https://eurosciencegateway.eu

This project was funded by the European Union's HORIZON-INFRA-2021- EOSC-01 under Grant Agreement number 101057388.

D 5.2 Publication of the usage of EuroScienceGateway by multiple communities

Work Package 5

Technical References

Project Acronym	*ESG*
Project Title	*EuroScienceGateway*
Project Coordinator	*Albert-Ludwigs-University Freiburg*
Project Duration	*September 2022 - August 2025*

Document	*D 5.2 Publication of the usage of EuroScienceGateway by multiple communities*
Work Package	*Work Package 5*
Task	*T5.1, T5.2, T5.3, and T5.4*
*Dissemination Level**	*PU*
Lead Beneficiary	*UiO*
Contributing Beneficiaries	*UiO, ALU-FR, CNRS, UNIFI, UKRI, EPFL, UP, BSC*

Funded by the European Union

Due Date of Deliverable	*31st August 2025*
Actual Submission Date	*20.08.2025*

* *PU = Public*

PP = Restricted to other programme participants (incl. the Commission Services)

RE = Restricted to a group specified by the consortium (incl. the Commission Services)

CO = Confidential, only for members of the consortium (incl. the Commission Services)

Version	*Date*	*Beneficiaries*	*Author*	*Approved*
1.0	*20-08-2025*	*UiO, ALU-FR, CNRS, UNIFI, UKRI, EPFL, UP, BSC*	*Armin Dadras (ALU-FR), Denys Savchenko (UP), Andrii Neronov (UP, EPFL), Volodymyr Savchenko (EPFL), Nikolay Vazov (UiO), Jean Iaquinta (UiO)*	*Eva Alloza (BSC), María Chavero-Díez (BSC), Anthony Bretaudeau (CNRS), Denys Savchenko (UP), Andrii Neronov (UP, EPFL), Volodymyr Savchenko (EPFL), Nikolay Vazov (UiO), Jean Iaquinta (UiO)*

Funded by the European Union

Acknowledgements

This project was funded by the European Union's HORIZON-INFRA-2021- EOSC-01 under Grant Agreement number 101057388.

Disclaimer

This work may rely on data from sources external to the members of the ESG project Consortium. Members of the Consortium do not accept liability for loss or damage suffered by any third party as a result of errors or inaccuracies in such data. The information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and neither the European Community nor any member of the ESG Consortium is liable for any use that may be made of the information.

Executive Summary

Scientific discovery is evolving rapidly, driven by increasingly complex data, interdisciplinary collaboration, and a growing demand for open, reproducible research. The EuroScienceGateway (ESG) project responded to this transformation by extending the Galaxy platform across diverse scientific domains and by lowering the barriers for new communities to engage with high-performance, FAIR-compliant computational workflows. From biodiversity and climate science to astrophysics, materials research, and biomedicine, ESG offered tools, training, and infrastructure tailored to real-world scientific challenges.

But bridging these communities wasn't simply a matter of software development. Each domain came with its own legacy systems, specialized data formats, and unique research practices. ESG had to meet scientists where they were and translate Galaxy's capabilities into workflows and environments that felt familiar, usable, and trustworthy. This required more than technical integration: it required dialogue, supervision, documentation, and the building of sustainable support structures that could evolve alongside each community's needs.

Therefore, ESG implemented a comprehensive approach that combined infrastructure development with community engagement. Over 800 new tools were integrated across Galaxy subdomains, while reusable workflows were published openly on WorkflowHub. Initiatives like the Galaxy Community onboarding framework and the Galaxy Codex ensured long-term sustainability and knowledge transfer. Demonstrators in sensitive data handling, image analysis, and federated workflows highlighted Galaxy's readiness for the next generation of open science challenges. Through active collaboration with EOSC, ELIXIR, NFDI, and other pan-European infrastructures, ESG helped position Galaxy as a central, trusted resource for scientific research across domains.

This deliverable (D5.2) captures the stories, strategies, and successes of ESG's cross-community engagement. It is both a reflection of what has been achieved and a guidepost for what comes next. By onboarding diverse scientific fields to Galaxy, ESG has built more than infrastructure. It has cultivated a culture of openness, reproducibility, and shared innovation that will continue to grow long after the project ends.

Funded by the European Union

List of Abbreviations

ATel: Astronomer's Telegram
CTAO: Cherenkov Telescope Array Observatory
EBP: Earth BioGenome Project
EGA: European Genome-phenome Archive
EPAC: Extreme Photonics Applications Centre
ERDDAP: Environmental Research Division Data Access Program
ESG: EuroScienceGateway
ESME: Earth System Modeling Environment
FAIR: Findable, Accessible, Interoperable, and Reusable
FITS: Flexible Image Transport System
GBIF: Global Biodiversity Information Facility
GIS: Geographical Information System
HEASoft: High Energy Astrophysics Software
HPC: High-Performance Computer
IVOA: International Virtual Observatory Alliance
JupyterGIS: Jupyter extension adding GIS visualization/analysis capabilities
MeerKAT: “meer” Karoo Array Telescope
MPI: Message Passing Interface
MSCP: Muon Spectroscopy Computational Project
MWA: Murchison Widefield Array
OBIS: Ocean Biodiversity Information System
QGIS: Quantum Geographic Information System
SKAO: Square Kilometre Array Observatory
UHI: Urban Heat Island
VGP: Vertebrate Genomes Project
XAS: X-ray Absorption Spectroscopy

Funded by
the European Union

Table of contents

Funded by
the European Union

1. Introduction

Scientific discovery is becoming increasingly data-driven, collaborative, and interdisciplinary. The sheer scale and complexity of modern datasets from biodiversity metrics and climate models to genomic sequences, astronomical observations, and biomedical images demand flexible, powerful, and reproducible computational platforms. And across Europe, a thriving ecosystem of research infrastructures and open science initiatives is emerging to meet this need. Among them, the Galaxy platform stands out as a mature and adaptable environment that allows researchers to build, share, and execute computational workflows with ease, while adhering to the principles of FAIR (Findable, Accessible, Interoperable, and Reusable) research.

But while Galaxy's potential has long been recognized in fields like genomics, its broader adoption across diverse scientific communities has often been delayed by fragmented infrastructures, domain-specific barriers, and a lack of accessible tools or training tailored to new user groups. Different communities possess different standards, data formats, and analytical practices. There was a gap, a missing bridge, between Galaxy's capabilities and the realities of how scientists in climate research, materials science, astrophysics, and biomedical informatics actually work. These were communities ready to benefit from FAIR workflows and reproducible science, but they lacked the means, guidance, and technical scaffolding to get there.

Therefore, the ESG project was launched to close that gap, bringing Galaxy to new communities, and new communities to Galaxy. Supported by the European Union under Horizon Europe, ESG united a network of partners to build workflows, integrate infrastructure, offer tailored onboarding, and establish lasting collaborative frameworks. ESG did not just expand Galaxy's technical treasure; it brought it into the daily practices of scientific fields undergoing digital transformation. Through continuous improvement, targeted support, and a commitment to usability and openness, ESG enabled climate scientists to forecast Arctic sea ice[1], helped biodiversity researchers annotate whole genomes[2], empowered materials scientists to simulate muon behavior[3], and allowed astrophysicists to run reproducible analyses of multi-wavelength observatory data all within Galaxy[4].

Deliverable D5.2 tells the story of that transformation. It traces the journey of how different research domains, each with its own language, culture, and computational needs, came together under a shared vision of open, reproducible science. It begins with the biodiversity and climate science communities, whose need for integrated, large-scale data analysis and compatibility with global initiatives like the Earth BioGenome Project

[1] https://galaxyproject.org/news/2024-04-28-egu2024/
[2] https://galaxyproject.org/news/2025-03-18-assembly-galaxy/
[3] https://galaxyproject.org/news/2025-03-24-muon-spectroscopy-paper/
[4] https://galaxyproject.org/news/2024-02-29-esg-initial-workflows-workflowhub/

Funded by the European Union

(EBP)[5] drove early innovation in tool development and workflow design. The narrative continues with materials science, where pioneering efforts in muon spectroscopy laid the groundwork for broader applications in experimental simulation and catalysis research. Astrophysics followed, bringing challenges of scale, data format complexity, and cross-institutional collaboration, ultimately showcasing Galaxy's power to support multi-messenger astronomy.

But the story does not end with specific disciplines. One of ESG's most impactful contributions lies in its approach to onboarding: a sustained effort to lower the threshold for participation in Galaxy-based research. Through workshops, training materials, hackathons, and the creation of the Galaxy Community onboarding cookbook (1), ESG developed a repeatable, scalable framework for bringing new communities on board. From cancer data analysis with EOSC4Cancer[6] to secure access to encrypted datasets in collaboration with European Genome-phenome Archive (EGA)[7] ESG demonstrated that onboarding is not just a technical task, it is a cultural process of co-creation, mentorship, and shared problem-solving.

Across all these efforts, ESG remained rooted in one essential belief: that reproducible, FAIR science should not be a privilege of those with deep computational expertise, but a right accessible to all researchers. And so, the project made careful choices to support under-resourced domains, to meet users where they are, and to build bridges, between disciplines, between infrastructures, and between scientists.

This deliverable synthesizes the outcomes of four major tasks of Work Package 5 (Task 5.1 through Task 5.4) and captures the real-world impact of ESG on scientific practice. It reflects the active efforts of partner institutions across Europe ranging from universities and national research centers to domain-specific consortia and the scientists who contributed their time, feedback, and ideas to shape Galaxy for their communities. The pages that follow provide a window into this collaborative experiment in making science more open and impactful.

ESG is not just a story of software development. It is a story of empowerment; of researchers gaining the tools and confidence to take ownership of their data, workflows, and discoveries. And it is a story that continues, as the resources, communities, and infrastructure built under ESG will persist and evolve long after the project concludes.

2. Biodiversity and Climate Science

The ESG project began its engagement with the biodiversity and climate science communities by addressing a shared challenge: the need to process and analyze

[5] https://www.earthbiogenome.org/
[6] https://eosc4cancer.eu/
[7] https://ega-archive.org/

Funded by the European Union

increasingly large and complex datasets from diverse sources. By leveraging the Galaxy platform, ESG provided these communities with unified, scalable access to critical data repositories (such as the Copernicus Data Space Ecosystem[8], Global Biodiversity Information Facility (GBIF)[9], Ocean Biodiversity Information System (OBIS)[10], Environmental Research Division Data Access Program (ERDDAP)[11], and Quantum Geographic Information System (QGIS) services[12]) laying the groundwork for cutting-edge research. This enabled researchers to explore complex relationships between biodiversity trends or climate variables in ways not previously possible within a single computational environment.

The project placed special emphasis on genome annotation, a cornerstone of biodiversity research. To support large-scale global initiatives like the EBP and Vertebrate Genomes Project (VGP)[13], ESG integrated a set of cutting-edge annotation tools into Galaxy. These included BRAKER3[14], Helixer[15], BUSCO[16], DeepSig[17], Miniprot[18], and Compleasm[19]. As an example, Helixer, which is an AI-powered GPU-enabled annotation tool, demonstrated the scalability of the infrastructure when deployed across the Pulsar network[20] (2), which provides on-demand access to GPU resources. This integration marked a milestone in making genome annotation more accessible and reproducible.

Beyond genome annotation, ESG supported the development of high-quality, reusable workflows that could be adopted across institutions. Workflows for tasks such as repeat masking (3), functional annotation (4,5), and lncRNA analysis[21] were shared on the Galaxy platform, all designed with reproducibility and FAIR principles in mind. On the Climate Science side, ESG enabled forecasting capabilities by incorporating workflows like IceNet for predicting Arctic sea ice[22] and FArLiG[23] for studying lichen browning, demonstrating the platform's value for both biodiversity monitoring and climate forecasting. UHI-Stream[24] was also made for people without technical background to

[8] https://dataspace.copernicus.eu/
[9] https://www.gbif.org/
[10] https://obis.org/
[11] https://oceanobservatories.org/erddap-server/
[12] https://qgis.org/
[13] https://vertebrategenomesproject.org/
[14] https://github.com/Gaius-Augustus/BRAKER
[15] https://github.com/weberlab-hhu/Helixer
[16] https://busco.ezlab.org/
[17] https://github.com/BolognaBiocomp/deepsig
[18] https://github.com/lh3/miniprot
[19] https://github.com/huangnengCSU/compleasm
[20] https://pulsar-network.readthedocs.io/en/latest/
[21] https://workflowhub.eu/workflows/1324
[22] https://galaxyproject.org/news/2024-04-28-egu2024/
[23] https://meetingorganizer.copernicus.org/EGU23/EGU23-2579.html
[24] https://indico.egi.eu/event/6441/contributions/19226/

Funded by the European Union

look at Urban Heat Island (UHI) effects and compare temperature changes between two points anywhere on Earth.

Supporting climate model execution across diverse system architectures became another priority as ESG matured. In response to community needs, particularly from NorESM users, ESG developed multiple versions of the Earth System Modeling Environment (ESME) for different Message Passing Interface (MPI) implementations (6). These were published on Bioconda[25] and other channels, allowing containerized models to run efficiently across high-performance computing platforms such as Betzy[26], Fram[27], and Lumi[28]. Complementing this, an interactive geospatial tool (JupyterGIS) was deployed on Galaxy Europe[29], with a version also being adapted for Norway's NIRD platform and OpenOnDemand for deployment on HPC to enhance user interactivity.

As the community expanded, so too did the platform's tool ecosystem. Over 250 tools were added to specialized Galaxy subdomains for biodiversity, earth systems, and climate science. ESG partners led a wide-reaching training and dissemination effort, producing targeted materials hosted on the Galaxy Training Network and organizing dedicated workshops. Community engagement was backed through ESG's active presence at major conferences and initiatives, including ELIXIR Biodiversity, EBP, ERGA, and JOBIM, ensuring that the tools and workflows reached the researchers who needed them most.

To organize and sustain this growing body of work, ESG launched the Galaxy Codex[30] initiative in 2025, a strategic community effort to curate Galaxy's scientific resources for specific domains. A biodiversity-specific Codex was initiated during a workshop at the ELIXIR All Hands Meeting 2025, aiming to serve as a centralized hub for community-curated tools, workflows, and training resources[31]. This will soon become part of a dedicated Biodiversity Galaxy Lab[32], providing structured access to everything from data preprocessing to full-scale analyses, tailored for biodiversity researchers.

Through its efforts, ESG did more than just provide tools; it actively endorsed the adoption of FAIR practices and built a culture of reproducibility within these scientific communities. By aligning technical development with domain-specific research needs, ESG helped lower barriers to entry for advanced analysis, supported cross-disciplinary

[25] https://anaconda.org/bioconda/esme_mpich_4_2_3
[26] https://documentation.sigma2.no/hpc_machines/betzy.html
[27] https://documentation.sigma2.no/hpc_machines/fram.html
[28] https://documentation.sigma2.no/hpc_machines/lumi.html
[29] https://usegalaxy.eu/root?tool_id=interactive_tool_jupytergis_notebook
[30] https://github.com/galaxyproject/galaxy_codex
[31] https://www.france-bioinformatique.fr/en/news/recap-of-the-elixir-all-hands-meeting-2025-in-thessaloniki/
[32] https://galaxyproject.org/news/2025-06-02-elixir-all-hands-meeting/

Funded by the European Union

collaboration, and enabled data reuse on a scale not previously possible in biodiversity and climate research.

As the project concludes, the biodiversity and climate science communities are better equipped than ever to address environmental challenges. With a mature ecosystem of tools, workflows, and training materials built on a sustainable infrastructure, Galaxy has become a trusted platform for these fields. ESG's legacy is not only in the software delivered, but in the communities that have formed. Researchers now have the means to collaborate effectively, share their work openly, and drive forward data-driven environmental science across Europe.

3. Materials Science

The materials science community's journey within the ESG project began with a strong focus on muon spectroscopy, specially through the Muon Spectroscopy Computational Project (MSCP)[33]. This effort led to the creation of MuonGalaxy[34], a dedicated Galaxy portal integrating specialized simulation tools like PyMuonSuite[35] and MuSpinSim[36]. Through close collaboration and iterative development, the team developed 11 Galaxy tools aimed at making muon-based computational modeling more accessible, reproducible, and FAIR-compliant. These tools were not only deployed on a UK-based instance tailored for the muon science community but were also integrated into the European Galaxy instance to enable broader collaboration and knowledge sharing.

As the platform matured, the team extended Galaxy's capabilities beyond muon spectroscopy to other materials science applications. At the PASC23 conference, they demonstrated how Galaxy could serve as a robust and flexible workflow environment for this broader domain[37]. Another important milestone was the development of prototype tools for analyzing X-ray Absorption Spectroscopy (XAS) data from catalysis experiments[38]. The early versions of these tools were presented at the UK Catalysis Hub Summer Conference 2023[39], promising an expansion into catalysis-focused materials science.

Training and community engagement were important parts of this community development. In 2024, the team led a session on the PyMuonSuite tools at the ISIS muon training school[40], illustrating how Galaxy workflows could be applied to identify muon

[33] https://muon-spectroscopy-computational-project.github.io/
[34] https://materialsgalaxy.stfc.ac.uk/
[35] https://github.com/muon-spectroscopy-computational-project/pymuon-suite
[36] https://muon-spectroscopy-computational-project.github.io/muspinsim/
[37] https://pasc23.pasc-conference.org/presentation/?id=msa156&sess=sess169
[38] https://galaxyproject.org/news/2025-03-25-catalysis-reproduction-paper/
[39] https://ukcatalysishub.co.uk/event/watch-now-uk-catalysis-hub-summer-conference-2023
[40] https://www.isis.stfc.ac.uk/Pages/March-Muon-training-school.aspx

Funded by the European Union

stopping sites which is an important step in spectroscopy studies[41]. Besides, a new tool called mudirac was introduced, enabling simulations of negative muon experiments, a rare but valuable capability for the community[42]. These developments helped demonstrate the practical impact of Galaxy-based solutions on experimental interpretation and training.

In parallel, the team developed a suite of seven Galaxy tools for XAS, building on the Larch Python library to support advanced analysis of experimental data[43]. To validate and promote these tools, they reproduced results from nine published catalysis studies using Galaxy workflows, all made publicly available via WorkflowHub (7)[44]. By capturing full data provenance using RO-Crate and leveraging Galaxy's built-in reproducibility features, they made a compelling case for how Galaxy can reduce the burden on researchers while ensuring transparent, high-quality outputs. These achievements were highlighted at the 2024 Galaxy Community Conference[45] and NoBugs conferences[46], further reinforcing the platform's value in materials science.

The community has since advanced to exploring next-generation applications. In collaboration with the Extreme Photonics Applications Centre (EPAC)[47], they are developing an end-to-end simulation system for modeling experimental setups using digital twin approaches. This allows users to replicate instrument behavior virtually, test different configurations, and optimize experimental conditions before actual runs. Such simulation workflows can improve efficiency and planning, offering a new level of control and insight in experimental materials science.

With this foundation in place, the materials science community within ESG has grown from a focused pilot in muon spectroscopy to a dynamic and forward-looking ecosystem embracing diverse experimental domains. Their work has not only enriched the Galaxy toolshed[48] but also showcased the platform's adaptability and scientific

[41] https://training.galaxyproject.org/training-material/topics/materials-science/tutorials/muon-stopping-sites-muairss-uep/tutorial.html
[42] https://toolshed.g2.bx.psu.edu/repository/view_repository?sort=name&operation=view_or_manage_repository&id=487662fd062ee564
[43] https://xraypy.github.io/xraylarch/
[44] https://workflowhub.eu/workflows/1798, https://workflowhub.eu/workflows/1799, https://workflowhub.eu/workflows/1800, https://workflowhub.eu/workflows/1801, https://workflowhub.eu/workflows/1802, https://workflowhub.eu/workflows/1803, https://workflowhub.eu/workflows/1804, and https://workflowhub.eu/workflows/1806, https://workflowhub.eu/workflows/1805
[45] https://galaxyproject.org/news/2024-07-19-gc-c2024-meeting-report/
[46] https://indico.esrf.fr/event/114/contributions/775/
[47] https://www.clf.stfc.ac.uk/Pages/EPAC.aspx
[48] https://toolshed.g2.bx.psu.edu/repository/browse_repositories_in_category?id=ab9d40c7264aa6

Funded by the European Union

relevance. By building tools, training users, and linking data to publications through reproducible workflows, the community has positioned Galaxy as a sustainable, long-term resource for materials science research. Their latest contributions were featured at the 2025 Galaxy and Bioconductor Conference[49], signaling both maturity and momentum for continued innovation.

4. Astrophysics

The ESG project's engagement with the astrophysics community began with a mission to integrate commonly used data formats, analytical tools, and FAIR principles into the Galaxy platform. This task aimed to serve the needs of astroparticle physicists and cosmologists, providing them with scalable computing and access to large astronomical datasets. Early collaboration with Work Package 2 ensured that Galaxy would support both domain-specific research needs and broader goal of reproducibility.

Initial technical developments focused on enabling Galaxy to handle key astrophysical data formats and tools. A dedicated Galaxy instance for astronomy was established[50], and core support for the Flexible Image Transport System (FITS) format, which is important in the astronomy field, was developed[51]. To enhance user interaction, a visualization plugin based on AladinLite[52] was developed, allowing researchers to view sky images and overlay catalog data directly in Galaxy. The complementary plugin[53] is developed to facilitate plotting of the table data from FITS files. Access to large astronomical archives was also simplified through tools interfacing with International Virtual Observatory Alliance (IVOA)-compliant data services[54], including the one specially developed[55] to expose the pre-processed data of the European Special Agency's (ESA) INTEGRAL[56] mission

Further tool development was supported by the Astroteam, a dedicated group within the Galaxy ToolShed[57]. They created new Galaxy tools using the Astropy Python library[58] to

44&message=&status=done and https://toolshed.g2.bx.psu.edu/repository/browse_repositories_in_category?id=9dca5b8f68c509af&message=&status=done

49 https://gbcc2025.bioconductor.org/program/scientific_program/

50 https://astronomy.usegalaxy.eu/

51 https://docs.galaxyproject.org/en/latest/lib/galaxy.datatypes.html#galaxy.datatypes.binary.FITS

52 https://aladin.cds.unistra.fr/#AladinLite

53 https://github.com/galaxyproject/galaxy-visualizations/tree/main/packages/fitsgraph

54 https://galaxyproject.org/news/2023-09-07-esg-wp5-astronomy-archives/

55 https://galaxyproject.org/news/2025-08-15-mmoda-tap-server/

56 https://www.esa.int/Science_Exploration/Space_Science/Integral

57 https://toolshed.g2.bx.psu.edu/repository?repository_id=a886a9af224fe78b&changeset_revision=f40d05521dca

58 https://www.astropy.org/

Funded by the European Union

manipulate and convert FITS files, making them compatible with existing Galaxy workflows. Legacy tools from projects like MMODA[59] from RenkuLab[60] were also integrated, and a prototype converter enabled Galaxy workflows to process data from a range of observatories, including LIGO[61]/Virgo[62], Gaia[63], and Fermi/LAT[64], thus expanding the platform's reach across different types of astronomical research.

With the technical groundwork laid, the team began active outreach to engage the broader astrophysical community. A highlight was the presentation at Swiss SKA Days 2023[65], which demonstrated the new tools and gathered feedback. This marked the transition from primary development to community collaboration, signaling Galaxy's growing credibility in the field.

During the second half of the project, the focus shifted toward scaling up tool development and improving sustainability. Responding to the community's existing practices, a tool for converting Jupyter notebooks into Galaxy workflows was developed and documented[66]. This helped bridge the gap between astronomers' preferred environments and the Galaxy, easing adoption and encouraging contributions.

More than 30 astronomy-specific tools were developed or co-developed with the community, supporting analyses across different telescope types such as Murchison Widefield Array (MWA)[67] and MeerKAT ("meer" Karoo Array Telescope)[68]. Another improvement was the integration of NASA's HEASoft (High Energy Astrophysics Software) tool suite[69], bringing hundreds of well-established workflows into Galaxy. These tools were made available through the astronomy staging instance and selected ones were promoted to the EU-wide Galaxy server, creating a dedicated entry point for astronomers: astronomy.usegalaxy.eu.

Several complete workflows were published on WorkflowHub, demonstrating end-to-end research processes. For example, one workflow traced the path from rapid micropublications (e.g., Astronomer's Telegram (ATels)) to peer-reviewed articles which is critical for transient astronomy, where timely action is key[70]. This workflow featured tools

59 https://www.astro.unige.ch/astroordas/mmoda
60 https://renkulab.io/
61 https://www.ligo.caltech.edu/
62 https://www.virgo-gw.eu/
63 https://www.esa.int/Science_Exploration/Space_Science/Gaia
64 https://ecap.nat.fau.de/indexphp/research/gamma-ray-astronomy/fermi-lat/
65 https://galaxyproject.org/events/2023-09-06/
66 https://galaxyproject.org/news/2025-06-16-jupyter-to-tool/
67 https://www.mwatelescope.org/
68 https://www.mpifr-bonn.mpg.de/pressreleases/2020/9
69 https://heasarc.gsfc.nasa.gov/docs/software/heasoft/
70 https://workflowhub.eu/workflows/1353

Funded by
the European Union

like AstroCOLIBRI[71] and was presented at a dedicated workshop before being shared publicly on WorkflowHub. The Voronoi segmentation for astronomy workflow[72] applies Voronoi segmentation techniques to LegacySurvey[73] imaging data, allowing galaxy detection and paving the way to conducting detailed morphological analyses of galaxies[74]. Another notable example is the workflow[75] that calculates the power spectrum of Stochastic Gravitational Wave Backgorund (SGWB) from a first-order cosmological phase transition, reproducing the results of a paper[76], it allows varying parameters of the calculation. This demonstrates the power of reproducible workflows as an addition to the classical journal publications.

As the community matured, Task 5.3 catalyzed spin-off projects and long-term collaborations. Galaxy became a key component in the OSCARS FAIR imaging initiative[77], Swiss AstroORDAS[78], and the ORD Solidipes project[79]. It was also formally integrated into ACME, a European collaboration for multimessenger astrophysics[80]. Finally, Galaxy was positioned as a potential analysis platform for next-generation observatories like Square Kilometre Array Observatory (SKAO)[81] and Cherenkov Telescope Array Observatory (CTAO)[82] highlighting its role in the future of astronomical research.

5. Other new communities and collaborations

The ESG project began its onboarding efforts to organize tailored outreach events across Europe. These initiatives focused on building connections with both local and emerging scientific communities, while aligning with broader European research infrastructures such as EOSC-Life[83], EOSC-Nordic[84], EOSC-Pillar[85], Aqua-Infra[86], EOSC Data

[71] https://astro-colibri.science/
[72] https://doi.org/10.48546/WORKFLOWHUB.WORKFLOW.1730.1
[73] https://legacysurvey.org
[74] https://galaxyproject.org/news/2025-06-11-voronoi-astronomy/
[75] https://doi.org/10.48546/WORKFLOWHUB.WORKFLOW.831.1
[76] https://doi.org/10.48550/arXiv.2405.07746
[77] https://oscars-project.eu/projects/fair-image-analysis-across-sciences
[78] https://www.astro.unige.ch/astroordas/
[79] https://www.epfl.ch/schools/enac/solidipes-curation-platform-2/
[80] https://www.mpifr-bonn.mpg.de/announcements/2024/6
[81] https://www.skao.int/en
[82] https://www.ctao.org/
[83] https://www.eosc-life.eu/
[84] https://eosc-nordic.eu/
[85] https://www.eosc-pillar.eu/
[86] https://aquainfra.eu/

Funded by the European Union

Commons[87], FAIR-EASE[88], OSCARS[89], BY-COVID[90], EOSC4Cancer[91], FAIR2ADAPT[92], Skills4EOSC[93], RELIANCE[94], ESCAPE[95], BioNT[96], HealthyCloud[97], European Genomic Data Infrastructure[98], AgroServ[99], and ELIXIR[100]. The objective was twofold: to introduce these communities to the Galaxy platform and to support the integration of their workflows and infrastructures into the ESG ecosystem. Additionally, the project recognized the growing need for interdisciplinary data analysis, aiming to create bridges across domains and encourage collaborative data analysis using shared Galaxy resources.

As part of this onboarding strategy was the documentation of real-world experiences from community members who were integrating Galaxy into their work[101]. By capturing frequent challenges, solutions, and best practices, ESG partners began forming a practical guide called the Galaxy Community onboarding cookbook (Deliverable 5.1 (1)). This living resource helped lower the barrier for new communities, offering structured support for getting started with Galaxy. In parallel, these efforts informed the development of a “maturity model”, which is a framework to guide new communities through different stages of engagement, from initial onboarding to long-term participation and sustainability within the Galaxy ecosystem (8).

New user groups received help with workflow adaptation, infrastructure setup, and Galaxy feature adoption. All onboarded communities were invited to join different Galaxy working groups, where they could receive mentorship, supervision, connect with peers, and request support. This approach provided a welcoming environment for new users and encouraged knowledge exchange across different research domains.

During the second half of the project (2023–2025), ESG’s onboarding strategy advanced further through collaborations, especially with the EOSC4Cancer project. This partnership led to showcasing Galaxy’s ability to manage and analyze sensitive clinical and biomedical data. A key development involved the integration of Galaxy with cBioPortal[102], which is a widely used platform in cancer genomics. A prototype was presented during

87 https://www.eosc-data-commons.eu/
88 https://fairease.eu/
89 https://oscars-project.eu/
90 https://by-covid.org/
91 https://eosc4cancer.eu/
92 https://fair2adapt-eosc.eu/
93 https://www.skills4eosc.eu/
94 https://www.reliance-project.eu/
95 https://projectescape.eu/
96 https://biont-training.eu/
97 https://healthycloud.eu/
98 https://gdi.onemilliongenomes.eu/
99 https://agroserv.eu/
100 https://elixir-europe.org/
101 https://galaxyproject.org/community/#a-few-success-stories
102 https://www.cbioportal.org/

Funded by the European Union

the EOSC4Cancer General Assembly, demonstrating how data could be transferred via API between cBioPortal and Galaxy, enabling automated analyses within Galaxy and returning results into the cBioPortal. This integration is the basis for researchers to perform federated multi-omics analyses across distributed data environments, which is an important step forward for personalized medicine workflows.

An important development related to ESG's goals was the introduction of secure data access to EGA[103] and its federated nodes from within Galaxy. During Biohackathon Europe 2024, ESG partners developed a secure mechanism for analyzing encrypted Crypt4GH datasets inside Galaxy without compromising data privacy[104]. The system ensured that decryption only occurred during computation and that encrypted results were delivered exclusively to the authorized user. This innovation extended previous work from the ELIXIR implementation study on strengthening data management and demonstrated Galaxy's potential to handle sensitive biomedical datasets in a secure and compliant way.

Another use case in EOSC4Cancer highlighted the integration of radiology image analysis with genomic and clinical data (9). This workflow combined Galaxy's computational capabilities with the visualization features of cBioPortal, allowing researchers to conduct image-based analyses alongside molecular and clinical datasets. The setup proved flexible, with the potential to substitute the image analysis pipeline with any compatible Galaxy workflow. Together, these demonstrators solidified ESG's role in enabling new communities to adopt Galaxy not only as a workflow platform but also as a bridge between diverse data modalities, research infrastructures, and scientific domains.

Within clinical research, ESG broadened its reach through a collaboration led by UMC Utrecht (CardiOmnics) with the University of Oslo, Avans University of Applied Sciences, and the UMC Utrecht myDRE team to deploy a GDPR-compliant Galaxy inside the myDRE digital research environment. The instance inherits myDRE's security controls (two-factor authentication, audit trails, regular penetration testing) and operates with restricted egress. Tool definitions are retrieved from the Galaxy ToolShed and container images from Quay only during short, whitelisted windows; they are then converted to Apptainer and stored locally so analyses proceed without persistent internet access. Sensitive data reside on a separate filesystem kept unmounted during maintenance, enabling tool management without moving or exposing patient data and replacing earlier procedures that required temporary data removal. Access is provided via myDRE's bastion/RDP workflow to the Galaxy web interface on the VM's private IP. Together, these practices provide a transferable template for ESG partners needing

[103] https://ega-archive.org/
[104] https://github.com/elixir-europe/biohackathon-projects-2024/blob/main/22.md

Funded by the European Union

institution-hosted Galaxy within enterprise DREs and demonstrate ESG's support for communities working with regulated clinical datasets.

6. Training and Community Onboarding

Preparation of training materials were important to ESG's success for community adoption. Workshops, training events, and hackathons provided opportunities for researchers, tool developers, and system admins to familiarize themselves with Galaxy's functionalities. Events such as the Galaxy Admin Training[105], hackathons, workshops and participation and presentation in domain specific conferences significantly increased user engagement, drawing over 130,000 registered users and over 6,000 active users per month on average[106]. These training sessions were supplemented by online resources and comprehensive documentation, enabling researchers of varying expertise levels to efficiently utilize Galaxy.

7. Conclusion and Future Outlook

The ESG project aimed to make scientific computing more accessible, reproducible, and community-driven and it did so by bridging diverse research domains through the shared infrastructure of the Galaxy platform. Over the course of three years, ESG supported researchers in biodiversity, climate science, materials science, astrophysics, biomedicine, and beyond to perform data analysis. ESG provided them with the tools, training, and collaborative environments needed to turn their data into insight. These communities now benefit from over 800 newly integrated tools, dozens of reusable workflows, secure data access protocols, and domain-specific Galaxy instances that are used by European scientists.

But ESG's success goes beyond technical outcomes. It demonstrated that meaningful adoption of open science infrastructure requires more than code; it demands trust, mentorship, and a shared narrative. Different communities brought different needs, practices, and points of view. The project had to balance innovation with inclusion, technical accuracy with usability, and domain expertise with generalizability. Many challenges emerged: adapting Galaxy for encrypted clinical data, aligning with FAIR principles across disciplines, onboarding new users with limited computational background, and maintaining the sustainability beyond the project's lifetime. Without careful attention to human factors such as training, documentation, and peer support, this momentum could have easily fractured or faded.

Therefore, ESG prioritized long-term engagement alongside infrastructure development. It created onboarding models like the Galaxy Community onboarding cookbook (1),

[105] https://galaxyproject.org/events/2023-admin-training/

[106] https://stats.galaxyproject.eu/d/bejvb2fpdi96ob/eu-factsheet?orgId=1&from=now-6h&to=now&timezone=browser&viewPanel=panel-16

domain-specific subdomains, public workflows on WorkflowHub[107], and initiatives such as the Galaxy Codex. All designed to outlive the project and serve as scalable foundations for future community growth. The project's strategic integration into broader European and national initiatives such as EOSC, NFDI, and ELIXIR ensures that its tools, processes, and people remain part of a living ecosystem. Whether it's an ecologist analyzing species distribution in a changing climate, an astrophysicist correlating data across observatories, or a medical researcher securely processing patient images, ESG has helped make Galaxy a trusted platform across scientific frontiers.

As ESG concludes, its legacy is not just what was built, but what was made possible. The project has proven that infrastructure can be inclusive, that FAIR workflows can cross disciplinary boundaries, and that open science can be both robust and human-centered. The future will bring new challenges such as bigger datasets, new forms of collaboration, rising demands for security and transparency. However, it will also bring communities better prepared to meet them. Because of ESG, they now have the tools, the support, and most importantly, the story to carry them forward.

8. References

1. Nikolay V. Community onboarding cookbook published. 2024 Feb 29 [cited 2025 Aug 1]; Available from: https://zenodo.org/doi/10.5281/zenodo.15826941

2. Dadras A, Tangaro MA. Publication on the Pulsar Network, integrated in workflow management systems. 2025 Aug 7 [cited 2025 Aug 14]; Available from: https://zenodo.org/doi/10.5281/zenodo.16764158

3. Bretaudeau A, Cormier A, Leroi L, Corre E, Robin S, Kreplak J. RepeatMasker [Internet]. WorkflowHub; 2025 [cited 2025 Aug 5]. Available from: https://workflowhub.eu/workflows/1538?version=2

4. Bretaudeau A. Functional annotation [Internet]. WorkflowHub; 2025 [cited 2025 Aug 5]. Available from: https://workflowhub.eu/workflows/1521?version=1

5. Bretaudeau A. Funannotate [Internet]. WorkflowHub; 2025 [cited 2025 Aug 5]. Available from: https://workflowhub.eu/workflows/1525?version=1

6. Iaquinta J, Fouilloux A. Making Earth System Modelling and Observation Workflows Reproducible and Portable Across Federated High-Performance Computer and Cloud Platforms with Conda-based Containers. 2025 June 23 [cited 2025 Aug 5]; Available from: https://zenodo.org/doi/10.5281/zenodo.15600666

7. Nieva-de-la-Hidalga A, Liborio L, Austin P, Devadasan S, Underwood T, Belozerov A, et al. Facilitating Reproducibility in Catalysis Research with Managed Workflows and RO-Crates: A Galaxy Case Study. ChemCatChem. 2025 May;17(10):e202401676.

[107] https://workflowhub.eu/projects/166#workflows

Funded by the European Union

8. Jain S. Galaxy Sustainability Report [Internet]. Zenodo; 2025 July [cited 2025 Aug 5]. Available from: https://zenodo.org/doi/10.5281/zenodo.16030329

9. Capella-Gutierrez S, Valencia A, Barrowdale D, Schulz C, Boiten JW, Navest R, et al. Roadmap towards a federated digital platform for advancing cancer research: Leveraging current efforts and projects for a sustainable ecosystem. 2025 May 30 [cited 2025 Aug 5]; Available from: https://zenodo.org/doi/10.5281/zenodo.15682470

Funded by
the European Union