

SpaceOps-2023, ID # 483

A New Age for Data Exploitation in Science and Mission Operations

Vicente Navarro^{a*}, Sara del Rio^b, Miguel Angel Diego^c, Marcos Lopez Caniego^c, Filip Marinic^a, Jan Reerink^a, Christophe Arviset^a

^a *Department of Science and Operations, European Space Agency (ESA), Camino Bajo del Castillo S/N, Villanueva de la Cañada, 28691, Madrid, Spain, vicente.navarro@esa.int, filip.marinic@esa.int, christophe.arviset@esa.int*

^b *Department of Science and Operations, Rhea for ESA, Camino Bajo del Castillo S/N, Villanueva de la Cañada, 28691, Madrid, Spain, sara.del.rio@ext.esa.int*

^c *Department of Science and Operations, Aurora for ESA, Camino Bajo del Castillo S/N, Villanueva de la Cañada, 28691, Madrid, Spain, miguel.angel.diego@ext.esa.int, marcos.lopez.caniego.alcarria@ext.esa.int*

* Corresponding Author

Abstract

At the European Space Astronomy Centre (ESAC) near Madrid, the ESAC Science Data Centre (ESDC) hosts ESA archives for Astronomy, Planetary and Heliophysics Space Science. Furthermore, the GNSS Science Support Centre (GSSC), with special attention to Galileo and EGNOS, consolidates an ESA archive for scientific exploitation of Global Navigation Satellite Systems (GNSS). The sheer amount of data derived from current development plans for Space Science and Navigation Systems, constitutes a unique opportunity for research. Nevertheless, this deluge of data triggers a depart from standard processes and architectures supporting archives exploitation. More than ever, science innovation requires to integrate multiple data sources, sophisticated computing algorithms and cross-domain collaboration tools to achieve its goals. Hence, motivated by this increasingly complex landscape, 'ESA Datalabs' and 'GSSC Now' platforms expand Space and Navigation Science archives respectively, with a brand-new palette of collaboration, exploration, and analysis capabilities. Designed from the ground-up with a focus on end-user productivity, these platforms accelerate time to market for science operations developments and increase their return on investment, unleashing the full potential of ESA Missions archives. Transparently to the user, the two platforms implement advanced network and computing capabilities to leave behind the discovery and download science era. 'ESA Datalabs' and 'GSSC Now' showcase a new era characterized by archives tight coupling with exploitation tools. These platforms leverage on mainstream big data, cloud, virtualisation and container technologies to create a software as a service (SaaS) computing framework. This work presents ESA vision and initiatives to provide a multi-mission environment for Science Data Exploitation that pivots around the paradigm shift characterised by the move of processing components to the data, rather than the move of data to the users.

Keywords: ML, Exploitation Platforms, HPC, Containers, Big Data, SaaS

Nomenclature

System Domain: high level subsystem responsible for the implementation of a user or technical functionality.

Information assets: data items located and managed in a system repository for preservation and processing purposes at different levels of processing.

Processing assets: software models, programmes, libraries and services in different packaging formats (VMs, containers ...).

Workspace: it refers to storage areas where users can read and write ESA Datalabs assets. A workspace includes persistent areas (assets repositories) and temporary areas (e.g. a scratch repository for temporary work). A workspace is associated with at least one user and every user has at least one private workspace.

Team Workspace: A team workspace is any workspace shared with multiple users. Users of a team workspace have individual roles and permissions that control what they are allowed to do inside a team workspace (e.g. read or write any asset). A team workspace is referred as project by other systems (set of assets managed by a user's group of users).

Acronyms/Abbreviations

Artificial Intelligence (AI), Common Workflow Language (CWL), ESAC Science Data Centre (ESDC), GNSS Science Support Centre (GSSC), High Performance Computing (HPC), Infrastructure as a Service (IaaS), James Webb Space Telescope (JWST), Machine Learning (ML), Platform as a Service (PaaS), Software as a Service (SaaS), Virtual Machine (VM)

1 Introduction

Nowadays, the sheer amount of data collected from space-borne and ground-based sensors, is changing past approaches towards data processing and storage. In the Information Technology domain, the rapid growth of data generation rates, expected to produce 175 zettabytes worldwide by 2025, has introduced new software architectures for High Performance Computing. Thus, current workload distribution, with 80% taking place remotely from data ‘on the cloud’, and 20% near the data ‘on the edge’, is likely to be inverted by 2025. This landscape has led to a new golden age of Machine learning (ML), able to extract knowledge and discover patterns between input and output variables given the sheer volume of available training data.

Operations of ESA missions includes Navigation, Earth Observation, Space Science and more. The significant and ever-increasing amount of data, generated by these missions both, from a scientific and operational viewpoint, calls for new technological solutions able to maximise data exploitation. In addition to existing data discovery capabilities, a brand-new palette of features shall be deployed to extract insights from these multi-mission, federated data sources. At the European Space Astronomy Centre (ESAC) near Madrid, the ESAC Science Data Centre (ESDC) hosts ESA archives for Astronomy, Planetary and Heliophysics Space Science. Furthermore, the GNSS Science Support Centre (GSSC), with special attention to Galileo and EGNOS, consolidates an ESA archive for scientific exploitation of Global Navigation Satellite Systems (GNSS). In the era of Big Data and Machine Learning, traditional data downloads from archives are being replaced by algorithm uploads and computation close to the data. Hence, “ESA Datalabs”[1], and “GSSC Now” [2] platforms provide seamless integration of ESA Archives with computing capabilities to deliver sophisticated collaboration, exploration, and analysis features.

Based on technologies such as, Docker containers, Kubernetes, and JupyterLab, these exploitation platforms aim to realise the full potential of ESA archives for Space and Navigation Science. Designed from the ground-up with a focus on end-user productivity, these platforms provide a catalogue of computing capabilities co-located with ESAC archives through an innovative software-as-a-service delivery model.

2 Multi-Domain Exploitation Platform

The aforementioned increase in data volumes and processing needs of space missions, has led to the identification of an integrated solution referred as Exploitation Platform [3]. While Thematic Exploitation Platforms are typically designed to support a research field, ESA Datalabs delivers a multi-mission and multi-disciplinary framework to address four major objectives:

- Scientific exploitation of spacecraft data. This objective, at the core of science operations, aims at implementing data processing systems to enable analysis and scientific discoveries.
- Long term data processing preservation. This objective aims at ensuring proper storage and access to mission data analysis tools and processors well beyond mission’s lifespan. When missions enter in legacy phase, their raw data and scientific products are preserved into science archives. Similarly, there is the need to preserve associated data analysis software to support future re-analysis of the data. For instance, the capability to run mission planning tools of old missions might be required to prepare new missions. Maintaining these systems in the long term represents a real challenge, that can be solved by “containerising” such software to make it easily available through the platform [4].

- Collaborative Research Environment. This enables scientists to work together on the platform itself, sharing data and software, initially into a restricted workspace and possibility later on publishing these publicly to all users accessing the platform.
- Enable cost effective approach towards mission operations. This objective is transversal to the previous three objectives and aims at maximising synergies and effectiveness across projects and organisational units.

In this context, the design of ESA Datalabs requires the adoption of a high-level abstraction viewpoint to define an appropriate set of features that will enable future development of mission specific extensions.

Therefore, the community of users in charge of specifying ESA Datalabs (see Fig.1), has aimed at achieving good coverage of multi-mission and multi-domain needs. Representatives from multiple ESA missions and the ESAC Science Data Centre (ESDC) [5], which hosts ESA archives for Astronomy, Planetary and Heliophysics Space Science, have provided a large multi-mission and multi-domain userbase. Additionally, the involvement of representatives from the GNSS Science Support Centre (GSSC) [6], which hosts ESA archive for the scientific exploitation of Global Navigation Satellite Systems, has expanded multi-domain’s understanding of ESA Datalabs beyond Space Science.

This ESA Datalabs community has been involved throughout the development in several workshops and formal activities for requirements gathering to ensure a comprehensive understanding of user needs. Contributions from these communities have led to the identification of system use cases (see Fig.2), capturing the most important goals and actors.



Fig. 1. Groups involved as part of the ESA Datalabs Community with “early adopters” in green

The main actors of the platform can be summarised as:

Scientific User: this actor is a consumer of the platform processing services and/or of the products stored and generated via the platform. These users do not perform any integration or development work within the platform but may start processing based on already integrated processing services. A Scientific User is not necessarily an expert on how to access ESA data holdings but is a domain-expert (e.g. astronomy, earth sciences, navigation, etc).

Operator: this actor is responsible for the execution of operational tasks to ensure correct execution and maintenance of the platform. The operator performs overall monitoring and tracking activities through dedicated interfaces for management purposes.

System Scheduler: this actor enables system-to-system interactions and routine procedure executions without intervention from the Operator.

The main goals of the platform can be summarised as:

- Interactive and collaborative capabilities for co-located data analysis (i.e. JupyterHub).
- Web based instantiation of legacy systems (i.e. data processing, mission planning, etc, ...).
- Execution of data processing pipelines on archived science data.
- Personal storage space in the platform private user data processing code close to the archives.
- High Performance Computing (HPC) capabilities for processing of massive, highly distributed datafeeds.
- Open Science and collaboration capabilities through a “Science App Store”.

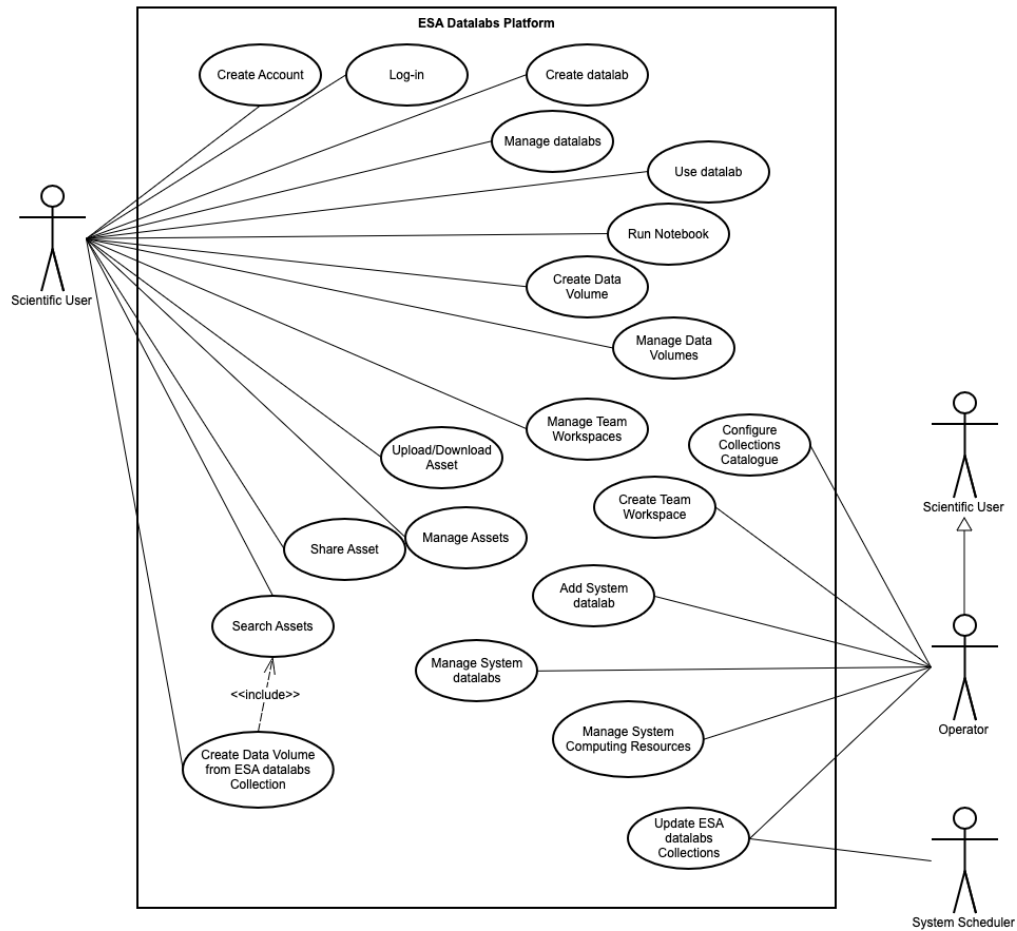


Fig. 2. ESA Datalabs Platform High Level Use Cases

3 Reference Architecture

A variety of activities, like SciServer [7], Escape [8], SWAN [9] or LSST [10] among others, enable the provision of advanced collaboration systems to execute data analysis processors deployed close to the data. Similar cyberinfrastructures, such as those found in Earth Observation programmes like Copernicus [11], are at the core of many efforts to manage and analyse large data sets [12]. As a framework, ESA Datalabs architecture provides a semi-complete expandable software infrastructure fully specified by its deployment context. The architecture of ESA Datalabs offers significant expansion capabilities that can accommodate multi-domain and multi-mission requirements through software plug-in and configuration mechanisms. Additionally, its interface extension features facilitate the integration of platforms like “GSSC Now”, scaling-up the solution to create a System of Systems [13] for science exploitation. From a Science Operations perspective, ESA Datalabs acts as an internal leverage agent,

offering a reusable and consistent solution that reduces time-to-market and minimises risks. Furthermore, ESA Datalabs promotes inter-organisational collaboration through federation, fostering external collaboration and teamwork.

ESA Datalabs High Level Architecture [14] resembles a typical N-tier architecture[15], organised into layers that group a set of system domains (see Fig 3):

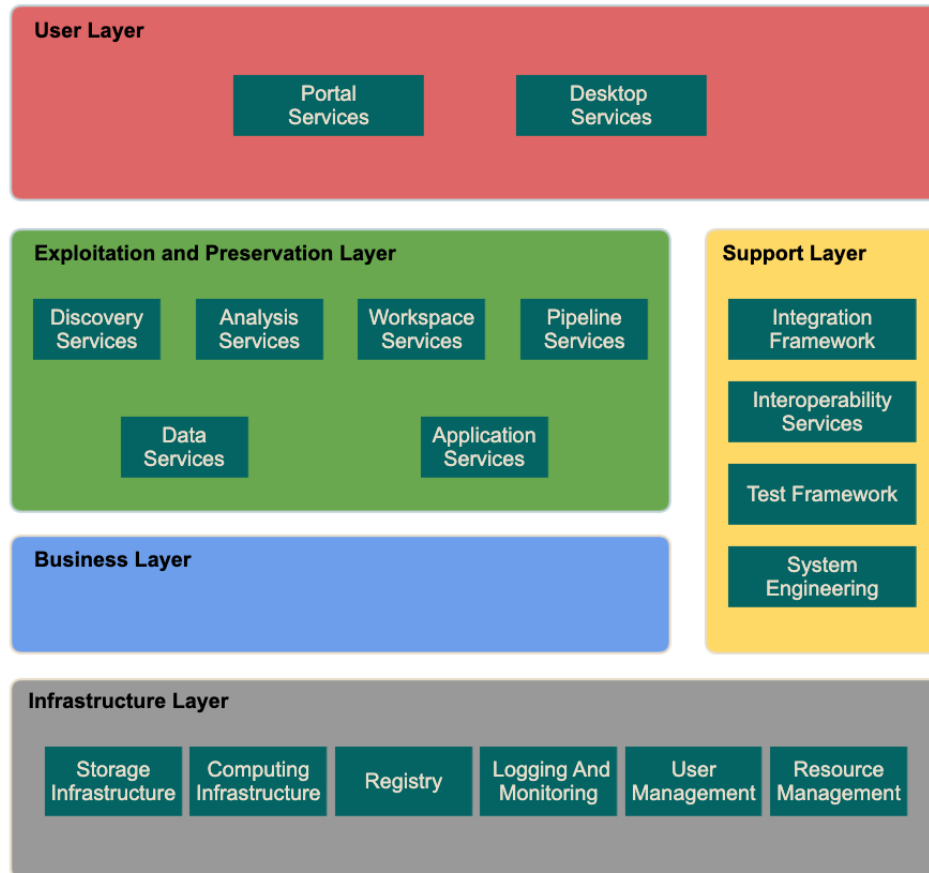


Fig 3. ESA Datalabs High Level Architecture

User Layer: this layer provides the HMIs for users and administrators to access all functionalities. The objective of this layer is to decouple presentation logic from business logic implemented by other layers. This layer allows smooth integration of HMI functionalities into a homogeneous look & feel through the provision of extension points.

Exploitation and Preservation Layer: this layer groups domains implementing generic, and user specific analysis functionalities. This layer groups domains implementing access to information and processing asset integrated in the platform. The system will rely on executable modules and data at different levels of processing which will be natively stored on the platform or federated to other systems. These two types of assets will be combined in the exploitation layer to deliver more complex services and products.

Business Logic Layer: this layer is key to provide support for multi-domain requirements. This area hosts extensions and configurations for specific data processing needs. The logic in this layer can be natively hosted in the platform or federated to other systems via the Interoperability Services.

Support Layer: this layer groups domains implementing components that include libraries providing common features required by multiple domains across the whole system. These libraries are likely to require glue-code to adapt to specific component needs.

Infrastructure Layer: this layer provides basic support for the implementation of the preservation and support layers. It is based on COTS that can be reused "as-is" with an integration pattern mainly based on the configuration of a set of parameters to adapt the behaviour of the COTS to the specific needs.

ESA Datalabs architecture is based on Kubernetes (k8s) as the orchestrator solution used for service management, scalability and availability. More concretely, the current environment deployed for ESA Datalabs comprises two k8s clusters provisioned through Rancher. In summary, ESA Datalabs system domains rely on a complex technology stack that spans across the cloud computing space: Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), and Infrastructure-as-a-Service (IaaS).

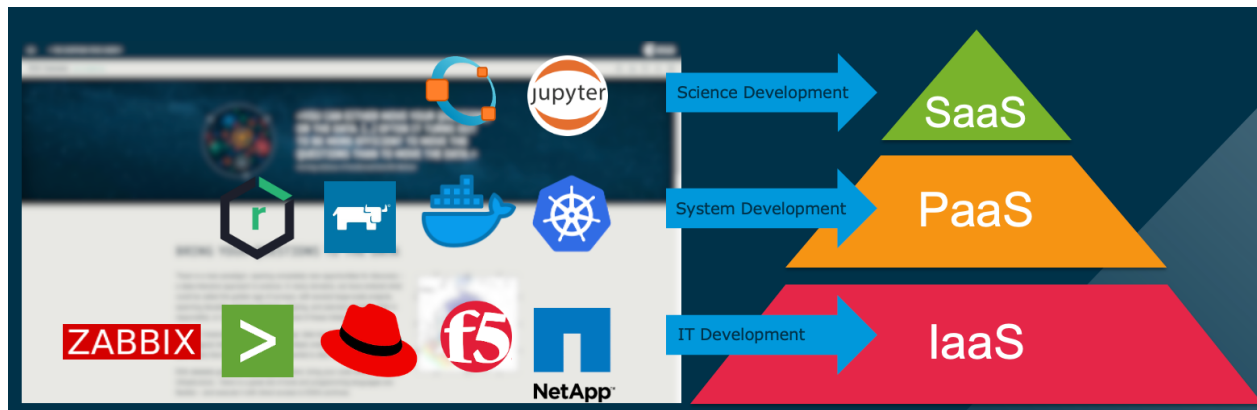


Fig. 4. Architecture and technology stack

At its core, ESA Datalabs essentially acts as a type of specialized orchestrator. Its main function is to keep track of users, teams, workspaces, and multiple types of digital assets (datasets and software), instantiating computing environments for users to carry out analysis and discovery tasks. Using a web-browser, the user is isolated from the complexities of bringing together computational resources and data access and sees only datasets, code, and results – personal, shared with and by others – as if working on their own personal computer. In a sense, all analysis occurs inside the user’s environments using tools they select.

The following is a short description of ESA Datalabs System Domains:

Portal Services: it provides a single access portal delivering an integrated web visualisation environment for user interfaces.

Desktop Services: it provides a set of tools that can be downloaded and set up in the client desktops or tools to support the portal services.

Discovery Services: it allows searching for assets based on their metadata and custom search language(s) in order to either download them or determine the ID(s) for their utilisation by other services (for example to carry out research on the asset through the Analysis Services). This system domain unifies the catalogue access to Data and Application Services.

Analysis Services: it provides frameworks to manipulate and analyse data integrated in the platform. Domain specific libraries and COTS required for data analysis are provided as part of this system domain, for example data format fits tables and fits maps in Healpix format (cfitsio, Healpix) in various programming languages (Fortran, C, C++, Python and IDL).

Workspace Services: it supports management of user personalised storage (scratch, persistent and public mounted areas) and execution environment. It allows users to publish and share their assets.

Pipeline Services: it offers pipelines for data integration, transformation and analytics based on processing assets developed by actors. The assets comprising the pipeline could be fully deployed in ESA Datalabs across multiple clouds. Pipelines could be integration pipelines, development pipelines and distributed pipelines.

Data Services: it offers access to the data sources. The most important is the Science Archives at ESAC.

Application Services: it provides visibility of available processing assets. This system domain provides key support for the future implementation of a "Science App Store".

Integration Framework: it includes libraries and tools to provide a uniform integration approach for internal communication across system domains, as well as external interfaces consumed and provided by ESA Datalabs to support System of Systems federation capabilities.

Interoperability Services: it provides access to the platform's interface available to external systems and gathers all external dependencies integrated in the platform (example, VO protocols support like TAP to allow access to the Science Archives). It supports federation implementing APIs, RestAPI services.

System Engineering: it provides the tooling supporting software engineering processes. This system domain is based on SOCCI reuse.

Test Framework: it provides the tooling supporting automated unit, integration and system test. This system domain is based on SOCCI reuse.

Storage Infrastructure: it provides distributed, scalable storage disk and backup infrastructure in support of the Application Services and Data Services system domains and to the Support layer.

Computing Infrastructure: it provides computing processing capabilities in support of the Application Services and Data Services system domains and to the Support layer. Platform resources to support redundancy, high availability, capacity, quality of services. It also provides the support to run the system across multiple clouds.

Registry: it provides a repository for System and User Processing assets to be stored, searchable and versioned. This system domain is in charge of managing and implementing the data model and serialisation approach for datalabs. This system domains in support of the Preservation and Support layers.

User Management: it centralises user management, authentication and authorisation. It also provides support for federated identity management.

Logging and Monitoring: it manages notifications regarding the behaviour of selected variables of the system domains as well as reports and insights from all the layers (Splunk, Zabbix, Matomo).

Resource Management: it coordinates access, deployment, instantiation, monitoring and reporting to execute processing assets. This system domain provides support for the execution of interactive and long term running processing tasks. Cloud computing, Virtualisation and Containers are at the core of this system domain to enable high scalability, redundancy and availability. This system domain provides a smooth integration of these technologies to manage information and processing assets. On one side the system domain is in charge of orchestrating the execution of processing assets as virtual machines or containers across local or external cloud computing platforms. On the other side it also orchestrates the storage of data and products consumed by the aforementioned processing assets, hosted locally, in ESA Datalabs Storage Infrastructure, or externally, in federated cloud storage services.

4 Main Lifecycle

ESA Datalabs sits next to astronomy, planetary, heliophysics and navigation science archives hosted at ESA. In its main lifecycle (see Fig. 5), users access ESA Datalabs, browsing a catalogue of available datalabs and archive data volumes. Once these two elements are selected, the instantiation of the Datalab is automatically performed with just one click and the user is forwarded to a fully autonomous working environment.

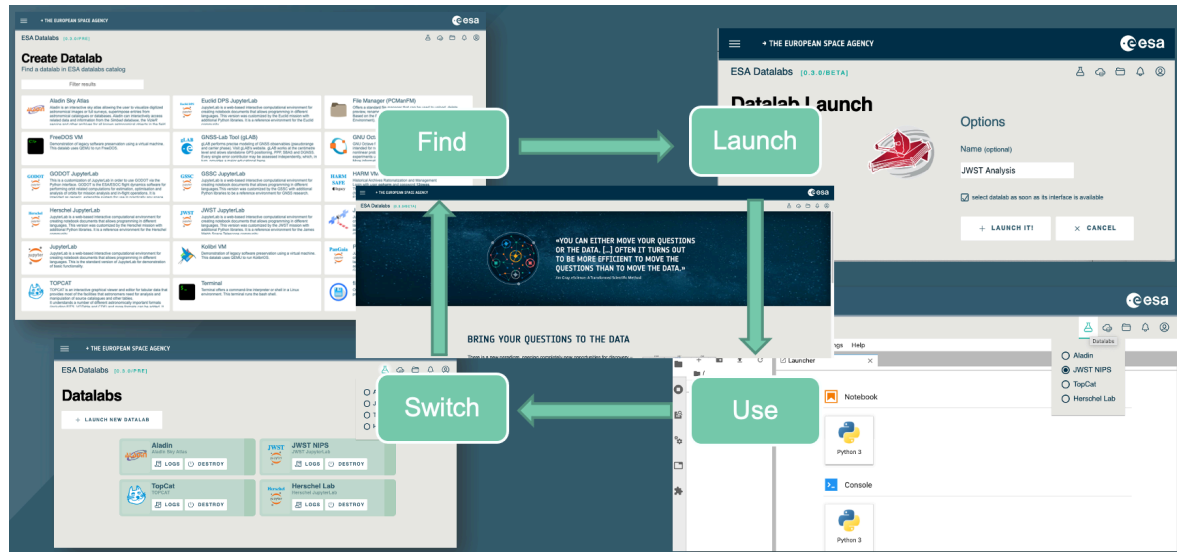


Fig. 5. ESA Datalabs Main Lifecycle

This main ESA Datalabs services implementing this main lifecycle are described in the following sub-sections.

4.1 Web and Desktop Datalabs for Interactive Data Analysis Services

ESA Datalabs provides on-demand, web-access to data analysis components referred as Datalabs. A Datalab is an application or software tool available for execution on ESA Datalabs cloud infrastructure from the user’s web browser. Datalabs range from general domain systems that are becoming de-facto standards for data processing, to domain-specific ones designed to tackle a particular problem. Prominent examples of general domain systems are JupyterLab [16], Octave, Zeppelin or R-Studio. Generalist environments like JupyterLab allow to develop algorithms in notebooks using a variety of programming languages. In addition to vanilla versions of these environments, ESA Datalabs is to become a hub for added-value notebooks further customised by ESA and the scientific community.

Beyond support for native web-based systems, ESA Datalabs ensures long-term preservation of legacy software, enabling re-execution of desktop-based data processing systems that encapsulate key product generation algorithms. ESA Datalabs solution repackages these systems into runnable web-based applications. Users can search, star, comment, bookmark, and run any Datalab in their catalogue. Datalabs listed in the catalogue can be published by developers from ESA missions.

Furthermore, external contributions can be added to the catalogue following a moderation process that includes automatic security scans. Once the user finds a Datalab of interest, a simple play icon allows to launch it immediately. As part of the launch wizard, the user can modify the default configuration of the Datalab set by its creator. At this point, it is possible to select a previous version of the Datalab, increase its computing resources, or adding data volumes from ESAC Archives. Upon completion, the user is automatically redirected to start working on the selected Datalab. Following this first run, the user can go back to the catalogue to run up to five Datalabs in parallel (default profile configuration). Finally, the user can access the dashboard to switch to another Datalab or destroy it.

4.2 Pipelines for Batch Data Processing Workflow Services

In a data-intensive domain, a finding is the outcome of a pipeline of computations, applied to some large existing dataset, or to a dataset created by researchers from multiple other large datasets, combined in unexpected and complex ways. Leveraging on the powerful infrastructure developed for the execution of Datalabs, the Pipeline catalogue provides capabilities for the creation and execution of complex batch data processing workflows based on a set of processing assets.

Hence, Pipelines represent an extension of the Datalab concept, defined as a set of an input area, a sequence of processing stages in between (steps), and an output area. The capabilities to perform data integration, pre-processing, transformation and analytics represent the entry point for Machine Learning Pipelines in ESA Datalabs.

Two major parts can be identified in this area: Pipeline Editor and Pipeline Executor.

The Pipeline Editor provides an integrated visual development environment for users to create a workflow of computing elements. The editor drives the user through the development cycle simplifying the creation process and transforming the graphical representation of the pipeline into its underlying code in CWL[17].

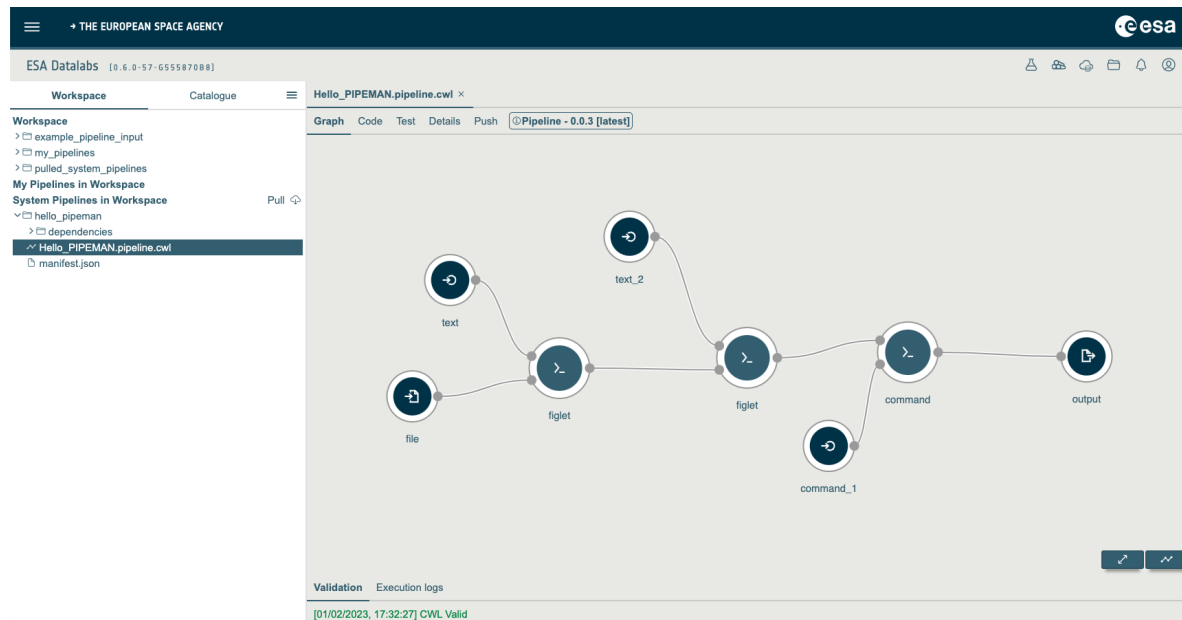


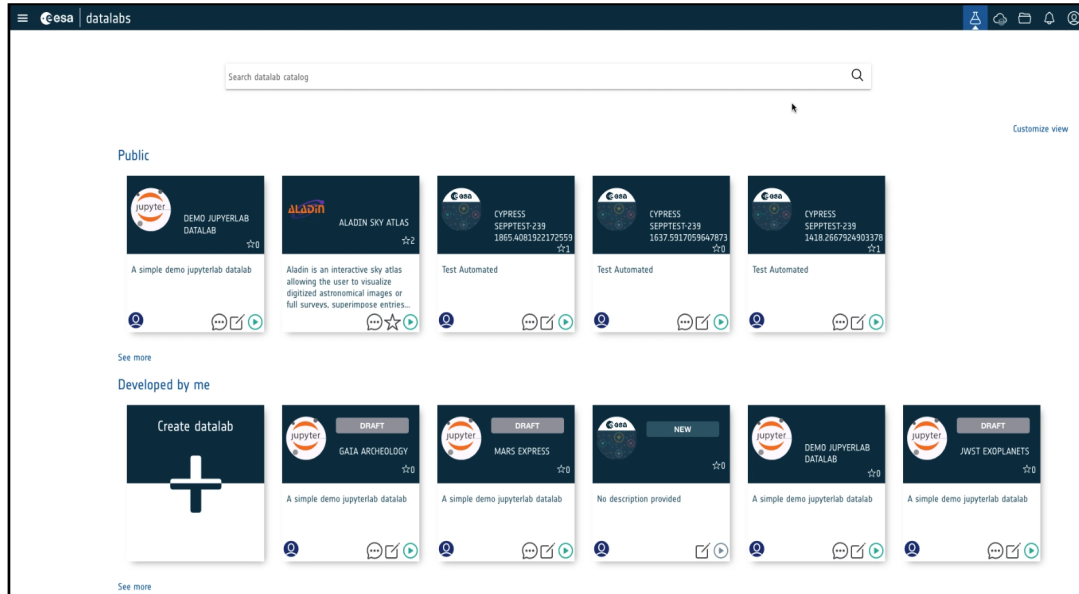
Fig. 6. ESA Datalabs Pipeline Editor

4.3 Collaboration Environment, Open Science and Datalabs Store

Leveraging on the workspace capability, ESA Datalabs users can generate Collaborative Environments to share the deployment and execution of complex software applications on large amounts of data. Hence, groups do not need to download multiple copies of the data, but bring elements into their workspace, which always has the environment setup with latest version.

With the Datalab Application Services, users can publish their Datalabs close to the scientific data, allowing the entire scientific community to discover new science applications and products in an open-science fashion. ESA Datalabs users with developer permission can create, build, test, validate, and export their own Datalabs. They can either keep them private, share them with selected individuals or groups, or make them accessible to the wider

community of ESA Datalab users. In order to have a Datalab featured in the main catalogue, it must pass through a moderation process transparent to the user. Once published, users are able to search for Datalabs of interest, rate them, leave feedback, bookmark and launch them. In a nutshell, Datalabs Application Services provides scientific developers with a friendly, fast and guided process to create a Datalabs Store.



5 Conclusions and future work

ESA Datalabs has joined the growing number of science exploitation platforms successfully completing its core development phase. At present, it is available at datalabs.esa.int [18] as a Private Beta.

In this work we have presented ESA Datalabs main features, describing how ESA Datalabs catalogues permeate through the Data Archives at ESAC, bringing new collaboration features derived from the possibility to share co-located computing elements and storage areas. Furthermore, ESA Datalabs capabilities to integrate with other thematic exploitation and discovery services enable reuse of its key processing features. This federated approach contributes to achieve strategic Science Operations objectives by reducing time-to-market and risks in the development of new systems.

ESA Datalabs architecture supports a set of capabilities to achieve following strategic innovation goals:

- **Data Access Innovation:** through a paradigm shift solution characterized by moving processing components to data, rather than moving data to users.
- **Data Processing Innovation:** through on-demand provision of innovative Data Analysis tools.
- **Open Collaboration Innovation:** through seamless access to a shared, user-extensible environment.

Implementation of Datalabs for multiple ESA missions such as Juice or JWST, has demonstrated its suitability to support cross-domain use cases. Researchers have fast access to pre-configured complex applications integrating ESA archives. ESA Datalabs next steps gear towards integration of parallel activities in the areas of GPU computing, machine learning pipelines and an integrated development environment for Datalabs.

The scientific and citizen-centric applications unleashed by these platforms are endless. ESA Archives users can already benefit from a set of JupyterLab based Datalabs for multiple missions (Gaia, Juice, JWST, XMM, Herschel). More advanced collaboration levels derive from team data volumes, designed for groups of users sharing files and analysis tools.

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.

References

- [1] V. Navarro, R. Alvarez, Sara del Río, S., Guerra, R., Martín-Rubio, V., Diego, M. Á., Ramos, N. A. V., Marques, A. L., Henriques, X. R., Pereira, A., et al. 2020, ESA Datalabs an e-Science Platform for Data Exploitation and Preservation at ESAC, Astronomical Data Analysis Software and Systems XXX
- [2] V. Navarro, Sara del Rio, E. Fraile, L. Mendes, J. Ventura-Traveset, GSSC Now: ESA GNSS Thematic Exploitation Platform for Data Science Innovation. 8th International Colloquium on Scientific and Fundamental Aspects of GNSS, Sofia, Bulgaria September 2022
- [3] V. Navarro et al., ESA Data Exploitation Platforms: Accelerating Space and Navigation Science; <https://doi.org/10.5281/zenodo.6570895>
- [4] Navarro, V., Dias, L., Constantino, A., Gabriel, C., Ibarra, A., Merin, B., ... & Arviset, C. (2016). One-click data analysis software for science operations. In 14th International Conference on Space Operations (p. 2399).
- [5] ESAC Science Data Centre, <https://www.cosmos.esa.int/web/esdc/home>, (accessed 31.01.23).
- [6] GNSS Science Support Centre, <https://gssc.esa.int/>, (accessed 31.01.23).
- [7] Taghizadeh-Popp, Manuchehr, et al. "SciServer: A science platform for astronomy and beyond." *Astronomy and Computing* 33 (2020): 100412.
- [8] Taffoni, Giuliano, et al. "Feasibility of access EGI resources through the ESCAPE developed ESFRI Science Analysis Platform." arXiv preprint arXiv:2301.09978 (2023).
- [9] Piparo, Danilo, et al. "SWAN: A service for interactive analysis in the cloud." *Future Generation Computer Systems* 78 (2018): 1071-1078.
- [10] Jurić, M., et al. "Lsst science platform vision document." LSE-319, LSST (2017).
- [11] Jutz, S., & Milagro-Pérez, M. P. (2020). Copernicus: the European Earth Observation programme. *Revista de Teledetección*, (56), V-XI.
- [12] Koubarakis, M., Bereta, K., Bilidas, D., Giannousis, K., Ioannidis, T., Pantazi, D. A., ... & Fleming, A. (2019). From copernicus big data to extreme earth analytics. *Open Proceedings*, 690-693.
- [13] Maier, M. W. (1998). Architecting principles for systems-of-systems. *Systems Engineering: The Journal of the International Council on Systems Engineering*, 1(4), 267-284.
- [14] V. Navarro, R. Alvarez, F. Pérez-López, C. Arviset, J. Ventura-Traveset, A. M. Furones. ESAC Science Exploitation and Preservation Platform Reference Architecture, 2019, *Astronomical Data Analysis Software and Systems XXVII*, 523, 285
- [15] Shan, Tony C., and Winnie W. Hua. "Solution architecture for n-tier applications." 2006 IEEE International Conference on Services Computing (SCC'06). IEEE, 2006.
- [16] Juneau, Stephanie, et al. "Jupyter-enabled astrophysical analysis using data-proximate computing platforms." *Computing in Science & Engineering* 23.2 (2021): 15-25.
- [17] Amstutz, Peter, Andeer, Robin, Chapman, Brad, Chilton, John, Crusoe, Michael R., Valls Guimerà, Roman, Carrasco Hernandez, Guillermo, Ivkovic, Sinisa, Kartashov, Andrey, Kern, John, Lee, Dan, Ménager, Hervé, Mikheev, Maxim, Pierce, Tim, Randall, Josh, Soiland-Reyes, Stian, Stojanovic, Luka, Tijanić, Nebojša (2016): Common Workflow Language, draft 3. figshare. <https://dx.doi.org/10.6084/m9.figshare.3115156.v1>
- [18] ESA Datalabs, <https://datalabs.esa.int/>, (accessed 31.01.23).