Received: 26 May 2026 | Revised: 3 August 2026 | Accepted: 25 August 2026

DOI: 10.1111/jmi.70165

Journal of Microscopy RMS

RAPID PUBLICATION

Bioimage management and analysis in Galaxy: Tools, workflows, training, and community practices

Leonid Kostrykin[1] | **Riccardo Massei**[2,3] | **Diana Chiang**[4] | **Thomas Wollmann**[1] | **Jean-Marie Burel**[5] | **David Lopez Tabernero**[4] | **Yi Sun**[6] | **Qi Gao**[1,7] | **Maarten W. Paul**[8] | **Pavankumar Videm**[4] | **Cameron Watson**[9] | **Daniel Franco-Barranco**[10,11,12] | **Anup Kumar**[4] | **Nadia Goué**[13,14] | **Reyhaneh Tavakoli Koopaei**[15] | **Vladimír Ulman**[16] | **Khaled Jum'ah**[17] | **Krzysztof Poterlowicz**[17] | **Martin Etzrodt**[18] | **Arrate Muñoz-Barrutia**[19] | **Sylvia E. Le Dévédec**[8] | **Josh Moore**[20] | **Jeremy Goecks**[21] | **Bjoern Gruening**[4] | **Karl Rohr**[1] | **Beatriz Serrano-Solano**[22]

[1]Biomedical Computer Vision Group, Heidelberg University, BioQuant, IPMB, Heidelberg, Germany

[2]Department of Monitoring and Environmental Technology (Research Data Management Team), Helmholtz Centre for Environmental Research GmbH (UFZ), Leipzig, Germany

[3]Department of Physiological Diversity, Working Group Imaging Flow Cytometry, German Centre for Integrative Biodiversity Research Halle-Jena-Leipzig (iDiv), Leipzig, Germany

[4]Bioinformatics Group, Department of Computer Science, University of Freiburg, Freiburg, Germany

[5]College of Life Sciences, University of Dundee, Dundee, Scotland, UK

[6]Data Science Centre, European Molecular Biology Laboratory (EMBL) Heidelberg, Heidelberg, Germany

[7]Core Facility Platform Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

[8]Division of Cell Systems and Drug Safety (CDS), Leiden Academic Centre for Drug Research (LACDR), Leiden Universiteit, Leiden, The Netherlands

[9]Department of Biomedical Engineering and Knight Cancer Institute, Oregon Health & Science University, Portland, Oregon, USA

[10]Neurobiology division, MRC Laboratory of Molecular Biology, Cambridge, England UK

[11]Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, England UK

[12]Donostia International Physics Center, San Sebastián, Spain

[13]AuBi Platform, Université Clermont Auvergne, Clermont-Ferrand, France

[14]IFB-core, Institut Français de Bioinformatique (IFB), CNRS, INSERM, INRAE, CEA, Villejuif, France

[15]Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran

[16]IT4Innovations, VSB – Technical University of Ostrava, Ostrava, Czech Republic

[17]Faculty of Health and Social Care, University of Bradford, Bradfort, England UK

[18]ISCC Foundation, Hengelo, The Netherlands

[19]Neuroscience and Life Sciences Department, Universidad Carlos III de Madrid, Madrid, Spain

[20]RSE Unit, German BioImaging – Gesellschaft für Mikroskopie und Bildanalyse e.V. (GerBI-GMB), Konstanz, Germany

[21]Department of Machine Learning, Moffitt Cancer Center, Tampa, Florida, USA

[22]Euro-BioImaging ERIC Bio-Hub, European Molecular Biology Laboratory (EMBL) Heidelberg, Heidelberg, Germany

Correspondence
Beatriz Serrano-Solano, Euro-BioImaging ERIC Bio-Hub, European Molecular Biology Laboratory (EMBL) Heidelberg, Heidelberg, Germany.
Email: beatriz.serrano.solano@eurobioimaging.eu

Funding information
Agence Nationale de la Recherche, Grant/Award Number: ANR-11-INBS-0013; BMFTR (Federal Ministry of Research, Technology and Space), Grant/Award Numbers: 031A538A, W-de.NBI-016; Deutsche Forschungsgemeinschaft, Grant/Award Numbers: IBSS-EXC-2189-ProjectID390939984, NFDI46/1-501864659; European Regional Development Fund; European Union, Grant/Award Number: 101129751; Ministerio de Ciencia, Innovación y Universidades and Agencia Estatal de Investigación, Grant/Award Numbers: PID2023-152631OB-I00, AIA2025-164165-C41; Ministry of Education, Youth and Sports of the Czech Republic, Grant/Award Number: 90254; Ministry of Science, Research and Technology of Iran; Ministry of Science, Research and the Arts Baden-Württemberg (MWK), Grant/Award Number: LiBiS; National Institutes of Health, Grant/Award Numbers: U24CA284167, U24HG010263; NWO National Roadmap for Large-Scale Research Infrastructure, Grant/Award Number: NWO184.036.012; UKRI, Grant/Award Numbers: BioFAIR, MR/V038966/1, UKRI/ST/B000299/1

ABSTRACT

Image analysis in the life sciences is constrained by fragmented software ecosystems, heterogeneous data formats, and limited reproducibility. These barriers hinder the reuse of image analysis methods and the sustainability of tools. In this article, we describe how the Galaxy platform enables FAIR (Findable, Accessible, Interoperable, and Reusable) image analysis by providing an integrated environment for data access, workflow execution, provenance tracing, and training.

We present Galaxy as a computational workbench that supports diverse image formats and integrates with public, institutional, and private repositories. We describe a reference structure for FAIR image analysis workflows and illustrate how this pattern supports reproducibility, interoperability, and reuse. We also describe community-driven training and sustainability practices that embed FAIR principles directly into executable tutorials and shared workflows.

Together, these foundations position Galaxy as a reproducible, scalable, and community-maintained platform for FAIR image analysis across the life sciences and beyond.

KEYWORDS
FAIR bioimage analysis, Galaxy, provenance, reproducibility, training, workflows

1 | INTRODUCTION

Bioimaging is a diverse and dynamic environment with no single file format that supports all applications and analysis. The necessity to use file format translators combined with (i) the diversity of computational tools whose scope, assumptions, and compatibility are often unclear, and (ii) the heterogeneous computational environments spanning from desktop applications to high-performance computing (HPC) systems, makes image analysis in life sciences challenging. While workflow management systems such as Nextflow[1] and Snakemake[2] offer a way to structure analyses,[3] their utilisation requires proficiency in programming or scripting languages. This creates additional barriers that make analyses tightly coupled to individual expertise and local environments, rather than being shareable or reproducible.

The fragmentation and technical challenges overwhelm researchers and significantly prevent using state-of-the-art tools for investigation, therefore impacting discoveries in a domain like bioimaging, where datasets are commonly large, multimodal, and closely tied to experimental metadata, and where meaningful analysis typically requires combining tools from multiple software ecosystems.[4]

Bioimaging data is also increasingly generated alongside other data types, such as transcriptomics or proteomics. This creates a need for analysis environments that can accommodate image analysis with established omics workflows, without requiring researchers to move data between platforms.

Together, tool fragmentation, heterogeneous computational environments, and increasing data multimodality highlight the lack of integrated environments that support FAIR[5, 6] (Findable, Accessible, Interoperable, and Reusable) image analysis in practice.

Applying the FAIR principles to image analysis workflows provides a framework for addressing these challenges. FAIR workflows rely on persistent identifiers for tools and workflows, machine-readable and standardised metadata, interoperable workflow definitions, explicit and exportable provenance, and long-term accessibility of both tools and data. These elements enable analyses that others can rerun, understand, adapt, and extend.

To cope with the growing complexity of the analyses, researchers increasingly rely on structured analysis workflows. Running an analysis workflow requires computational infrastructure that orchestrates a sequence of analysis steps to transform raw image data into meaningful, interpretable, and preferably quantitative results.[7] Such workflows often include data import, pre-processing, image analysis, statistical analysis, visualisation, and deposition into a data archive. By organising analyses into workflows, researchers capture the relationships between steps, link results to the original data, and lay the foundation for reproducible and reusable analyses.

While registries and repositories improve the discoverability of different software tools, those resources do not guarantee that they can integrate into coherent and reproducible workflows. Addressing this gap requires an execution environment that brings together different tools for image analysis, enforces structured workflow execution, and automatically captures provenance and versioning, conditions that are essential for FAIR image analysis in practice.

Galaxy, a well-established web-based computational workbench,[8] proposes to fill this gap by providing a common workflow environment in which the FAIR principles are embedded by design for users without programming experience. Galaxy emphasises accessibility, provenance tracking, interactive workflow construction, and integrated sharing of workflows and histories (an example workflow invocation is shown in Figure 1). Like Nextflow and Snakemake, Galaxy supports containerised tools through technologies such as Docker and Singularity/Apptainer, facilitating reproducibility across different computational environments. Galaxy integrates natively with high-performance computing clusters, providing transparent scalability for large-scale analyses, further ensuring reproducible and efficient resource utilisation. A thorough comparison of Galaxy with other workflow management systems can be found in Loach et al.[9]

Galaxy addresses the key barriers to FAIR image analysis workflows by providing:

1. a user-friendly interface that makes data and software tools findable and accessible,
2. interoperability across data formats and between software tools,
3. provenance of analyses to ensure reproducibility,
4. integration with research data management (RDM) systems, and
5. scalable execution environments.

In addition, human-in-the-loop workflows can combine headless tools with interactive applications, enabling both automation and exploratory analysis within a single environment.

A key feature of Galaxy lies in its open access to scalable HPC resources. Users can run compute-intensive workflows on institutional clusters through public servers at no cost (https://galaxyproject.org/use/?platform_group=public-servers), benefiting from parallelisation and reproducibility without managing infrastructure.

The Galaxy ecosystem is sustained through a community-driven model of development, training,[14] and curation of tools[15] and workflows. Contributions from researchers and developers worldwide ensure continuous updates, quality control, and alignment with emerging standards. This collaborative foundation makes Galaxy not only a reliable workbench but also a sustainable ecosystem for FAIR image analysis.

This article describes how the Galaxy computational workbench enables FAIR image analysis, robust workflow provenance tracing, and a rich community-sustained training ecosystem. We demonstrate how Galaxy enables image analysis workflows to be reusable and interoperable, providing example workflows that illustrate these principles in practice.

2 | FAIR IMAGE DATA MANAGEMENT AND ANALYSIS IN GALAXY

Galaxy's user interface provides findability and accessibility by enabling anyone, regardless of their technical expertise, to run analyses using thousands of different software tools. It captures every analysis step; represents multi-tool workflows in both human- and machine-readable formats; curates tools with metadata; and enables tool and workflow execution on scalable, public infrastructure without requiring local installation or programming expertise.

Galaxy supports FAIR image analysis by acting as a computational workbench that integrates different data repositories and storage solutions, software ecosystems for both automated and interactive analyses, and HPC resources within a single environment. Users can combine

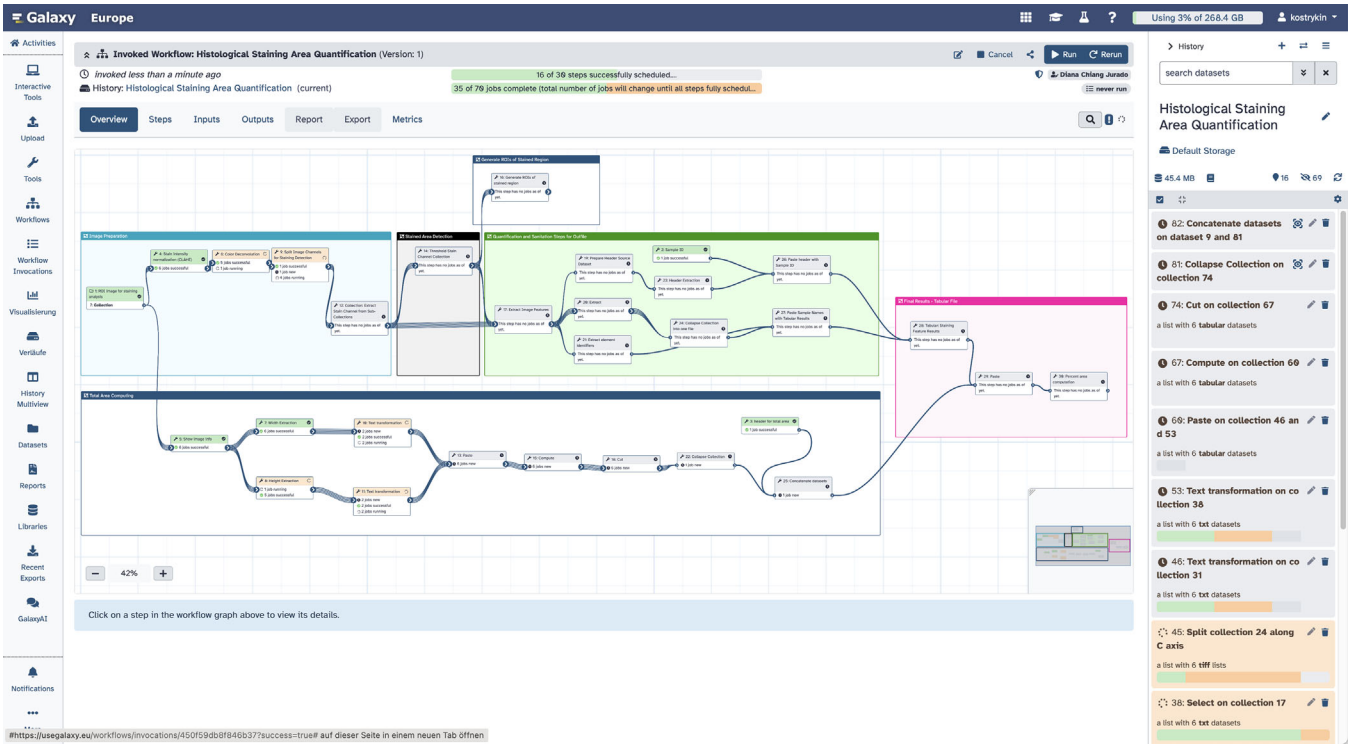

FIGURE 1 Example invocation of a workflow for quantification of staining areas in histology[10] for an input of six brightfield microscopy images. Each box of the workflow invocation graph represents a tool. The box colour changes over time and indicates the current state of the tool run (grey for queued, orange for in-progress, green for completed). The large boxes are groups of tools that represent the key steps of the workflow, comprising pre-processing[11, 12] ('Image Preparation'), image analysis via intensity thresholding[13] ('Stained Area Detection'), visualisation ('Generate ROIs of Stained Region'), and feature quantification ('Total Area Computing', 'Quantification and Sanitation Steps for Outfile', and 'Final Results – Tabular File').

methods from different domains while complying with the FAIR principles.

A common task in image analysis is the need for human input for validation, annotation, or visual inspection of intermediate results. Galaxy supports interactive, human-in-the-loop workflows that combine automated and manual tasks. When running such workflows, execution is automatically paused so that intermediate results can be visually inspected or annotated using Galaxy interactive tools such as CellProfiler,[16] QuPath,[17] ilastik,[18] napari,[19] AnyLabeling,[20] Vitessce,[21] or Jupyter Notebooks. For instance, YOLO segmentation workflows[22] integrate interactive annotation with automated model training. When the user finishes manual corrections and validation, the workflow automatically resumes, converting annotations into the required format and performing model training (https://gxy.io/GTN:T00550). This human-in-the-loop execution model integrates desktop-like user interactions and HPC resources within a single environment. Compared to fully manual workflows, it reduces tedious, repetitive, and error-prone tasks while achieving better model performance than fully automatic workflows.

The Galaxy computational workbench supports interoperability between individual image analysis tools, allowing researchers to combine them into a single reproducible workflow. This approach makes workflows easily shareable, as all steps, from pre-processing and segmentation to analysis and visualisation, can be executed within the same integrated environment (Figure 2).

2.1 | Data access and formats in Galaxy

Bioimaging generates large amounts of data that necessitate quality control, curation, and thorough annotation, even before processing and eventual publication.[29] The Galaxy ecosystem addresses these challenges by enabling access to a wide range of file formats, public and private repositories, and integrated storage solutions, while supporting interoperability between tools and workflows. By connecting data access with workflow execution and automated provenance tracing, Galaxy ensures that both raw data and all intermediate results are captured in a reproducible way. This helps researchers to apply FAIR

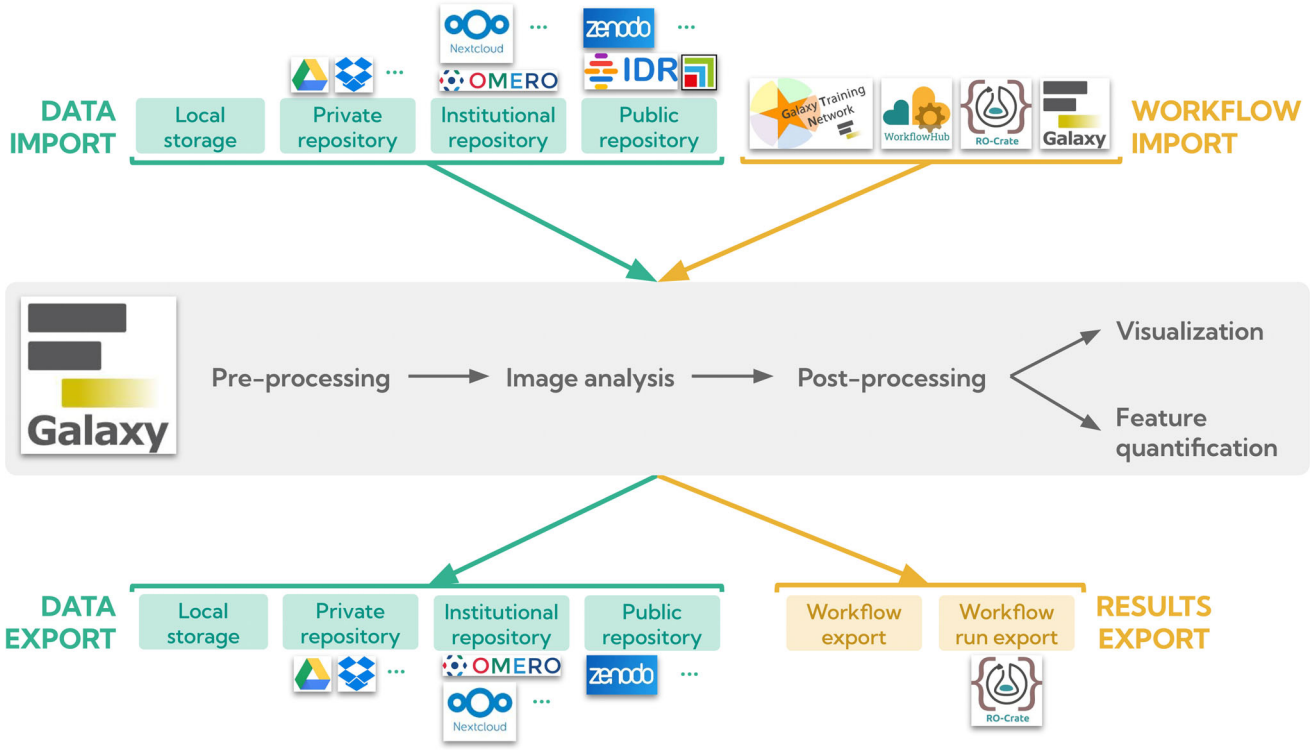

FIGURE 2 Galaxy-supported FAIR image analysis workflow reference structure. Data can be imported from local storage, private repositories (e.g., Dropbox, Google Drive), institutional systems (e.g., OMERO,[23] Nextcloud), and public resources (e.g., IDR,[24] Zenodo,[25] BIA[26]). Workflows can be discovered, imported, and executed from community registries such as the WorkflowHub[27] in RO-Crate[28] format or from training tutorials available at the Galaxy Training Network.[14] Within Galaxy, analyses are organised into structured workflows covering pre-processing, image analysis, and post-processing steps, followed by visualisation and feature quantification. Output data, workflows, and workflow runs can be exported to storage systems or shared as FAIR digital objects (e.g., RO-Crate).

principles and facilitate reproducible, shareable, and reusable image analysis across a wide range of modalities.

2.1.1 | Supported file formats

Image analysis workflows often require support for a wide range of file formats. Many commercial imaging devices write data in a Proprietary File Format (PFF), which varies by vendor and even imaging modality. Researchers, therefore, rely on file format converters to translate PFFs into open and standard file formats. For example, the Bio-Formats library[30] from the Open Microscopy Environment (OME) can translate over 160 PFFs (https://bio-formats.readthedocs.io/en/stable/supported-formats.html) and map their metadata into a common model.

In Galaxy, the Bio-Formats library is integrated as a tool that enables conversion of many image formats supported by Bio-Formats into TIFF, a format supported by a majority of image analysis tools in Galaxy. For visualisation purposes (e.g., teaching, publication, or visual inspection), images can also be exported in PNG format, which is widely supported by most tools but generally not intended for quantitative analysis.

Not all image formats supported by Bio-Formats can be converted using the file format conversion tools in Galaxy. The broader imaging community is actively working with vendors to ensure that they can adopt OME-Zarr,[31] a community-driven next-generation file format spearheaded by the OME initiative. OME-Zarr is an emerging cloud-optimised file format designed to efficiently handle large, multi-dimensional datasets; OME-Zarr enables selective remote access to multi-terabyte datasets using web standards, thereby making image data and subsets thereof accessible over HTTP/S via a URL. This allows tools and users to fetch portions of the data on demand (e.g., image sections, individual frames, or z-slices) without requiring full local downloads, either for (i) chunk-wise processing of the image data or for (ii) narrowing of the analyses to specific regions of interest. OME-Zarr thus permits efficient processing of very large image data, and Galaxy offers the ability to convert images using its conversion tool based on Bio-Formats. The number of Galaxy analysis tools supporting OME-Zarr is growing, therefore enabling FAIR-compliant access and processing within

Galaxy workflows. Several public repositories, such as the BioImage Archive (BIA)[26] or the Image Data Resource (IDR),[24] make newly published datasets available in the OME-Zarr format.

2.1.2 | Access to public, institutional, and private repositories

Galaxy is designed to be storage-agnostic and provides a wide range of options for storing and accessing data for the user. For example, on the European Galaxy server (https://usegalaxy.eu), users can freely access 250 GB of personal storage plus 2 TB of scratch storage. Data can be directly uploaded from local storage or retrieved remotely from any URL. OME-Zarr datasets can also be downloaded on-demand during job execution (namely, deferred datasets), supporting (i) adherence to Galaxy storage quotas and (ii) analyses of large files by streaming the image data (or subsets thereof) from remote storage (instead of acquiring large local copies). In addition, Galaxy offers a direct connection to multiple external cloud storage services, including Amazon S3, Microsoft Azure, and OneDrive, as well as S3-compatible institutional storage and European infrastructures. This flexibility is vital for maintaining data sovereignty and compliance with regional data protection regulations (e.g., GDPR) when handling sensitive bioimaging datasets.

Furthermore, Galaxy provides access to public deposition resources such as Zenodo[25] or public repositories hosting curated bioimage datasets and analysis results associated with peer-reviewed publications (e.g., IDR or BIA). Users with institutional OMERO instances can transfer image data directly into Galaxy workflows. Processed outputs and linked metadata can then be automatically saved back to these management systems, ensuring that stored data remains structured, reducing the risk of manual errors.[32]

Finally, Galaxy also supports linked documentation via electronic lab journals and laboratory management systems such as RSpace (https://www.researchspace.com) or eLabFTW,[33] embedding experimental context within the image analysis workflow. By integrating with these external experiment recording systems, Galaxy ensures that data and metadata are linked. This connectivity allows users to import, analyse, and, when appropriate, deposit datasets while preserving provenance.

2.2 | FAIR image analysis workflows

In Galaxy, users can create analysis workflows by freely combining analysis tools. The ability to tailor the sequence of steps allows for the transformation of raw image data into interpretable and, preferably, quantitative results. This flexibility facilitates reproducibility and reuse. The generated workflows are therefore adaptable to different imaging modalities, dimensionalities, or domain-specific questions.

In this section, we describe a reference structure for image analysis workflows in Galaxy that serves as a framework and facilitates modularity and interoperability. Earlier reference structures for image analysis workflows in Galaxy were designed for specific analysis tasks, such as cellular phenotyping.[34] Our reference structure encompasses a broader range of applications and comprises four steps described below.

While this reference structure distinguishes individual steps of image analysis for clarity and modularity, not all tools map to a single step. Some modern tools (e.g., Cellpose,[35, 36] SuperDSM[37, 38]) intentionally span multiple steps by integrating pre-processing, image analysis, post-processing, and, in some cases, quantitative output within a single workflow step. BiaPy[39] also provides configurable deep-learning pipelines that combine data preparation, model training or inference, post-processing, and optional extraction of object-level measurements within a single workflow step (see the BiaPy end-to-end workflow tutorial: https://gxy.io/GTN:T00571). The rest of this section considers the more general case of using individual tools for individual workflow steps.

2.2.1 | Pre-processing

Pre-processing comprises data preparation like file format conversion as well as pixel-level operations such as channel selection, intensity normalisation, background removal, and filtering (e.g., Gaussian or median filtering for image denoising) that are applied to improve overall image quality to ensure suitability for downstream analysis. Images may also be cropped (e.g., to select regions of interest) or resampled (e.g., to establish isotropic voxel sizes).

2.2.2 | Image analysis

The core analysis step transforms raw or pre-processed image intensities into structured biological representations. In practice, this includes detecting or segmenting objects using either supervised methods like PlantSeg,[40] Cellpose, YOLOv8 or unsupervised methods such as RFOVE,[41, 42] SuperDSM, or thresholding- and watershed-based approaches[43, 44] to generate segmentation masks or point detections. From these intermediate representations, for example, tracking methods can be employed by linking

objects over time into trajectories[45] while registration tools align images across channels, time points, samples, or modalities.[46]

By standardising on a small set of intermediate data representation types (e.g., segmentation masks, centroid tables, transformation matrices), Galaxy enables interchangeable combinations of tools within the same workflow. For instance, tools for processing, analysis, and interactive visualisation of multiplex tissue imaging datasets in Galaxy[47] adhere to the Minimum Information about Highly Multiplexed Tissue Imaging (MITI) data levels, which define common data types and file formats for each step in image processing.[48] In this framework, segmentation, detection, tracking, and registration operations are treated as part of a unified analysis stage that produces standardised intermediate representations that can be directly consumed by downstream tools in Galaxy.

2.2.3 | Post-processing and visualisation

Post-processing steps refine segmentation and tracking results, for example, by removing spurious objects or trajectories, refining trajectories or segmentation masks, or splitting falsely merged objects, typically using simple rules or forward computations. These steps may include rule-based filtering (e.g., removal of small objects), morphological operations, or consistency checks between channels or time points. Treating these corrections as explicit workflow components avoids manual, undocumented editing and supports reproducible tuning of analysis parameters.

Results can be visually inspected using different visualisation approaches and tools. While images can be directly inspected in Galaxy, users can also leverage interactive tools such as napari or Jupyter Notebooks for a more feature-rich visualisation environment. Galaxy also offers specialised tools like Vitessce for flexible visualisation of single-cell datasets and multimodal data. Additionally, Vizarr,[49] a tailored viewer for Zarr[50] files, has recently been added to Galaxy to enable interactive exploration of OME-Zarr files.

2.2.4 | Feature quantification

Workflows may also convert masks, points, or trajectories into quantitative measurements. Typical examples include morphological features (area, shape descriptors) or intensity-based statistics of objects or image regions,[34] co-localisation metrics, and motion descriptors (displacements, diffusion coefficients). These features are exported as annotated tables that remain linked to their provenance through Galaxy histories and can be further processed using statistical or machine learning-based tools within the same environment (e.g., SCIMAP,[51] Squidpy[52]), thereby preserving reproducibility and traceability.

By adopting this reference structure, that is, preprocessing, image analysis (e.g., segmentation, tracking, registration), post-processing, and feature quantification, Galaxy provides a template that can be reused and specialised across experiments. Different tools can be swapped in at each step without breaking the overall workflow. The resulting analyses and workflows can be registered, shared, and adapted by others. This approach does not eliminate the diversity of image analysis tasks, but it offers a practical level of standardisation that facilitates FAIRness, training, and interoperability across the image analysis community.

2.3 | Galaxy features supporting FAIRness

Galaxy sustains FAIR principles by design, providing a framework in which workflows, tools, and data remain findable, accessible, interoperable, and reusable. Its architecture ensures that every analysis step is captured, enabling reproducibility and traceability across diverse datasets and tools.

2.3.1 | Metadata and semantic annotation of tools

Galaxy records, at every step of the analysis, rich metadata for tools and workflows that can both be annotated with EDAM[53] ontology terms (https://edamontology.org) or bio.tools[54] identifiers. Limited support for BioImage Informatics Index (BIII, https://biii.eu) identifiers is also available.

Currently, tools for image processing and image analysis in Galaxy are annotated with the EDAM Operation Term 3443 ('Image Analysis'), supporting their findability within the large Galaxy toolset. The EDAM Bioimaging ontology (https://bioportal.bioontology.org/ontologies/EDAM-BIOIMAGING), currently under development, is an extension of the EDAM ontology with terms specific to bioimage analysis that allows a more fine-grained annotation of the individual tools and workflows according to the specific analysis tasks.

2.3.2 | Reproducibility, workflow provenance, and versioning

Workflow execution in Galaxy is traced through persistent histories and RO-Crate exports[28] to preserve provenance

and metadata for both intermediate and final results. Workflows are human- and machine-readable, fully versioned, and can be shared through public workflows repositories like WorkflowHub,[27] enabling precise reproducibility and reuse.

The Galaxy Training Network[14] (GTN) further reinforces best practices by automatically depositing workflows into WorkflowHub as part of tutorial development. This embeds workflow sharing and versioning directly into the creation of FAIR training materials, ensuring consistency between documented methods and executable analyses, but also makes the workflows easily searchable and citable.

In addition, to support reproducibility, the ISCC-SUM tool suite can be used in Galaxy. This provides content-based fingerprints that enable reproducibility checks by validating final or intermediate results by implementing the International Standard Content Code[55] (ISCC). ISCC-SUM (https://sum.iscc.codes) is file-type agnostic and not limited to bioimaging data; it can be extended to various heterogeneous data types, such as omics data. This content-based validation complements Galaxy's workflow provenance tracing, providing format-independent reproducibility assurance aligned with FAIR principles, which can be applied for validation when reproducing analyses (see the ISCC-SUM tutorial: https://gxy.io/GTN:T00572).

3 | COMMUNITY, TRAINING, AND IMPACT

The long-term sustainability of a platform like Galaxy cannot depend solely on its technical capabilities and funded resources (e.g., compute, development team), but it also depends on the strength of its developer and user communities.[56] Community building, engagement, and training are critical to raising awareness, promoting adoption, establishing common practices, and ensuring the integration and maintenance of tools and workflows.

This section describes the community-building strategies, training and outreach activities, and impact measurements for the image analysis ecosystem in Galaxy.

3.1 | Community building

As is often the case when developing software solutions, the early adoption of Galaxy was constrained by the availability of integrated tools. Over the last few years, the toolset has grown thanks to the sustained contributions from the community, including developers, image analysts, workflow authors, facility staff, and trainers. This collective effort has expanded the range of image analysis tasks that can be performed within Galaxy. Importantly, tools and workflows developed by the community benefit all users as contributions are shared and can be deployed on all Galaxy servers worldwide.

For the past few years, the FAIR Image Data Workflows Expert Group (https://www.eurobioimaging.eu/expert-groups/fair-image-data-workflows-expert-group) has run monthly meetings to coordinate the community efforts. Since 2023, three community projects have been accepted at the BioHackathon Europe, annually organised by ELIXIR (https://biohackathon-europe.org), enrolling new contributors every time. These projects have resulted in three preprints,[57–59] which include community conventions for annotation of tools to ensure internal consistency. The community also organised two dedicated hackathons (April 2024 and April 2025) to improve the codebase, develop tutorials, contribute to the ecosystem, and initiate collaborations with initiatives like EDAM, BIA, the AI4Life project,[60] OME, BIOMERO,[61] and Fractal.[62, 63]

Training in image analysis using Galaxy has been provided through various events in the past few years. Over 60 participants attended in-person courses on Microscopy Image Analysis held at Heidelberg University in 2016, 2019, and 2023. More recently, two workshops were run at the annual OME 2026 meeting in Düsseldorf, focusing on the integration of Galaxy with data management tools like OMERO and attended by more than 30 participants. To reach out to a broader audience, trainings were also delivered via webinars, including the GloBIAS Bioimage Analysis Seminar Series (2025), the EMBL-EBI Course on Microscopy Data Analysis and the BioImage Archive (2025, 2026), and the Euro-BioImaging Image Data Community Days (2025, 2026), attended by more than 200 participants worldwide.

The community maintains a dedicated section on the Galaxy Community Hub (https://galaxyproject.org/community/sig/image-analysis). In addition, Galaxy Labs[64] has been employed to create a community-tailored landing page that offers users a consistent interface with curated tools, workflows, and training materials that can be deployed on any Galaxy server. A production deployment is already available on the French Galaxy server (https://imaging.usegalaxy.fr) and is scheduled to be deployed on other public servers.

3.2 | Training materials at the Galaxy Training Network

Training activities are largely made possible by the Galaxy Training Network (GTN), which provides openly accessible and versioned tutorials. Those materials serve as step-by-step guides and documentation of workflows with

use cases. All image analysis tutorials are reviewed and maintained by the community; the material is available on GitHub, enabling collaborative updates.[65] GTN tutorials comply with Bioschemas[66] standards and are annotated with a custom set of terms that closely aligns with the latest available version of the EDAM Bioimaging ontology (cf. Section 2.3.1), making them easier to discover.

The entry point for tutorials on image analysis in Galaxy is the GTN landing page for the 'imaging' topic (https://training.galaxyproject.org/training-material/topics/imaging), which displays a continuously growing set of curated tutorials covering tasks such as image pre-processing, segmentation, and feature extraction, while also introducing advanced analyses like tissue multiplexing, parameter optimisation, and the integration of machine learning models. Tutorials have been created to cater to users with different levels of expertise. For newcomers, the tutorial 'Where to start with bioimage analysis in Galaxy' (https://gxy.io/GTN:T00573) provides a foundational overview of image data representations, data types, tools, and analysis possibilities within the Galaxy ecosystem, including identifying entry points via Bio-Formats or OMERO. Building on these fundamentals, tutorials such as 'Introduction to image analysis' (https://gxy.io/GTN:T00181) offer hands-on guidance for pre-processing filters and segmentation strategies. For more experienced researchers, the platform provides specialised tutorials for high-throughput workflows, including tracking and multi-step segmentation and feature extraction using CellProfiler, application of BioImage.IO models (https://bioimage.io) for AI-driven image analysis. This ensures a transition from basic data handling to complex, multi-step analysis within a unified, FAIR-compliant environment.

3.3 | Impact and sustainability

Community building and training efforts are reflected in the expansion of available tools, tutorials, and usage.

To assess the evolution of the image analysis toolset in the Galaxy ecosystem, we evaluated the number of tools and contributor activities. To do so, we focused on four GitHub repositories: bmcv/galaxy-image-analysis, bgruening/galaxytools, goeckslab/tools-mti, galaxyproject/tools-iuc, and lldelisle/tools-lldelisle, where most tools and tool integrations for image analysis in Galaxy are developed. Figure 3A shows a continuous increase in the number of available tools for image analysis in Galaxy, distributed across the four mentioned GitHub repositories, as can be seen in Figure 3B. The number of unique contributors has also increased since 2016, in particular since 2020, as shown in Figure 3C. This clearly indicates community engagement and reduced reliance on a core group of developers, contributing to improved sustainability through the distributed nature of maintenance and expertise.

We also assessed usage metrics for image analysis tools on the European Galaxy server, the server currently used by the majority of the tool users. The number of registered users has grown to over 600, with a steadily increasing adoption since 2022 (Figure 4A), which is also reflected in the number of tool runs (Figure 4B). Note that the number of users and tool runs started to be collected in late 2018, which was several months after the European Galaxy server was first launched.

The Galaxy Training Network conveniently provides yearly cumulative activity statistics. We extracted this data for the 'imaging' topic to assess how our training resources evolved. Figure 5 shows the continuous growth of training resources on image analysis at the GTN. Each tutorial is usually accompanied by at least one corresponding workflow, which is also reflected in the cumulative growth of the training resources.

Beyond documenting quantitative growth, our assessments suggest that sustainability is being reinforced through diversification of contributions, an increased number of workflows and tutorials, and growing user adoption. They also highlight the need for continued investment in the adoption of new tools, alignment with evolving standards, and sustained community engagement to maintain the long-term sustainability of the ecosystem.

The data and code used for our assessments are made publicly available: https://usegalaxy.eu/u/videmp/h/galaxy-imaging-community-stats-15-05-26.

4 | ROADMAP AND FUTURE DIRECTIONS

We have discussed how Galaxy can advance FAIR image analysis workflows through robust data management, adoption of community data formats, implementation of interoperable and reproducible workflows, provenance tracing, and community-driven training. Our vision for Galaxy is to serve as an open and integrated platform where image analysis workflows, metadata, training materials, and computational resources converge. By integrating these components within a single environment, Galaxy strengthens FAIRness by preserving explicit links between data, analysis steps, and execution context. Rather than replacing specialised desktop tools or domain-specific software, Galaxy complements them by integrating tools and software into structured, reproducible workflows embedded within the broader scientific ecosystem.

Galaxy has key strengths that make it well-suited for image analysis. The versatility of Galaxy has been demon-

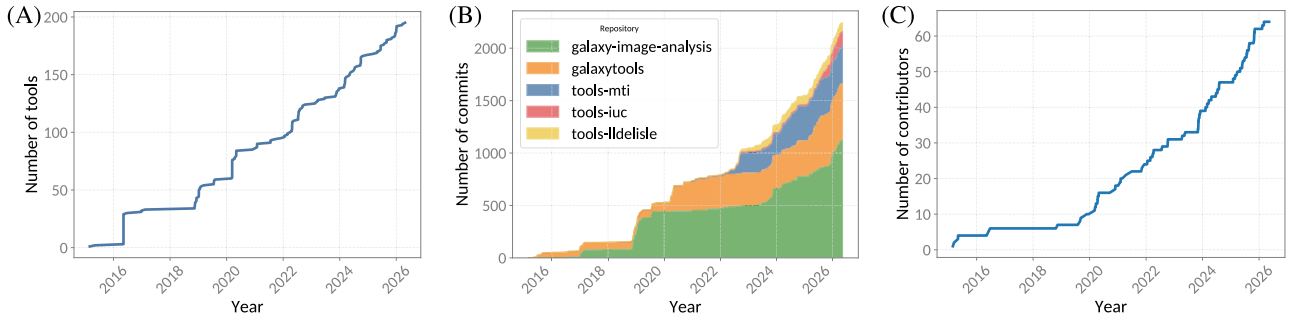

FIGURE 3 Assessment of growth of the tool count and contributor activity. (A) Number of tools for image analysis in Galaxy over time. (B) Number of Git commits in the different GitHub repositories over time (the repositories where most tools and tool integrations for image analysis in Galaxy are developed). (C) Number of unique tool integration contributors over time.

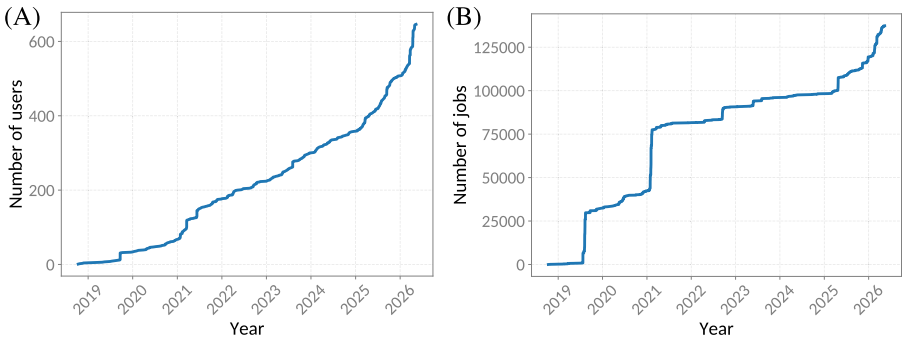

FIGURE 4 Assessment of metrics of bioimaging tool usage on the European Galaxy server. (A) Cumulative number of unique users over time. (B) Cumulative number of jobs executed over time (each job corresponds to a tool run).

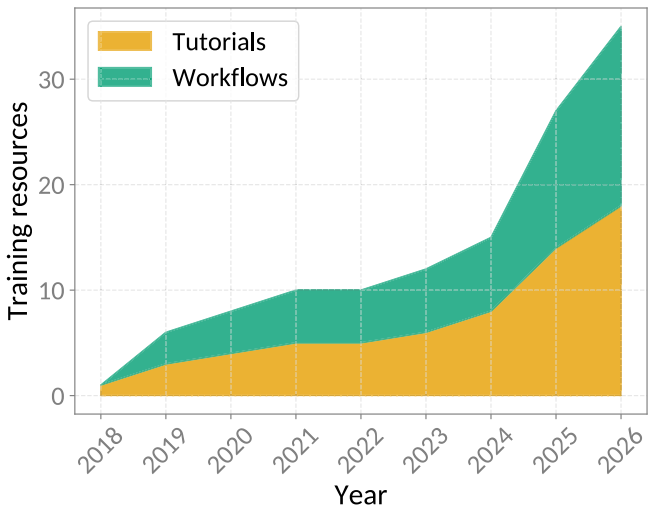

FIGURE 5 Cumulative number of training resources for image analysis (tutorials and workflows) at the Galaxy Training Network over time.

strated in diverse computational science consortia projects such as COVID-19 monitoring[67] and large-scale genome assembly across vertebrate animals.[68] These projects have showcased Galaxy's capabilities for reproducibility, scalability, and distributed reuse across users and institutions.

Galaxy is particularly well-suited for analyses that benefit from explicit workflow structure, provenance tracing, and integration with shared computational and data infrastructures. The modular architecture of workflows in Galaxy allows the combination of automated analysis with human-in-the-loop interaction and integrating tools originating from different scientific communities.

While Galaxy has traditionally been used in genomics, it now supports a wide range of interactive, automated, and exploratory analyses across different domains.[69] Being a shared platform by different communities makes cross-disciplinary workflows increasingly feasible, allowing methods developed in one field to be adapted and reused in others. Ongoing projects, such as FIESTA-OSCARS (https://oscars-project.eu/projects/fair-image-analysis-across-sciences), further explore this potential by investigating the reuse of image analysis workflows across bioimaging, environmental imaging, and astronomy.

Looking ahead, the image analysis community in Galaxy continues to expand support for emerging standards and practices. In particular, OME-Zarr is becoming central to handling large, multi-dimensional image data, and future Galaxy development will ensure that both new and existing tools increasingly support this format natively. For example, future efforts could include expand-

ing support for preclinical image data, building on already existing training resources (e.g., tutorial on segmentation of anatomical structures in 3-D images: https://gxy.io/GTN:T00574), or the adoption of the SpatialData[70] framework, which builds upon OME-Zarr to provide a FAIR format for unified storage of image data and downstream-derived datasets (e.g., annotations, label masks, and quantified feature tables), further improving provenance tracing, particularly for complex spatial omics datasets. In parallel, work is ongoing to strengthen Galaxy's interoperability with external repositories and research data management systems, including tighter integration with resources such as the BIA, to better support end-to-end FAIR workflows from data ingestion to deposition.

While Galaxy already provides strong support for FAIR image analysis, several areas offer opportunities for further growth. Image analysis workflows often involve rapidly evolving tools, heterogeneous metadata practices, and domain-specific assumptions that complicate standardisation. Community discussions have highlighted the need for improved metadata harmonisation and clearer conventions for workflow design. Addressing these challenges requires continued collaboration between tool developers, infrastructure providers, and user communities.

By design, Galaxy is not the most suitable environment for highly interactive exploration of image data. Tasks that demand rapid, low-latency visual feedback or intensive graphical interaction, such as exploratory inspection of raw image stacks or fine-tuning segmentation parameters, are often more efficiently performed in standalone desktop applications. Similarly, the shared and scheduled nature of public Galaxy servers, as with any HPC or cloud infrastructure, can introduce waiting times that obstruct real-time interaction. In practice, many users adopt a complementary approach: they explore data and optimise parameters locally, then translate the resulting procedures into reproducible Galaxy workflows for scalable execution, sharing, and long-term reuse.

The continued growth of demand for image analysis in the life sciences requires platforms that combine flexibility with reproducibility, scalability, and FAIR data practices. The Galaxy computational workbench is well-equipped to fulfil this role as an integrated platform for bioimage data and software analysis tools, workflows, and training resources. Its open architecture, active global community, and established infrastructure provide a sustainable foundation for FAIR image analysis.

ACKNOWLEDGEMENTS

We warmly thank Bérénice Batut, Jean-Karim Hériché, Lucille Delisle, Solenne Correard, Thomas Chaussepied, Titusz Pan, Judith Lacoste, and Virginia Silio for their valuable support, feedback, and contributions throughout this work. We also gratefully acknowledge all contributors to the 'imaging' topic of the Galaxy Training Network (GTN) for their efforts in developing and maintaining community training resources.

This work was partially funded by the BMFTR (Federal Ministry of Research, Technology and Space), Heidelberg Center for Human Bioinformatics (HD-HuB), within the German Network for Bioinformatics Infrastructure (de.NBI) and ELIXIR-DE (W-de.NBI-016), and by the MWK (Baden-Württemberg Ministry of Science, Research and Arts) within LiBiS (Baden-Württemberg Institute for Bioinformatics Infrastructure); Ministerio de Ciencia, Innovación y Universidades and Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033/) under grants PID2023-152631OB-I00 and AIA2025-164165-C41, co-financed by the European Regional Development Fund (ERDF), 'A way of making Europe'; the OSCARS 1st Open Call for Open Science Projects and Services under grant agreement No. 101129751 through the BIOCODES ('Enhancing AI-Readiness of Bioimaging Data with Content-Based Identifiers') and FIESTA ('Image Analysis across Sciences') projects; the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (CIBSS – EXC-2189 – Project ID 390939984) and through the National Research Data Infrastructure programme NFDI4BIOIMAGE (NFDI 46/1 – 501864659); the BMFTR (Federal Ministry of Research, Technology and Space) through de.NBI-RBC (031 A538A) and LiBiS/de.NBI Freiburg; NL-Bioimaging infrastructure, NWO National Roadmap for Large-Scale Research Infrastructure (NWO 184.036.012); BioFAIR, a UKRI-funded federated digital research infrastructure; UKRI grant MR/V038966/1 and the ELIXIR Spatial2Galaxy grant; and the UKRI Digital Research Skills Catalyst funded by the UKRI Digital Research Infrastructure Programme (Project reference: UKRI/ST/B000299/1); the Clermont-Auvergne Mesocenter infrastructure (UCA) and the UCA Research Council; the HPC facilities of the National Network of Computing Resources of the Institut Français de Bioinformatique (IFB), funded by the Programme d'Investissements d'Avenir (PIA) grant ANR-11-INBS-0013; the United States National Institutes of Health grants U24CA284167 and U24HG010263; the Ministry of Education, Youth and Sports of the Czech Republic through e-INFRA CZ (ID:90254); and the Ministry of Science, Research and Technology of Iran together with the Freiburg City Fellowship.

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interest.

ORCID

Leonid Kostrykin https://orcid.org/0000-0003-1323-3762
Riccardo Massei https://orcid.org/0000-0003-2104-9519
Diana Chiang https://orcid.org/0000-0002-5857-1477
Thomas Wollmann https://orcid.org/0000-0002-4741-3844
David Lopez Tabernero https://orcid.org/0000-0002-9541-3961
Yi Sun https://orcid.org/0000-0002-7636-0200
Maarten W. Paul https://orcid.org/0000-0002-7990-6010
Pavankumar Videm https://orcid.org/0000-0002-5192-126X
Cameron Watson https://orcid.org/0000-0002-6942-2469
Daniel Franco-Barranco https://orcid.org/0000-0002-1109-110X
Anup Kumar https://orcid.org/0000-0002-2068-4695
Nadia Goué https://orcid.org/0000-0003-2750-1473
Reyhaneh Tavakoli Koopaei https://orcid.org/0000-0002-8964-3023
Vladimir Ulman https://orcid.org/0000-0002-4270-7982
Khaled Jum'ah https://orcid.org/0000-0001-5481-8893
Krzysztof Poterlowicz https://orcid.org/0000-0001-6173-5674
Martin Etzrodt https://orcid.org/0000-0003-1928-3904
Arrate Muñoz-Barrutia https://orcid.org/0000-0002-1573-1661
Sylvia E. Le Dévédec https://orcid.org/0000-0002-0615-9616
Josh Moore https://orcid.org/0000-0003-4028-811X
Jeremy Goecks https://orcid.org/0000-0002-4583-5226
Bjoern Gruening https://orcid.org/0000-0002-3079-6586
Beatriz Serrano-Solano https://orcid.org/0000-0002-5862-6132

REFERENCES

1. Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nature Biotechnology*, *35*(4), 316–319. https://doi.org/10.1038/nbt.3820
2. Mölder, F., Jablonski, K. P., Letcher, B., Hall, M. B., van Dyken, P. C., Tomkins-Tinch, C. H., Sochat, V., Forster, J., Vieira, F. G., Meesters, C., Lee, S., Twardziok, S. O., Kanitz, A., VanCampen, J., Malladi, V., Wilm, A., Holtgrewe, M., Rahmann, S., Nahnsen, S., & Köster, J. (2025). Sustainable data analysis with snakemake. *F1000Research*, *10*, 33. https://doi.org/10.12688/f1000research.29032.3
3. Wratten, L., Wilm, A., & Göke, J. (2021). Reproducible, scalable, and shareable analysis pipelines with bioinformatics workflow managers. *Nature Methods*, *18*(10), 1161–1168. https://doi.org/10.1038/s41592-021-01254-9
4. Dobson, E. T. A., Cimini, B., Klemm, A. H., Wählby, C., Carpenter, A. E., & Eliceiri, K. W. (2021). ImageJ and CellProfiler: Complements in Open-Source bioimage analysis. *Current Protocols*, *1*(5), e89. https://doi.org/10.1002/cpz1.89
5. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Gonzalez-Beltran, A. (2016). The FAIR guiding principles for scientific data management & stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18
6. Wilkinson, S. R., Gustafsson, J., Bacall, F., Belhajjame, K., Capella, S., Gonzalez, J. M. F., Tande, J. F., Gadelha, L., Garijo, D., Grubel, P., Grüning, B., Khan, F. Z., Kanwal, S., Leo, S., Owen, S., Pireddu, L., Pouchard, L., Rodríguez-Navas, L., Serrano-Solano, B., … Soiland-Reyes, S. (2025). An ecosystem of services for FAIR computational workflows. *arXiv preprint*, arXiv:2505.15988. https://doi.org/10.48550/arXiv.2505.15988
7. Schmied, C., Nelson, M. S., Avilov, S., Bakker, G.-J., Bertocchi, C., Bischof, J., Boehm, U., Brocher, J., Carvalho, M. T., Chiritescu, C., Christopher, J., Cimini, B. A., Conde-Sousa, E., Ebner, M., Ecker, R., Eliceiri, K., Fernandez-Rodriguez, J., Gaudreault, N., Gelman, L., … Grunwald, D. (2024). Community-developed checklists for publishing images and image analyses. *Nature Methods*, *21*(2), 170–181. https://doi.org/10.1038/s41592-023-01987-9
8. Galaxy Community. (2024). The galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Research*, *52*(W1), W83–W94. https://doi.org/10.1093/nar/gkae410
9. Loach, M., Smith, K., & Bacon, W. (2026). A scoping review of approaches to evaluating workflow management systems for bioinformatics users. *Briefings in Bioinformatics*, *27*(4). https://doi.org/10.1093/bib/bbag396
10. Jurado, D. C. (2026). Histological staining area quantification. Zenodo. https://doi.org/10.5281/zenodo.20433204
11. Ruifrok, A. C., & Johnston, D. A. (2001). Quantification of histochemical staining by color deconvolution. *Analytical and Quantitative Cytology and Histology*, *23*(4), 291–299.
12. Zuiderveld, K. (1994). Viii.5. – Contrast limited adaptive histogram equalization. In Heckbert, P. S. (Ed.), *Graphics gems* (pp. 474–485). Academic Press, 1st edition. https://doi.org/10.1016/b978-0-12-336156-1.50061-6
13. Li, C., & Lee, C. (1993a). Minimum cross entropy thresholding. *Pattern Recognition*, *26*(4), 617–625. https://doi.org/10.1016/0031-3203(93)90115-D
14. Hiltemann, S., Rasche, H., Gladman, S., Hotz, H.-R., Larivière, D., Blankenberg, D., Jagtap, P. D., Wollmann, T., Bretaudeau, A., Goué, N., Griffin, T. J., Royaux, C., Le Bras, Y., Mehta, S., Syme, A., Coppens, F., Droesbeke, B., Soranzo, N., Bacon, W., …, Psomopoulos, F. (2023). Galaxy training: A powerful framework for teaching! *PLOS Computational Biology*, *19*(1), e1010752. https://doi.org/10.1371/journal.pcbi.1010752
15. Blankenberg, D., Von Kuster, G., Bouvier, E., Baker, D., Afgan, E., Stoler, N., Taylor, J., & Nekrutenko, A. (2014). Dissemination of scientific software with galaxy ToolShed. *Genome Biology*, *15*(2), 403. https://doi.org/10.1186/gb4161

16. Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., Guertin, D. A., Chang, J. H., Lindquist, R. A., Moffat, J., Golland, P., & Sabatini, D. M. (2006). CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, *7*(10), R100. https://doi.org/10.1186/gb-2006-7-10-r100
17. Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., James, J. A., Salto-Tellez, M., & Hamilton, P. W. (2017). QuPath: Open source software for digital pathology image analysis. *Scientific Reports*, *7*(1), 1–7. https://doi.org/10.1038/s41598-017-17204-5
18. Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J. I., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Koethe, U., Hamprecht, F. A., & Kreshuk, A. (2019). ilastik: Interactive machine learning for (bio)image analysis. *Nature Methods*, *16*(12), 1226–1232.
19. Chiu, C.-L., & Clack, N. (2022). napari: A python multi-dimensional image viewer platform for the research community. *Microscopy and Microanalysis*, *28*(S1), 1576–1577. https://doi.org/10.1017/S1431927622006328
20. Wang, W. (2023). Advanced auto labeling solution with added features. Retrieved from https://github.com/CVHub520/X-AnyLabeling
21. Keller, M. S., Gold, I., McCallum, C., Manz, T., Kharchenko, P. V., & Gehlenborg, N. (2024). Vitessce: integrative visualization of multimodal and spatially resolved single-cell data. *Nature Methods*, *22*(1), 63–67. https://doi.org/10.1038/s41592-024-02436-x
22. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). You only look once: Unified, real-time object detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. https://doi.org/10.1109/CVPR.2016.91
23. Allan, C., Burel, J.-M., Moore, J., Blackburn, C., Linkert, M., Loynton, S., MacDonald, D., Moore, W. J., Neves, C., Patterson, A., Porter, M., Tarkowska, A., Loranger, B., Avondo, J., Lagerstedt, I., Lianas, L., Leo, S., Hands, K., Hay, R. T., …, Patwardhan, A. (2012). OMERO: Flexible, model-driven data management for experimental biology. *Nature Methods*, *9*(3), 245–253. https://org/10.1038/nmeth.1896
24. Williams, E., Moore, J., Li, S. W., Rustici, G., Tarkowska, A., Chessel, A., Leo, S., Antal, B., Ferguson, R. K., Sarkans, U., Brazma, A., Carazo Salas, R. E., & Swedlow, J. R. (2017). Image data resource: A bioimage data integration and publication platform. *Nature Methods*, *14*(8), 775–781. https://doi.org/10.1038/nmeth.4326
25. European Organization For Nuclear Research and OpenAIRE. (2013). Zenodo. https://www.zenodo.org
26. Hartley, M., Kleywegt, G. J., Patwardhan, A., Sarkans, U., Swedlow, J. R., & Brazma, A. (2022). The BioImage archive – Building a home for life-sciences microscopy data. *Journal of Molecular Biology*, *434*(11), 167505. https://doi.org/10.1016/j.jmb.2022.167505
27. Gustafsson, O. J. R., Wilkinson, S. R., Bacall, F., Soiland-Reyes, S., Leo, S., Pireddu, L., Owen, S., Juty, N., Fernández, J. M., Brown, T., Ménager, H., Grüning, B., Capella-Gutierrez, S., Coppens, F., & Goble, C. (2025). WorkflowHub: A registry for computational workflows. *Scientific Data*, *12*(1), 837.
28. Soiland-Reyes, S., Sefton, P., Crosas, M., Castro, L. J., Coppens, F., Fernández, J. M., Garijo, D., Grüning, B., La Rosa, M., Leo, S., Carragáin, E. Ó., Portier, M., Trisovic, A., Groth, P., & Goble, C. (2022). Packaging research artefacts with RO-Crate. *Data Science*, *5*(2), 97–138.
29. Schmidt, C., Boissonnet, T., Dohle, J., Bernhardt, K., Ferrando-May, E., Wernet, T., Nitschke, R., Kunis, S., & Weidtkamp-Peters, S. (2024). A practical guide to bioimaging research data management in core facilities. *Journal of Microscopy*, *294*(3), 350–371. https://doi.org/10.1111/jmi.13317
30. Linkert, M., Rueden, C. T., Allan, C., Burel, J.-M., Moore, W., Patterson, A., Loranger, B., Moore, J., Neves, C., Macdonald, D., Tarkowska, A., Sticco, C., Hill, E., Rossner, M., Eliceiri, K. W., & Swedlow, J. R. (2010). Metadata matters: access to image data in the real world. *Journal of Cell Biology*, *189*(5), 777–782. https://doi.org/10.1083/jcb.201004104
31. Moore, J., Basurto-Lozada, D., Besson, S., Bogovic, J., Bragantini, J., Brown, E. M., Burel, J.-M., Casas Moreno, X., de Medeiros, G., Diel, E. E., Gault, D., Ghosh, S. S., Gold, I., Halchenko, Y. O., Hartley, M., Horsfall, D., Keller, M. S., Kittisopikul, M., Kovacs, G., … Küpcü Yoldaş, A. (2023). OME-Zarr: A cloud-optimized bioimaging file format with international community support. *Histochemistry and Cell Biology*, *160*(3), 223–251. https://doi.org/10.1007/s00418-023-02209-1
32. Massei, R., Busch, W., Serrano-Solano, B., Bernt, M., Scholz, S., Nicolay, E. K., Bohring, H., & Bumberger, J. (2025). High-content screening (HCS) workflows for FAIR image data management with OMERO. *Scientific Reports*, *15*(1), 16236. https://doi.org/10.1038/s41598-025-00720-0
33. Hewera, M., Hänggi, D., Gerlach, B., & Kahlert, U. D. (2021). eLabFTW as an open science tool to improve the quality and translation of preclinical research. *F1000Research*, *10*, 292. https://doi.org/10.12688/f1000research.52157.3
34. Wollmann, T., Erfle, H., Eils, R., Rohr, K., & Gunkel, M. (2017). Workflows for microscopy image analysis and cellular phenotyping. *Journal of Biotechnology*, *261*, 70–75. https://doi.org/10.1016/j.jbiotec.2017.07.019
35. Pachitariu, M., Rariden, M., & Stringer, C. (2025). Cellpose-sam: Superhuman generalization for cellular segmentation. *BioRxiv*. https://doi.org/10.1101/2025.04.28.651001
36. Stringer, C., & Pachitariu, M. (2025). Cellpose3: One-click image restoration for improved cellular segmentation. *Nature Methods*, *22*(3), 592–599. https://doi.org/10.1038/s41592-025-02595-5
37. Kostrykin, L., & Rohr, K. (2023). Superadditivity and convex optimization for globally optimal cell segmentation using deformable shape models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(3), 3831–3847. https://doi.org/10.1109/TPAMI.2022.3185583
38. Kostrykin, L., & Rohr, K. (2024). Robust graph pruning for efficient segmentation and cluster splitting of cell nuclei using deformable shape models. In *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI)* (pp. 1–4). IEEE. https://doi.org/10.1109/ISBI56570.2024.10635542
39. Franco-Barranco, D., Andrés-San Román, J. A., Hidalgo-Cenalmor, I., Backová, L., González-Marfil, A., Caporal, C., Chessel, A., Gómez-Gálvez, P., Escudero, L. M., Wei, D., Muñoz-Barrutia, A., & Arganda-Carreras, I. (2025). BiaPy: Accessible deep learning on bioimages. *Nature Methods*, *22*(6), 1124–1126. https://doi.org/10.1038/s41592-025-02699-y

40. Wolny, A., Cerrone, L., Vijayan, A., Tofanelli, R., Barro, A. V., Louveaux, M., Wenzl, C., Strauss, S., Wilson-Sánchez, D., Lymbouridou, R., Steigleder, S. S., Pape, C., Bailoni, A., Duran-Nebreda, S., Bassel, G. W., Lohmann, J. U., Tsiantis, M., Hamprecht, F. A., Schneitz, K., … Maizel, A. (2020). Accurate and versatile 3D segmentation of plant tissues at cellular resolution. *eLife*, *9*, e57613. https://doi.org/10.7554/eLife.57613
41. Panagiotakis, C., & Argyros, A. (2018). Cell segmentation via region-based ellipse fitting. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)* (pp. 2426–2430). IEEE. https://doi.org/10.1109/ICIP.2018.8451852
42. Panagiotakis, C., & Argyros, A. (2020). Region-based fitting of overlapping ellipses and its application to cells segmentation. *Image and Vision Computing*, *93*, 103810. https://doi.org/10.1016/j.imavis.2019.09.001
43. Li, C., & Lee, C. (1993b). Minimum cross entropy thresholding. *Pattern Recognition*, *26*(4), 617–625. https://doi.org/10.1016/0031-3203(93)90115-D
44. Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. https://doi.org/10.1109/TSMC.1979.4310076
45. Wang, D., Gao, Q., Schaefer, I., Moerz, H., Hoheisel, U., Rohr, K., Greffrath, W., & Treede, R.-D. (2022). TRPM3-mediated dynamic mitochondrial activity in nerve growth factor-induced latent sensitization of chronic low back pain. *Pain*, *163*(11), e1115–e1128. https://doi.org/10.1097/j.pain.0000000000002642
46. Föll, M. C., Moritz, L., Wollmann, T., Stillger, M. N., Vockert, N., Werner, M., Bronsert, P., Rohr, K., Grüning, B. A., & Schilling, O. (2019). Accessible and reproducible mass spectrometry imaging data analysis in Galaxy. *GigaScience*, *8*(12), giz143. https://doi.org/10.1093/gigascience/giz143
47. Creason, A., Watson, C., Gu, Q., Persson, D., Sargent, L., Chen, Y.-A., Lin, J.-R., Sivagnanam, S., Wünnemann, F., Nirmal, A. J., Chin, K., Feiler, H., Coussens, L., Schapiro, D., Grüning, B. A., Sorger, P., Sokolov, A., & Goecks, J. (2022). A web-based software resource for interactive analysis of multiplex tissue imaging datasets. *BioRxiv*. https://doi.org/10.1101/2022.08.18.504436
48. Schapiro, D., Yapp, C., Sokolov, A., Reynolds, S. M., Chen, Y.-A., Sudar, D., Xie, Y., Muhlich, J., Arias-Camison, R., Arena, S., Taylor, A. J., Nikolov, M., Tyler, M., Lin, J.-R., Burlingame, E. A., Human Tumor Atlas Network, Chang, Y. H., Farhi, S. L., Thorsson, V., Venkatamohan, N., … Drewes, J. L. (2022). MITI minimum information guidelines for highly multiplexed tissue images. *Nature Methods*, *19*(3), 262–267. https://doi.org/10.1038/s41592-022-01415-4
49. Manz, T., Gold, I., Patterson, N. H., McCallum, C., Keller, M. S., Herr, B. W., Börner, K., Spraggins, J. M., & Gehlenborg, N. (2022). Viv: Multiscale visualization of high-resolution multiplexed bioimaging data on the web. *Nature Methods*, *19*(5), 515–516. https://doi.org/10.1038/s41592-022-01482-7
50. Newman, D. J. (2024). Zarr storage specification. Retrieved from https://doi.org/10.5067/DOC/ESCO/ESDS-RFC-048V1
51. Nirmal, A. J., & Sorger, P. K. (2024). SCIMAP: A python toolkit for integrated spatial analysis of multiplexed imaging data. *Journal of Open Source Software*, *9*(97). https://doi.org/10.21105/joss.06604
52. Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A. C., Kuemmerle, L. B., Rybakov, S., Ibarra, I. L., Holmberg, O., Virshup, I., Lotfollahi, M., Richter, S., & Theis, F. J. (2022). Squidpy: A scalable framework for spatial omics analysis. *Nature Methods*, *19*(2), 171–178. https://doi.org/10.1038/s41592-021-01358-2
53. Ison, J., Kalaš, M., Jonassen, I., Bolser, D., Uludag, M., McWilliam, H., Malone, J., Lopez, R., Pettifer, S., & Rice, P. (2013). Edam: An ontology of bioinformatics operations, types of data and identifiers, topics and formats. *Bioinformatics*, *29*(10), 1325–1332. https://doi.org/10.1093/bioinformatics/btt113
54. Ison, J., Rapacki, K., Ménager, H., Kalaš, M., Rydza, E., Chmura, P., Anthon, C., Beard, N., Berka, K., Bolser, D., Booth, T., Bretaudeau, A., Brezovsky, J., Casadio, R., Cesareni, G., Coppens, F., Cornell, M., Cuccuru, G., Davidsen, K., …, Vedova, G. D. (2016). Tools and data services registry: A community effort to document bioinformatics resources. *Nucleic Acids Research*, *44*(D1), D38–47.
55. ISO 24138:2024. (2024). Information and documentation – International Standard Content Code (ISCC). Retrieved from https://www.iso.org/standard/77899.html
56. Jain, S. (2025). Galaxy sustainability report. Zenodo. https://doi.org/10.5281/zenodo.16030329
57. Chiang, D., Videm, P., Tabernero, D. L., Paul, M. W., Heidari, A., Etzrodt, M., Serrano-Solano, B., & Kostrykin, L. (2026). Evolving fair image analysis in galaxy for cross-domain and AI-ready applications. BioHackrXiv. https://doi.org/10.37044/osf.io/tsxby_v1
58. Kostrykin, L., Depoortere, B., Chiang, D., Klumpe, S., Massei, R., Kalaš, M., Fouilloux, A., & Serrano-Solano, B. (2025). Development of FAIR image analysis workflows and training in Galaxy. BioHackrXiv. https://doi.org/10.37044/osf.io/hbw6m_v1
59. Kostrykin, L., Woller, T., Kalaš, M., Özdemir, B., Buono, R. A., Sun, Y., Kumar, A., Goué, N., Gruening, B., & Serrano-Solano, B. (2024). Enhancing the image analysis community in galaxy. BioHackrXiv.org. https://doi.org/10.37044/osf.io/w8dsz
60. Muñoz-Barrutia, A., Serrano-Solano, B., Fuster Barcelo, C., Pape, C., Russell, C. T., Lundberg, E., Gómez-de Mariscal, E., Jug, F., Deschamps, J., Hidalgo-Cenalmor, I., Ferreira, M. G., Hartley, M., Seifi, M., Henriques, R., Zulueta-Coarasa, T., Galinova, V., & Ouyang, W. (2025). *Artificial intelligence for image data analysis in the life sciences*. European Union. https://doi.org/10.3030/101057970
61. Luik, T. T., Rosas-Bertolini, R., Reits, E. A., Hoebe, R. A., & Krawczyk, P. M. (2024). BIOMERO: A scalable and extensible image analysis framework. *Patterns*, *5*(8). https://doi.org/10.1016/j.patter.2024.101024
62. Lüthi, J. (2024). Fractal: An open-source framework for reproducible bioimage analysis at scale using OME-Zarrs. In *BIO Web of Conferences* (Vol. *129*, p. 10026). EDP Sciences. https://doi.org/10.1051/bioconf/202412910026
63. Lüthi, J., Cerrone, L., Comparin, T., Hess, M., Hornbachner, R., Tschan, A., Glasner de Medeiros, G. Q., Repina, N. A., Cantoni, L. K., Steffen, F. D., Bourquin, J.-P., Liberali, P., Pelkmans, L., & Uhlmann, V. (2026). Fractal: Towards fair bioimage analysis at scale with ome-zarr-native workflows. bioRxiv. https://doi.org/10.64898/2026.03.05.709921
64. Nasr, E., Amato, P., Bhardwaj, A., Blankenberg, D., Brites, D., Cumbo, F., Do, K., Ferrari, E., Griffin, T. J., Gruening, B., Hiltemann, S., Jagtap, P., Mehta, S., Métris, K. L. K., Momin, S., Oba, A., Pavloudi, C., Pechlivanis, N., Péguilhan,

R., … Psomopoulos, F. (2024). microGalaxy: A gateway to tools, workflows, and training for reproducible and FAIR analysis of microbial data. *bioRxiv*. https://doi.org/10.1101/2024.12.23.629682

65. Batut, B., Hiltemann, S., Bagnacani, A., Baker, D., Bhardwaj, V., Blank, C., Bretaudeau, A., Brillet-Guéguen, L., Čech, M., Chilton, J., Clements, D., Doppelt-Azeroual, O., Erxleben, A., Freeberg, M. A., Gladman, S., Hoogstrate, Y., Hotz, H.-R., Houwaart, T., Jagtap, P., …, Larivière, D. (2018). Community-driven data analysis training for biology. *Cell Systems*, *6*(6), 752–758.e1.
66. Gray, A. J., Goble, C. A., & Jimenez, R. C., (2017). Bioschemas: From potato salad to protein annotation. In *International Semantic Web Conference (Posters, Demos & Industry Tracks)*. RWTH Aachen University.
67. Maier, W., Bray, S., van den Beek, M., Bouvier, D., Coraor, N., Miladi, M., Singh, B., De Argila, J. R., Baker, D., Roach, N., Gladman, S., Coppens, F., Martin, D. P., Lonie, A., Grüning, B., Pond, S. L. K., & Nekrutenko, A. (2021). Ready-to-use public infrastructure for global SARS-CoV-2 monitoring. *Nature Biotechnology*, *39*(10), 1178–1179. https://doi.org/10.1038/s41587-021-01069-1
68. Larivière, D., Abueg, L., Brajuka, N., Gallardo-Alba, C., Grüning, B., Ko, B. J., Ostrovsky, A., Palmada-Flores, M., Pickett, B. D., Rabbani, K., Antunes, A., Balacco, J. R., Chaisson, M. J. P., Cheng, H., Collins, J., Couture, M., Denisova, A., Fedrigo, O., Gallo, G. R., & Giani, A. M. (2024). Scalable, accessible and reproducible reference genome assembly and evaluation in galaxy. *Nature Biotechnology*, *42*(3), 367–370. https://doi.org/10.1038/s41587-023-02100-3
69. Bacon, W. A., Batut, B., Srikakulam, S. K., Zierep, P., Bretaudeau, A., Grüning, B., Le Corguillé, G., Hecht, H., Davis, J. Y., Hotz, H.-R., & Serrano-Solano, B. (2026). Ten common misconceptions about galaxy (and why they are wrong!). *PLOS Computational Biology*, *22*(2), e1013869. https://doi.org/10.1371/journal.pcbi.1013869
70. Marconato, L., Palla, G., Yamauchi, K. A., Virshup, I., Heidari, E., Treis, T., Vierdag, W.-M., Toth, M., Stockhaus, S., Shrestha, R. B., Rombaut, B., Pollaris, L., Lehner, L., Vöhringer, H., Kats, I., Saeys. Y., Saka, S. K., Huber, W., Gerstung, M., …, Moore, J. (2025). SpatialData: An open and universal data framework for spatial omics. *Nature Methods*, *22*(1), 58–62. https://doi.org/10.1038/s41592-024-02212-x

How to cite this article: Kostrykin, L., Massei, R., Chiang, D., Wollmann, T., Burel, J.-M., Tabernero, D. L., Sun, Y., Gao, Q., Paul, M. W., Videm, P., Watson, C., Franco-Barranco, D., Kumar, A., Goué, N., Koopaei, R. T., Ulman, V., Jum'ah, K., Poterlowicz, K., Etzrodt, M., … Serrano-Solano, B. (2026). Bioimage management and analysis in Galaxy: Tools, workflows, training, and community practices. *Journal of Microscopy*, 1–15. https://doi.org/10.1111/jmi.70165