

ESA Datalabs: Expanding Space Archives with Innovative Analysis and Collaboration Capabilities

V. Navarro, F. Marinic, S. del Rio, M. A. Diego, M. Lopez-Caniego, A. Anku, C. Arviset

PSIDA 2022, 21.06.2021

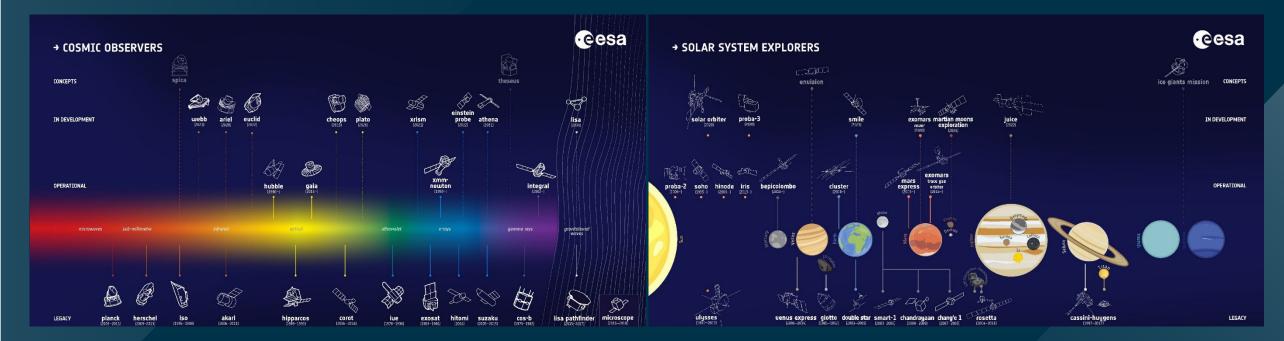
@ivicentenavarro

ESA UNCLASSIFIED – Releasable to the Public



Space Science Context





 \mathbf{k}

+

*

ESA Datalabs – Cross Domain Data Exploitation Platform .

→ THE EUROPEAN SPACE AGENCY

ESA Datalabs [0.3.0/BETA]

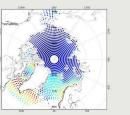


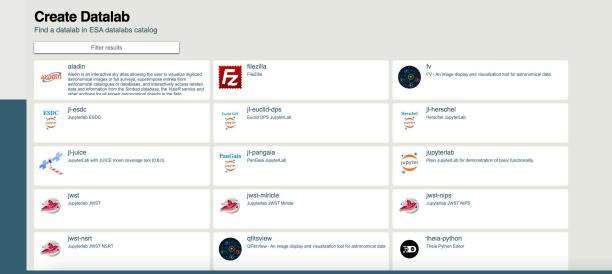
BRING YOUR QUESTIONS TO THE DATA

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.

There is, however, a new profound change: data has become a major technological challenge. Increases by multiple orders of magnitude in dataset size means that transferring data to a scientist is often unfeasible.

ESA datalabs gives you a privileged position; bring your code directly to ESA's infrastructure – there is a great set of tools and programming languages are flexible – and execute it with direct access to ESA's archives.





datalabs.esa.int

→ THE EUROPEAN SPACE AGENCY

*

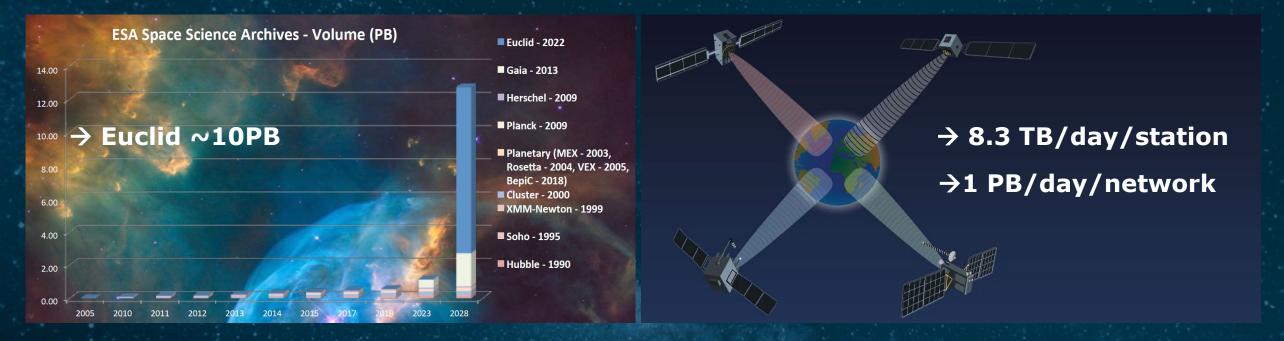
eesa

A Q D A Q



Data Exploitation Paradigm Shift



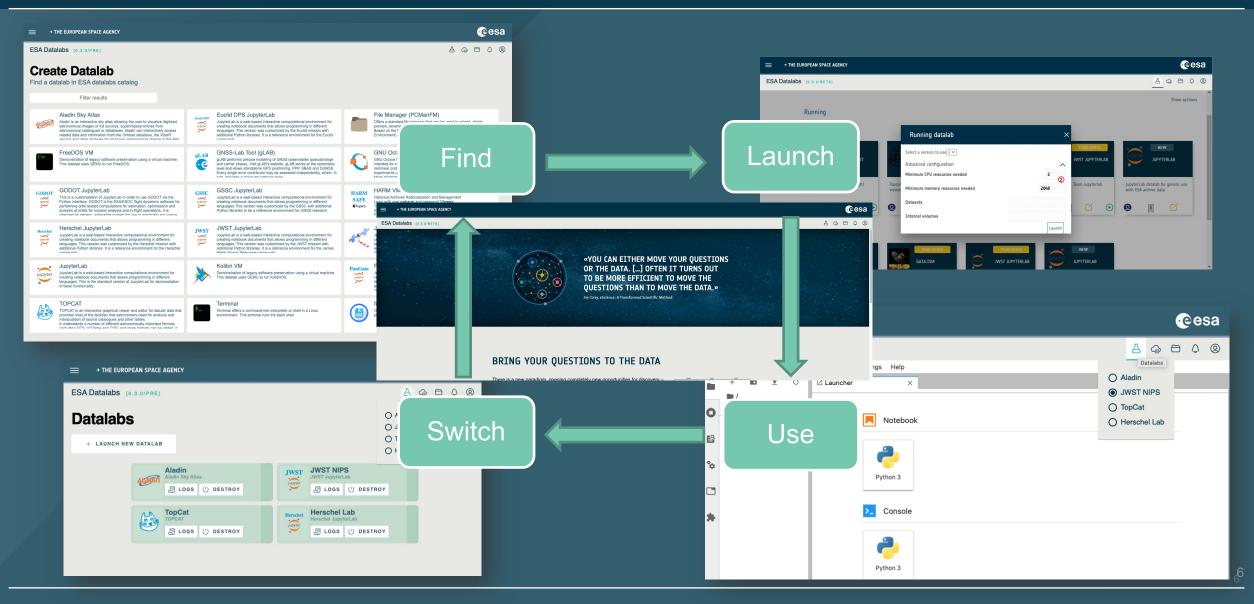


From bring the data to the user To bring the user to the data

💳 📕 🚼 🧫 🔚 🔚 🗮 🔚 📕 🔚 🔚 🔤 👬 🔤 🔤 🚱 📴 📕 🛃 🛃 📩 🖬 🖉

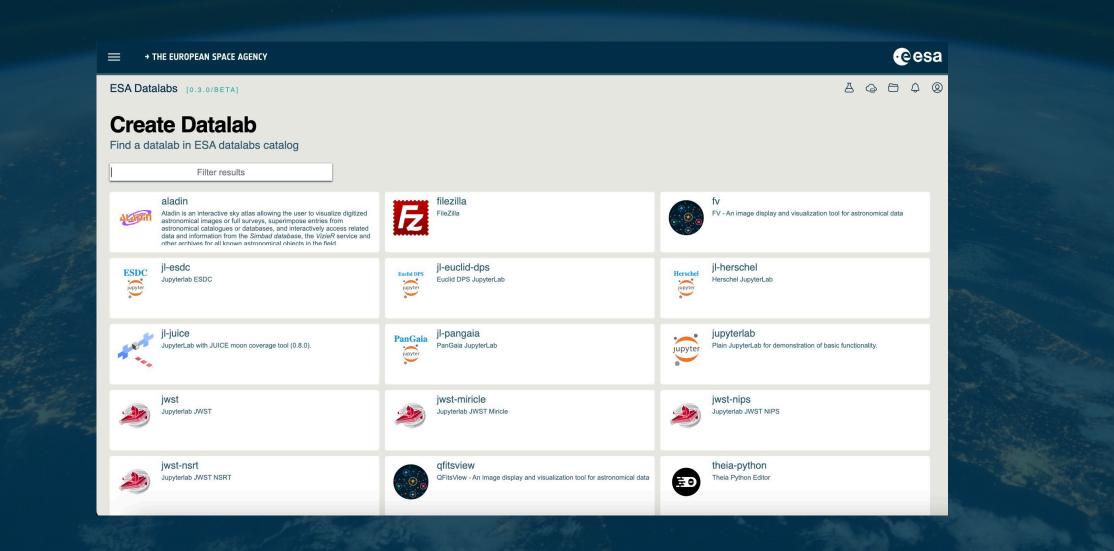
Datalabs – Standard Flow





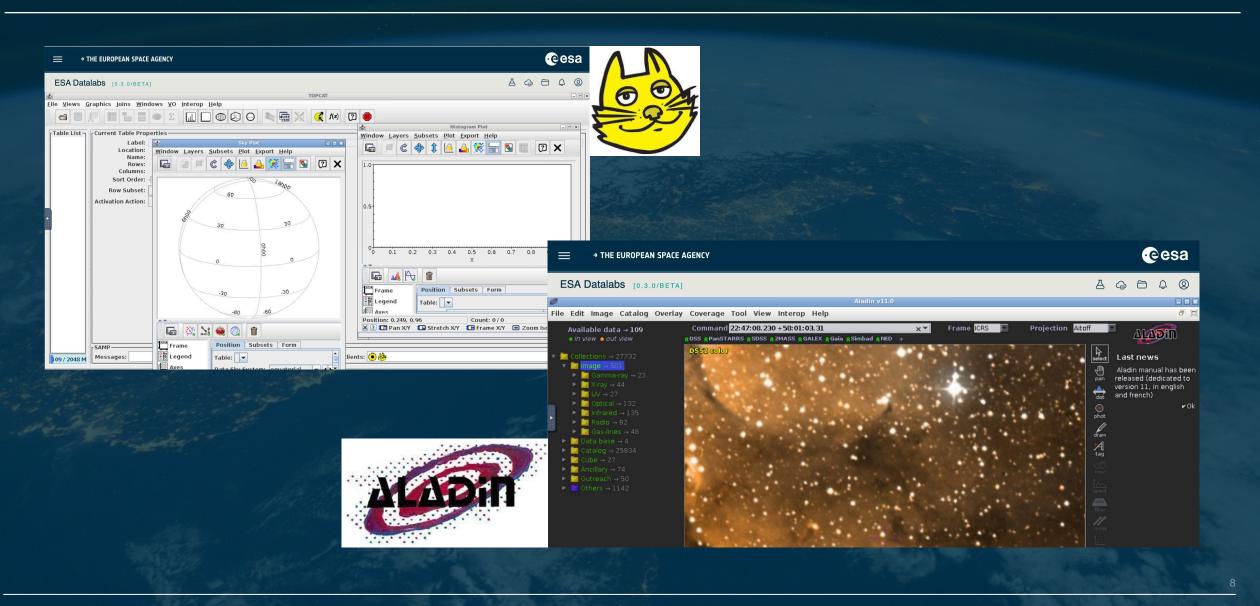
FIND





FIND – Desktop Apps





FIND – Web Apps (i.e. JupyterLab)



eesa

A @ B A @

🕸 Python 3 (ipykernel) 🔇

False

False

False False

-1 False

dataproducttype calibrationlevel public

eesa

ESASKY

imag

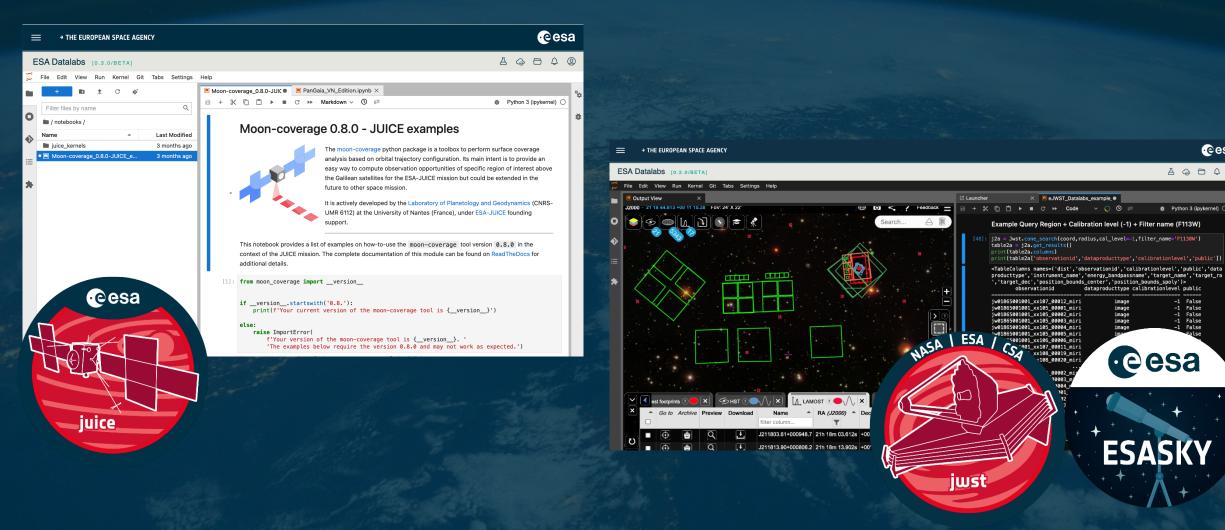
image image

image image

× 🖪 eJWST_Datalabs_example_ ●

observationid

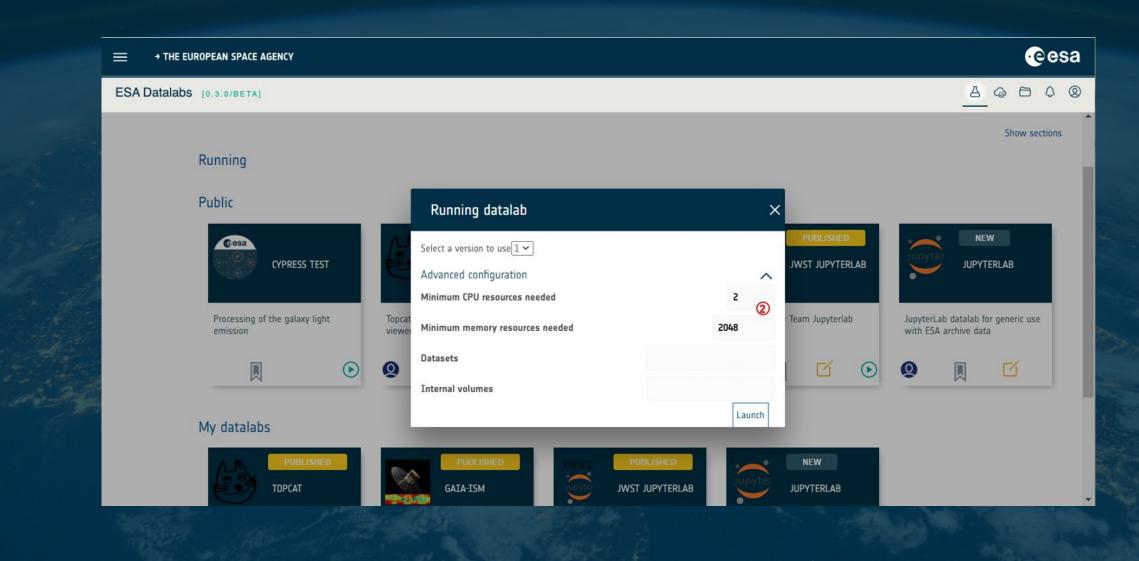
108_00020_mir 00002 10003



₩ + 0 |

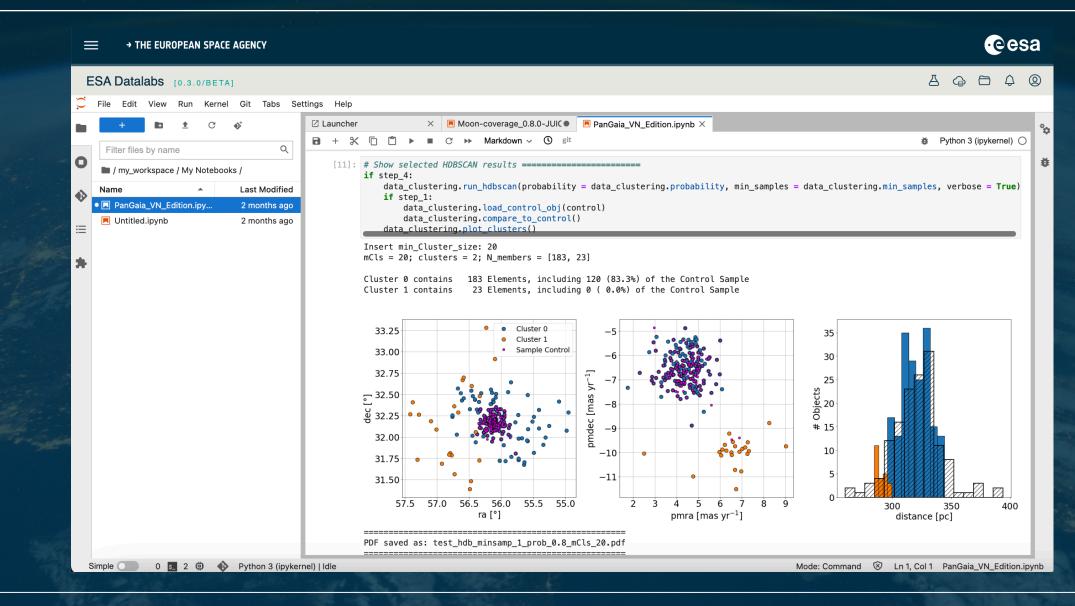
LAUNCH





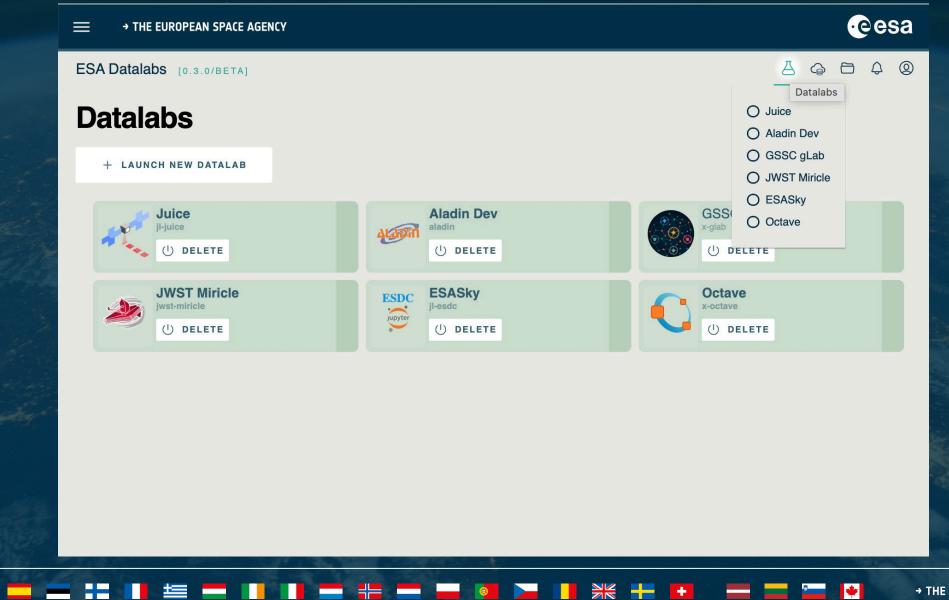
━ ━ ■ ■ # = = # # ■ # = ■ ■ # = # ● ■ ■ ■ # = = = = ■ • ●





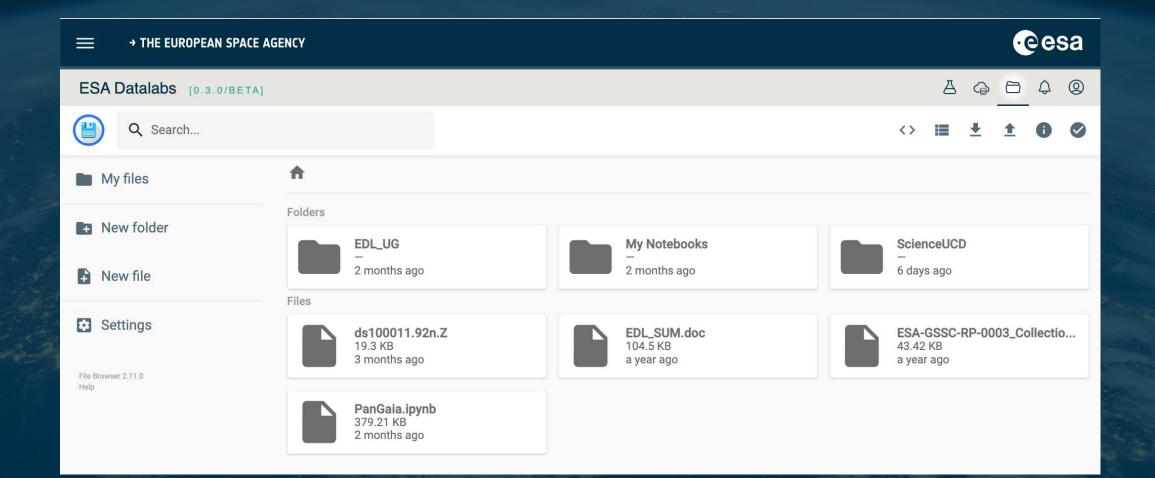
SWITCH





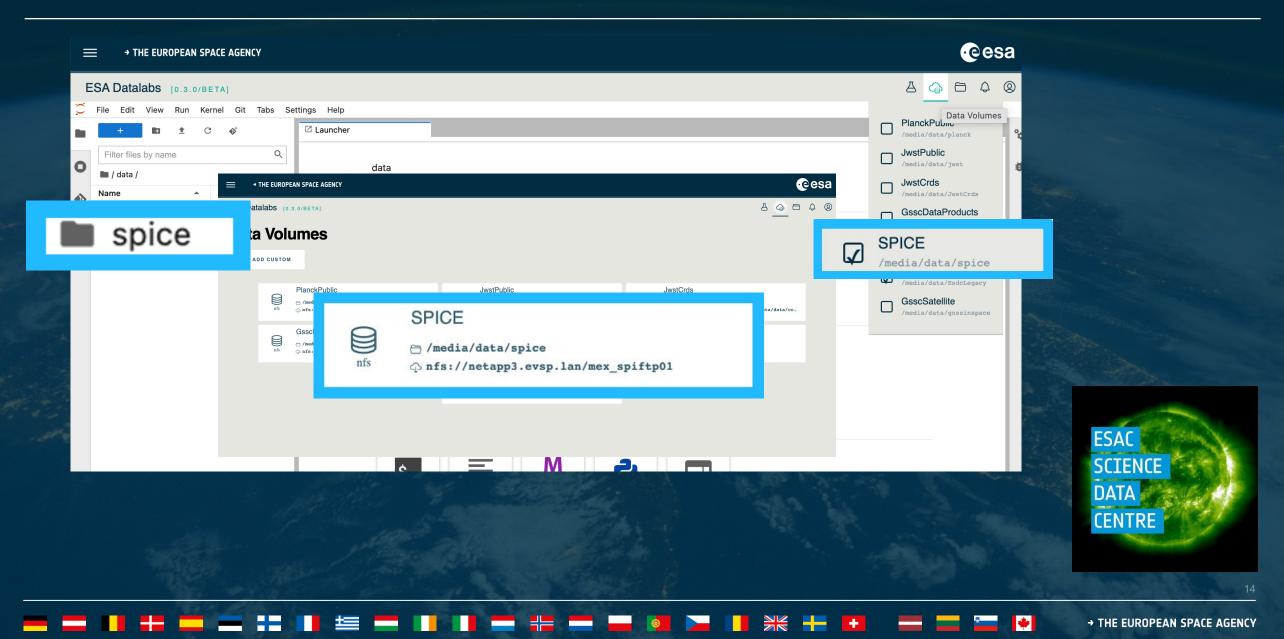
Personal Workspace





Data and Computing Colocation - Coarse Grain



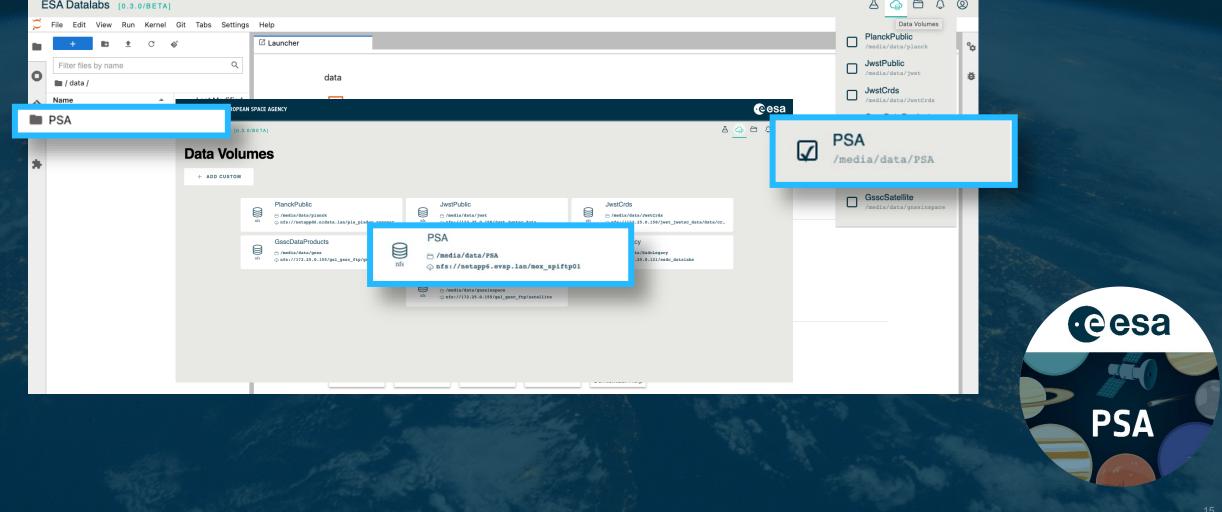


Coarse Grain Integration with PSA (evolution opt)

 \equiv

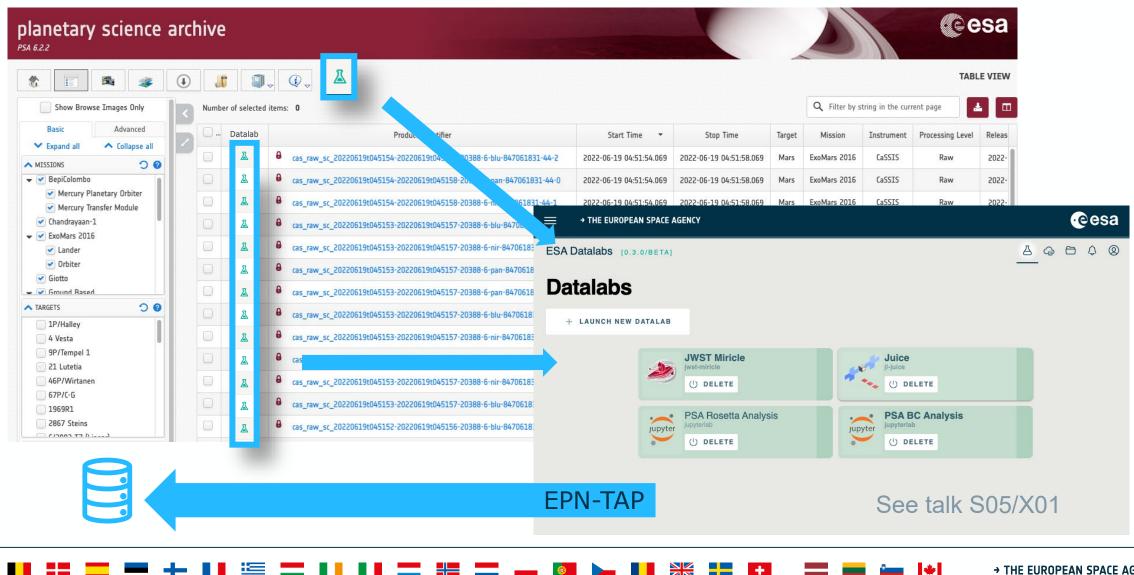
→ THE EUROPEAN SPACE AGENCY





Fine Grain Integration with PSA (evolution opt)





→ THE EUROPEAN SPACE AGENCY

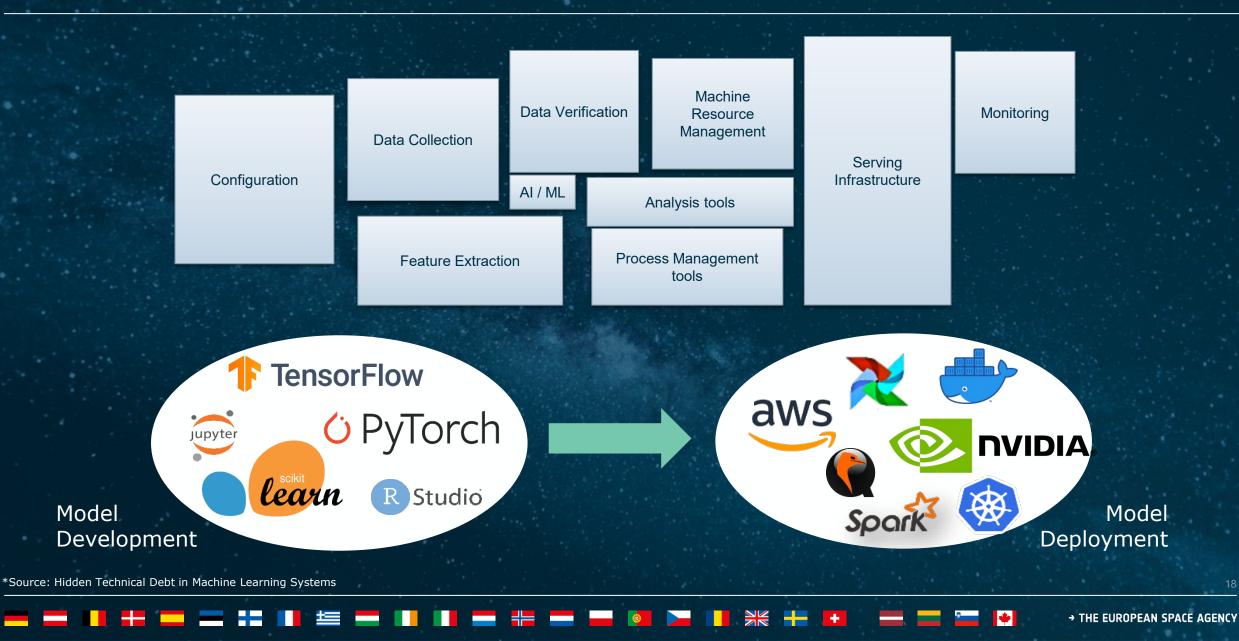


Data Processing Innovation

European Space Agency

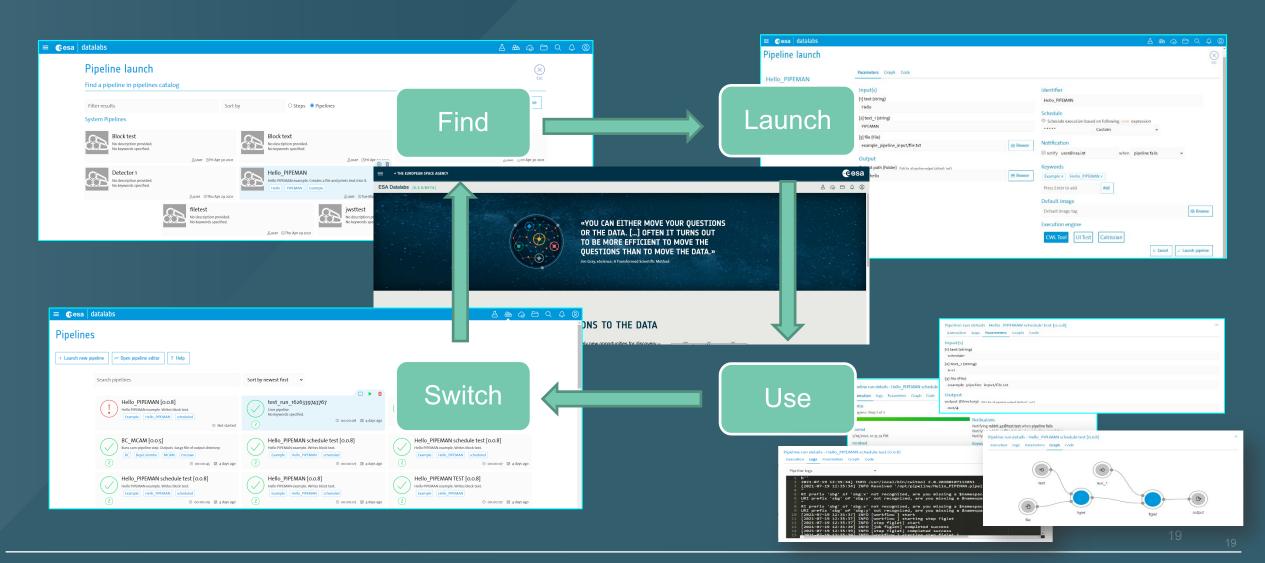






Pipelines – Standard Flow

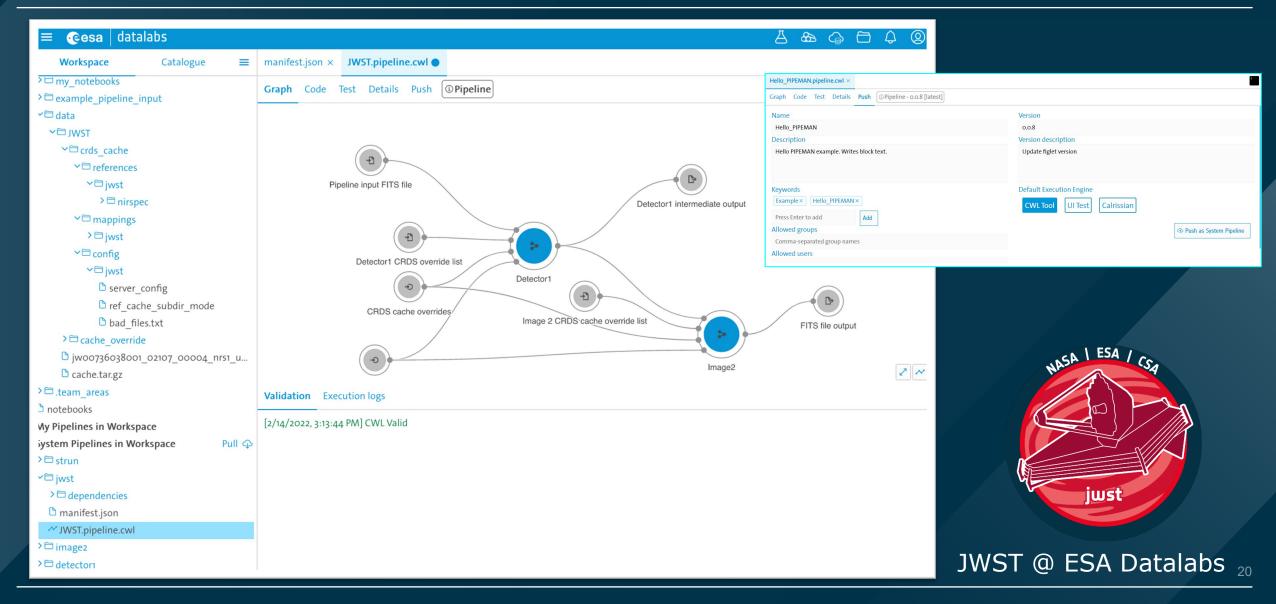




💳 💶 📕 🚼 💳 🔚 📕 🗮 🔚 📕 🚛 📕 📲 📥 📲 🔤 🛶 🧖 🖕 📲 👫 🕂 🖬 📰 🔤 🔤 🙀 Yhe European space agency ce agency

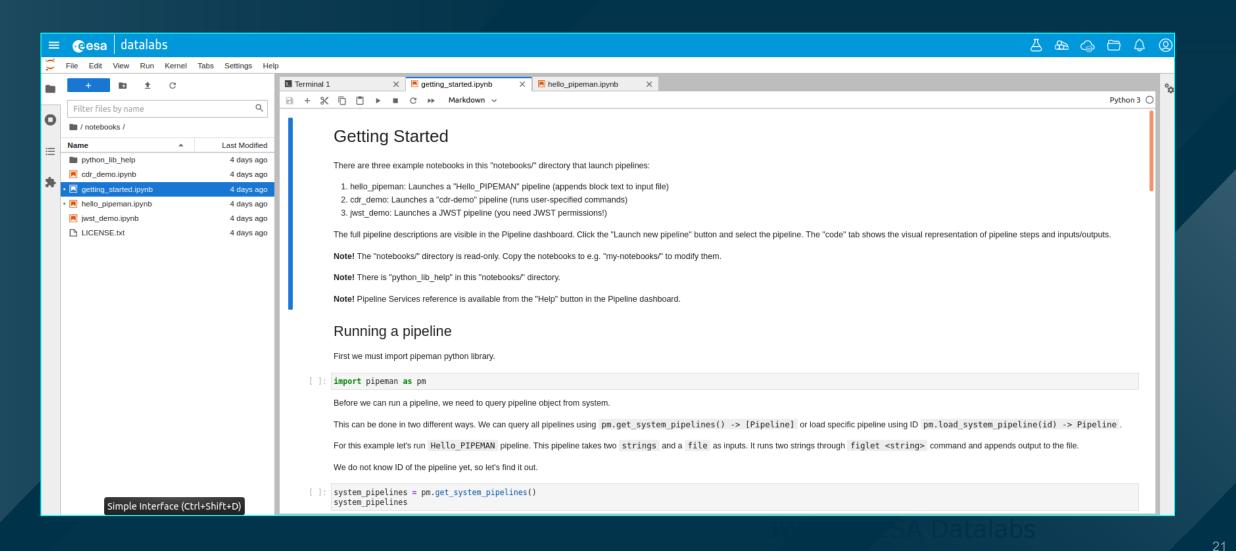
Pipeline Editor





European Space Agency

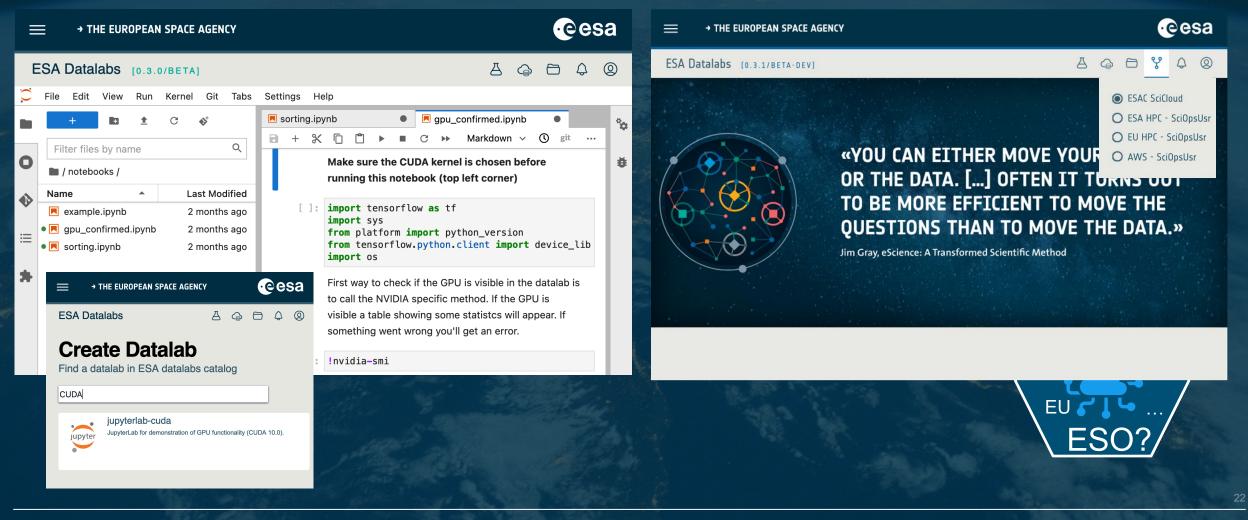
JupyterLab Pipeline Management



GPU Computing and Federated Scalability



→ THE EUROPEAN SPACE AGENCY



At vero eos et accusamus et iusto odio dignissimos ducimus qui blanditiis praesentium voluptatu deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non pro dent, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga. narum quidem rerum facilis est et expedita distinctio. Nam loero tempore, cum soluta nobis eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, om roluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut offic lebitis aut rerum necessitatibus saepe eveniet u et roluptates repudiandae sint et molestiae n ecusandae. Itaque earum rerum hic tenets e piente delectus, ut aut reiciendis voluptatib naiores alias consequarur aut perferendit curve pus asperiores repellat.



European Space Agency

Data Centric Collaboration - Team Workspaces



0

¢

ŧ

*

Ξ	∃ → THE EUROPEAN SPACE AGENCY		esa
E	SA Datalabs [0.3.0/BETA]	A	@ <mark>=</mark> 4 @
	File Edit View Run Kernel Git Tabs Settings + • • C • · · · Filter files by name - · · · · · • - · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · · · · • · · · · · <	Help Moon-coverage_0.8.0-JUIC●	ilitate the at are apabilities sted in ae analysis ew members chine
		added to the queried table (like e.g. the distance, computed as	the inverse of

added to the queried table (like e.g. the distance, computed as the inverse of

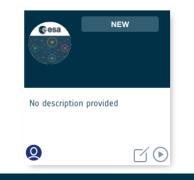
+

Application Centric Collaboration - AppStore



≡ @esa datalabs										Ą	
		Search datalab c	atalog					Q New datalab			
Public											Customize view
	≥ 219	55 SEPPTEST- 571690539014		CYPRESS SEPPTEST- 219 1111.6848610854097	Cesa • • • •	CYPRESS SEPPTEST- 219 1419.7853655034294	Cesa • • • •	CYPRESS SEPPTEST- 219 1713.2724871003359	Cesa O O O	CYPRESS SEPPTEST- 219 1673.9987705736833	
Test .	Automated		Test Automat	ted	Test Automa	ted	Test Automa	ted	Test Automa	ated	
_		♫☆⊙		$\bigstar {ullet}$		$\Box \& igodot$		$\Box \& igodot$		$\Box \& igodot$	
See more	e										

Developed by me



Datalabs Editor

 \equiv

 ${}^{\textcircled{\baselineskip}}$



📀esa datalabs



(i)

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.

Modification date: n/a
Datalab version:
Author: Vicente Navarro

Where are your files?		File upload	
Select your source You can select or drag and drop the files or archive of the exported datalab that you want to use in your next datalab. To upload a complex file structure please use a zip	Analysis.ip	12/05/2022 10:18	256.9KE

+ → THE EUROPEAN SPACE AGENCY

A G C A Ø

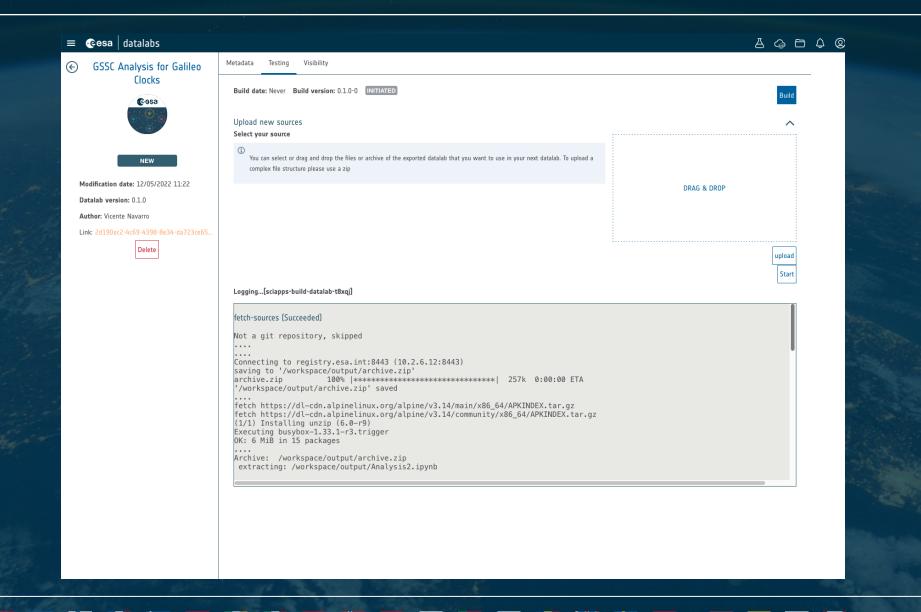
Datalabs Editor - Metadata



⊙esa datalabs			4964
Cosa	Metadata Testing Visibility		
• • •	Mandatory information		^
NEW	Title	Title/name of the datalab	
dification date: 12/05/2022 11:01	Alternate Name	Short name or acronym for the datalab	
alab version: hor: Vicente Navarro	Abstract	Abstract (a short description)	
: 2d190ec2-4c69-4398-8e34-da7	Version	Version information	
Delete	Recommended information		^
	Creator Name	Name of the organisation authoring the datalab	
	Description		h.
	Instrument	Instrument	
	Mission	Mission	
	Thumbnail Url	URL to an icon (for thumbnail)	
	Optional information		^
	Associated File Types	List of associated filetypes - if applicable/pertinent (e.g. FITS, VOTable, GeoTIFF, netCDF)	
	Audience Type	Intended audience for the datalab	
Citatio	n	Citation for the datalab (e.g. article DOI)	
	Contact Point 10	UKE of the contact point	

Datalabs Editor - Build





Datalabs Editor - Share



≡ •@esa datalabs	A Q A
GSSC Analysis for Galileo	Metadata Testing Visibility
Clocks	(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: 12/05/2022 11:23 Build version: 0.1.0-0 SUCCESS Sharing audience Public, everyone can use it Private, specify a list of users
DRAFT	Publish
Modification date: 12/05/2022 11:40 Datalab version: 0.1.0 Author: Vicente Navarro	License Apache License 2.0
Link: 2d190ec2-4c69-4398-8e34-da723ce65	Export
	() Clicking on this button will allow you to get an archive of the current datablab information that can be imported then in another datalab
	History 12/05/2022 11:23: Build started 12/05/2022 11:25: Build ended with status BUILD_SUCCESS

Take away messages



•eesa

ESA DATALAB

ESA Science strategy support

- Increased science return from its missions
- Increased science operations efficiency

Innovation traits

- Science data exploitation coupled with data
- Science pipelines for current and future needs
- Collaborative research & citizen science

—— — II == — II ±= — II ±= — II = = = = 0 = 0 = ...

Thank You!



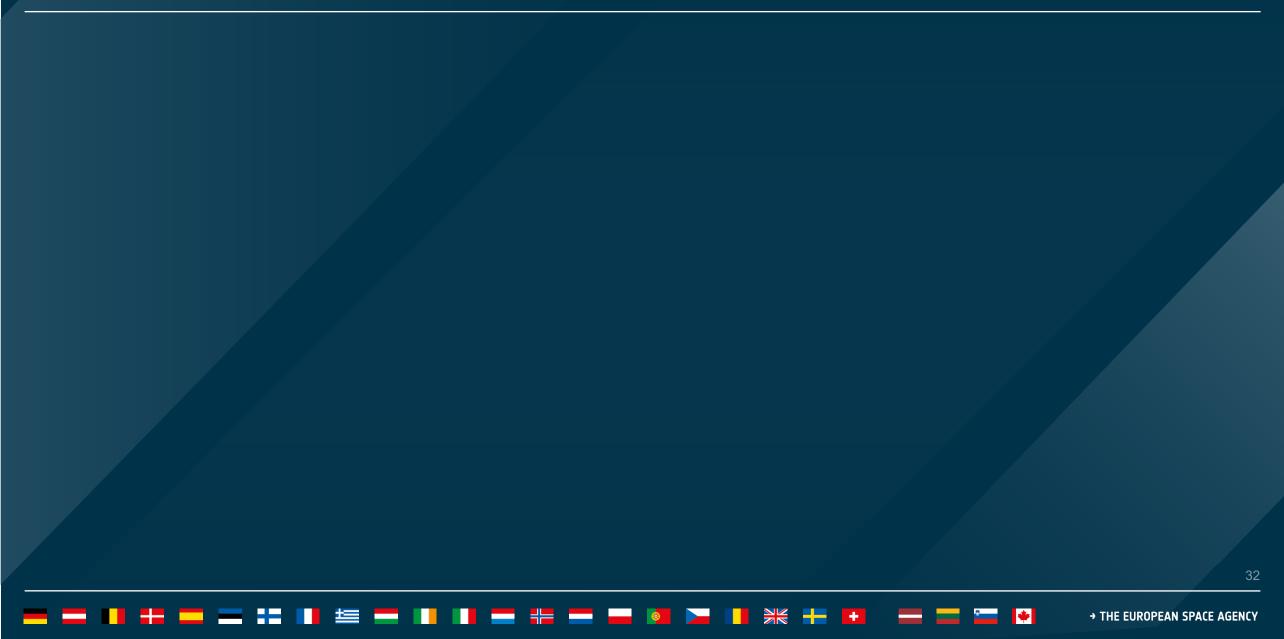


datalabs.esa.int

... special thanks to ESA Datalabs Industry Teams and User's Group (+50 members)

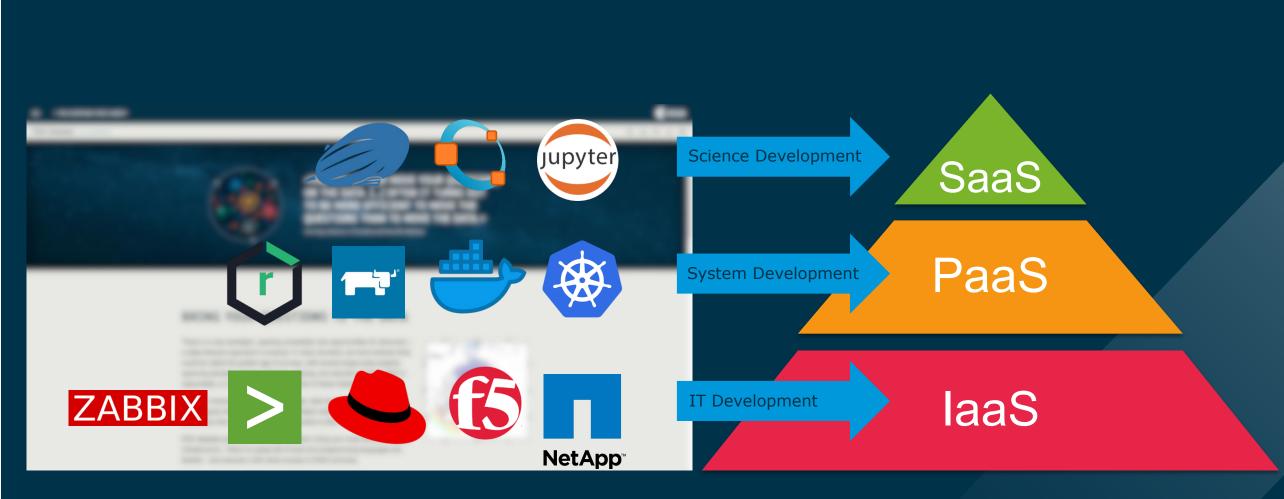
Bonus Slides





IT Paradigms





━ ━ ■ ■ ₩ ₩ = = # # ■ ■ # = # ■ ■ ₩ ₩ += ₩ = = = ₩

33

European Space Agency

IT Monitoring

±==





+

_

→ THE EUROPEAN SPACE AGENCY

*

IT capacity



- Pre-production Environment: <u>https://datalabs.esa.int</u>
 - System Cluster:
 - 6 nodes
 - 56 vCPUs
 - 100 GB RAM
 - Currently, 70 containers
 - User Cluster:
 - 6 nodes
 - 48 vCPUs
 - 256 GB RAM
 - Currently, 216 containers
 - Storage:
 - 500GB for system persistence
 - 1TB for user persistence

- Development & Acceptance environments:
 - System cluster: running on shared cluster, where other systems are located
 - User cluster:
 - 4 nodes, one with GPU
 - 20 vCPUs
 - 74 GB RAM
 - Currently, 73 containers
 - Shared by DEV and E2E envs.
 - Storage: shared storage with other projects.