

ESA Datalabs

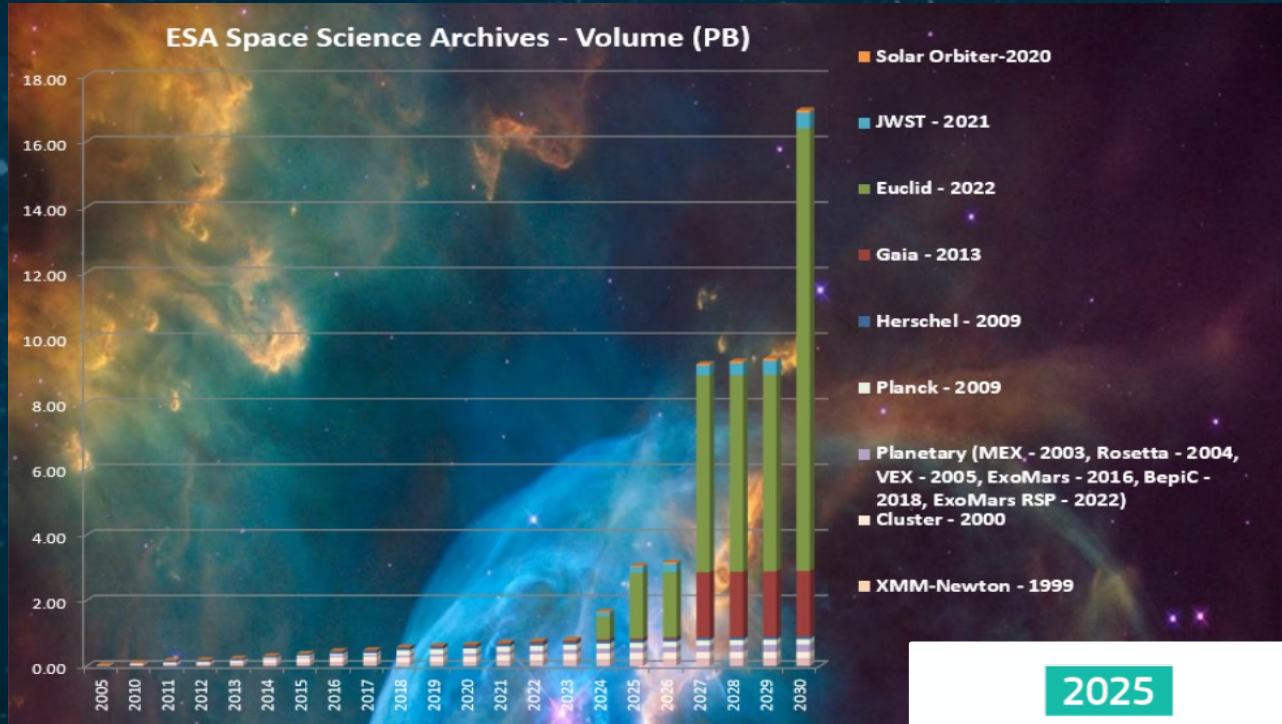
Digital Innovation in Space Science and Beyond

Vicente Navarro

IHDEA Meeting, JHU/APL in Laurel MD & Virtual - Oct 12-13, 2023

12/10/2023

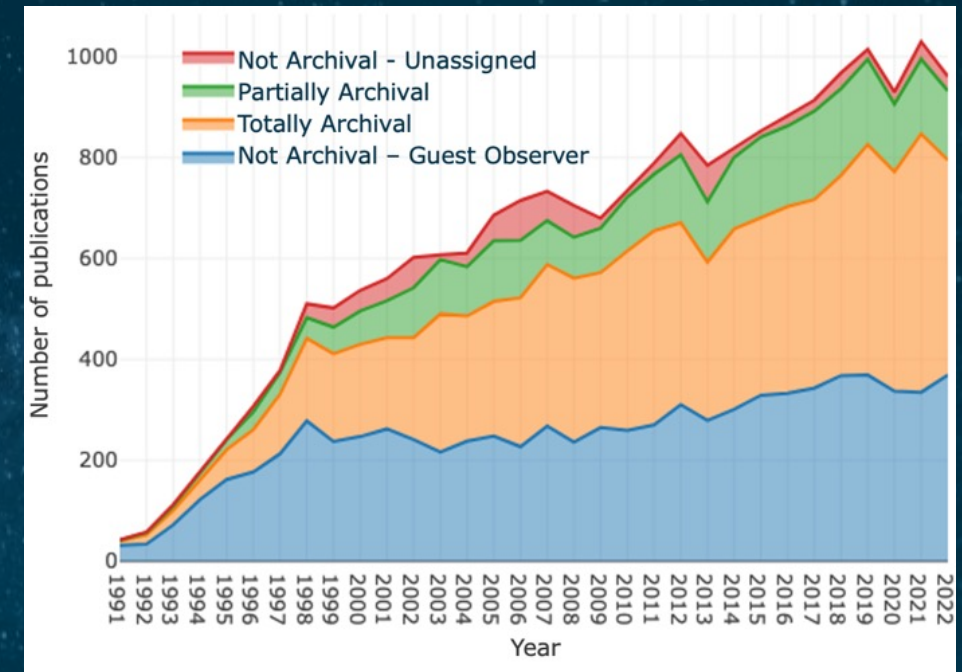
Data Driven Science for a Data Informed World



Data → Insights

NASA & ESA

Stored on 512 GB tablets, it would form five towers reaching the moon



ESAC Science Data Center



→ THE EUROPEAN SPACE AGENCY

ESA Datalabs [0.6.1/BETA]

Log in

ESA Datalabs is available as "Public Moderated Beta"
If you wish to apply for access, please **submit your request here**.

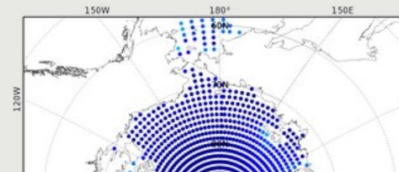


TO BE MORE EFFICIENT TO MOVE THE QUESTIONS THAN TO MOVE THE DATA.»

Jim Gray, eScience: A Transformed Scientific Method

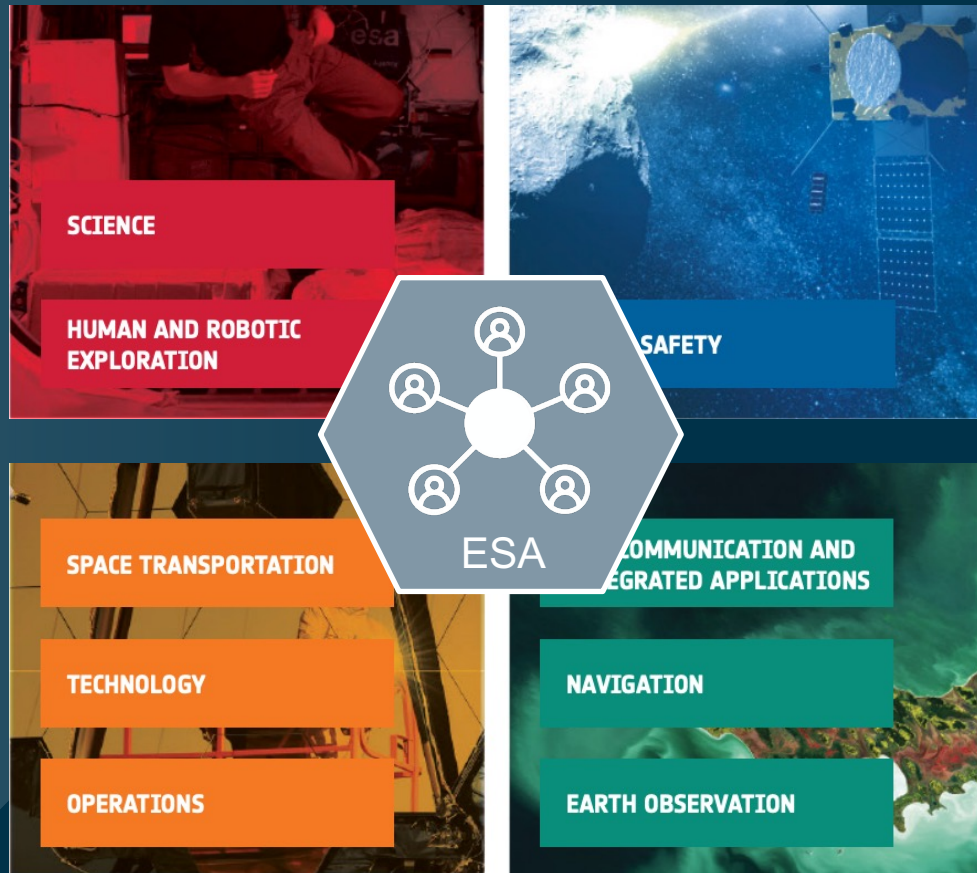
BRING YOUR QUESTIONS TO THE DATA

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.

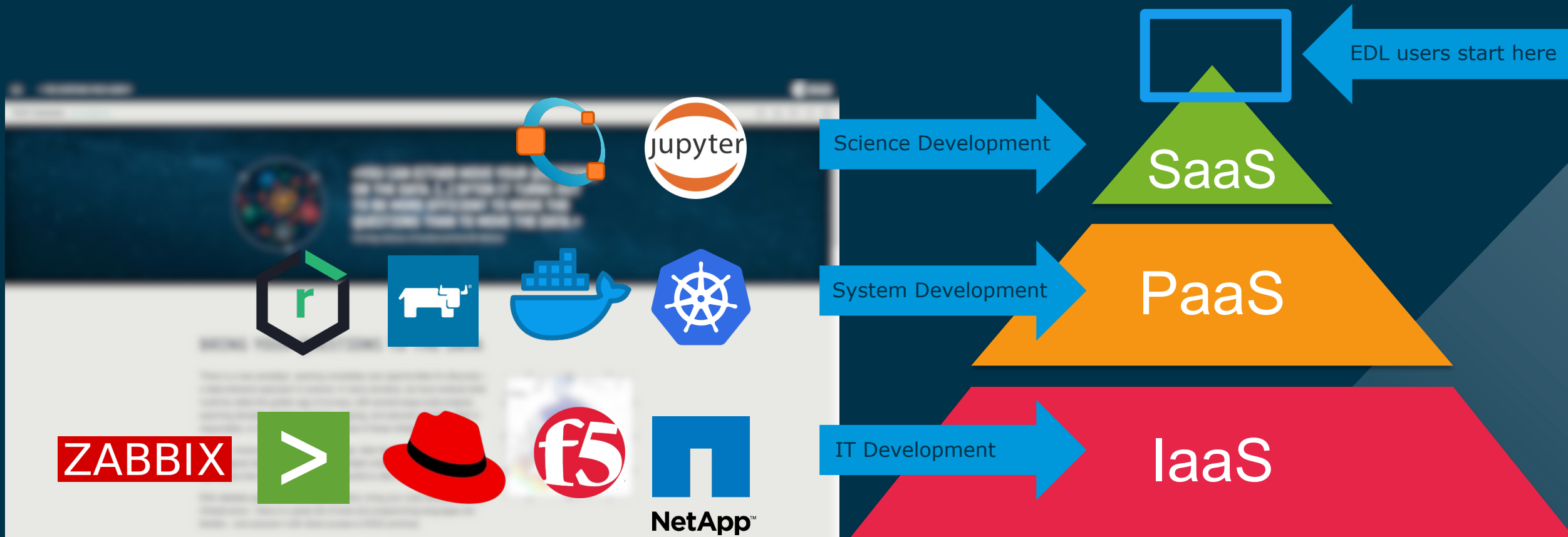


datalabs.esa.int

Multi-Domain & Multi-Mission Digital Platform

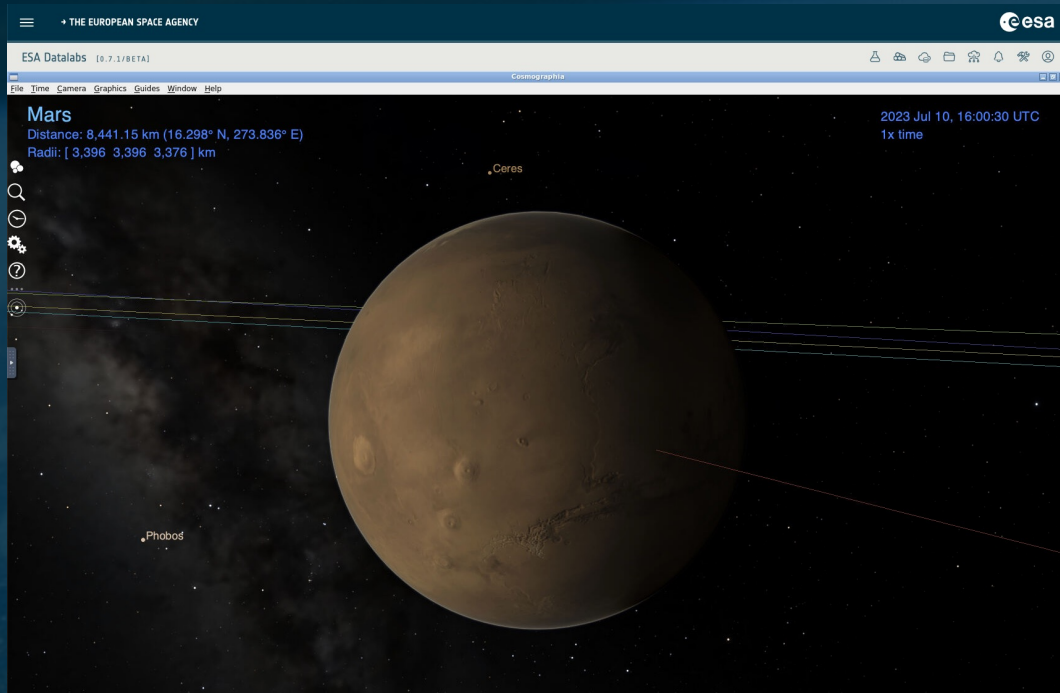


A Platform Designed to Boost Research Productivity

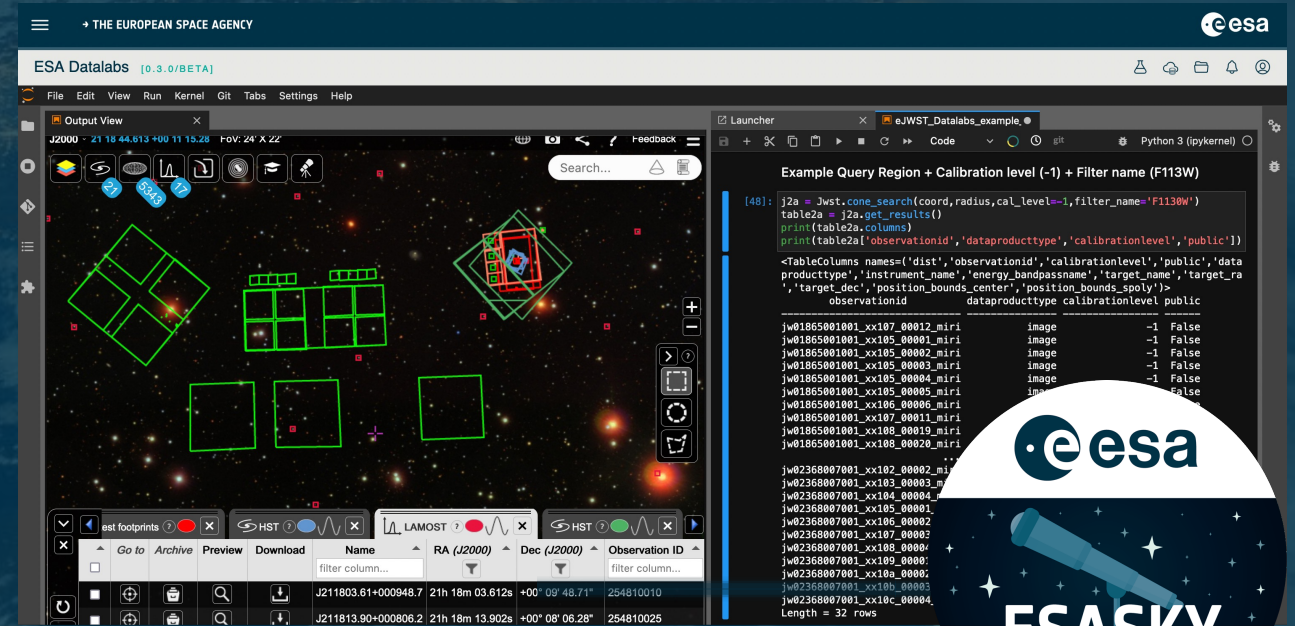


Software as a Service solution, low-code platform based on highly scalable microservice architecture

Datalabs for Interactive Analysis

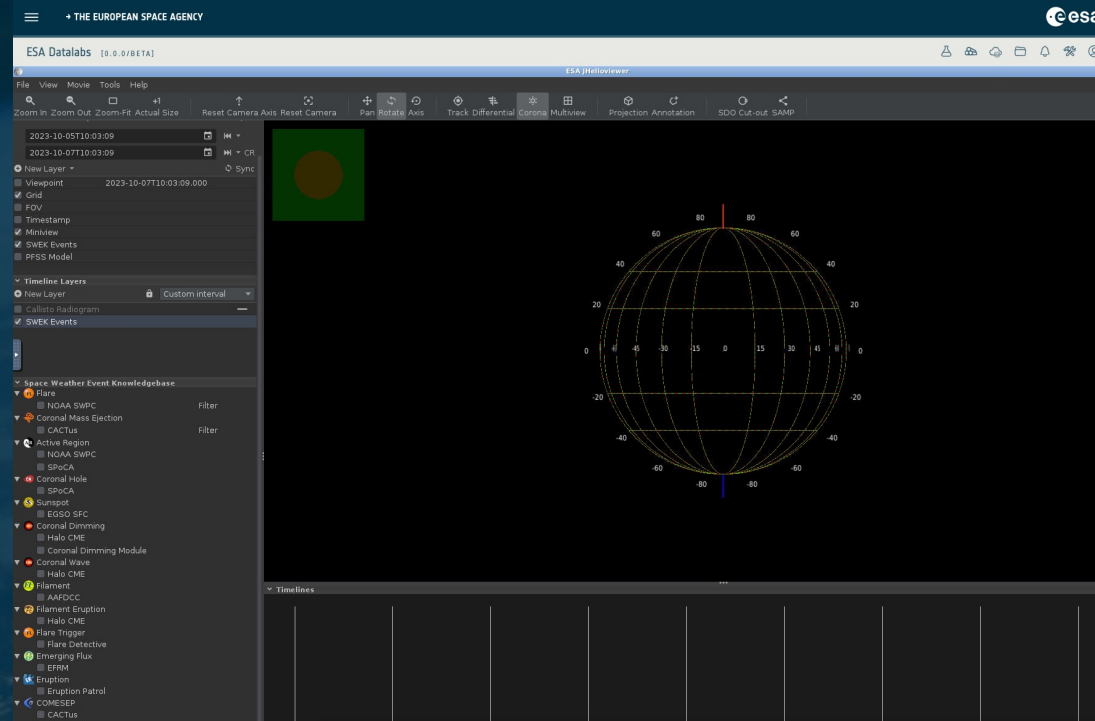


Desktop-Based



Web-Based





J-HelioViewer

JupyterLab Helio

The screenshot shows a JupyterLab notebook interface. At the top, it says 'ESA Datalabs [0.0.0/BETA]'. Below that is a notification banner: 'ESA Datalabs will be upgraded to version 0.9.0 on Monday, September 9 starting from 10 am until 1 pm EST. During this time you may experience unexpected behaviour while using the platform. Should you have any questions please contact the ESA Datalabs Service Desk.' The notebook content includes:

- Tutorial to plot the Solar Orbiter orbit position**
- Import necessary libraries**


```
[2]: import matplotlib.pyplot as plt
from matplotlib import dates
import pandas as pd
from astropy import units as u
from sunpy.coordinates import frames, get_horizons_coord
from astropy.coordinates import SkyCoord
from astropy.time import Time
import astropy
from datetime import datetime, date, time
```
- Define the timing of interest**






```
[4]: date = '2023-07-15T14:00:00'
# Convert string to datetime format
date1 = pd.to_datetime(date)
times = Time(date1)
# print day, month and year separately from the to_datetime output
# print("Day: ", date1.day)
# print("Month", date1.month)
# print("Year", date1.year)
# print(date1.second)
```
- Load the kernels**

```
[6]: kernels = astropy.registry.get_kernels("solar orbiter", "predict")
```
- Load the coordinates for the objects to be plotted (here, SO, Earth and the Sun)**

```
[9]: solo_coords = astropy.generate_coords("SOLAR ORBITER", times).heliographic_stonyhurst
# solo_coords[0]
earth_coords = astropy.generate_coords("earth", times).heliographic_stonyhurst
sun_coords = astropy.generate_coords("sun", times).heliographic_stonyhurst
```
- Generate the figures**


















→ THE EUROPEAN SPACE AGENCY 

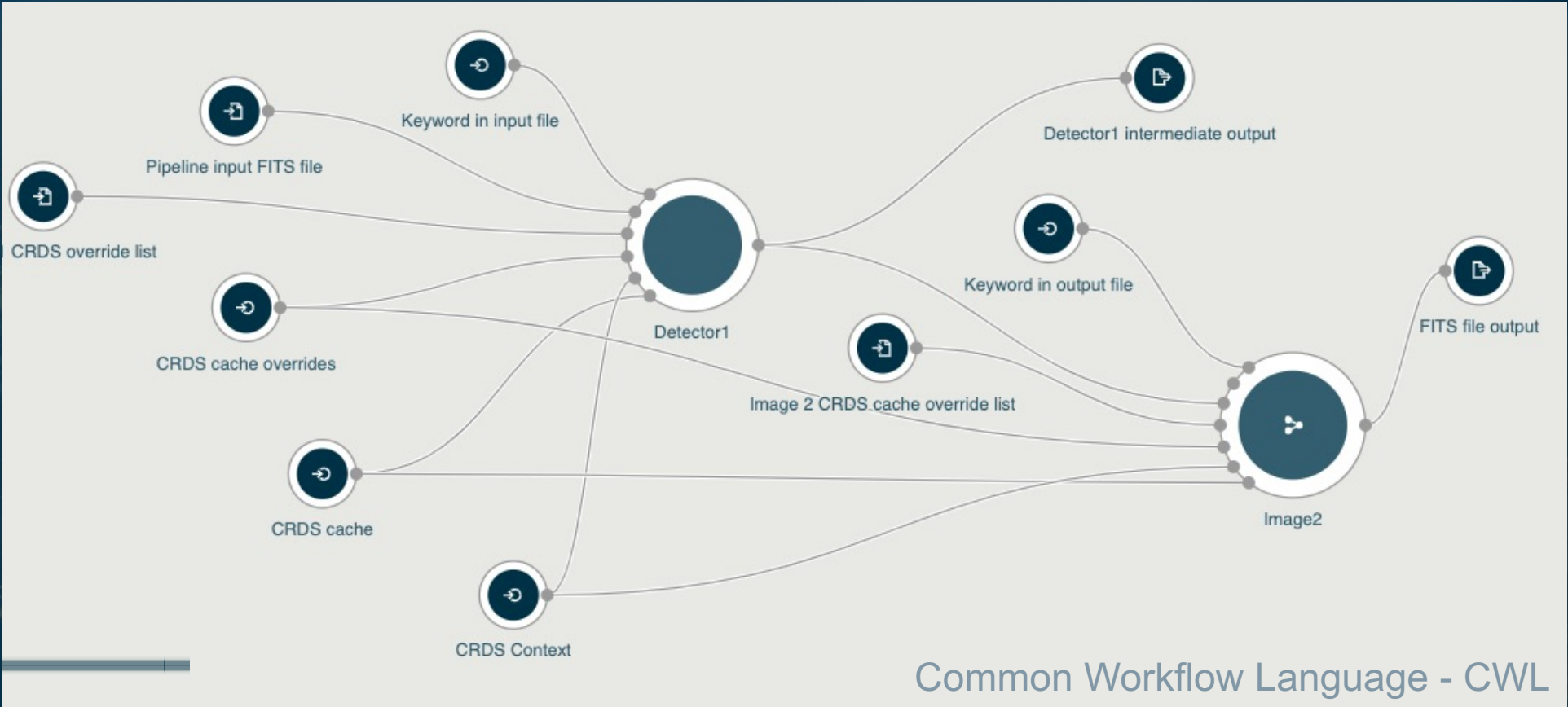
ESA Datalabs [0.3.0/BETA]     

Create Datalab

Find a datalab in ESA datalabs catalog

 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 <i>Simbad database</i> , the <i>VizieR</i> 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-miricle Jupyterlab JWST Miricle	 jwst-nips Jupyterlab JWST NIPS
 jwst-nrst Jupyterlab JWST NSRT	 qfitsview QFitsView - An image display and visualization tool for astronomical data	 theia-python Theia Python Editor

Pipelines for Batch Processing Analysis



→ THE EUROPEAN SPACE AGENCY

ESA Datalabs [0.7.0-3-G7FECB2D]

Pipeline launch

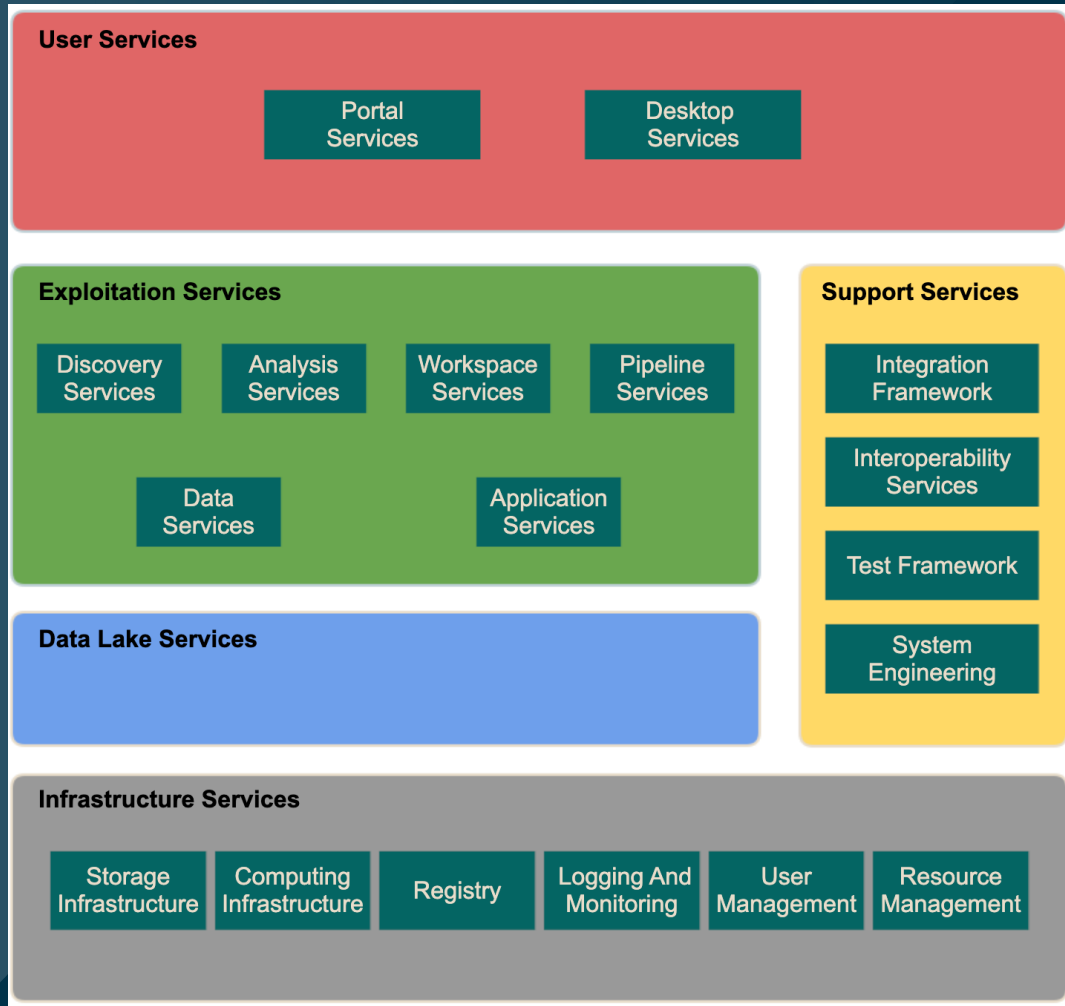
Find a pipeline in pipelines catalog [Browse User pipeline](#)

Filter results Sort by last modified Steps Pipelines

System Pipelines

BC_MCAM This pipeline contains 2 steps: tm2raw_mcam and pds4_packager. bc example 👤 jkuhi 🕒 Wed May 10 2023	JWST This pipeline runs 'jwst.pipeline.Detector1Pipeline' and then processes the output on 'jwst.pipeline.Image2Pipeline' on 'strun' command. example jwst 👤 jkuhi 🕒 Wed May 10 2023	Image2 This pipeline runs 'jwst.pipeline.Image2Pipeline' on 'strun' command. example jwst 👤 jkuhi 🕒 Wed May 10 2023
Detector1 This pipeline runs 'jwst.pipeline.Detector1Pipeline' on 'strun' command. example jwst 👤 jkuhi 🕒 Wed May 10 2023	timestamp This one-step pipeline has two inputs. The first input is a text file. The output of the pipeline is the copy of the input file with current time appended. The second input to the pipeline is an integer that specifies sleep time. The pipeline step sleeps the specified number of seconds. example 👤 jkuhi 🕒 Wed May 10 2023	scatter-gather This pipeline contributed by INTEGRAL demonstrates the CWL scatter feature. In the scattering pipeline a step produces an array output of files and the next step is spawned for each array element. The entry-point file is 'scatter-gather.cwl'. The 'data' input directory takes text files. This is an advanced pipeline. example scatter-gather 👤 jkuhi 🕒 Wed May 10 2023
cdr-demo This pipeline chains two 'command' steps. There is intermediate output after the first step and the final output after the second. example 👤 jkuhi 🕒 Wed May 10 2023	Hello_PIPEMAN This pipeline has two steps. The input to the pipeline is a text file. The two steps append "Hello" and "PIPEMAN" in block letters to the copy of the input file. The output of the pipeline is a copy of the input file with the appended "Hello PIPEMAN". This example uses a 'figlet' pipeline step that needs to exist as a step in Pipeline Catalogue. 'figlet' is a unix command that writes text in block letters. example 👤 jkuhi 🕒 Wed May 10 2023	simple Very basic pipeline. example 👤 jkuhi 🕒 Wed May 10 2023

Innovative Architecture – Technology Stack



For Problem Domain Science End-Users



Find

Launch

Switch

Use

«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.»

BRING YOUR QUESTIONS TO THE DATA

Running datalab

Available data → 109

Command 22:47:08.230 + 58:01:03.31

Frame ICRS Projection Altoff

Aladin v11.0

Last news

Aladin manual has been released (dedicated to version 11, in english and french)



For Problem Domain Science End Users++

The screenshot displays the ESA Datalabs creator interface. At the top, it says "THE EUROPEAN SPACE AGENCY" and "esa". The main header is "ESA Datalabs [0.8.0/BETA]". A sidebar on the left shows the ESA logo and metadata: "Modification date: n/a", "Datalab version:", "Author: Vicente Navarro", and "Popularity: ☆0". The main content area has a welcome message: "Welcome to the datalab creator! To start please select from the pulldown menu below the type of datalab you want to create and select the files (e.g. Jupyter notebooks) you want to use. You can use files from a git repository (public URL, no login required), your ESA Datalabs workspace or your own computer. To include your favourite python packages upload a requirements.txt file listing all our packages. Click on next to continue and add metadata for your datalab. If you include a meta-data.yaml file we will prepare your lab based on it." Below this is a form with fields: "Select your datalab type" (with a dropdown menu showing "Jupyter Lab", "VNC based X server", and "Git repository"), "Where are your files?", "Enter public Git repository", and "Enter branch to use" (with "master" entered). At the bottom are "Cancel" and "next" buttons. On the right, a "License" section is visible with a dropdown menu open, listing various licenses under categories: "Popular Licenses", "European Space Agency", and "Others". The "Popular Licenses" list includes 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, and MIT License. The "European Space Agency" list includes several versions of the ESA Community License and Public License. The "Others" list includes 3dfx Glide License, AMD's plpa_map.c License, and ANTI-R Software Rights Notice. A "Visibility" section is partially visible above the license dropdown, showing options for "Public, everyone can use it" and "Private, specify a list of users". A build status bar at the bottom right shows "22 21:19 Build version: 1.4.0-0 SUCCESS" and "21:19: Build started", "21:21: Build ended with status BUILD_SUCCESS".

768 Registered Users

584 Beta Users


ESA Datalabs – JWST Workshop, Heidelberg 4 - 8 Sept 2023


ESA Datalabs – Euclid Workshop, Copenhagen 21 - 22 Jun 2023


ESA Datalabs Workshop – ESAC 24 - 25 Nov 2022

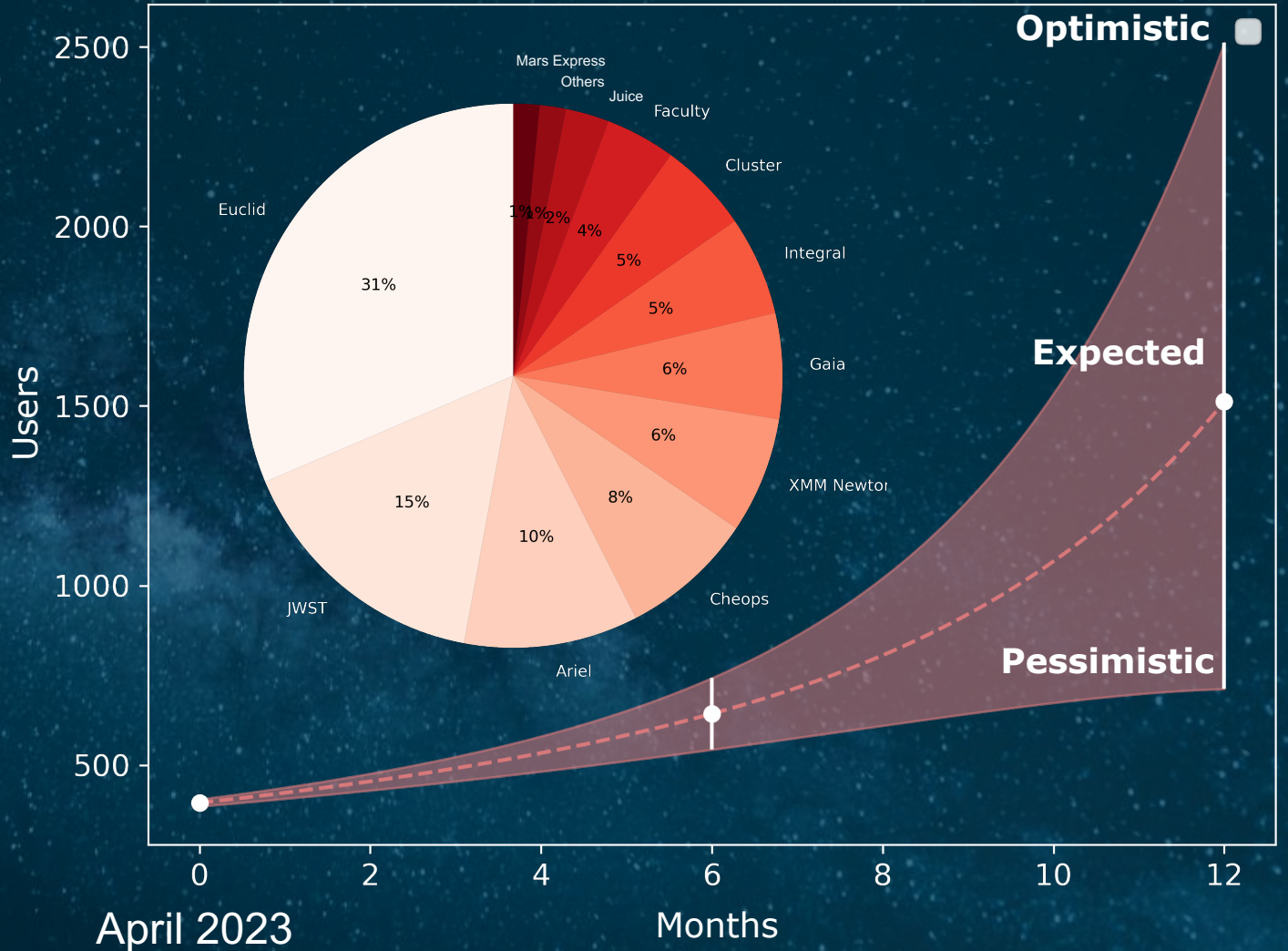


Evolution

 **Optimistic:** onboarding of all ESA Space Science interested missions and collaboration opportunities

 **Expected:** onboarding of Space Science Euclid and JWST missions as currently planned

 **Pessimistic:** onboarding of existing registration requests



ESA Datalabs Community - ESA



Platforms

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

ESDC - Home

ESAC SCIENCE DATA CENTRE

Monthly Users (*) 24 676

Monthly downloaded (*) 123.1 TB

Archive Total Size 608.0 TB

THE EUROPEAN SPACE AGENCY

earth online

EARTH ONLINE

MISSIONS DATA NEWS EVENTS TOOLS

THE EUROPEAN SPACE AGENCY

Copernicus Open Access Hub

Welcome to the Copernicus Open Access Hub

38,892 jobs published in the last 24h

338,550 downloads in the last 24h

THE EUROPEAN SPACE AGENCY

GSSC Now

Fostering science collaboration in GNS5

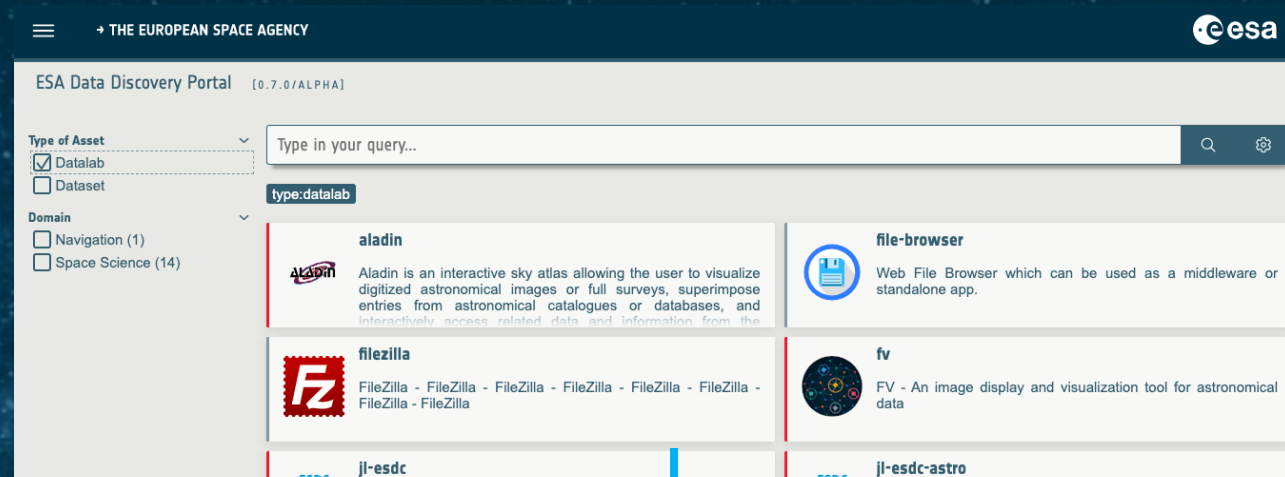
THE EUROPEAN SPACE AGENCY

HRE Data Archive - Human and Robotic Exploration Data Archive

Antares Orion SpaceX Mars Moon Star Line

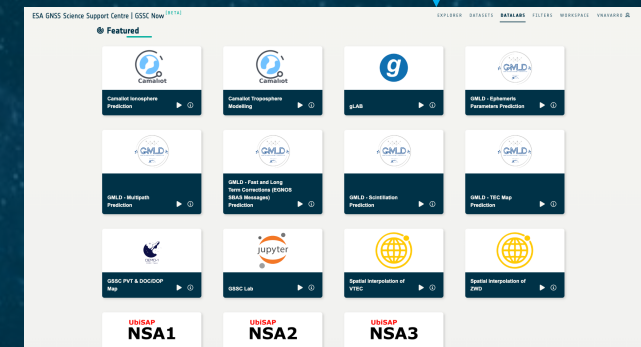
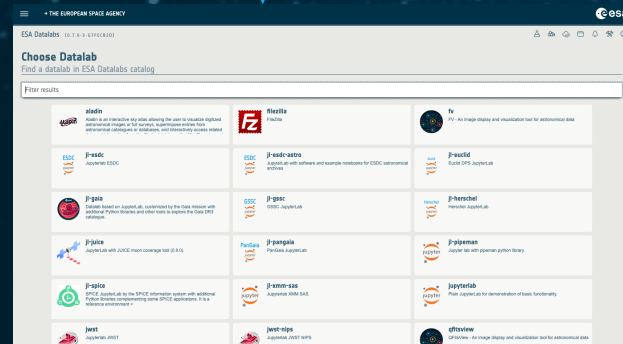


ESA Data Discovery Portal – One Portal To Discover Them All



Platforms

Datasets



SCI, NAV, OPS, EOP ...

datalabs.esa.int

gssc.esa.int



Discovery of Exploitation Platforms

The screenshot displays the ESA Data Discovery Portal interface. On the left, a sidebar contains filters for 'Type of Asset' (Datalab, Dataset), 'Domain' (Navigation, Space Science), 'Instrument', 'Thematic Area', 'Mission', and 'Properties' (Query Tool, Visualization Tool, Analysis Tool, Data Volume). The main search results area shows a list of assets, with the top result being the 'GOCE Mission' (Gravity field and Ocean Circulation Explorer). A detailed view of this mission is shown on the right, including a description and a section titled 'Launch a datalab to analyze the asset'. This section lists several exploitation platforms: x-glab, x-octave, jupyterlab, and GSSC Now, each with a brief description and a 'datalabs.esa.int' link.

Platform and Community Release Roadmap

Q1 Private Beta
2022



Q1 Public Beta
Moderated
2023



Q4 Public Beta
Unmoderated
2023



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!



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.

Find out more in datalabs.esa.int