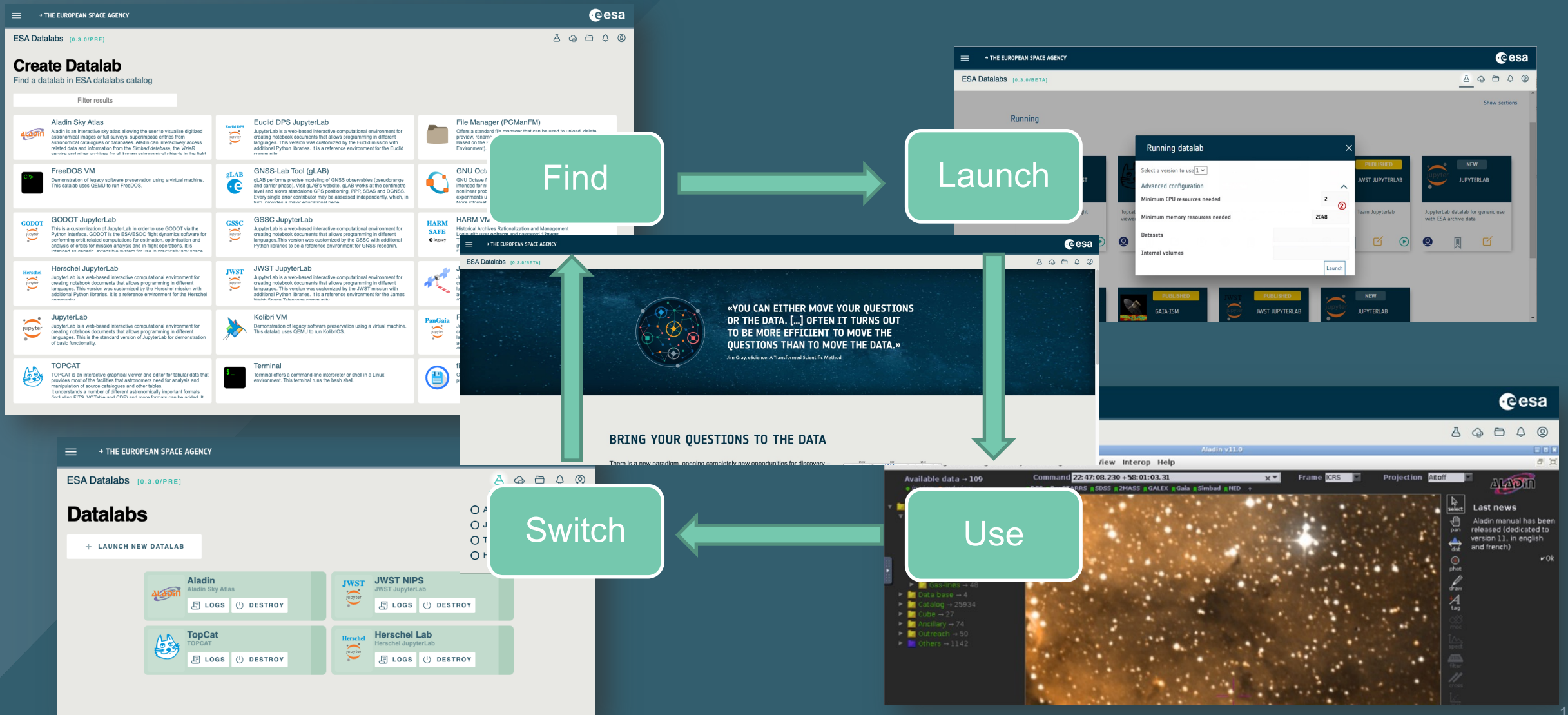


The screenshot shows the ESA Datalabs [0.3.0/BETA] desktop application. The interface includes a menu bar with 'File', 'Views', 'Graphics', 'Joins', 'Windows', 'VO', 'Interop', and 'Help'. A toolbar with various icons is located below the menu. The main workspace is divided into several panels: 'Table List' on the left, 'Current Table Properties' in the center-left, and a 'Sky Plot' window displaying a celestial sphere with coordinate axes. To the right of the Sky Plot is a 'Histogram Plot' window showing a graph with a y-axis from 0 to 1.0 and an x-axis from 0 to 0.8. The ESA logo and a yellow cartoon cat icon are visible in the top right corner of the application window.

The screenshot displays the Aladin v11.0 interface. The top menu bar includes 'File', 'Edit', 'Image', 'Catalog', 'Overlay', 'Coverage', 'Tool', 'View', 'Interop', and 'Help'. The main window shows a star field visualization with a command line at the top: 'Command 22:47:08.230 +58:01:03.31'. A sidebar on the left lists 'Available data → 109' with categories like 'Collections → 27732', 'Image → 5811', 'Gamma-ray → 23', 'X-ray → 44', 'UV → 27', 'Optical → 132', 'Infrared → 135', 'Radio → 92', 'Gas-lines → 48', 'Data base → 4', 'Catalog → 25934', 'Cube → 27', 'Ancillary → 74', 'Outreach → 50', and 'Others → 1142'. The main visualization area shows a star field with a 'DSS2 color' overlay. A toolbar on the right contains icons for 'select', 'pan', 'dist', 'phot', 'draw', 'tag', 'moc', 'spect', 'filter', and 'cross'. A 'Last news' section on the far right mentions 'Aladin manual has been released (dedicated to version 11, in english and french)'. The Aladin logo is in the top right corner.



# Datalab – Utilisation Lifecycle



**Data Volume Catalog**

Planck

**Data Volume Settings**

**Name**  
Planck

**Path to mount in datalab**  
/data/user/ pla\_datalabs

**Details**  
Hide technical details

+ Add to my volumes    × Cancel

# Datalab & Volume Integration



The screenshot displays the ESA Datalabs interface. At the top, it says 'THE EUROPEAN SPACE AGENCY' and 'ESA Datalabs [0.3.0/BETA]'. The interface is divided into several sections:

- File Explorer (Left):** Shows a directory structure under '/ data /'. A folder named 'pla\_legacy' is highlighted with a blue box. A search bar above it says 'Filter files by name'. A table below the search bar shows 'Last Modified' dates, with '5 years ago' visible for the 'pla\_legacy' folder.
- Launcher (Center):** Shows a 'data' directory with options for 'Notebook', 'Console', and 'Other'. Each option has a corresponding icon (Python 3 (ipykernel) for Notebook and Console, and a shell icon for Other).
- Data Volumes (Right):** A dropdown menu titled 'Data Volumes' is open, listing several data volumes with checkboxes. The 'Planck Legacy Archive' is checked and highlighted with a blue box. Other volumes include 'JwstPublic', 'JwstCrds', 'GsscDataProducts', 'EsdcLegacy', 'GsscSatellite', 'spice', 'PSA', and 'Soho'. Each volume has a path listed below it, such as '/media/data/pla\_legacy' for Planck Legacy Archive.

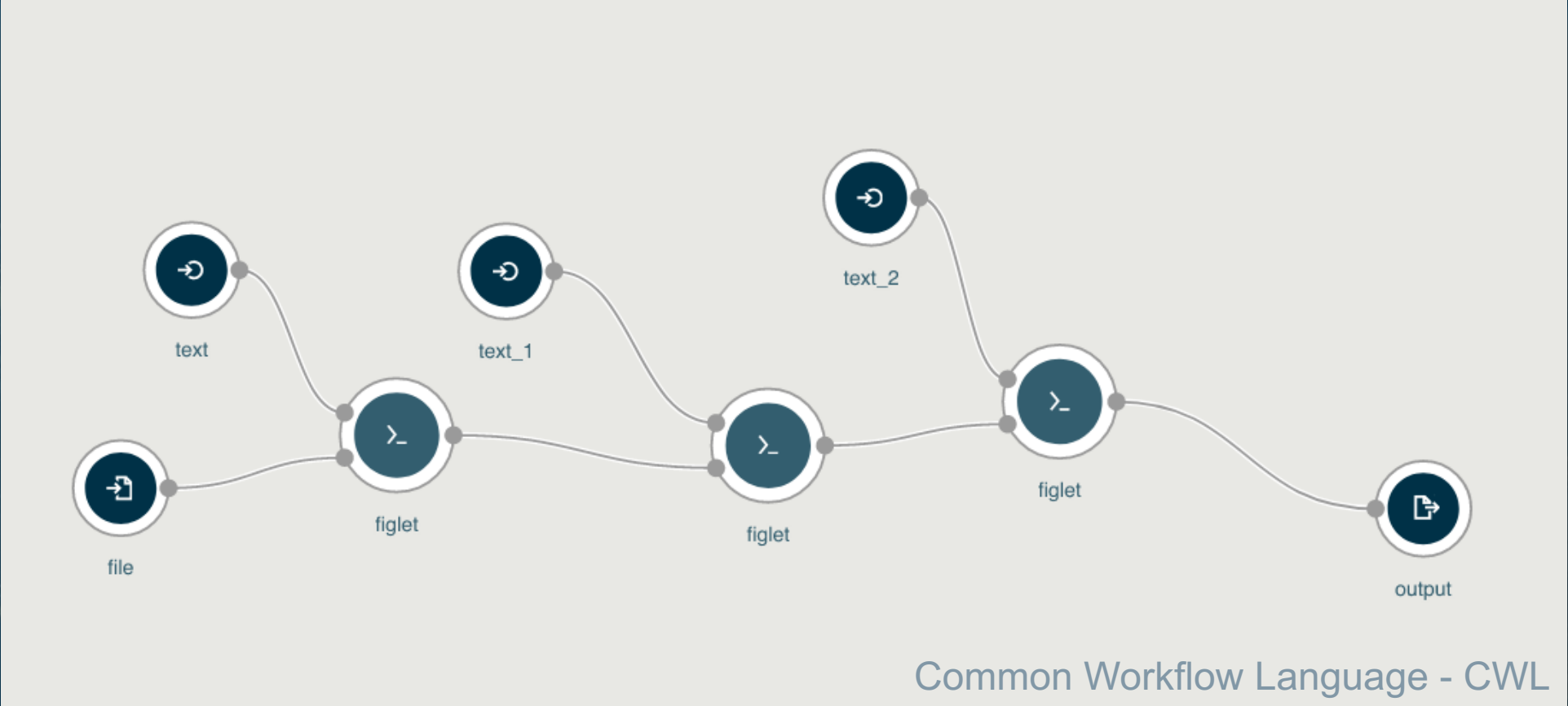
At the bottom of the interface, there is a URL bar showing 'https://datalabs.esa.int/datavol-manager' and a 'Launcher' button.



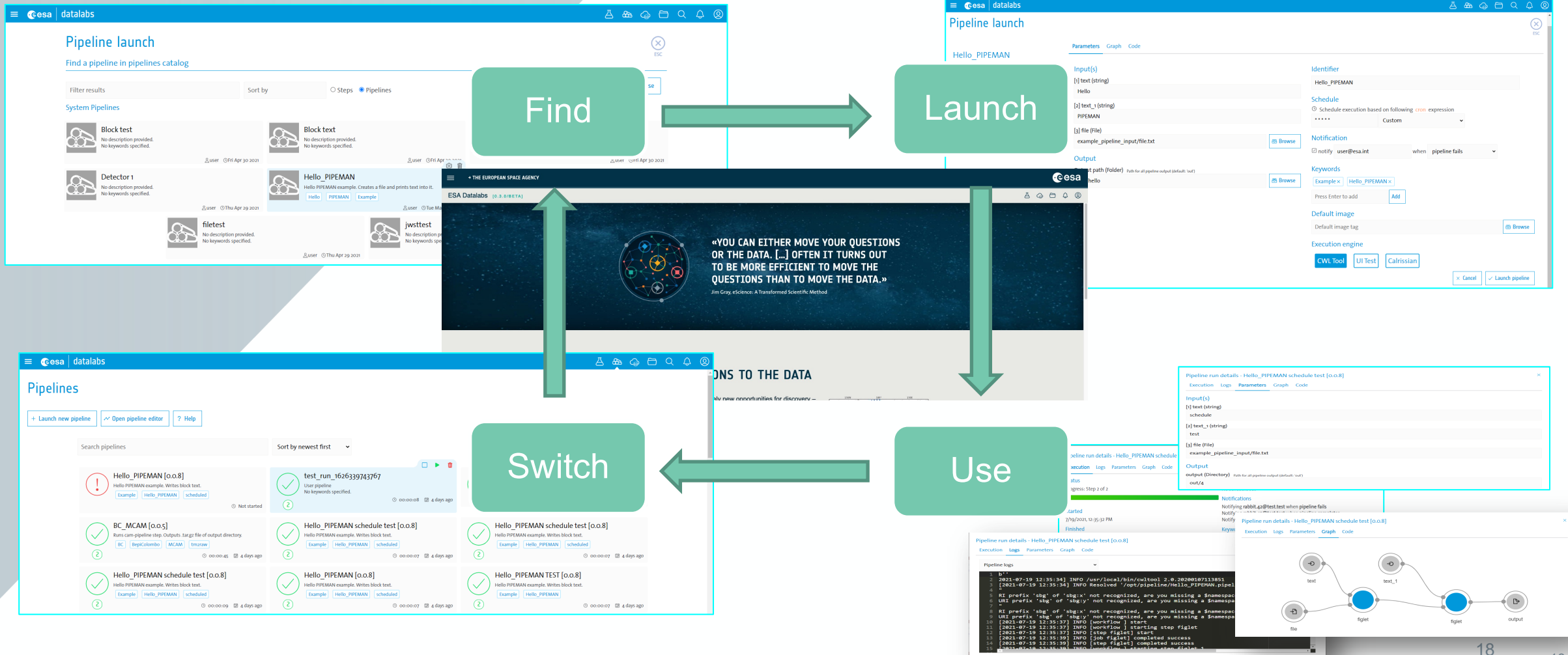
# Data Processing Innovation



# Pipelines on Datalabs: Batch Processing Analysis



# Pipelines Catalogue & Utilisation Lifecycle



# Pipeline Development Environment

The screenshot displays the Pipeline Development Environment (PDE) interface. At the top, it shows the ESA Datalabs logo and the text "THE EUROPEAN SPACE AGENCY". The main workspace is titled "Hello\_PIPEMAN.pipeline.cwl" and shows a pipeline graph with nodes: "file", "text", "figlet", "figlet", "command\_1", "command", and "output". A details panel on the right shows the pipeline's metadata, including Name ("Hello\_PIPEMAN"), Description ("Hello PIPEMAN example. Writes block text."), Version ("0.0.8"), and Keywords ("Example", "Hello\_PIPEMAN"). The panel also includes a "Push as System Pipeline" button. The bottom of the interface shows a validation status: "[01/02/2023, 17:32:27] CWL Valid".

At vero eos et accusamus et iusto odio dignissimos ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga. Et harum quidem rerum facilis est et expedita distinctio. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omni voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum hic tenetur a sapiente delectus, ut aut reiciendis voluptatibus maiores alias consequatur aut perferendis doloribus asperiores repellat.

# Collaboration Innovation

INNOVATION  
SOLUTION  
BRANDING  
IDEAS  
MARKETING  
SUCCESS  
MANAGEMENT  
ANALYSIS

→ THE EUROPEAN SPACE AGENCY

ESA Datalabs [0.3.0/BETA]

File Edit View Run Kernel Git Tabs Settings Help

Filter files by name

Name	Last Modified
data	7 hours ago
my_workspace	7 hours ago
notebooks	7 hours ago
team_workspaces	seconds ago

## Overview

PanGaia simplifies the access, exploration, and clustering analysis of the Gaia DR2 [catalogue](#). This toolkit has been developed with the aim to facilitate the research of astromers who are *familiar* with the Gaia archive and that are interested in Star Formation. However, because of its design and capabilities PanGaia might be useful for a broad audience of researchers interested in exploring large astrometric catalogues. This code closely follows the analysis described by [Canovas et al. 2019](#), where more than 150 potential new members of the  $\rho$ -Ophiuchus Star Forming Region were identified using Machine Learning algorithms applied to the Gaia DR2.

### PanGaia in a Nutshell:

- Data Access:** An [ADQL](#) cone-search (e.g. [link](#)) in the Gaia DR2 archive is performed using the [astroquery.gaia](#) package. Several extra columns are added to the queried table (like e.g. the distance, computed as the inverse of

esa | datalabs

Search datalab catalog

Customize view

### Public

- CYPRESS SEPPTEST-219  
1994.7571690539014  
Test Automated
- CYPRESS SEPPTEST-219  
1111.6848610854097  
Test Automated
- CYPRESS SEPPTEST-219  
1419.7853655034294  
Test Automated
- CYPRESS SEPPTEST-219  
1713.2724871003359  
Test Automated
- CYPRESS SEPPTEST-219  
1673.9987705736833  
Test Automated

See more

### Developed by me

- NEW**  
No description provided

The screenshot shows the Datalabs IDE interface for a project named 'Gaia archeology'. The interface is divided into several sections:

- Left Sidebar:** Contains the 'Gaia archeology' title, a 'jupyter' logo, a 'DRAFT' button, and metadata including 'Modification date: 15/06/2022 21:18', 'Datalab version: 1.4.0', 'Author: sciapps developer', 'Popularity: ☆0', and a link '53a11c69-4982-4bc7-8115-a83...'. A 'Delete' button is also present.
- Top Navigation:** Includes 'Metadata', 'Testing', and 'Visibility' tabs.
- Main Content Area:**
  - A warning message: 'Use this form to share with some users as a private access, only them will have access to your lab. Or you can share with everyone using the publish flow. Your datalab will be reviewed by a moderator'.
  - 'Build date: 15/06/2022 21:19' and 'Build version: 1.4.0-0' with a 'SUCCESS' status.
  - 'Sharing audience' options: 'Public, everyone can use it' (selected) and 'Private, specify a list of users'.
  - 'Publish' section with a 'License' dropdown menu open, showing a list of licenses including 'Apache License 2.0', 'BSD 3-Clause No Nuclear Warranty', 'Creative Commons Attribution Non Commercial 3.0 Germany', 'GNU General Public License v2.0 only', 'GNU General Public License v3.0 or later', 'GNU Library General Public License v2.1 or later', 'MIT License', 'European Space Agency' licenses, and 'Others'.
  - 'Export' section with a 'Submit' and 'Export' button.
  - 'History' section showing build logs: '15/06/2022 21:19: Build started' and '15/06/2022 21:21: Build ended with status BUILD\_SUCCESS'.



# The Future?



# Platform and Community Release Roadmap

**Q1** Private Beta  
2022

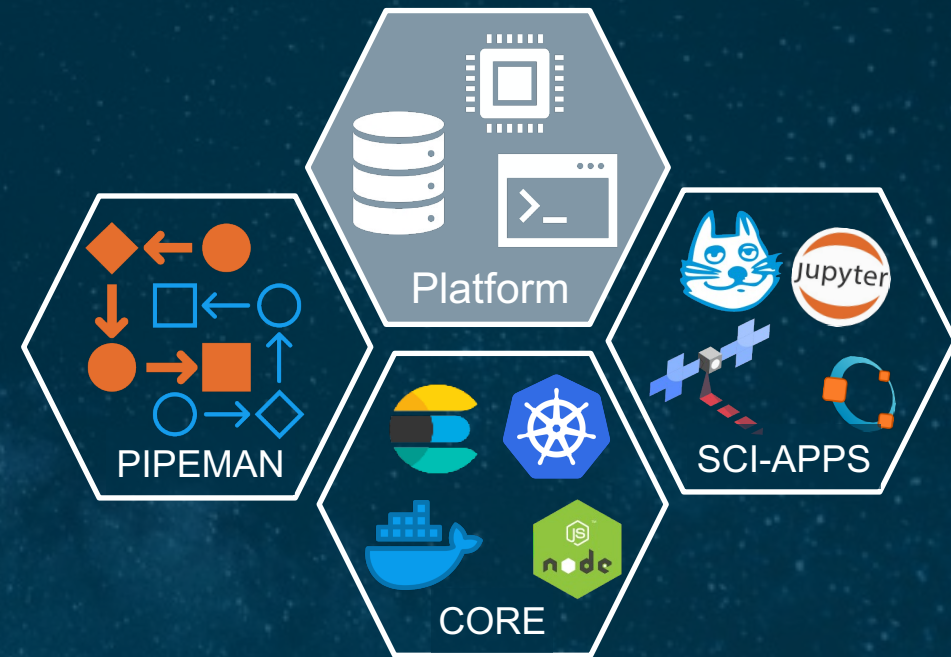


**Q1** Public Beta  
Moderated  
2023



**Q4** Public Beta  
Unmoderated  
2023

...



**218** Beta Users

**168** Waiting List

ESA Datalabs Workshop, ESAC 24-25 Nov 2022

# Interoperability Across IT Infrastructures PoC

→ THE EUROPEAN SPACE AGENCY

ESA Datalabs [00.0/BETA] Current datazone: SCI-CLOUD

### Datalabs

Manage your running datalabs

**situ** jupyterlab [Delete]

**jupyterlab-cuda** [Delete]

**jupyterlab** [Delete]

→ THE EUROPEAN SPACE AGENCY

ESA Datalabs [00.0/BETA] Current datazone: GO.ESA

### Datalabs

Manage your running datalabs

**x-ds9** [Delete]

**jl-juice** [Delete]

**Datazones**

- SCI-CLOUD
- GO.ESA
- EO-CLOUD

<https://datalabs.esa.int>

<https://datalabs.esa.int>

→ THE EUROPEAN SPACE AGENCY

ESA Datalabs

## Create Datalab

Find a datalab in ESA datalabs catalog

**jupyterlab-cuda**  
JupyterLab for demonstration of GPU functionality (CUDA 10.0).

→ THE EUROPEAN SPACE AGENCY

ESA Datalabs [0.3.0/BETA]

File Edit View Run Kernel Git Tabs Settings Help

Filter files by name

/ notebooks /

Name	Last Modified
example.ipynb	2 months ago
gpu_confirmed.ipynb	2 months ago
sorting.ipynb	2 months ago

sorting.ipynb

Make sure the CUDA kernel is chosen before running this notebook (top left corner)

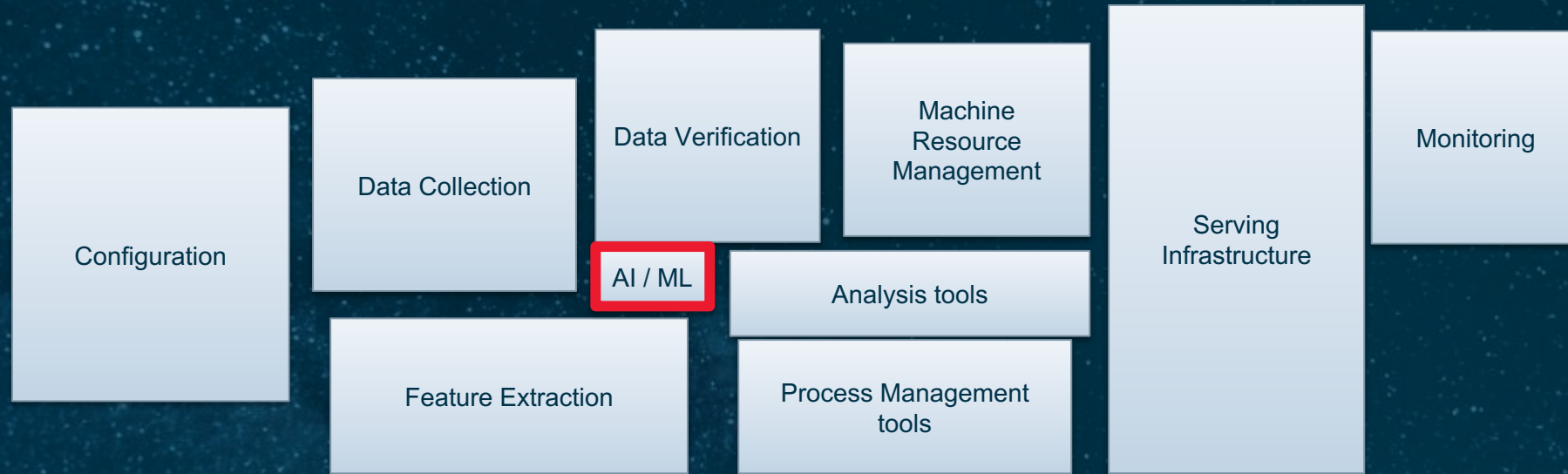
```
[ ]: import tensorflow as tf
import sys
from platform import python_version
from tensorflow.python.client import device_lib
import os
```

First way to check if the GPU is visible in the datalab is to call the NVIDIA specific method. If the GPU is visible a table showing some statistics will appear. If something went wrong you'll get an error.

```
[ ]: !nvidia-smi
```

Time measured	numpy + no gpu	numpy + gpu	cupy + gpu
Total execution time	≈ 44m47s	≈ 14min26s	≈ 3min2s

# Data-Intensive Approach to Science Exploitation



\*Source: Hidden Technical Debt in Machine Learning Systems

## ESA Datalabs - Multi Domain

ESA Datalabs [0.3.0/BETA]

### Create Datalab

Find a datalab in ESA datalabs catalog

Filter results

- aladin**: Aladin is an interactive sky atlas allowing the user to visualize digitized astronomical images or full surveys, superimpose entries from astronomical catalogues or databases, and interactively access related data and information from the Simbad database, the XGA/R service and other archives for all known astronomical objects in the field.
- filezilla**: FileZilla
- fv**: FV - An image display and visualization tool for astronomical data
- jl-esdc**: Jupyterlab ESDC
- jl-euclid-dps**: Euclid DPS JupyterLab
- jl-herschel**: Herschel JupyterLab
- jl-juice**: JupyterLab with JUICE moon coverage tool (0.8.0).
- jl-pangaia**: PanGaia JupyterLab
- jupyterlab**: Plain JupyterLab for demonstration of basic functionality.
- jwst**: Jupyterlab JWST
- jwst-miracle**: Jupyterlab JWST Miricle
- jwst-nips**: Jupyterlab JWST NIPS
- jwst-nsrt**: Jupyterlab JWST NSRT
- qfitsview**: QFitsView - An image display and visualization tool for astronomical data
- theia-python**: Theia Python Editor

[datalabs.esa.int](https://datalabs.esa.int)

## GSSC Now - Navigation Science Domain

ESA GNSS Science Support Centre | GSSC Now [BETA]

### Running

- Datablab GSSC Lab**
- Datalab Spatial Interpolation of VTEC**
- UbiSAP NSA2**: Datalab UbiSAP RINEX PPK Processor

### Featured

- CamalioT**: CamalioT Troposphere Modelling
- gLAB**
- GMLD - Ephemeris Parameters Prediction**
- GMLD - Fast and Long Term Corrections (EGNOS SBAS Messages) Prediction**
- GMLD - TEC Map Prediction**
- GSSC PVT & DOC/DOP Map**
- GSSC Lab**
- Spatial Interpolation of VTEC**

[gssc.esa.int/now](https://gssc.esa.int/now)

# Interoperability Across Exploitation Platforms



→ THE EUROPEAN SPACE AGENCY esa

ESA Data Discovery Portal [0.4.0/ALPHA]

Type of Asset  
 Datalab  
 Dataset

Properties  
 Query Tool  
 Visualization Tool  
 Analysis Tool  
 Data Volume

Domain  
 Earth Observation (1,766)  
 Navigation (21)  
 Space Science (21,818)

Instrument

Thematic Area

Mission  
 Astronomy (15,365)  
 Gaia (5)  
 Herschel (723)  
 Hubble Space Telescope (9,745)  
 ISO (1,515)  
 Lisa Pathfinder (6)  
 Planck (4)  
 XMM-Newton (3,367)  
 Earth Observation Satellite (1,088)  
 ADEOS (1)

Type in your query...

<b>aladin</b> Aladin is an interactive sky atlas allowing the user to visualize digitized astronomical images or full surveys, superimpose entries from astronomical catalogues or databases, and interactively access related data and information from the <i>Simbad</i> database, the <i>VizieR</i> service and other archives.	<b>file-browser</b> Web File Browser which can be used as a middleware or standalone app.
<b>filezilla</b> FileZilla - FileZilla - FileZilla - FileZilla - FileZilla - FileZilla - FileZilla	<b>fv</b> FV - An image disp
<b>jl-esdc</b> Jupyterlab ESDC - Jupyterlab ESDC - Jupyterlab ESDC - Jupyterlab ESDC	<b>jl-euclid-dps</b> Euclid DPS Jupyterlab - Euclid DPS Jupyterlab
<b>jl-herschel</b> Herschel JupyterLab - Herschel JupyterLab - Herschel JupyterLab - Herschel JupyterLab	<b>jl-juice</b> JupyterLab with JUICE
<b>jl-pangaia</b> PanGaia JupyterLab - PanGaia JupyterLab - PanGaia JupyterLab - PanGaia JupyterLab	<b>jl-xmm-sas</b> Jupyterlab XMM SAS - Jupyterlab XMM SAS
<b>jupyterlab</b>	<b>jupyterlab-cuda</b>

**aladin** Version 0.3.1  
Launch datalab

Aladin is an interactive sky atlas allowing the user to visualize digitized astronomical images or full surveys, superimpose entries from astronomical catalogues or databases, and interactively access related data and information from the *Simbad* database, the *VizieR* service and other archives for all known astronomical objects in the field.





## The ESA Space Science Exploitation Platform

- SCI Data available for researches to work on it, made easy

## Increase Space Science Operations Efficiency

- Reusable for fast implementation of Scientific Processing Pipelines
- Reusable for fast implementation of Scientific Analysis and Visualisation Tools

## Enable Collaboration and Open Science

- Share complex processing tools and data with your team (ala JWST)
- Share your contributions with the community in SCI 's AppStore

# Thank You!



[datalabs.esa.int](https://datalabs.esa.int)

Acknowledgements: The diverse skillset required for the development of ESA Datalabs and GSSC Now, requires involvement of different multi-disciplinary groups. Hence, we would like to thank our industrial partners, Edisoft, Uninova, CGI, ACRI-ST, GMV and Ideorum. Thanks also to the Science and Operations Technical IT Unit at ESAC, the ESA Datalabs User's Group, the members of ESA Space Science missions, and the ESAC Science Data Centre.

# Back-Up Slides

---



# Architecture – Technology Stack

