



ESA DATALABS

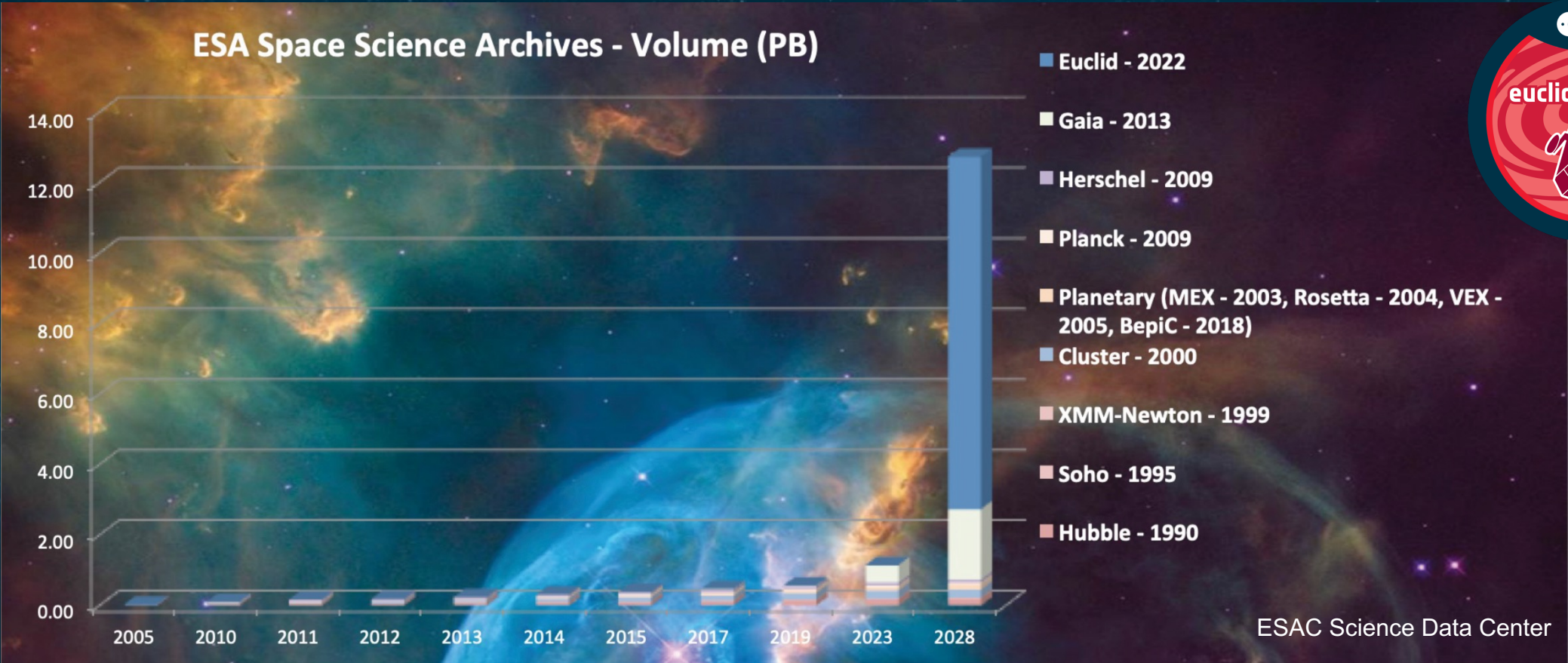
ESA Datalabs

Sandor Kruk

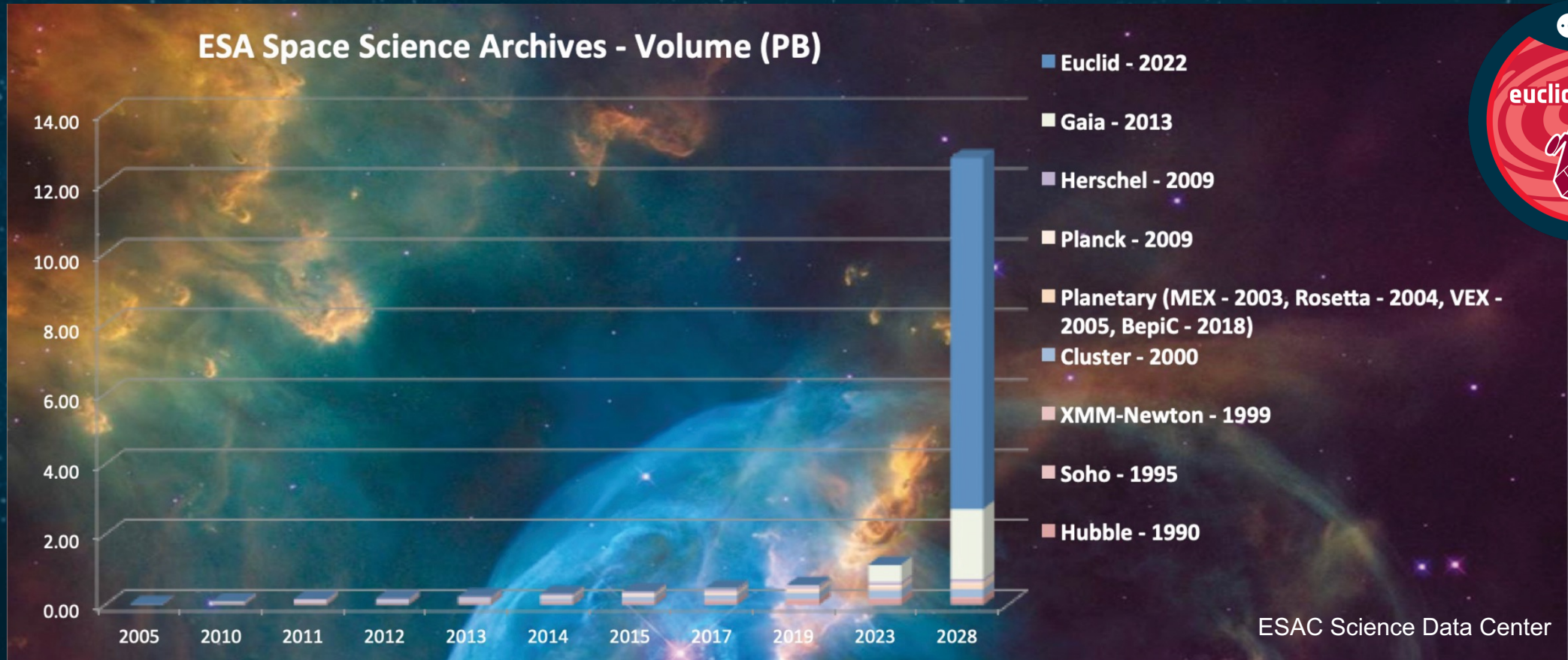
on behalf of the Data Science Team
Archives and Data Science Division
European Space Agency
ESAC, Madrid, Spain

INTEGRAL Workshop ESAC
22 October 2024

Increasing data volumes in the ESA Science Archives



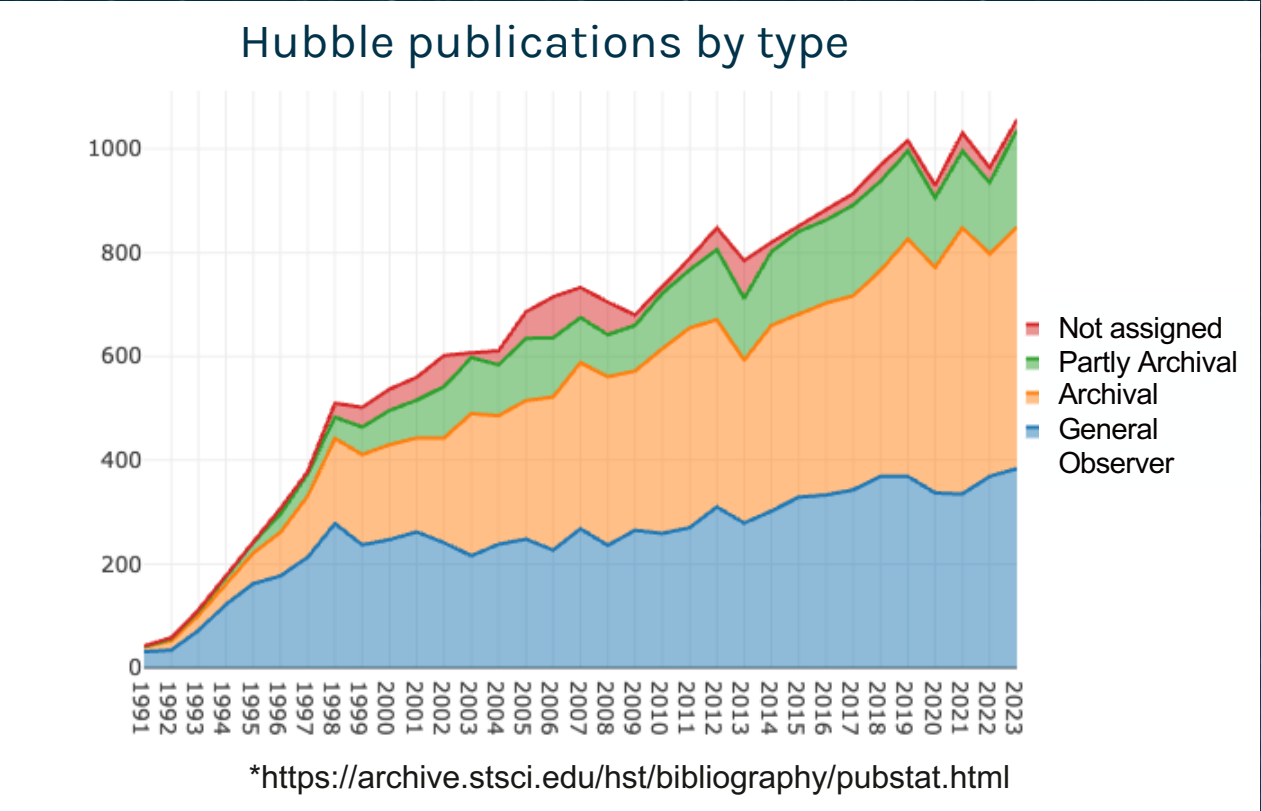
Increasing data volumes in the ESA Science Archives



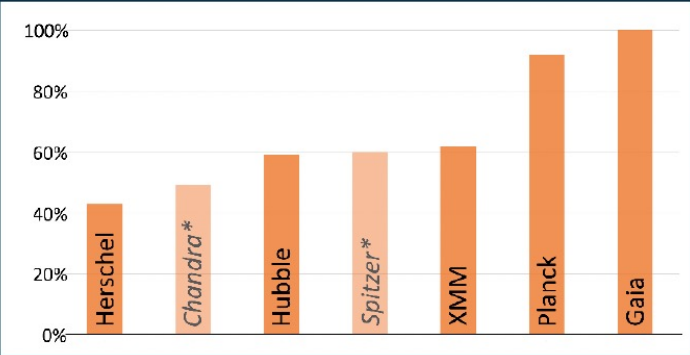
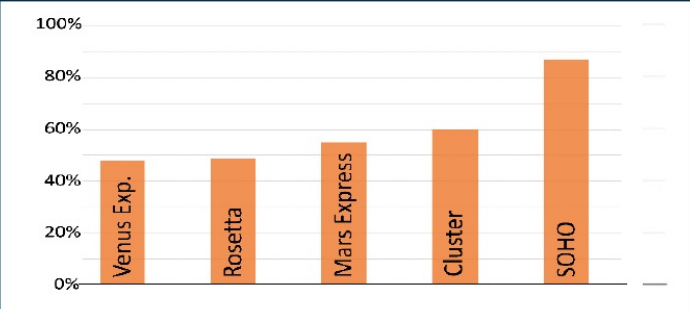
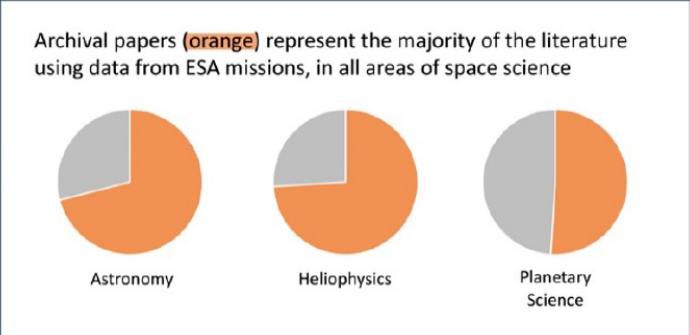
From bringing the data to the user

To bringing the user to the data

Increasing importance of archival research



Hubble Space Telescope publications statistics




de Marchi & Parmar, 2024

→ 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](#).

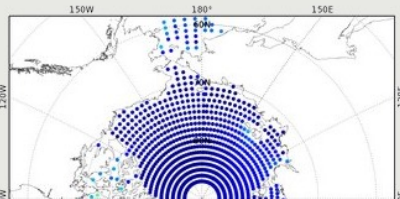



«YOU CAN EITHER MOVE YOUR QUESTIONS
OR THE DATA. [...] OFTEN IT TURNS OUT
TO BE MORE EFFICIENT TO MOVE THE
QUESTIONS THAN TO MOVE THE DATA.»

Jim Gray, eScience: A Transformed Scientific Method

BRING YOUR QUESTIONS TO THE DATA

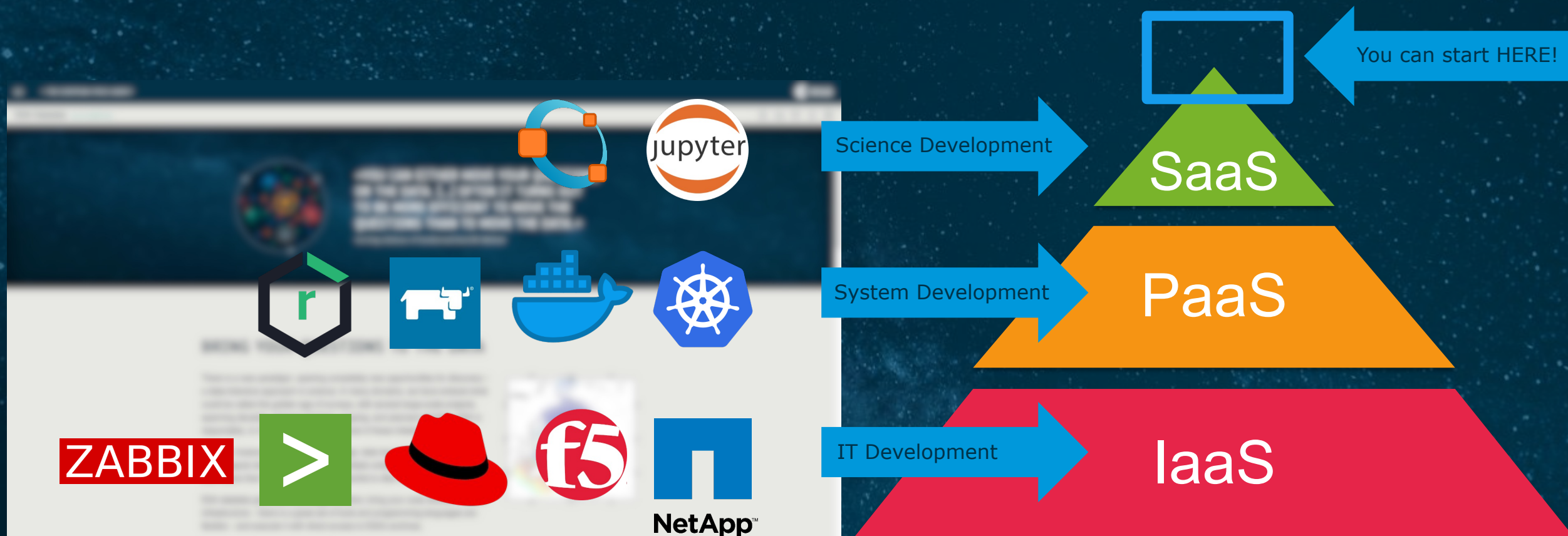
There is a new paradigm, opening completely new opportunities for discovery – a data-intensive approach to science. In many domains, we have entered what could be called the golden age of surveys, with several large-scale projects, spanning decades, between finished, ongoing, and planned activities.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



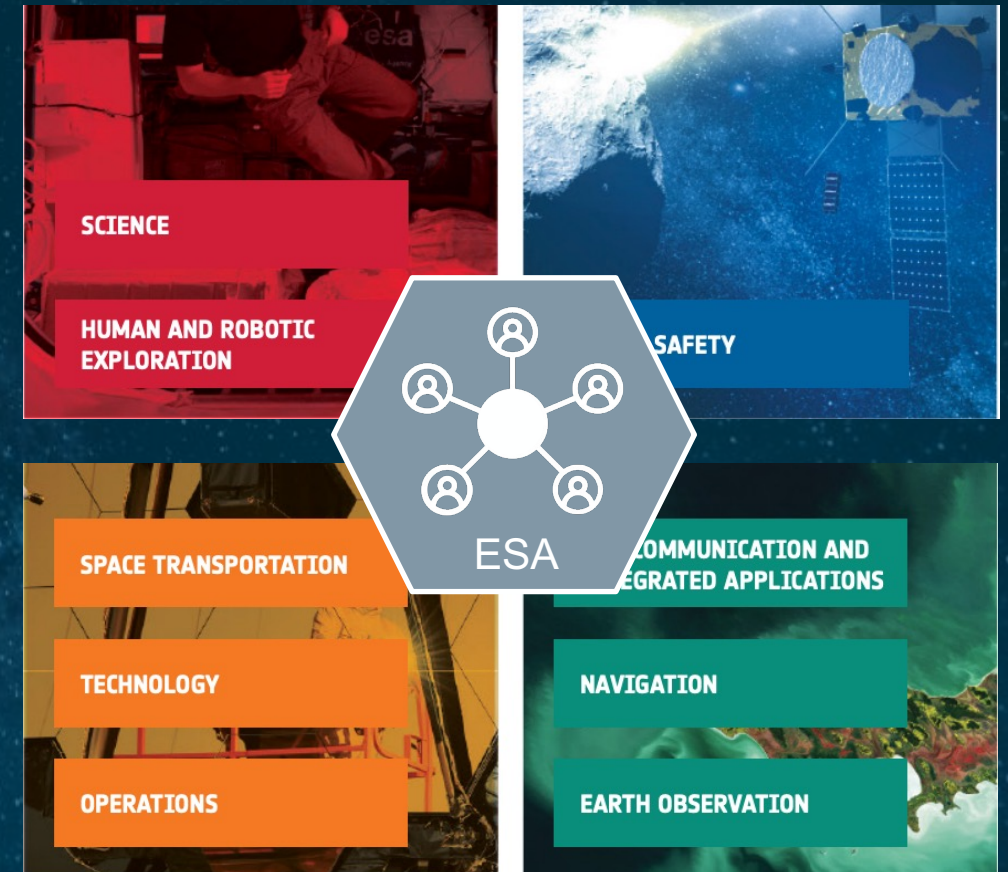
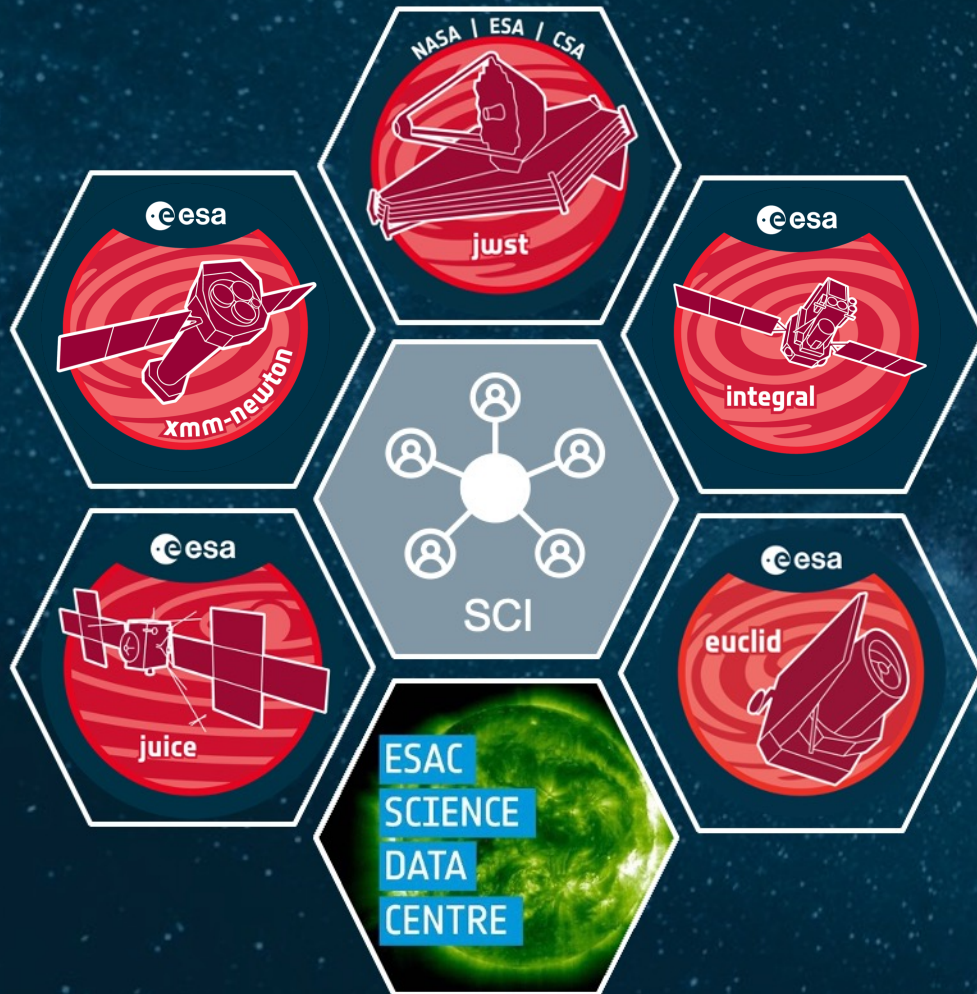


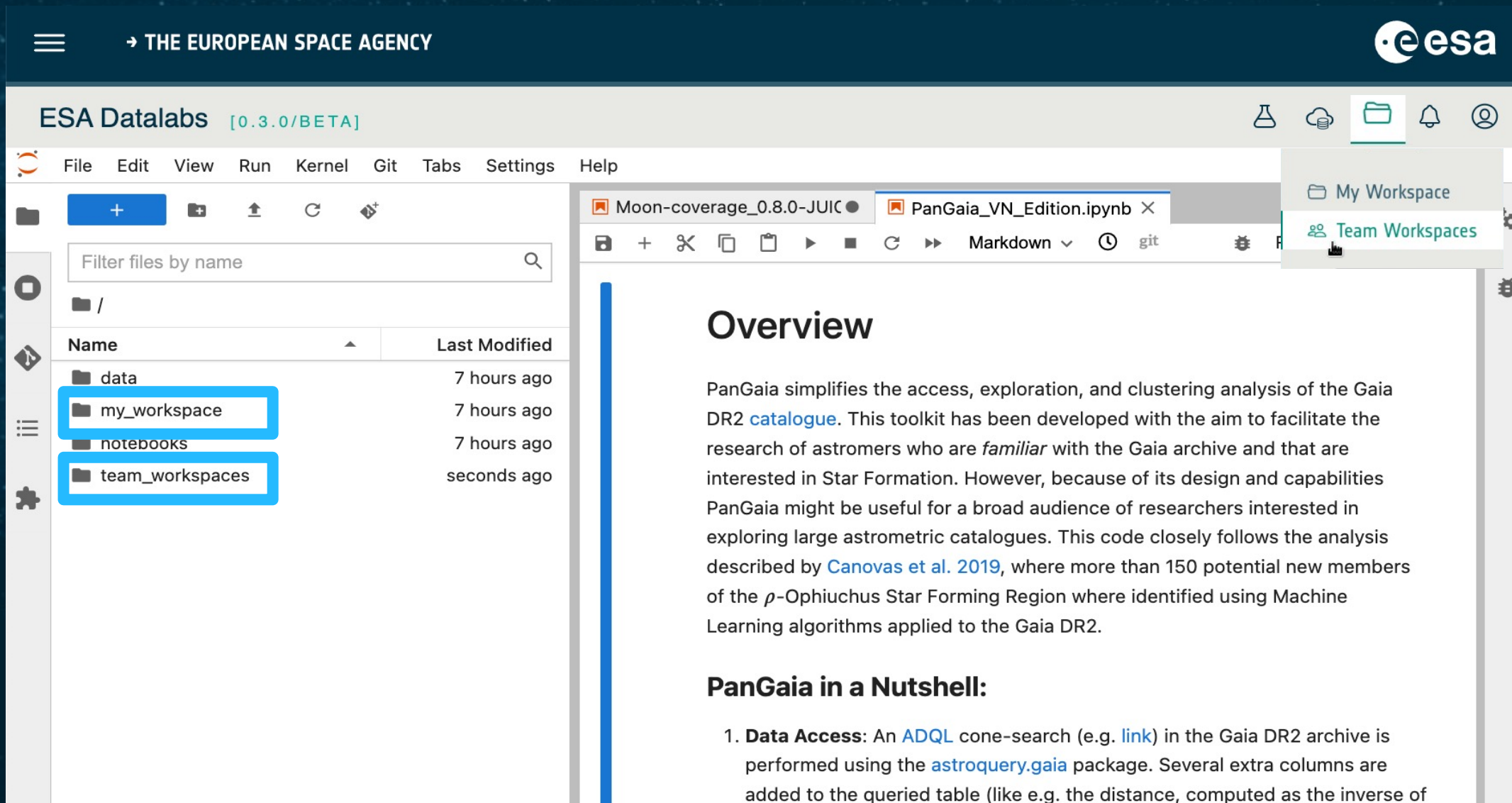
→ THE EUROPEAN SPACE AGENCY

5



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





The screenshot displays the ESA Datalabs interface, version 0.3.0/BETA. The top navigation bar includes the ESA logo and the text "THE EUROPEAN SPACE AGENCY". The interface is divided into two main sections: a file explorer on the left and a notebook editor on the right.

File Explorer (Left Panel):

- Search bar: "Filter files by name"
- File list table:

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

Notebook Editor (Right Panel):

- Open tabs: "Moon-coverage_0.8.0-JUIC" and "PanGaia_VN_Edition.ipynb"
- Current notebook: "PanGaia_VN_Edition.ipynb"
- Section:

Overview
- Text: "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 where identified using Machine Learning algorithms applied to the Gaia DR2."
- Section:

PanGaia in a Nutshell:
- List item: "1. **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

THE EUROPEAN SPACE AGENCY

ESA Datalabs

Data Volume Catalog

Domain

☐ Space Science (1)

Euclid

Euclid ERO

The Euclid Early Release Observations (ERO), led by the Euclid Science Team and ESA, showcase the Euclid mission's capabilities, focusing on legacy science. ERO DR3 includes image stacks and validation catalogues for 16 fields from six selected projects. Detailed dataset information is available <https://www.cosmos.esa.int/web/euclid/ero-data-release>

Perseus Cluster

Horsehead Nebula

IC432

Euclid's Early Release Observations

Credit: ESA/Euclid/Euclid Consortium/NASA, image processing by J.-C. Cuillandre (CEA Paris-Saclay), G. Anselmi

Computing & Data Colocation: Accessing Data Volumes



→ THE EUROPEAN SPACE AGENCY

esa

ESA Datalabs [0.9.0/BETA]

File Edit View Run Kernel Git Tabs Settings Help

Filter files by name

/ ... / 591 / 54 /

Name	Last Modified
j59154010_asn.fits.gz	3 months ago
j59154010_drc_thumb.jpg	3 months ago
j59154010_drc.fits.gz	3 months ago
j59154010_drc.jpg	3 months ago
j59154010_drz_thumb.jpg	3 months ago
j59154010_drz.fits.gz	3 months ago
j59154010_drz.jpg	3 months ago
j59154010_jif.fits.gz	3 years ago
j59154010_jit.fits.gz	3 years ago
j59154010_log.txt	3 months ago
j59154010_spt.fits.gz	3 months ago
j59154010_trl.fits.gz	3 months ago
j59154geq_flg_thumb.jpg	3 months ago
j59154geq_flg.fits.gz	3 months ago
j59154geq_flg.jpg	3 months ago
j59154geq_flg_hlet.fits.gz	3 months ago
j59154geq_flg_thumb.jpg	3 months ago
j59154geq_flg.fits.gz	3 months ago
j59154geq_flg.jpg	3 months ago
j59154geq_log.txt	3 months ago
j59154geq_raw_thumb.jpg	3 years ago
j59154geq_raw.fits.gz	3 months ago
j59154geq_raw.jpg	3 years ago
j59154geq_spt.fits.gz	3 months ago
j59154geq_trl.fits.gz	3 months ago

Launcher

j59154010_drc_thumb.jpg




































j59154010_drc.jpg

j59154010_drc.jpg 0






































Choose Datalab

Find a datalab in ESA Datalabs catalog

Filter results				
 Demoda A simple demo jupyterlab datalab	 Euclid_test This is a JupyterLab demo datalab	 G-Tomo The G-Tomo scientific data application (SDA) created by the EXPLORE project allows users to extract 1d profiles and 2d maps from the latest 3d dust extinction maps based on Gaia eDR3 and 2MASS data (Lallement et al. 2022).	 GODOT JupyterLab GODOT is an ESA/ESOC flight dynamics software. This is a customization of JupyterLab in order to use GODOT via the Python interface.	 INTEGRAL OSA JL INTEGRAL Jupyterlab with the OSA 11 analysis environment.
 INTEGRAL OSA VNC INTEGRAL VNC based datalab with the OSA 11 analysis environment and its GUI for interactive analysis.	 JWST_General JupyterLab is a web-based interactive computational environment. This version was customized by the JWST mission with additional Python libraries. It is a reference environment for the JWST community.	 NCTEST Test	 PSA JupyterLab A JupyterLab based DataLab custom built for working with planetary science data in the Planetary Science Archive (PSA).	 Pyxel datalab Pyxel is a novel, open-source, modular Python software framework designed to host and pipeline models (analytical, numerical, statistical) simulating different types of detector effects on images produced by Charge-Coupled Devices (CCD), Monolithic, and Hybrid CMOS imaging sensors.
 TESTMOD1 INTEGRAL Jupyterlab with the OSA 11 analysis environment.	 XMM-SAS21.0.0 XMM-Newton SAS v21.0.0	 YPE24-Euclid Jupyterlab datalab for the YPE Euclid workshop	 YPE_Galaxy_Class Jupyterlab Datalab for the Galaxy classification workshop at YPE2024	 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 service</i> 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-astro JupyterLab with software and example notebooks for ESDC astronomical archives	 jl-esdc-helio JupyterLab is a web-based interactive computational environment for creating notebook documents that allow programming in different languages. This version was customized by the ESDC with additional Python libraries to be a reference environment for Heliophysics research.	 jl-euclid Euclid JupyterLab
 jl-euclid-c11g Euclid Commissioning JupyterLab	 jl-gaia Datalab based on JupyterLab, customized by the Gaia mission with additional Python libraries and other tools to explore the Gaia DR3 catalogue.	 jl-gssc GSSC JupyterLab	 jl-herschel Herschel JupyterLab	 jl-juice JupyterLab with JUICE planetary coverage tool (1.0.0).
 jl-pangaia PanGaia JupyterLab	 jl-pipeman Jupyter lab with pipeman python library.	 jl-pyhc-test Testing pyhc environment	 jl-spice SPICE JupyterLab by the SPICE information system with additional Python libraries complementing some SPICE applications. It is a reference environment >	 jl-xmm-sas Jupyterlab XMM SAS
 jl_democonf This is a JupyterLab demo datalab	 jupyterlab Plain JupyterLab for demonstration of basic functionality.	 jwst Jupyterlab JWST	 jwst-gto-miri Jupyterlab JWST Miricle	 jwst-high-mem copy of Jupyterlab JWST, high memory version for internal use

Choose Datalab

Find a datalab in ESA Datalabs catalog

Filter results									
	Demoda A simple demo jupyterlab datalab		Euclid_test This is a JupyterLab demo datalab		G-Tomo The G-Tomo scientific data application (SDA) created by the EXPLORE project allows users to extract 1d profiles and 2d maps from the latest 3d dust extinction maps based on Gaia eDR3 and 2MASS data (Lallement et al. 2022).		GODOT JupyterLab GODOT is an ESA/ESOC flight dynamics software. This is a customization of JupyterLab in order to use GODOT via the Python interface.		INTEGRAL OSA JL INTEGRAL Jupyterlab with the OSA 11 analysis environment.
	INTEGRAL OSA VNC INTEGRAL VNC based datalab with the OSA 11 analysis environment and its GUI for interactive analysis.		JWST_General JupyterLab is a web-based interactive computational environment. This version was customized by the JWST mission with additional Python libraries. It is a reference environment for the JWST community.		NCTEST Test		PSA JupyterLab A JupyterLab based DataLab custom built for working with planetary science data in the Planetary Science Archive (PSA).		Pyxel datalab Pyxel is a novel, open-source, modular Python software framework designed to host and pipeline models (analytical, numerical, statistical) simulating different types of detector effects on images produced by Charge-Coupled Devices (CCD), Monolithic, and Hybrid CMOS imaging sensors.
	TESTMOD1 INTEGRAL Jupyterlab with the OSA 11 analysis environment.		XMM-SAS21.0.0 XMM-Newton SAS v21.0.0		YPE24-Euclid Jupyterlab datalab for the YPE Euclid workshop		YPE_Galaxy_Class Jupyterlab Datalab for the Galaxy classification workshop at YPE2024		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-astro JupyterLab with software and example notebooks for ESDC astronomical archives		jl-esdc-helio JupyterLab is a web-based interactive computational environment for creating notebook documents that allow programming in different languages. This version was customized by the ESDC with additional Python libraries to be a reference environment for Heliophysics research.		jl-euclid Euclid JupyterLab
	jl-euclid-c11g Euclid Commissioning JupyterLab		jl-gaia Datalab based on JupyterLab, customized by the Gaia mission with additional Python libraries and other tools to explore the Gaia DR3 catalogue.		jl-gssc GSSC JupyterLab		jl-herschel Herschel JupyterLab		jl-juice JupyterLab with JUICE planetary coverage tool (1.0.0).
	jl-pangaia PanGaia JupyterLab		jl-pipeman Jupyter lab with pipeman python library.		jl-pyhc-test Testing pyhc environment		jl-spice SPICE JupyterLab by the SPICE information system with additional Python libraries complementing some SPICE applications. It is a reference environment >		jl-xmm-sas Jupyterlab XMM SAS
	jl_democonf This is a JupyterLab demo datalab		jupyterlab Plain JupyterLab for demonstration of basic functionality.		jwst Jupyterlab JWST		jwst-gto-miri Jupyterlab JWST Miricle		jwst-high-mem copy of Jupyterlab JWST, high memory version for internal use

→ THE EUROPEAN SPACE AGENCY

ESA Datalabs [0.13.0/BETA]

HELP

Introduction

Login to ESA Datalabs

Logout from ESA Datalabs

Datalabs

Data Volumes

Workspaces

Pipelines

Using multiple Data Centres in ESA Datalabs

Service Desk and contacts

FREQUENTLY ASKED QUESTIONS

ACCESS TO ESA DATALABS

ASKING FOR HELP OR REPORTING AN ISSUE

USING ESA DATALABS

COMMON ISSUES AND WORKAROUNDS

SECURITY AND ARCHITECTURE

HELP

Here you will find support resources and documentation on using ESA Datalabs. Please browse through the chapters on the left side menu for instructions on how to use and navigate ESA Datalabs.

Below you can find the answers to Frequently Asked Questions. In case your question hasn't been answered or you wish to submit feedback or report an issue, please contact us in the [Service Desk](#).

FREQUENTLY ASKED QUESTIONS

WHAT IS ESA DATALABS?

ESA Datalabs is an e-science platform that allows users to bring their code to ESA's infrastructure and have direct access to ESA's archives. ESA Datalabs are full computational environments and the catalogue of ESA Datalabs ranges from new tools that have become de-facto standard for analysis, to complex legacy systems repackaged to run via a web browser. ESA Datalabs' underlying architecture is domain agnostic; it fosters research and innovation through the integration of transversal access to big data, notebook technologies and domain specific software. For example, customised JupyterLab environments are readily available for astronomers, scientists in Earth Observation related fields, or researchers in global navigation. Moreover, ESA Datalabs support for development environments such as Octave, or reference tools such as TopCat in astronomy that enable reuse of existing code baselines. In addition, ESA Datalabs will soon include full support for pipelines and support for developing datalabs, both of which are currently limited to a smaller group of users for beta testing.

WHAT IS THE CURRENT STATUS OF ESA DATALABS?

ESA Datalabs is under active development as of early 2023. Users should not rely on the platform for operational use since data loss and unplanned downtime may occur.

ACCESS TO ESA DATALABS


HOW CAN I ACCESS ESA DATALABS / HOW DO I CREATE AN ACCOUNT?


To access ESA Datalabs you will need a Cosmos account first. If you do not have a Cosmos account please see the instructions in the [Cosmos FAQ](#). Usually, a Cosmos account is provided because your institution/company is part of a collaboration. If you know the credentials of your Cosmos account, you must first self-register using the [ESA Datalabs Self-Registration](#) tool. If you wish to recover your Cosmos account credentials read the FAQ item on recovering Cosmos credentials.

ESA DATALABS ASKS FOR A COSMOS ACCOUNT. WHAT IS A COSMOS ACCOUNT AND HOW DO I GET ONE?

Cosmos offers the single sign-on service that provides access to several secure ESA websites and services. If you do not have a Cosmos account please see the instructions in the [Cosmos FAQ](#). Usually, a Cosmos account is provided because your institution/company is part of a collaboration.

Last updated on 18/06/2024, 14:34





→ THE EUROPEAN SPACE AGENCY

14

ESA Datalabs Community – Today



1200 Registered users

950 Beta Users

60 daily users

ESA Datalabs Hackathon, Jan 2025
NLP for Space Science Workshop, Sep 2024
.Astronomy 13, ESAC 23-26 April 2024
Masterclass on JWST, ESAC 12 Dec 2023
JWST Summer School, Heidelberg 4 - 8 Sept 2023
Euclid Workshop, Copenhagen 21 - 22 Jun 2023
ESA Datalabs Workshop, ESAC 24 - 25 Nov 2022

