



ESA Datalabs, an Open Digital Platform for Innovation and Collaboration in Space Science

IVOA interoperability meeting, Bologna, Science Platforms II – Vicente Navarro

11.05.2023, Virtual



→ THE EUROPEAN SPACE AGENCY 

ESA Datalabs [0.7.0-3-G7FECB2D]  Log in

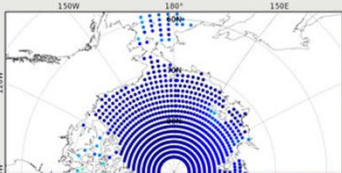
ESA Datalabs is available as "Public Moderated Beta"
If you wish to apply for access, please [submit your request here](#).

**«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.»**

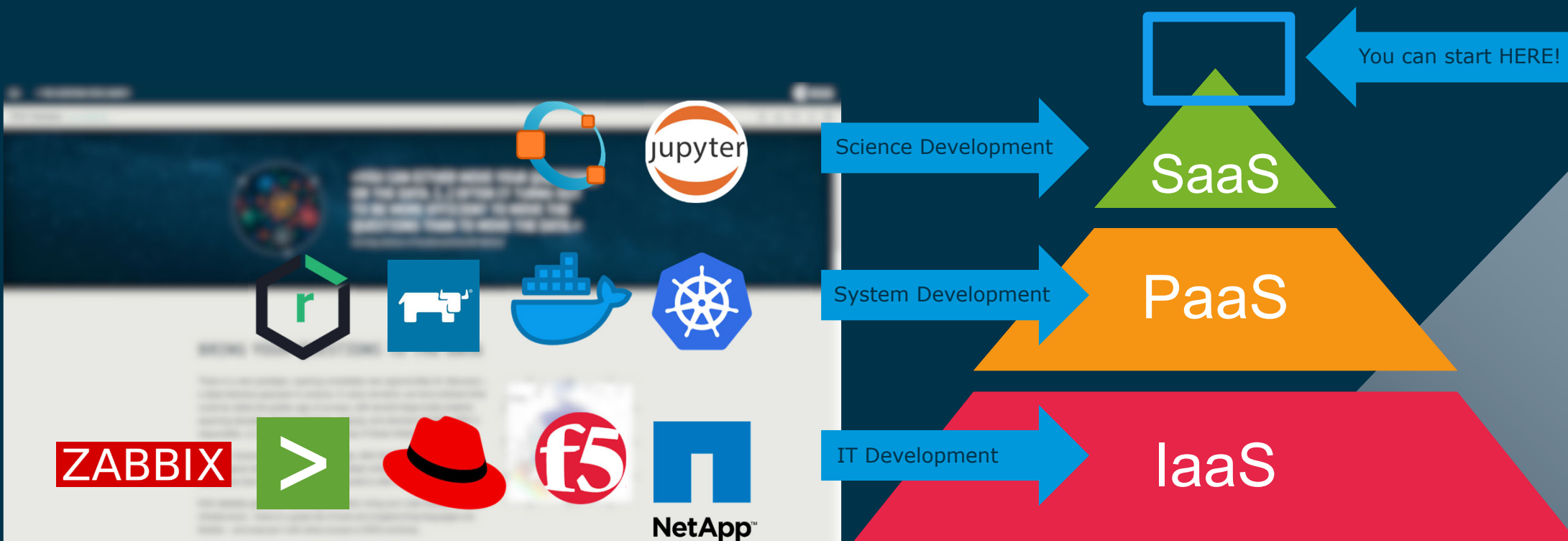
Jim Gray, eScience: A Transformed Scientific Method

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.



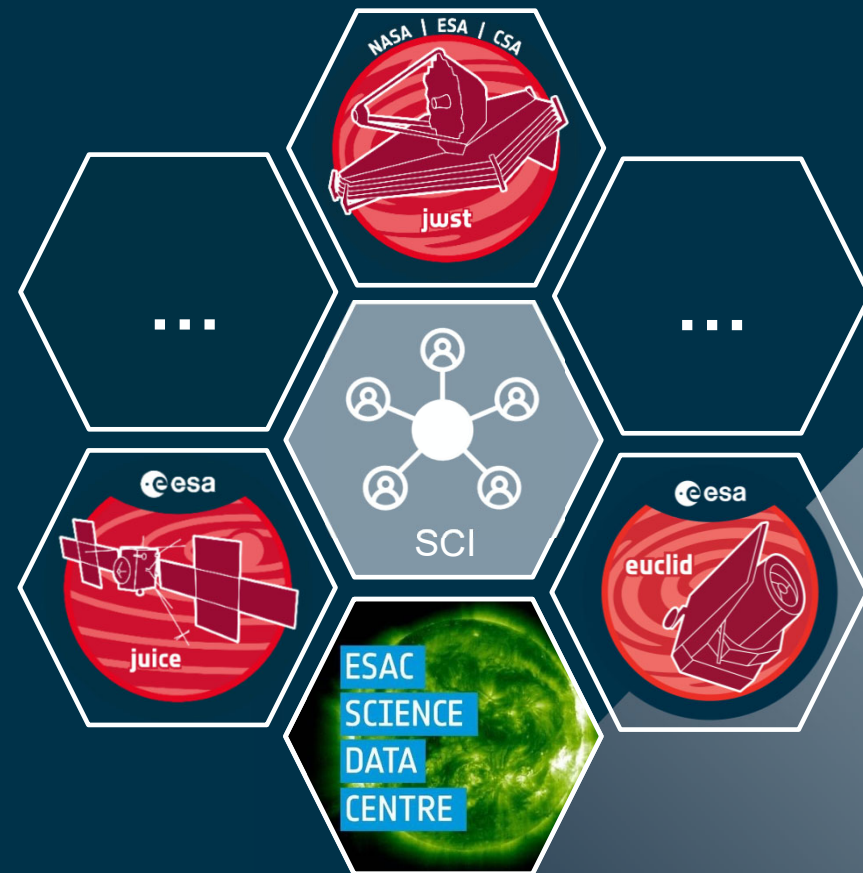
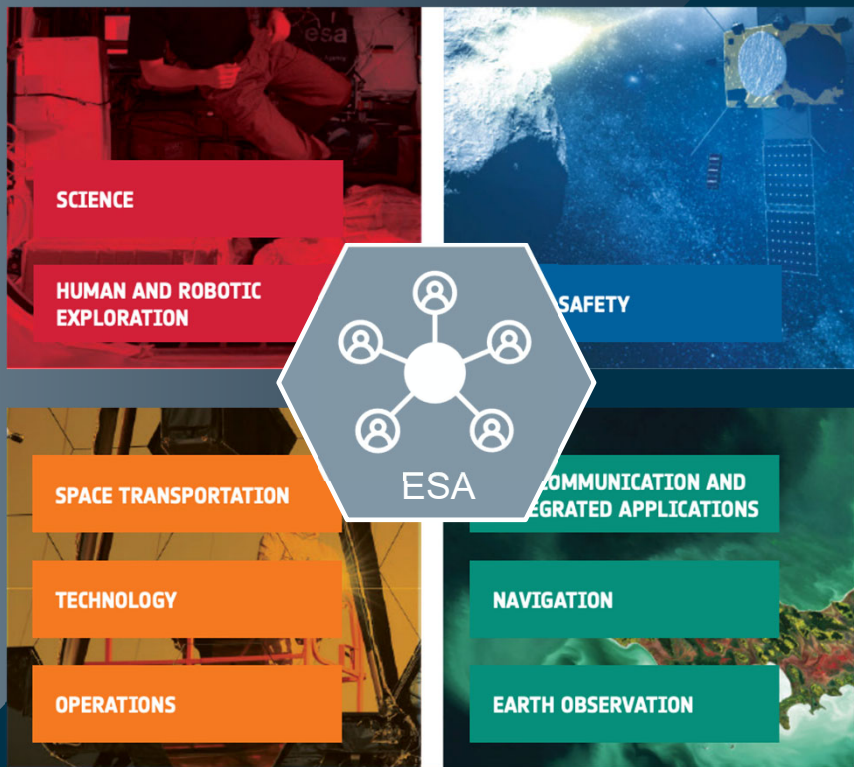
A Platform Designed to Boost Research Productivity



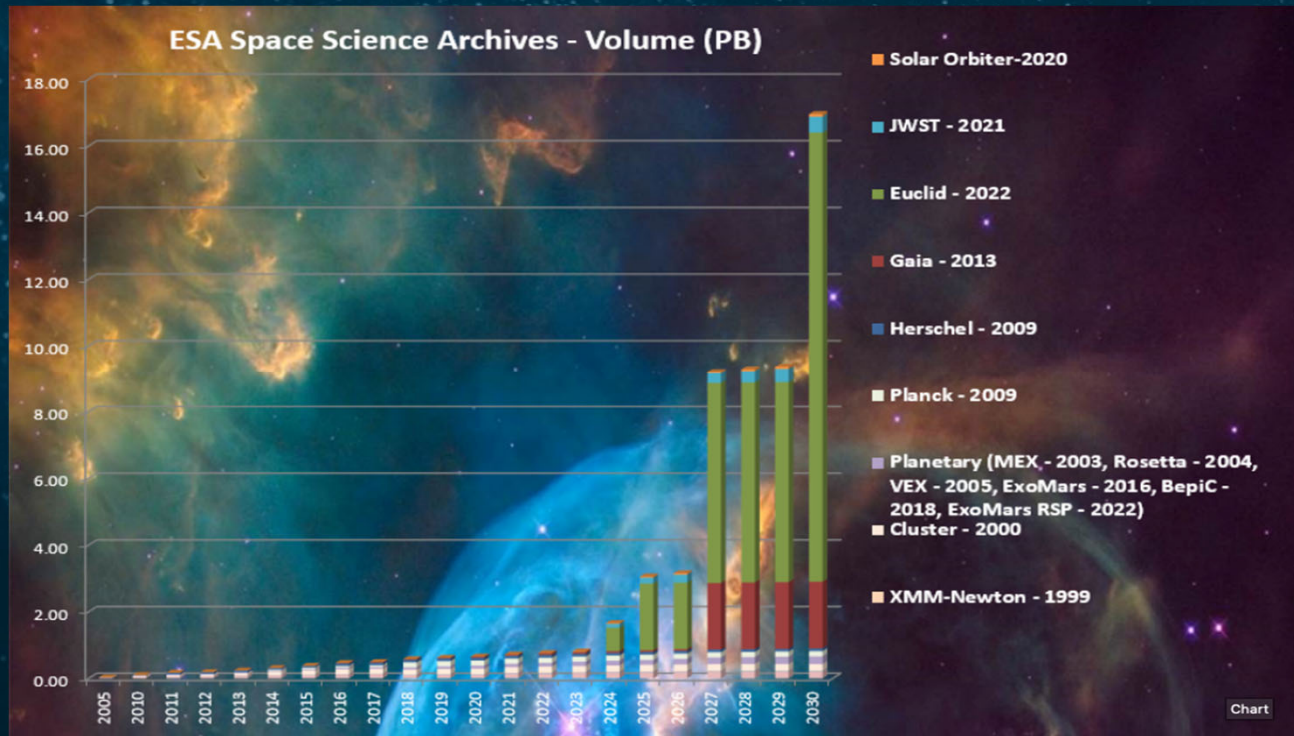
“The tools we use have a profound (and devious) influence on our thinking habits, and, therefore, on our thinking abilities”
Edsger W. Dijkstra



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



Space Data Colocation as a Differential Factor



ESAC Science Data Center

From bring the data to the user

To bring the user to the data



Datalabs Catalogue: Interactive Analysis


















→ THE EUROPEAN SPACE AGENCY esa

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.
 jlwst Jupyterlab JWST	 jlwst-miricle Jupyterlab JWST Miricle	 jlwst-nips Jupyterlab JWST NIPS
 jlwst-nsrt Jupyterlab JWST NSRT	 qfitsview QFitsView - An image display and visualization tool for astronomical data	 theia-python Theia Python Editor



Pipelines Catalogue : Batch Processing Analysis



→ THE EUROPEAN SPACE AGENCY esa

ESA Datalabs [0.7.0-3-G7FE(B20)] 🔍 📄 📁 📧 ⚙️

Pipeline launch

Find a pipeline in pipelines catalog Browse User pipeline

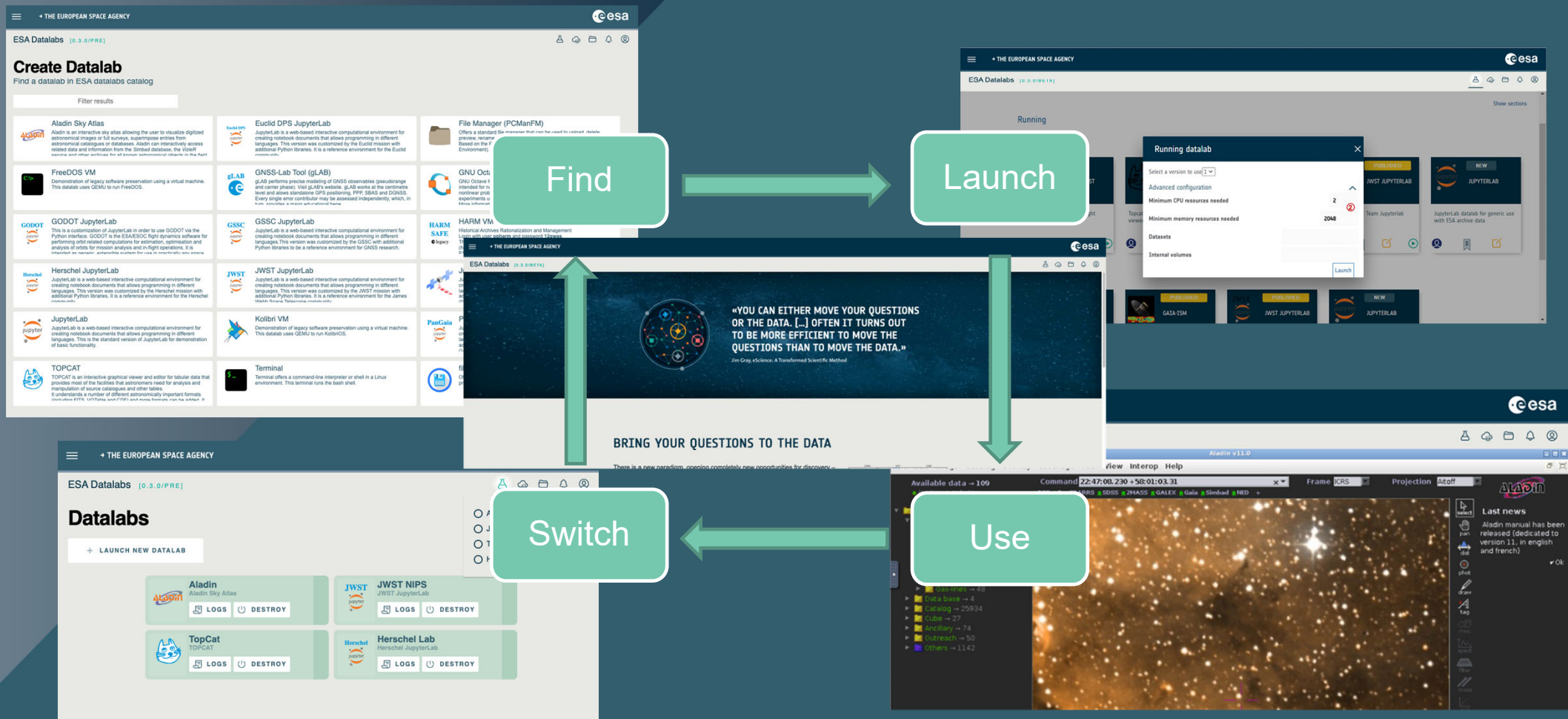
Filter results Sort by last modified Steps Pipelines

System Pipelines

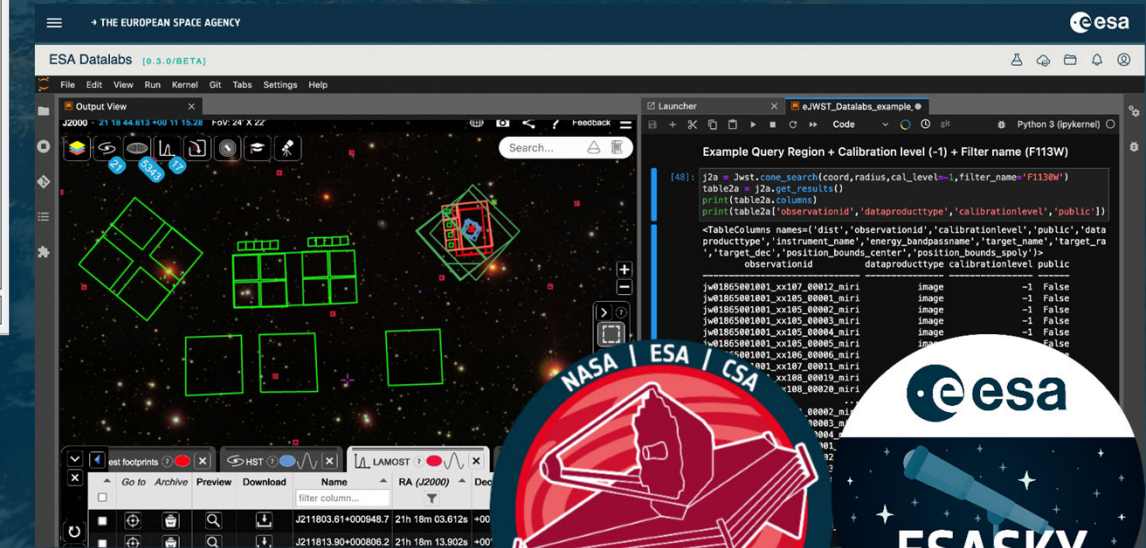
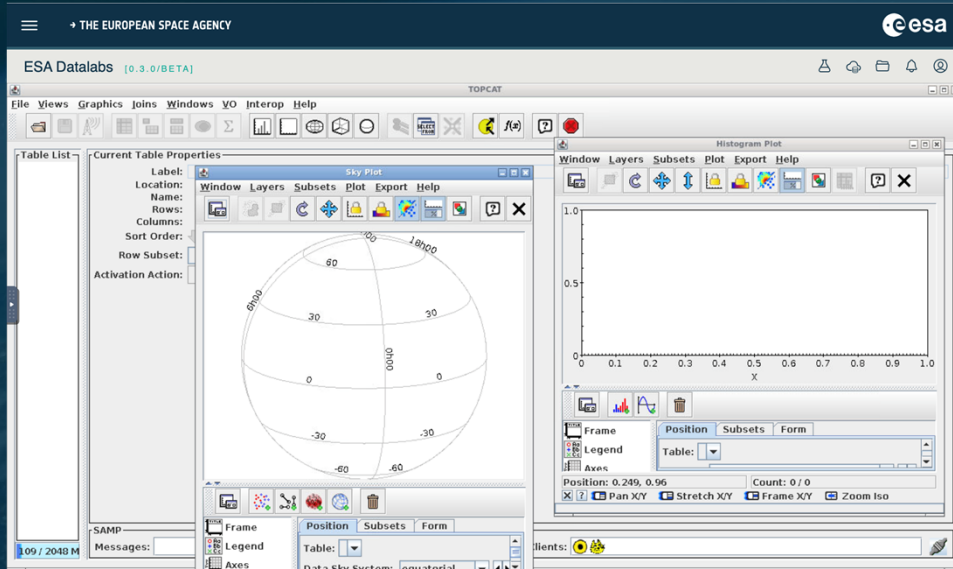
BC_MCAM This pipeline contains 2 steps: tm2raw_mcam and pds4_packager. <input type="button" value="bc"/> <input type="button" value="example"/> <small>👤 jkuhi 🕒 Wed May 10 2023</small>	JWST This pipeline runs 'jwst.pipeline.Detector1Pipeline' and then processes the output on 'jwst.pipeline.Image2Pipeline' on 'strun' command. <input type="button" value="example"/> <input type="button" value="jwst"/> <small>👤 jkuhi 🕒 Wed May 10 2023</small>	Image2 This pipeline runs 'jwst.pipeline.Image2Pipeline' on 'strun' command. <input type="button" value="example"/> <input type="button" value="jwst"/> <small>👤 jkuhi 🕒 Wed May 10 2023</small>
Detector1 This pipeline runs 'jwst.pipeline.Detector1Pipeline' on 'strun' command. <input type="button" value="example"/> <input type="button" value="jwst"/> <small>👤 jkuhi 🕒 Wed May 10 2023</small>	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. <input type="button" value="example"/> <small>👤 jkuhi 🕒 Wed May 10 2023</small>	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. <input type="button" value="example"/> <input type="button" value="scatter-gather"/> <small>👤 jkuhi 🕒 Wed May 10 2023</small>
cdr-demo This pipeline chains two 'command' steps. There is intermediate output after the first step and the final output after the second. <input type="button" value="example"/> <small>👤 jkuhi 🕒 Wed May 10 2023</small>	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. <input type="button" value="example"/> <small>👤 jkuhi 🕒 Wed May 10 2023</small>	simple Very basic pipeline. <input type="button" value="example"/> <small>👤 jkuhi 🕒 Wed May 10 2023</small>



Datalab / Pipeline – Utilisation Lifecycle



Web-Based & Desktop-Based Datalabs



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 Datalab interface (version 0.3.0/BETA) with a menu bar (File, Edit, View, Run, Kernel, Git, Tabs, Settings, Help) and a top navigation bar (THE EUROPEAN SPACE AGENCY, ESA logo). The main workspace is titled 'Launcher' and shows a file browser on the left with a search bar and a list of files in the '/ data /' directory. The 'pla_legacy' folder is highlighted with a blue box. The main area contains a 'data' section with a 'Notebook' icon, two 'Python 3 (ipykernel)' icons, and an 'Other' section. A 'Data Volumes' panel on the right lists several volumes, with 'Planck Legacy Archive' highlighted by a blue box. The URL 'https://datalabs.esa.int/datavol-manager' is visible at the bottom left, and 'Launcher' is at the bottom right.



Datalabs Integrated Development Environment



esa | datalabs

← Gaia archeology

DRAFT

Modification date: 15/06/2022 21:18
Datalab version: 1.4.0
Author: sciapps developer
Popularity: ☆0
Link: 53a11c69-4982-4bc7-8115-a83...

Delete

Metadata Testing **Visibility**

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 Build version: 1.4.0-0 **SUCCESS**

Sharing audience Public, everyone can use it Private, specify a list of users

Publish

License

Please select a license

- Popular Licenses**
- 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**
- European Space Agency Community License – v2.4 Permissive (Type 3)
- European Space Agency Community License – v2.4 Strong Copyleft (Type 1)
- European Space Agency Community License – v2.4 Weak Copyleft (Type 2)
- European Space Agency Public License (ESA-PL) Commentary – v2.3
- European Space Agency Public License – v2.4 – Permissive (Type 3)
- European Space Agency Public License – v2.4 – Strong Copyleft (Type 1)
- European Space Agency Public License – v2.4 – Weak Copyleft (Type 2)
- Others**
- 3dfx Glide License
- AMD's plpa_map.c License
- ANTIR Software Rights Notice

Submit

Export

Export

Clicking on this button will allow you to get an archive

History

```
15/06/2022 21:19: Build started
15/06/2022 21:21: Build ended with status BUILD_SUCCESS
```



Pipeline Integrated Development Environment



→ THE EUROPEAN SPACE AGENCY

ESA Datalabs [0.6.0-57-G55587088]

Workspace Catalogue Hello_PIPEMAN.pipeline.cwl ×

Graph Code Test Details Push @Pipeline - 0.0.3 [latest]

Workspace
example_pipeline_input
my_pipelines
pulled_system_pipelines
My Pipelines in Workspace
System Pipelines in Workspace
hello_pipeman
dependencies
Hello_PIPEMAN.pipeline.cwl
manifest.json

file → figlet → figlet → command → output

text → figlet → figlet → command → output

text_2 → figlet → figlet → command → output

command_1 → figlet → figlet → command → output

Common Workflow Language - CWL

Validation Execution logs
[01/02/2023, 17:32:27] CWL Valid

Graph Code Test Details Push @Pipeline - 0.0.8 [latest]

Name: Hello_PIPEMAN
Description: Hello PIPEMAN example. Writes block text.
Keywords: Example × Hello_PIPEMAN ×
Allowed groups: Comma-separated group names
Allowed users:

Version: 0.0.8
Version description: Update figlet version
Default Execution Engine: CWL Tool UI Test Calrissian

Push as System Pipeline



JWST @ ESA Datalabs 13



→ 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

My Workspace
Team Workspaces

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 ρ -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



Application Centric Collaboration – Datalabs Marketplace



esa | datalabs

Search datalab catalog

Customize view

Public

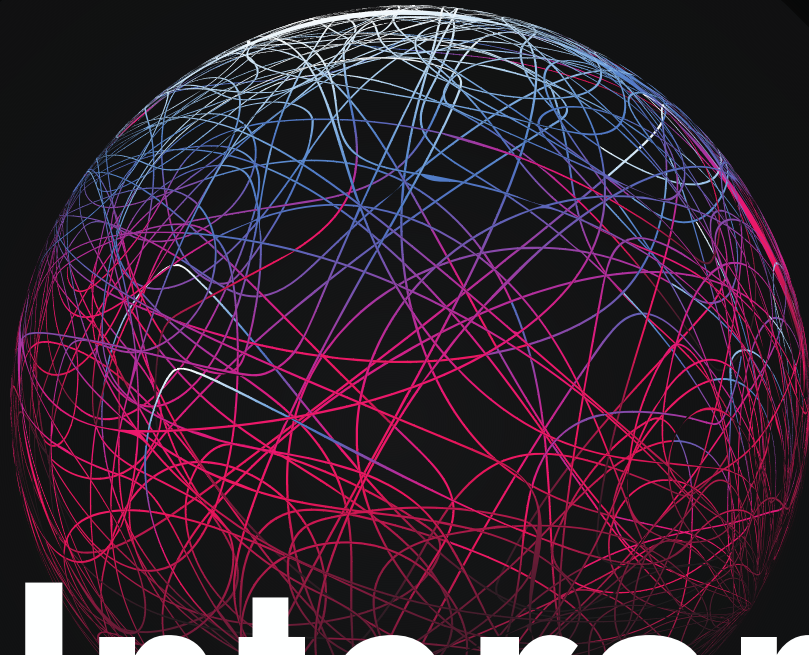
<p>CYPRESS SEPPTTEST-219 1994.7571690539014</p> <p>Test Automated</p> <p>🗨️ ⭐️ ▶️</p>	<p>CYPRESS SEPPTTEST-219 1111.6848610854097</p> <p>Test Automated</p> <p>⭐️ ▶️</p>	<p>CYPRESS SEPPTTEST-219 1419.7853655034294</p> <p>Test Automated</p> <p>🗨️ ⭐️ ▶️</p>	<p>CYPRESS SEPPTTEST-219 1713.2724871003359</p> <p>Test Automated</p> <p>🗨️ ⭐️ ▶️</p>	<p>CYPRESS SEPPTTEST-219 1673.9987705736833</p> <p>Test Automated</p> <p>🗨️ ⭐️ ▶️</p>
---	--	---	---	---

See more

Developed by me

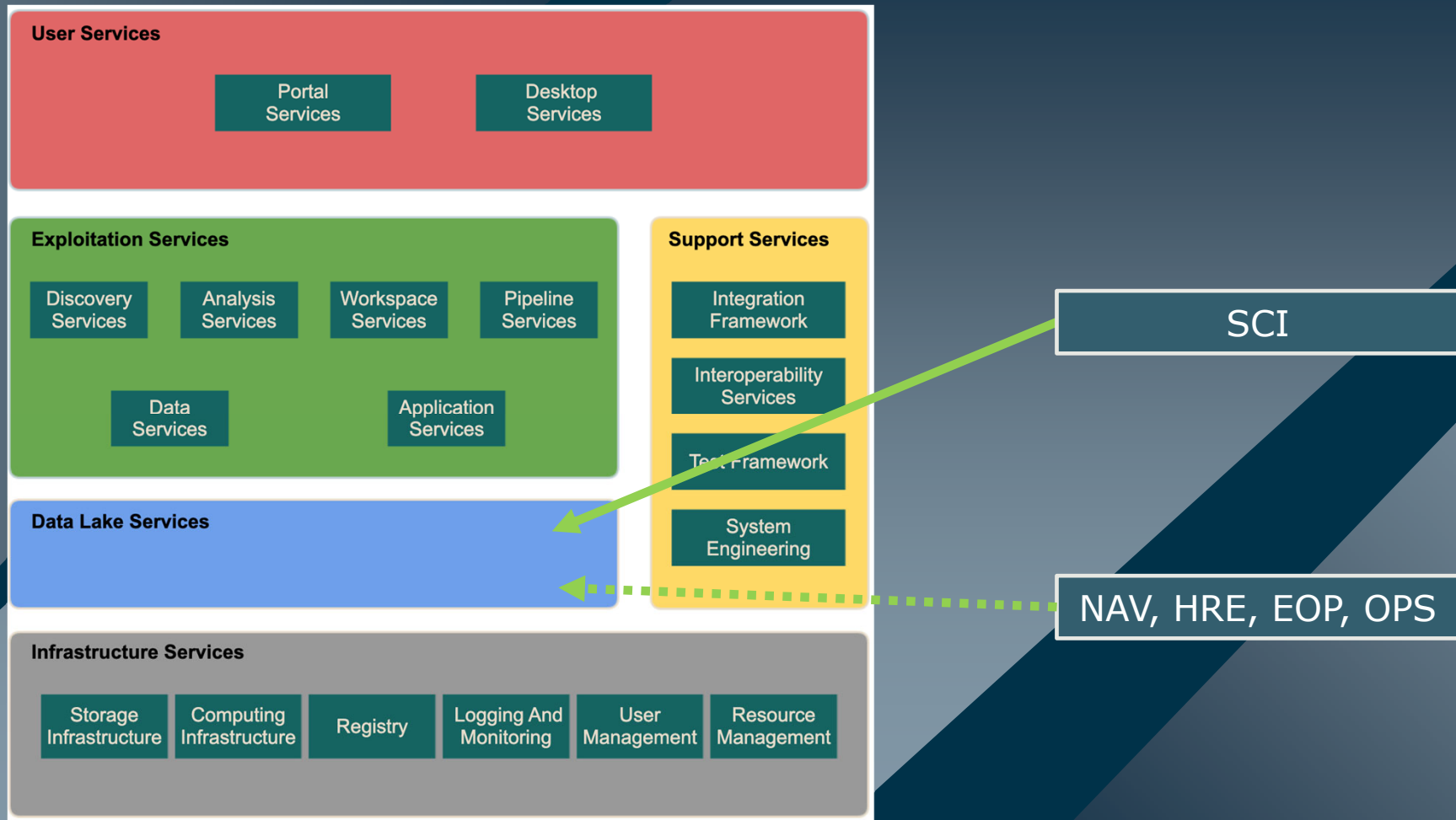
<p>NEW</p> <p>No description provided</p> <p>🔍 📄 ▶️</p>



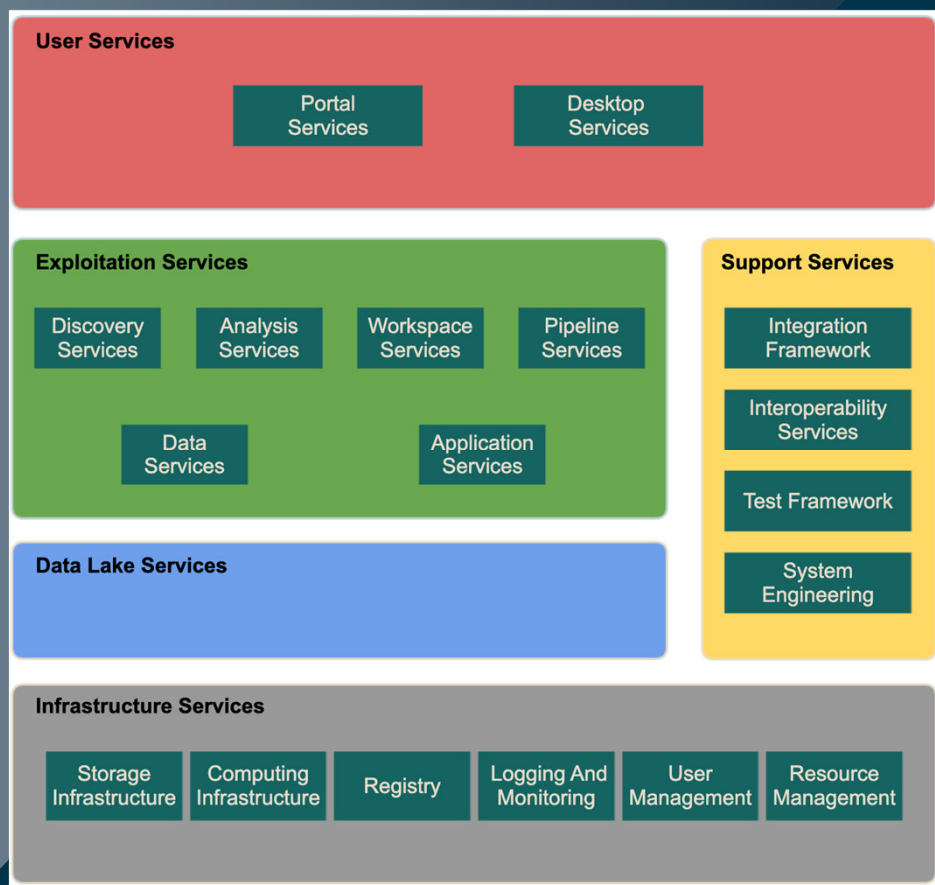


Interoperability

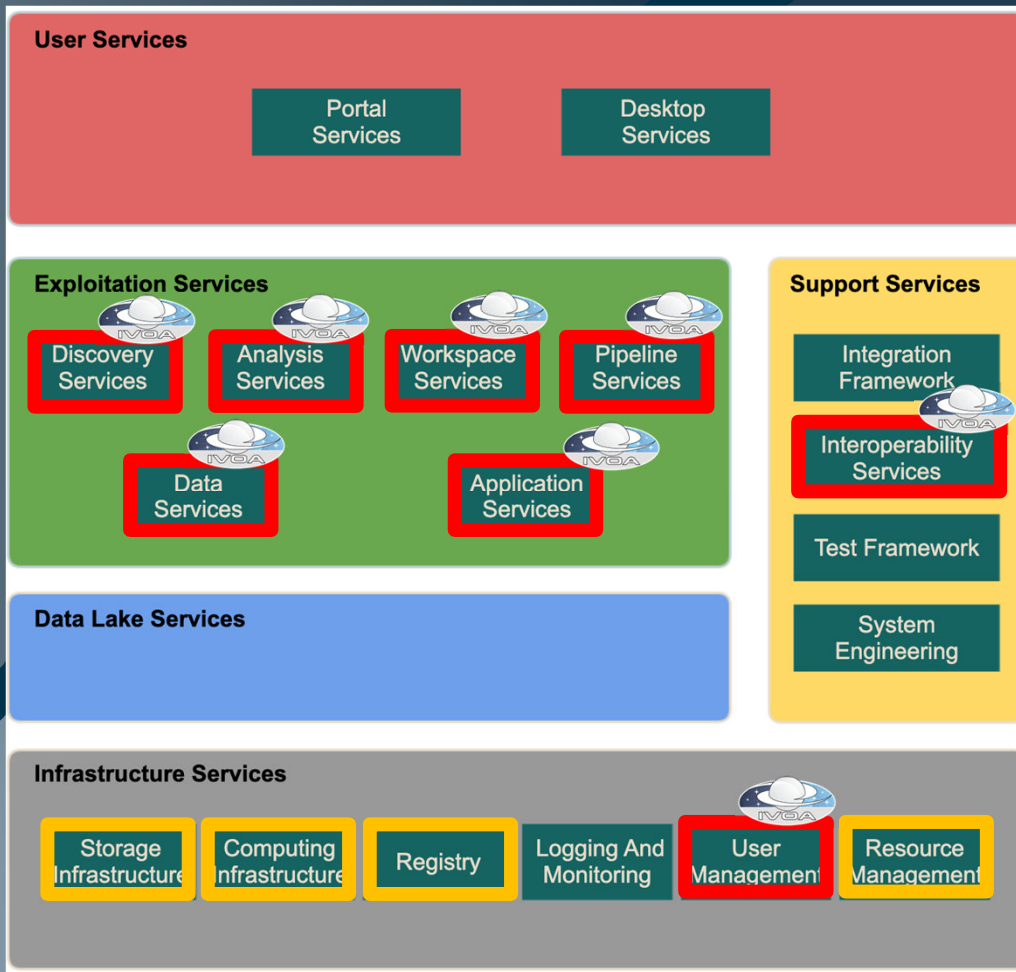
ESA Datalabs Blueprint



ESA Datalabs Technology Stack



IVOA and Interoperability

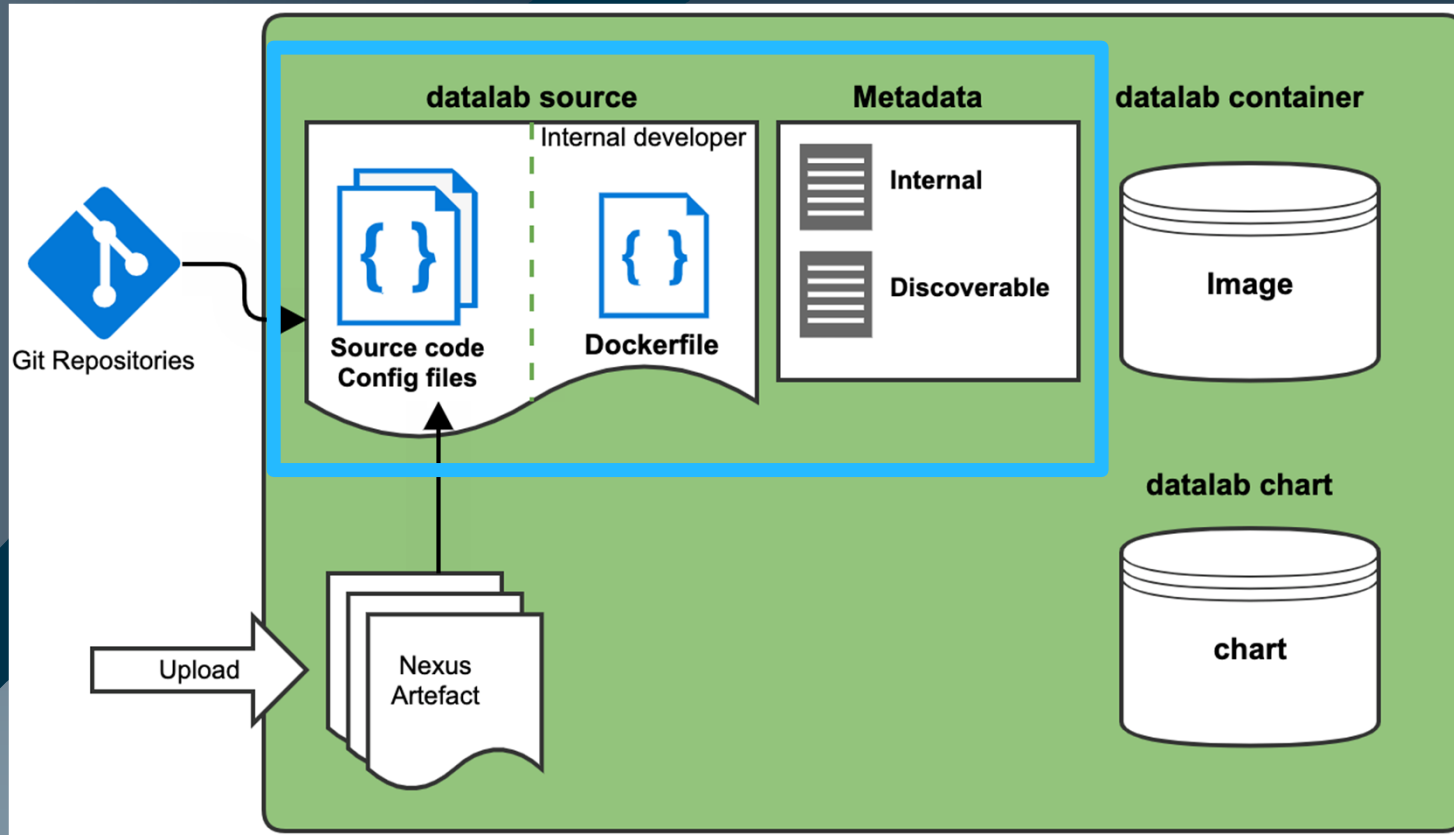


FUNCTIONAL

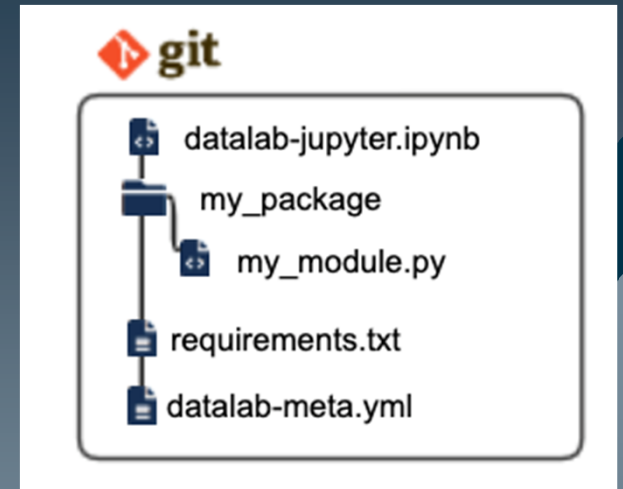
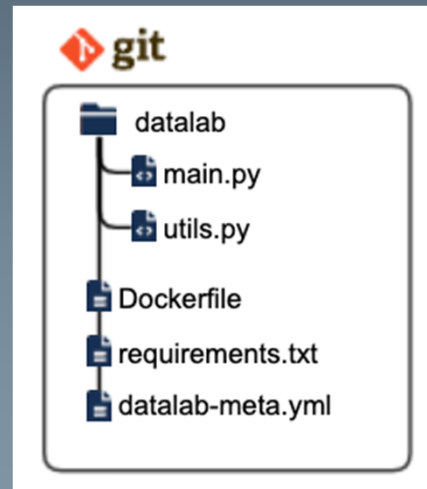
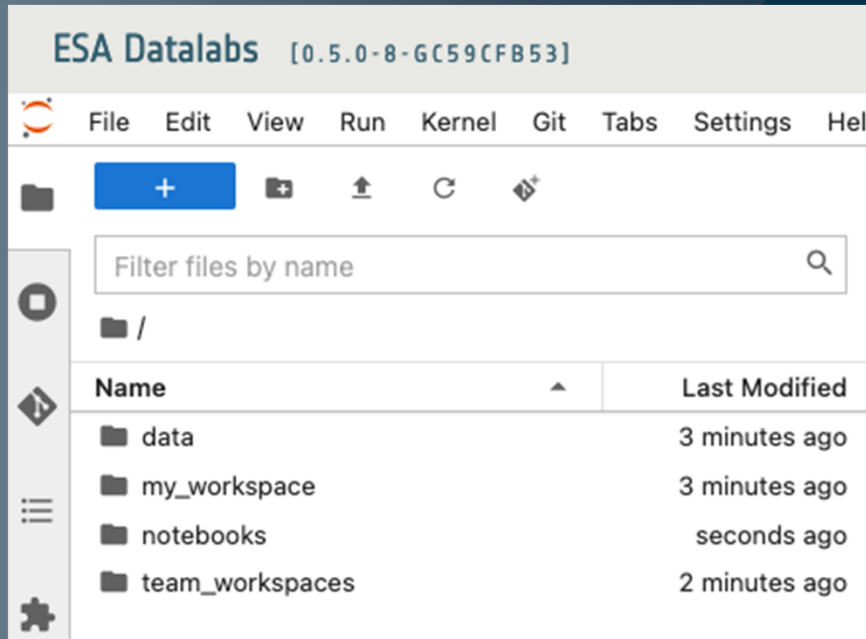
TECHNOLOGICAL



Application Services: Datalab Package Specification



Application Services: Datalab Structure



Discovery Services: Datalab Metadata

ESA DDP Term	Namespace:Term	Description	Visibility (private / public)	Write access (developer / moderator / sciapps)	Data type
abstract	sdo:abstract	Abstract (a short description)	PUBLIC	DEVELOPER	Text
alternateName	sdo:alternateName (pipeman: displayName)	Short name or acronym for the datalab	PUBLIC	DEVELOPER	Text
associatedFileType	esado:associatedFileType (new)	List of associated filetypes - if applicable/pertinent (e.g. FITS, VOTable, GeoTIFF, netCDF)	PUBLIC	DEVELOPER	Text
ESAOfficial	esado:ESAOfficial	flag indicating that the datalab was created by ESA	PUBLIC	DEVELOPER	Boolean
audienceType	sdo:audienceType	Intended audience for the datalab	PUBLIC	DEVELOPER	Text
citation	sdo:citation	Citation for the datalab (e.g. article DOI)	PUBLIC	DEVELOPER	Text
datalabImage.platform	sdo:version	ESA datalabs platform version	PRIVATE	APPLICATION...	Text
datalabImage.id	sdo:identifier	Identifier of the datalab image	PRIVATE	APPLICATION...	Text
datalabImage.keyword	dcat:keyword	Keyword tag of the current datalab image	PRIVATE	DEVELOPER	Text
dateCreated	sdo:dateCreated	Date the datalab was created	PRIVATE	APPLICATION...	Date

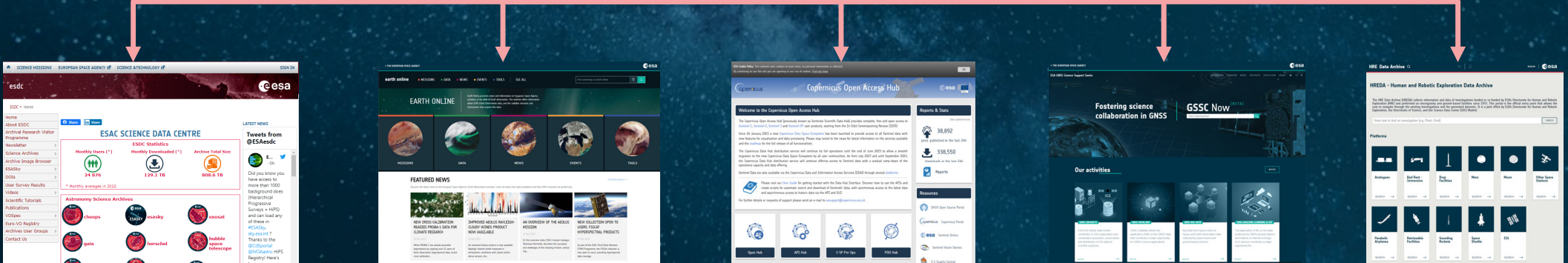
PUBLIC

PRIVATE

Leveraging on existing vocabularies and ontologies: sdo, dcat, skos, foaf ... esado



Leveraging on ESA's Digital Ecosystem of Archives

Registry/DCP Session: Christophe Arviset - ESA Data Discovery Portal, link to ESA datasets DOIs and to Google Dataset Search.



Leveraging on ESA's Digital Ecosystem of Platforms



datalabs.esa.int

gssc.esa.int



Federation of ESA Exploitation Platforms



ESA Data Discovery Portal [0.7.0 (LPHA)]

Type of Asset
 Datalab
 Dataset

Domain
 Navigation (1)
 Space Science (15)

Search: GOCE

type:datalab property:Analysis Tool clear all

GOCE Mission
Gravity field and Ocean Circulation Explorer (GOCE)
GNSS data. Access policy is Public and license is ESA Data Policy

Cycle 18 doublePoint UVIS Bowtie Monitor
Ground testing revealed an intermittent hysteresis type effect in the UVIS keyOpen both CCDskeyClose at the level of virgul1% comma last days. Initially found via an unexpected bowtieshaped feature in the comma subsequent lab tests on similar e2v devices have since shown also present as clearly as overall effect across the entire CCD sensor.

Cycle 17 doublePoint UVIS Bowtie Monitor
Ground testing revealed an intermittent hysteresis type effect in the UVIS keyOpen both CCDskeyClose at the level of virgul1% comma last days. Initially found via an unexpected bowtieshaped feature in the comma subsequent lab tests on similar e2v devices have since shown also present as clearly as overall effect across the entire CCD sensor.

Mrk590 doublePoint A Disappearing AGN questionMark
Recent optical spectroscopy reveals that the active nucleus of the Seyfert Mrk590 keyOpen NGC863keyClose has gone into an extremely faint broad emission lines characteristic of Type 1 AGNs and the F_{cont} continuum are no longer visible in optical spectra comma and even emission lines appear to have faded to about a factor of two since the previous observations.

Using a Cepheid based Distance to Test the Large Peculiar M...
We propose to use the HST and WFPC2 to obtain periods and luminosities of Cepheids in a spiral galaxy in the Centaurus cluster in order to determine distance. A Cepheid distance to Centaurus is possible with HST as a major step in the comparison of observed peculiar velocities with the expected velocities of the cluster field from the spatial distribution of the cluster members.

Confirming the binarity of Kuiper Belt Object 2015 RR245dou...
Binary systems are prevalent throughout the Kuiper Belt comma with systems known to date. The question of how these binary systems form is of great interest comma as each proposed mechanism has its own implications for the conditions in the early Solar System during planetesimal growth.

Properties
 Query Tool
 Visualization Tool
 Analysis Tool
 Data Volume

GOCE Mission

Overview Details Query Visualize Analyze Download

Launch a datalab to analyze the asset

- <https://datalabs.esa.int/datalab-launch-wizard/x-glab>
- <https://datalabs.esa.int/datalab-launch-wizard/x-octave>
- <https://datalabs.esa.int/datalab-launch-wizard/jupyterlab>
- <https://gssc.esa.int/portal?redirect=datalabs>



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

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.

