PyReconstruct

A fully opensource, collaborative successor to *Reconstruct*.

Michael A. Chirillo*, Julian N. Falco*, Michael D. Musslewhite, Larry F. Lindsey, Kristen M. Harris

Center for Learning and Memory, The University of Texas at Austin, Austin, Texas 78712

*These authors contributed equally to this work.

Abstract

As the serial section community transitions to volume electron microscopy, tools are needed to balance rapid segmentation efforts with documenting the fine detail of structures that support cell function. New annotation applications should be accessible to users and meet the needs of the neuroscience and connectomics communities while also being useful across other disciplines. Issues not currently addressed by a single, modern annotation application include: 1) built-in curation systems with utilities for expert intervention to provide quality assurance, 2) integrated alignment features that allow for image registration on-the-fly as image flaws are discovered during annotation, 3) simplicity for non-specialists within and beyond the neuroscience community, 5) a system to store experimental meta-data with annotation data in a way that researchers remain masked regarding condition to avoid potential biases, 6) local management of large datasets, 7) fully open-source codebase allowing development of new tools, and more. Here, we present PyReconstruct, a modern successor to the Reconstruct annotation tool. PyReconstruct operates in a field-agnostic manner, runs on all major operating systems, breaks through legacy RAM limitations, features an intuitive and collaborative curation system, and employs a flexible and dynamic approach to image registration. It can be used to analyze, display, and publish experimental or connectomics data. PyReconstruct is suited for generating ground truth to implement in automated segmentation, outcomes of which can be returned to PyReconstruct for proofreading and quality control.

Significance statement

In neuroscience, the emerging field of connectomics has produced annotation tools for reconstruction that prioritize circuit connectivity across microns to centimeters and farther. Determining the strength of synapses forming the connections is crucial to understand function and requires quantification of their nanoscale dimensions and subcellular composition. PyReconstruct, successor to the early annotation tool Reconstruct, meets these requirements for synapses and other structures well beyond neuroscience. PyReconstruct lifts many restrictions of legacy Reconstruct and offers a user-friendly interface, integrated curation, dynamic alignment, nanoscale quantification, 3D visualization, and more. Extensive compatibility with third-party software provides access to the expanding tools from the connectomics and imaging communities.

Introduction

Reconstruct was released more than 20 years ago as one of the first open-source image annotation applications that could be run on a personal computer (1–3), also cite SynapseWeb download cite here. It offers a robust set of annotation features that are user-defined and field-agnostic. Since its release, additional annotation applications have followed, many written with connectomics-level analyses in mind (4–12). These modern applications have prioritized features that allow for rapid imaging and long-range segmentation of cell membranes to identify circuit connectivity (13–16). However, features needed for the more detailed annotation and curation of synapses and subcellular features in the context of circuit-level analyses are downplayed.

Despite the many compelling features offered by the alternatives, Reconstruct continues to be used by a large and diverse group of researchers from a variety of fields. Reconstruct has been used to produce data in hundreds of publications ranging from neuroscience (17–23) to plant biology (24, 25), entomology (26–28), human anatomy (29), herpetology (30), and materials science (31), to name a few. We attribute its widespread use in part to Reconstruct's simple interface and flexible approach to annotation. Image stacks are imported into the application, and with little effort, researchers can quickly begin annotating serial sections, export data for analysis, and produce publication-quality 3D visualizations, all in a single intuitive application.

Despite significant advances in EM annotation tools, several fundamental challenges persist. Many require considerable programming expertise to access full annotation details and advanced features. This requirement limits accessibility to technical specialists rather than serving the broader research community. Most annotation tools lack robust, integrated systems for quality control and data curation. Many do not accommodate post-hoc realignment of serial sections, making it difficult to correct alignment errors discovered during annotation or analysis. Finally, most tools separate meta-data from image data, forcing researchers to maintain parallel data management systems and complicating long-term data preservation. These limitations can significantly impede annotation workflows and compromise data integrity, particularly in large-scale projects.

Reconstruct has long been overdue for a major overhaul. It only ran natively on Windows machines, datasets were RAM-restricted in size, a complex system of transformations hampered new alignment strategies, and minimal interaction with outside tools restricted its expansion. Limitations like these required users to implement kludgy workarounds that complicated their workflows. Despite these limitations, Reconstruct continues to be used for its simplicity and efficiency.

With these limitations in mind, we have produced a generalizable and streamlined annotation tool in the spirit of Reconstruct that integrates readily into any serial imaging pipeline (Fig. 1). We have named this new system PyReconstruct, a modern, user-friendly successor to what we now refer to as "legacy" Reconstruct. PyReconstruct was written in Python, retaining legacy Reconstruct's intuitive interface and addressing the shortcomings necessary to transition from serial section to the volume electron microscopy needed for circuit analyses. It features an integrated curation system for maintaining data quality, flexible alignment capabilities,

and unified data storage keeping meta-data and image data together. PyReconstruct runs on most operating systems, employs lightweight data structures, implements dynamic alignment, and provides multi-resolution scaling. It lifts RAM restrictions allowing users to scale up to large volumes. PyReconstruct is also backward compatible with data annotated in legacy Reconstruct, giving legacy users the ability to port their data into the modern annotation environment. Importantly, PyReconstruct is fully opensource and users are at liberty to customize the source code to their own use cases.

User experience

We designed the core functionalities and features of PyReconstruct to facilitate working with serial images in an unopinionated manner, to maintain data integrity, and to promote a collaborative workflow. In the sections below, we describe the annotation and segmentation interface, object organization capabilities, curation tools, alignment systems, and data visualization features.

Segmentation and object organization

PyReconstruct offers an uncomplicated graphical interface for registration, segmentation, and meshing of image data and allows for multiple points of entry and exit to and from third-party software (Fig. 2). By leveraging Python wrappers for openCV (4.8.1), stacks of serial 2D images can be loaded into PyReconstruct in multiple formats (.tiff, .jpeg, .png, etc.). Though PyReconstruct directly imports standard images without conversion, users can also benefit from its support of precomputed volume formats, the advantages of which are outlined below.

The user's primary workspace contains a single image from the series (the "section") displayed in the main window (Fig. 3A). The trace palette, tool bar, section navigator, and brightness/contrast sliders are superimposed over the image and are moveable around the field, such that the user can personalize their workspace as they annotate. Detachable windows display quantitative 2D (the trace list) and 3D annotation data (the object list) as it is being collected in the main window (Fig. 3B). Users are at liberty to choose trace name, color, tag, and appearance. The trace list allows users to view and edit traces on a section and see quantitative measurements, while the object list allows users to view and edit object data and see quantitative measurements in 3D based on the calibrated section thickness (2).

In PyReconstruct, 3D "objects" represent collections of 2D "contours" over the image stack that belong to a single feature in the series. An object's contour on a single section may consist of one or several separate "traces" (Fig. 4A). Segmentation data in PyReconstruct is stored as lists of x and y positions that make up individual traces, and objects can be organized and classified in a number of ways. Hierarchical and nested objects groups can be created by assigning "hosts" that point to other objects. For example, "synapse" may point to "dendrite" and/or "axon". Users can also create custom categories to assign objects qualitative measures. For example, "synapse" from the above example could also be categorized by "synapse type" as

"excitatory" or "inhibitory". For cases that fall outside host-inhabitant relationships and categorical variables, users can create simple, custom-named groups. Because these classification schemes are customizable, users can define semantically meaningful groups and relationships that are specific to their use-cases.

Rich, free-form annotations and event tracking

PyReconstruct allows for rich annotations that might otherwise be stored externally and risk being separated from the original series. Annotations that provide meta-data and context to series features can be applied to individual traces (tags), entire objects (groups and comments), and points of interest in the field (flags) (Fig. 4B-C). For example, traces denoting object profiles distorted by image artifacts might be tagged to indicate they have been interpolated (or inferred by the user) from traces on adjacent intact sections. Users may flag objects in a series that serve as probands or objects the annotator wishes to revisit later. Objects can be grouped to track their use in specific publications or projects, while comments provide additional explanatory notes. Free-form annotations like these enable users to create and customize their own annotation system based on project needs.

Experimentalists often revisit previously annotated series to test new hypotheses, and keeping track of the history of actions performed in a series may serve to guide subsequent annotation. In the past, this information was stored externally. PyReconstruct provides automation by logging all actions that add new or modify existing annotation data logged. Log entries contain brief descriptions of annotation events, including date, time, user, objects edited, and the section or sections on which the action was performed. Users can view and filter the full log to display entries relevant to one or more objects.

Collaborative curation for quality control

Annotating large amounts of serial section data is generally performed among teams of annotators and PyReconstruct was developed with this approach in mind. Annotations performed by a team member are subsequently passed to a more expert annotator to be proofread and curated for quality assurance (Fig. 5). To this end, PyReconstruct includes a simple curation system, and users can be assigned curation tasks through the object list (Fig. 6A). Curation history (user, date assigned, completion status, etc.) is tracked with the series, so that project leads can verify if annotation criteria is applied uniformly among team members and correct tracing errors. Curation data and history is readily available and displayed through the object list. Free-form annotations (see Fig. 4B-C) provide further context to aid collaboration among annotation teams. Field flags also include a running commentary that can be used to alert present and future annotators to information about locations in the series (Fig. 6B). Flag meta-data including the creation time, username, and related notes are tracked and stored with series data. A conversation-style comment system allows users to maintain a running discussion for each flagged item. Tags applied to traces on a particular section can be accessed in a trace list (Fig. 6C), which when filtered and sorted, provides a rapid means to identify the status of a trace.

PyReconstruct runs on a local machine and does not currently implement a client-server solution to allow multiple users to work simultaneously on a single series. Trace conflicts might therefore arise when series annotated simultaneously by different users are subsequently merged. For example, two annotators tracing the same object will lead to non-identical, overlapping traces representing the same feature on a section. Merging large, heavily annotated series can result in many duplicate traces that pollute the workspace and lead to quantification errors.

To deal with this problem, a trace import system has been implemented in PyReconstruct that identifies and alerts users to merge conflicts that must be resolved. Conflicting traces are determined based on the Jaccard similarity coefficient, trace history, and annotator preference. Users can determine an overlap threshold above which traces are considered “identical enough”, which gives PyReconstruct leeway to discard all but one overlapping trace. This provides some automation to the tedious tasks of curating multiple traces for a single feature. Remaining conflicts are flagged for resolution, and users can walk through each conflict systematically.

Dynamic alignments applied on demand

During processing, sectioning, and imaging, planar specimens are subjected to a variety of physical alterations (e.g., compression, shrinkage, heating) and technical complications (e.g., stage and specimen drift, and non-uniform scan and lens characteristics) that introduce linear and non-linear distortion into the final images (37–39). Image distortion can be mitigated through choice of processing protocol (40–43), imaging modality (44, 45), and post-hoc computational processing, but in many cases must ultimately be corrected through image registration. Registration methods are rapidly improving (46–57), and choice in strategy depends on distortion type and personal preference. We therefore sought to implement a flexible and dynamic registration strategy to allow users 1) to import alignments generated externally, 2) to perform simple registration inside PyReconstruct, and 3) to store multiple alignments that can be applied to stacks of images on the fly (Fig. 7).

An “alignment” in PyReconstruct is represented by a list of affine transformations (Fig. 7A). Each section in the series is associated with a single affine transformation, which is stored as a set of six numbers that correspond to the first two rows of the 2D transformation matrix. This single transformation is applied to both image pixels and contour data of a particular section, which are then displayed in the field when the section is called by the user. In this way, users whose images require affine transformations alone need not work with images whose alignments have been previously “baked in” (i.e., transformations applied directly to stored images).

This strategy offers several advantages. First, the only image data stored with a project are the unaligned images that were initially produced at the microscope and imported into PyReconstruct, dramatically reducing the size of a project. (Users can still of course use third-party software like Fiji, TrakEM2, and SWiFT-IR to import images with baked-in alignments and perform additional local alignments.) Second, stored contour points reflect the x and y position of the points on the unaligned images (not the transformed images), making it possible for users to apply external transformations to those points when exporting contour data from

PyReconstruct. Third, users can rapidly tweak a section's alignment should they note alignment flaws while annotating, which is a common occurrence when working with serial images.

Finally, the simplicity of this strategy means users are at liberty to store multiple alignment profiles specific to an object, region of interest, or context, which can be applied to the series on demand (Fig. 7B). This is especially useful when annotating objects in areas of section deformation, for example, on either side of image artifacts (breaks, folds, tears, etc.), which are frequent in serial section EM. These artifacts would otherwise require users to use third-party image editing software to correct. Thus, a complete history of alignments for each series is stored. Objects in the series can be assigned to specific alignments, ensuring the object's quantitative measurements and 3D reconstructions are accurate.

PyReconstruct supports simple manual alignment natively and users can interactively translate, rotate, scale, and shear images in the main window. Users can also directly edit a section's affine transformation through the menu bar (Fig. 7C). Affine transformations can be estimated by identifying fiducial markers in the field, such as cell structures that span multiple sections. Transformations performed on single sections can be applied throughout the remaining sections, allowing for a correction in alignment to be propagated throughout the series. PyReconstruct supports importing alignments from other PyReconstruct projects as well as from external applications, for example, directly from SWiFT-IR project files (33, 56) or in the form of simple .txt files, each line representing an affine transformation.

Data visualization and analysis

PyReconstruct's object list represents a serialized view of objects, their measurements, and meta-data, and allows users to export data for analysis and render objects in 3D (Fig. 8). 2D contours are voxelized and represented as NumPy arrays, which are translated through a matrix-to-marching-cubes algorithm provided by the trimesh library (58) to generate watertight triangle meshes for visualization and quantification (Fig. 9A). The trimesh library provides a number of in-place smoothing filters (59, 60) that can be applied to meshes in PyReconstruct. 3D meshes are visualized in a customized 3D scene built on top of the Python scientific visualization library vedo (61). The meshing strategies employed natively in PyReconstruct are modularized, such that users proficient in Python can implement more complex meshing algorithms accessible from the user interface should they choose to do so. 3D meshes can be exported from the object list in several formats (obj, ply, stl, etc.), which can be imported into external 3D modeling software, such as Blender, for further editing.

Segmentation issues are often not apparent when viewing objects as 2D profiles. Misalignments, poor segmentation, and misplaced traces become more evident when objects are rendered in 3D, and annotation quality is improved when users switch between 2D and 3D views while working. PyReconstruct's 3D scene includes several features to facilitate these actions. Double-clicking on a point in the 3D scene focuses on that point in the main window, providing a rapid way to identify 2D locations from the 3D space (Fig. 9B). Objects rendered in the 3D scene can be translated and rotated, allowing the user to customize the 3D view, for example, to visualize all objects of interest in a single glance (Fig. 9C). Objects from other series can also be

imported into the current 3D scene. Customized 3D views can be saved as "scenes" that store object attribute and location information, which can be opened across annotation sessions. Scenes, in addition to aiding users in annotation, mitigate the need to export meshes to external 3D modeling software when making simple figures for publications.

Under the hood

Interface framework, data structures, and legacy compatibility

PyReconstruct's user interface is powered by Python bindings for Qt6, an opensource and platform-independent framework for developing graphical user interfaces (62). Users can therefore run PyReconstruct on most major modern operating systems (Windows, macOS, and many Linux distros) right out of the box.

Annotation data in legacy Reconstruct was stored over multiple "trace files", one for each section, making moving series data between computers and users cumbersome and prone to data loss. Storing annotation data in a single, locally stored file facilitates data sharing amongst users. Annotation data in PyReconstruct is collected and stored in a single JSON-structured file with the extension .jser (portmanteau of "JSON" and "series", pronounced "jay-sir", Fig. 10). Image data are stored separately in a format and location defined by the user.

The jser file is divided into data pertaining to individual sections (section-specific data) and data pertaining to the entire series (series-specific data). Section-specific data contain contour information and annotations, transformations, and the section's image attributes (magnification, brightness, contrast, and image file pointer). Series-specific data store information that spans multiple sections, including, for example, series meta-data, object attributes, current alignment, user information, and user preferences. All annotation data is accessible graphically through the application and through PyReconstruct's Python API.

PyReconstruct implements a composite file pattern approach when loading and saving data: Data stored in the single jser file is decomposed at startup into discrete temporary hidden files that hold information pertaining to individual sections and the series. This approach optimizes RAM utilization while processing contour data and provides fault tolerance should the program unexpectedly close, or the user's system fails.

Labs from a variety of fields continue to employ legacy Reconstruct as the primary tool used to annotate serial section EM data manually due in part to its ease of use. We therefore sought to make porting data from legacy Reconstruct into PyReconstruct simple. PyReconstruct is backward compatible with legacy structures: Series that were previously annotated in Reconstruct and saved as XML can be loaded into PyReconstruct by simply pointing to the legacy structures. No external conversion is necessary. Importantly, this process is bidirectional: Annotation data collected in PyReconstruct can also be exported as legacy XML structures in case users have established workflows that rely on legacy structures remaining intact, for example, with neuropil tools (36).

Features only compatible with PyReconstruct are saved in the jser file. Users therefore retain full access to the legacy application should they choose to do so and can still benefit from PyReconstruct's modern annotation features.

Lifting RAM restrictions through multi-resolution scaling

Datasets loaded in legacy Reconstruct could exceed RAM; however, two sections were loaded into memory simultaneously. This meant large images needed to be cropped to a size compatible with the user's RAM specs. Modern, ultra-large data viewers typically solve this issue by employing data structures that store chunked, compressed N-dimensional arrays at multiple resolution (or "scales"), such as with HDF5 or Zarr. Precomputed image datasets can be manipulated through a variety of wrappers such as h5py (63), PyTables (64), Z5 (65), and zarr-python (66). Instead of mapping entire images from disk into RAM, moving compressed chunks of data dramatically reduces loading times and makes navigating serial sections a much quicker and more pleasant experience. Saving the image data at multiple resolutions further reduces the amount of data loaded when users view low magnification views of a section.

To lift legacy Reconstruct's RAM restrictions, we have implemented a multi-resolution Zarr approach to loading and viewing image data in PyReconstruct. Zarr was chosen as it was designed specifically for Python, supports robust multi-threading, provides access to customizable compressor and filter classes, and is used extensively in machine-assisted segmentation. Unlike applications that consolidate serial sections into a volumetric, 3D Zarr, PyReconstruct implements a 2D layer approach: each section being stored as a separate annotated group, which allows users to switch between the original images and Zarr at will.

Storing and retrieving image data in a chunk-wise manner means PyReconstruct users are in practice no longer RAM-limited when annotating series. Users are now virtually unbounded by series size, restricted only by the disk space available on their local machines, which can be augmented with external and remote storage. This strategy permits near-instantaneous loading of images in the field, as only chunks at the lowest possible resolution based on zoom are loaded during viewing. PyReconstruct maintains its flexible alignment strategy by applying affine transformations non-destructively, only after image data is loaded. A multi-resolution Zarr approach does increase the overall size of data stored per project, but PyReconstruct's scaling strategy means Zarr datasets are never heavier than 1.33 times the size of the original image stack.

Interacting with external large-image applications and datasets

Volumetric datasets are much larger than they used to be, even recently (67–72). This is due in part to the transition from volumes constructed from small-field TEM images to those from high-throughput methods such as focused ion beam SEM (73, 74), serial block face SEM (75, 76), and tape and array methods (77–80). Relying on manual segmentation alone is therefore increasingly unfeasible. However, many labs (including our own) continue to employ manual segmentation as a primary technique. We therefore sought to provide a

simple platform that annotators can use to begin to interact with the many automated segmentation tools being developed by the connectomics community.

Neuroglancer is a WebGL-based application for viewing, annotating, and querying large, multi-resolution volumetric data (81) and is the application of choice in several connectomics projects. Porting data to and from Neuroglancer offers one path through which automated segmentation efforts can be tapped into. PyReconstruct provides a graphical interface where users can export image data and contours as labeled Zarrs compatible with Neuroglancer. Data annotated in Neuroglancer can also be imported and converted to contour data for use in PyReconstruct.

Discussion

PyReconstruct provides solutions that enhance manual segmentation of serial images in a user-friendly and familiar interface and offers several advantages over legacy Reconstruct. The target audience includes researchers from a variety of largely unrelated fields, who rely on manual annotating serial sections and volume EM data. Thus, PyReconsruct is a non-specialized tool that offers platform-independence, simplified data structures, a unique alignment strategy, and functionality to interact with a variety of external tools.

Integrated data quality control is often overlooked in annotation software, requiring users to develop their own external tracking systems. PyReconstruct has a robust and built-in curation system that addresses this critical gap. By embedding curation data directly within project files, PyReconstruct provides a streamed approach to team coordination and data integrity. PyReconstruct is fully opensource under a GNU General Public License v3 and users are encouraged to scrutinize and improve its source code, which is publicly available at GitHub (https://github.com/synapseweb/pyreconstruct). Opensource tools like PyReconstruct empower users to create custom solutions that can be integrated into the main product to the benefit the broader scientific community. This open environment also makes it possible to develop data conversion strategies that will allow PyReconstruct to accept annotations and segmentations from other systems.

As our understanding of alignment evolves, image registration strategies are advancing (46–57). The ability to re-align a series means users can adapt and refine their work at all stages, rather than being constrained to a single baked-in alignment. Image flaws resulting in alignment errors are often not noticed until the annotation stage, when users are viewing sections at higher magnifications. The ability to tweak existing alignments rapidly and create new alignments when required therefore facilitates annotation and improves the quality of the 3D outcomes. In PyReconstruct, users are at liberty to align the series, import alignments from external software, and store multiple alignments that can be switched to and from on demand. Importantly, this strategy enables users to employ simple, non-elastic alignments in series that suffer from image artifacts that warp portions of a section (see Fig. 7). Multiple alignment stacks for a single series and the ability to realign on demand also provides ground truth for developing more sophisticated automatic alignment methods.

Simplicity and generalizability come with tradeoffs. Fully segmenting, annotating, and curating EM volumes is time-consuming. For example, the complete manual annotation of a single, roughly 180 μm^3 volume of hippocampal tissue—an exceedingly small volume by today's connectomics standards—took a team of expert annotators more than a year to complete and curate using legacy Reconstruct (82). The process of scaling up is therefore impractical (if not impossible) when annotators rely on manual segmentation alone. The push by many teams to produce semi- and fully automated segmentation pipelines has led to the segmentation of breathtakingly large datasets (67–69, 71). Nevertheless, these pipelines remain largely relegated to specialist communities. Therefore, a need remains to adopt automated segmentation approaches in simple user interfaces for non-specialist users.

One key challenge in adopting automated segmentation routines lies in balancing technical sophistication with accessibility. Modern automated segmentation tools require computational infrastructure and expertise in programming. The benefits of expanding access to these tools are immense. Moving forward, the developers of PyReconstruct are evolving its API and plugin framework to support the embedding of customized automated segmentation features into the software. This modular approach will enable researchers to leverage cutting-edge routines without the need for deep technical expertise, bridging the gap between specialist and non-specialist communities.

The evolution of imaging modalities over the past several decades has led to microscopy datasets that are larger than they have ever been (67–72, 83). These massive datasets have predominantly required server-based storage. Annotation tools that have grown up around these projects have therefore prioritized this approach (4, 9, 10). At present, PyReconstruct is a fully standalone application without server-client dependencies. Its users are not dependent on an internet connection, server maintenance, or remote data stores. While this approach offers simplicity and autonomy, users may need to generate manageable subvolumes from their data, especially when working with extremely large datasets.

Previously, researchers faced the onerous task of manually cropping individual serial images. Today, the widespread adoption of chunk-wise data storage formats and the increasing availability of standardized, open-source databases for large image data (84–88) have significantly mitigated these limitations. Users can efficiently extract subvolumes from remotely stored datasets through Python libraries like CloudVolume, which offer APIs to access precomputed volumes (89). These tools can be readily integrated into a workflow to sample regions from larger volumes and use the suite of annotation, curation, quantification, and visualization features to produce high quality data output from PyReconstruct.

Funding

This work was supported by NIH Grants 1R56MH139176-01 and NSF Grants 1707356, 2014862, and 2219894 to K.M.H.

Acknowledgements

We would like to thank Harris lab members, GitHub users, members of our regular PyReconstruct user meetings, and students in NEU466G at the University of Texas at Austin for beta testing PyReconstruct and for their suggestions. Their input was vital in shaping the interface. We also would like to thank James Carson at the Texas Advanced Computing Center (TACC) for his help in integrating PyReconstruct with TACC's online access points and Vijay Venu Thiyagarajan for his suggestions regarding data structures.

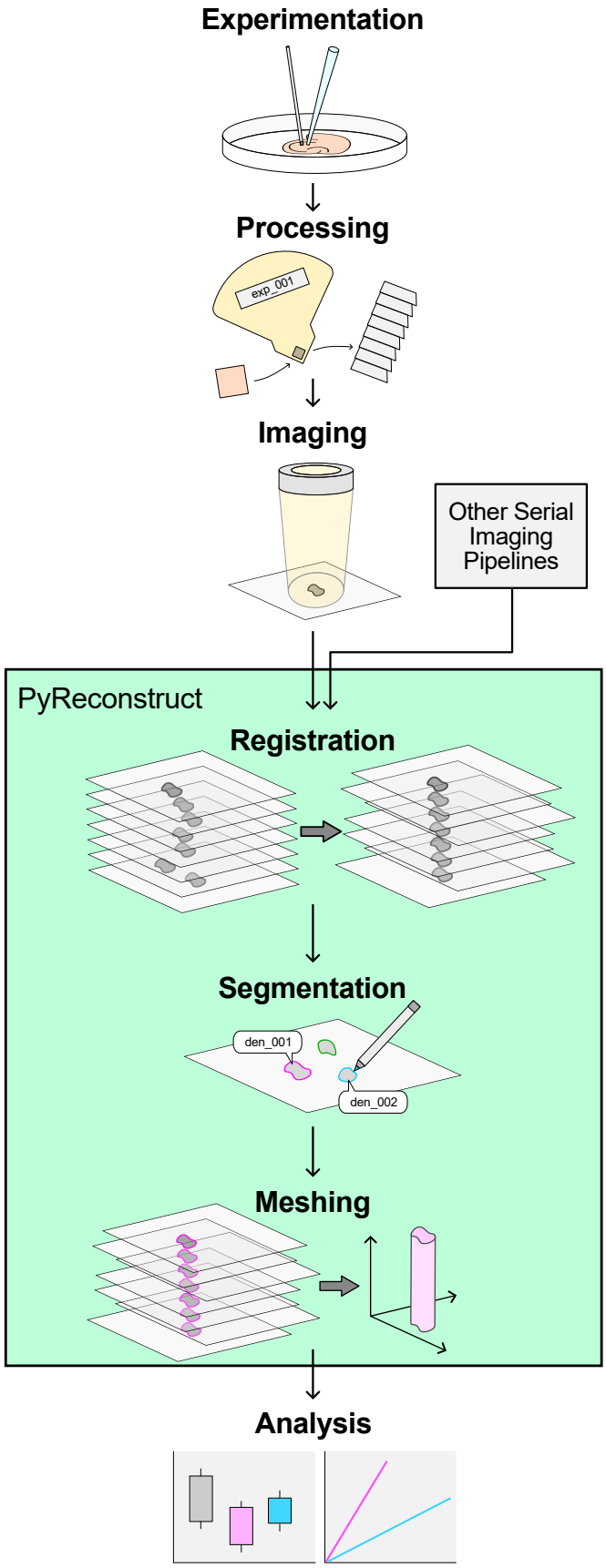


Figure 1. Example serial imaging pipeline from experimentation through data analysis. PyReconstruct encompasses the most time-consuming steps (green box) in pipelines that rely on manual segmentation or proof reading of serial images to produce analyzable data. Here, a serial section EM pipeline is shown. Image registration, segmentation, and contour meshing can be performed in a single application. Images from any source can be used in this pipeline, including series or single sections.

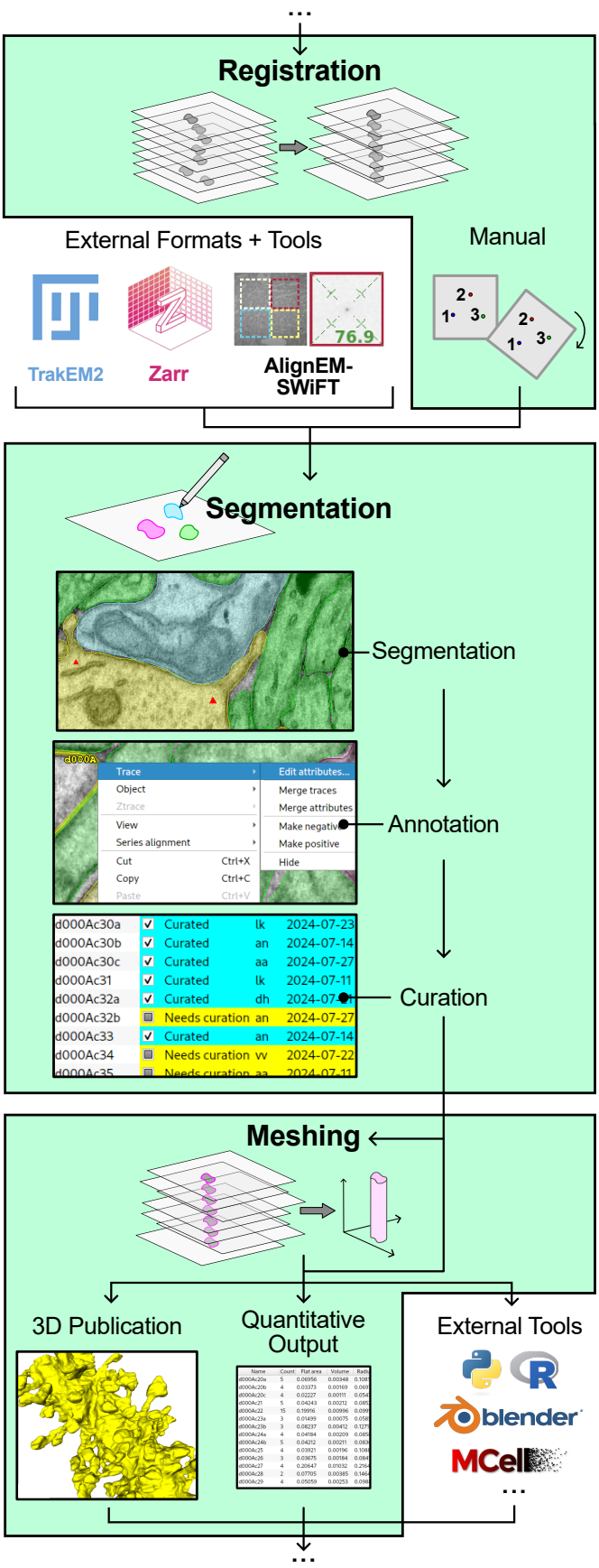


Figure 2. PyReconstruct readily interacts with third-party software. Core PyReconstruct functions are highlighted in green and include native manual registration through the application interface, while also supporting importation of pre-registered images and transformation matrices (6, 32, 33). PyReconstruct stores detailed segmentation data as 2D contours and features a checklist-based curation system for quality assurance. Users can generate, view, and export quantitative data and meshes directly in PyReconstruct to create publication-ready figures or for use externally in programming languages such as Python and R. Exported meshes can be further edited in advanced 3D platforms like Blender for biophysical modeling using programs like CellBlender and Mcell (34-36).

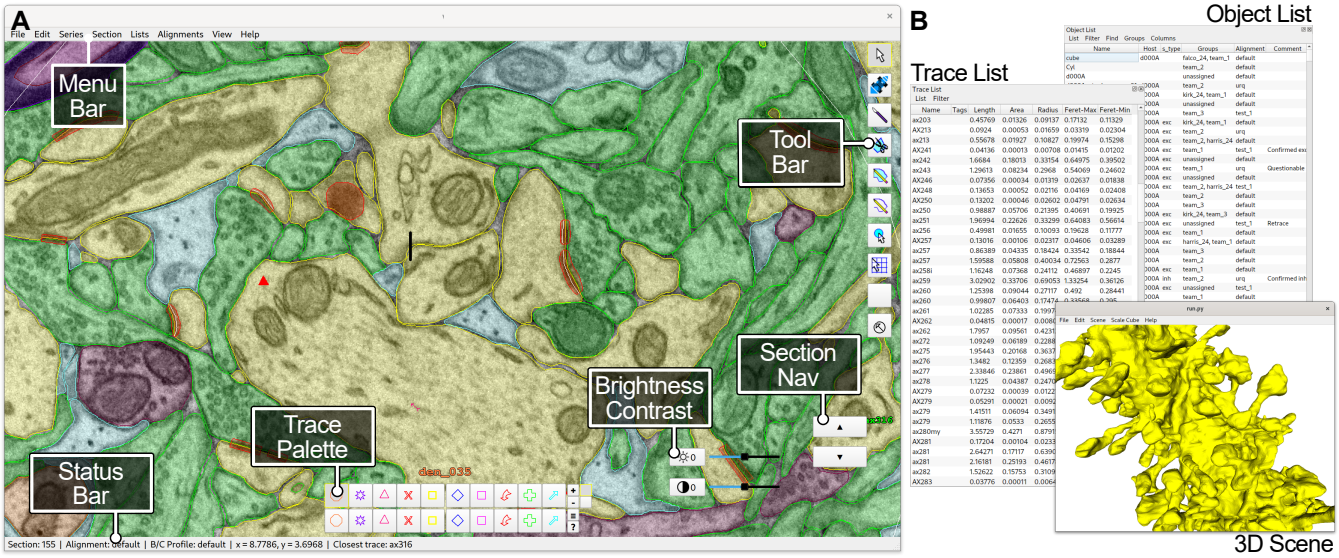


Figure 3. Windowing schema in PyReconstruct. (**A**) The main window contains images and widgets (tools, trace palette, etc.) that can be moved around the frame and toggled on and off, allowing users to customize their primary workspace. (**B**) Users interact with 2D and 3D data as it is being collected through ancillary, detachable windows, such as the trace list, object list, and 3D scene.

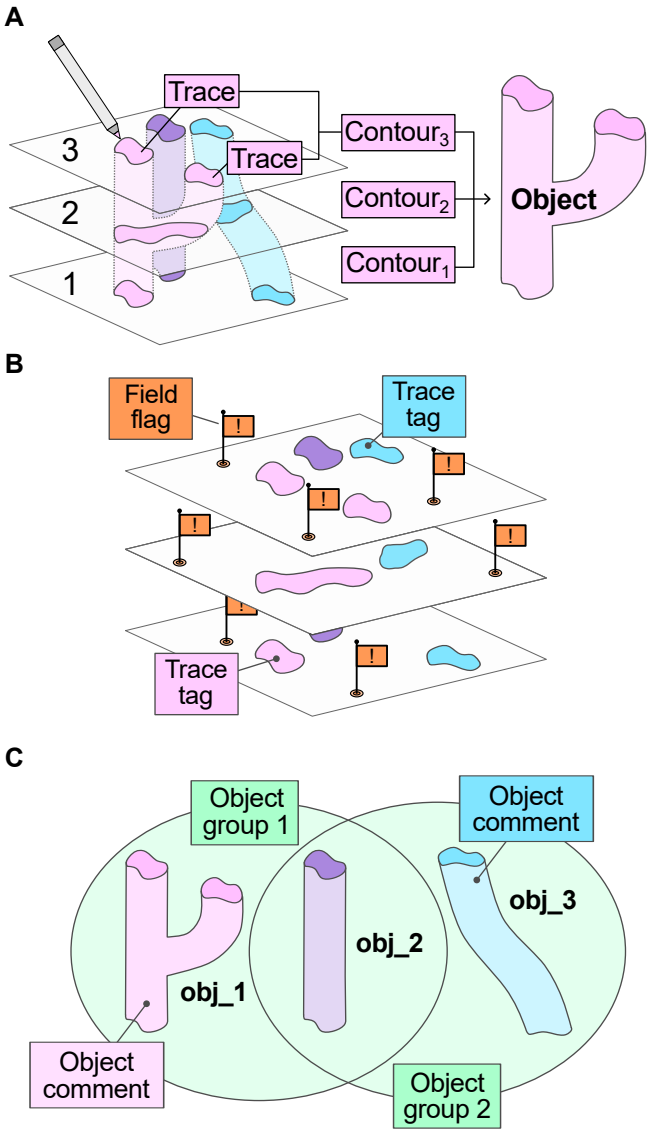


Figure 4. PyReconstruct stores both free-form and hierarchically organized annotation data. (A) Object profiles on individual sections are identified and their outlines traced by annotators. An object's 2D contour consists of all traces on a section belonging to the object. 3D objects are constructed from multiple contours across serial images. (**B**) A trace tagging system allows users to ascribe additional information concerning 2D segmentations. (**C**) Annotations applied to objects include groups and comments, which can be accessed from the object list.

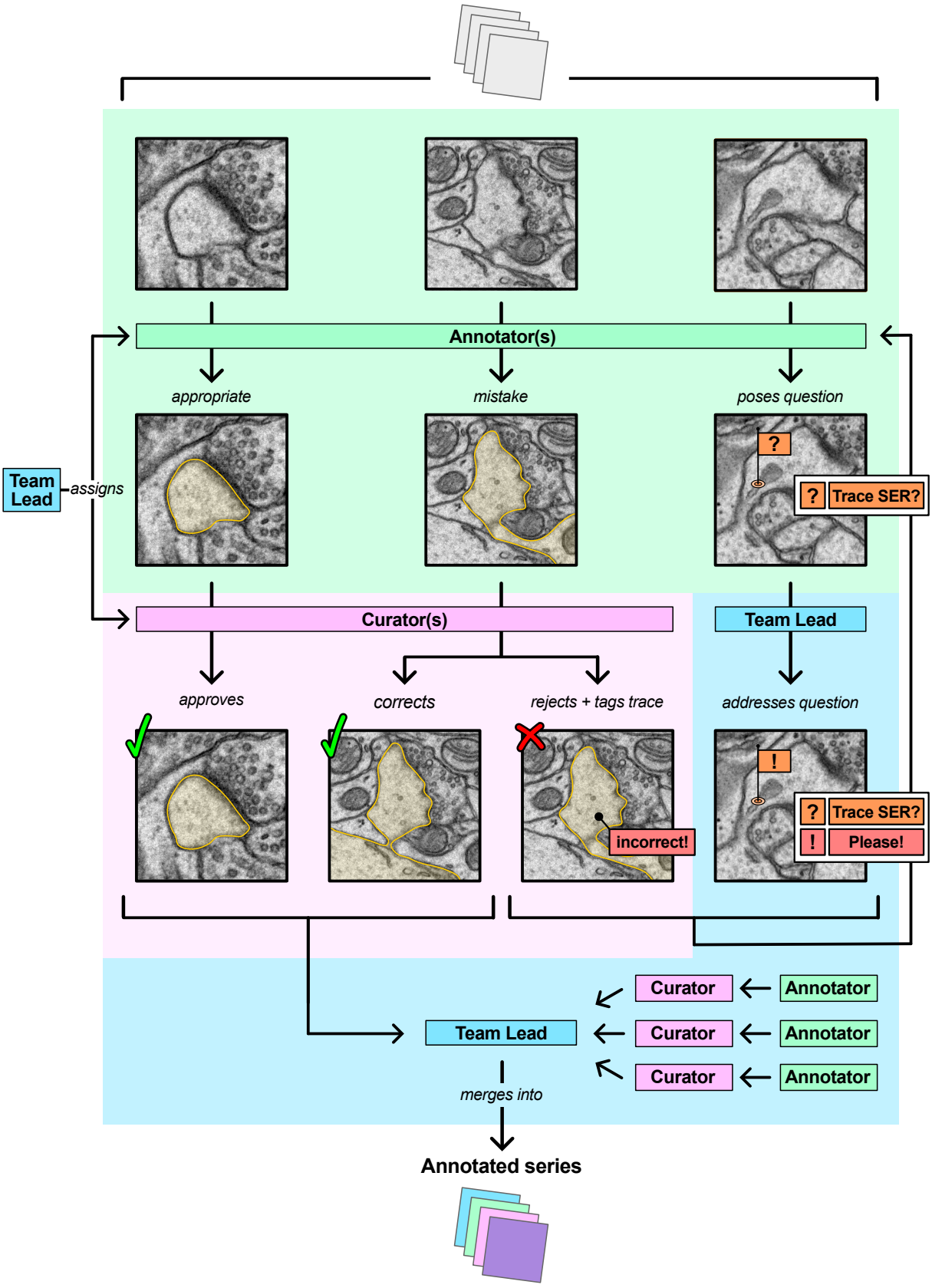


Figure 5. Implementation of a team-based annotation and curation scheme to ensure quality control. A team lead (blue) assigns annotation (green) and curation tasks (pink) to team members working simultaneously on copies of a single series. Curators approve, correct, or reject traces, which can be tagged for review. Annotators can pose questions to team leaders through a system of flags with commentary that can be answered and returned. As annotations are curated, the partially annotated series can be passed back to the team lead. Curated series are merged into a single annotated dataset for analysis. PyReconstruct provides features at multiple steps in this pathway to facilitate curation and quality control.

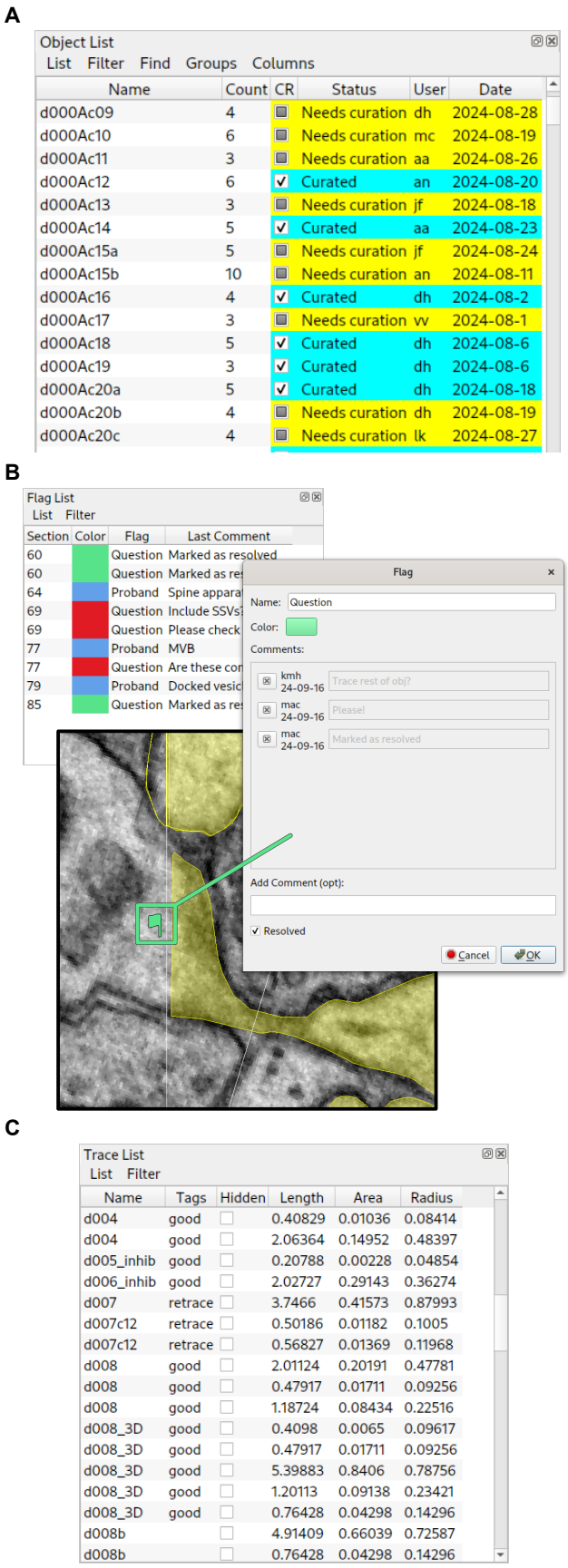
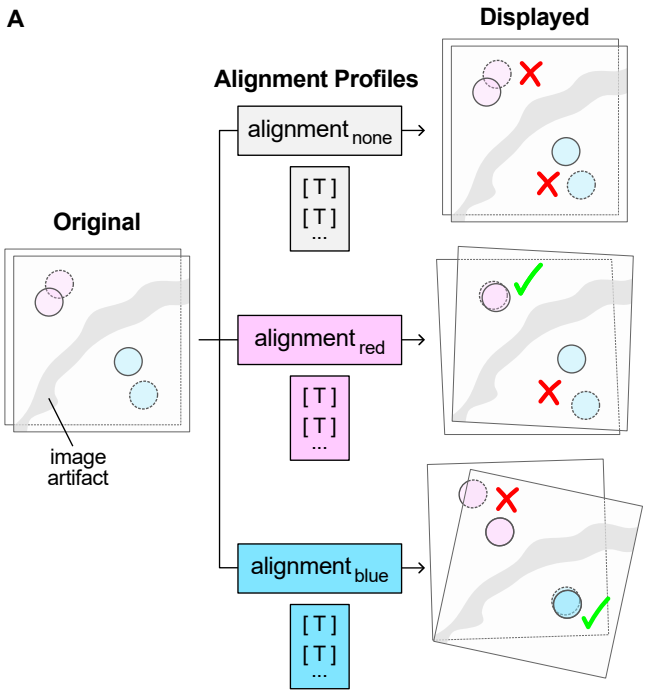
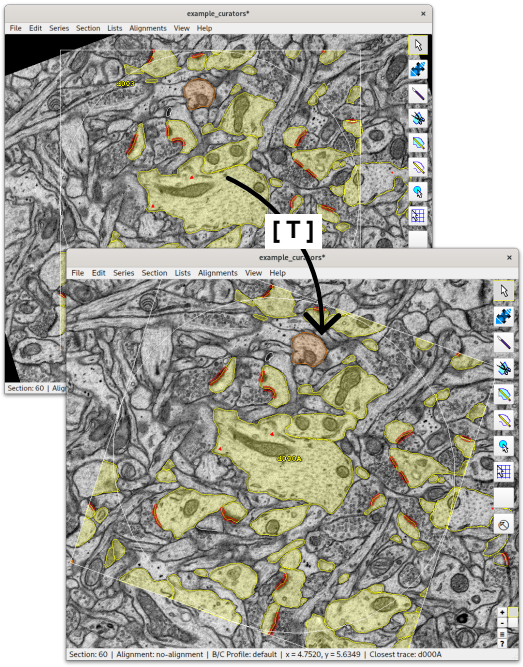


Figure 6. PyReconstruct provides automation for curating segmentation data. (**A**) Tracking of curation tasks is stored with series data in PyReconstruct. Curation tasks and status can be assigned interactively to objects from the object list. Multiple filtering options allow users to view pending and completed curation tasks assigned to them and to others. Here, users were randomized to curation tasks through PyReconstruct's Python API. (**B**) Flags placed in the field are shown in a flag list, which can be sorted in various ways. Running commentary allows team members to pose questions and discuss issues that arise during annotation. (**C**) Traces can be tagged with information that is accessed from a sortable trace list.





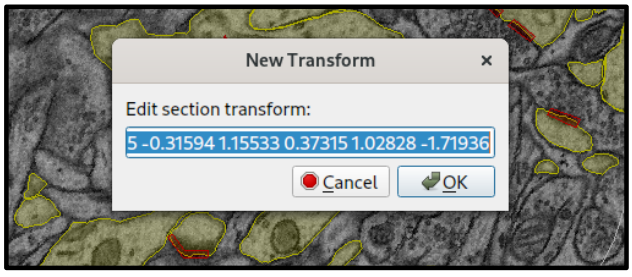


Figure 7. Multiple alignment profiles applied to serial sections on demand. (**A**) PyReconstruct stores image and contour data in their untransformed state. Alignments are represented as lists of affine transformations ([T]), one for each section. Multiple alignments are stored as profiles that can be applied independently on demand. Image artifacts (resulting from tears, folds, etc. in serial sections) often warp serial sections and produce local misalignments. The ability to tweak, create, and store multiple alignments facilitates annotation. (**B**) Alignment profiles can be accessed and modified through the user interface. The current alignment is displayed by name in the main window's status bar (bottom of main window, see Fig. 3A). (**C**) New alignments can be created in PyReconstruct and imported from external applications or section transformations can be edited directly in the interface, allowing users fine control over series alignments.

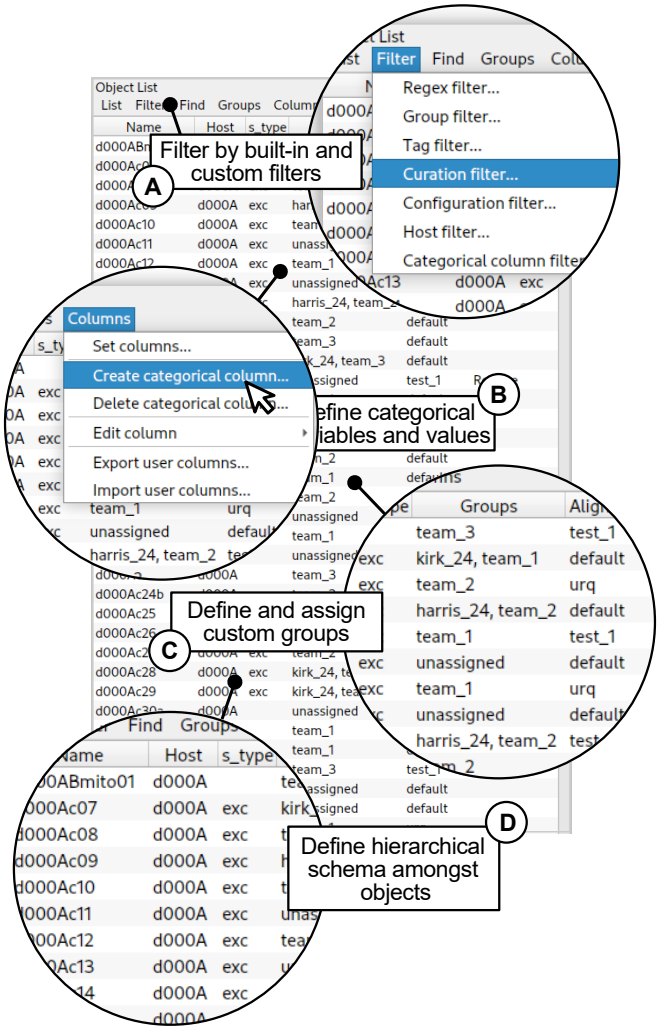


Figure 8. Three-dimensional data is stored in the object list. Data concerning 3D objects constructed from 2D contours is displayed in the object list. Users can sort by built-in and custom filters (**A**), define categorical variables and assign values (**B**), define and assign custom groups (**C**), and define hierarchical schema amongst objects by assigning objects as hosts (**D**).

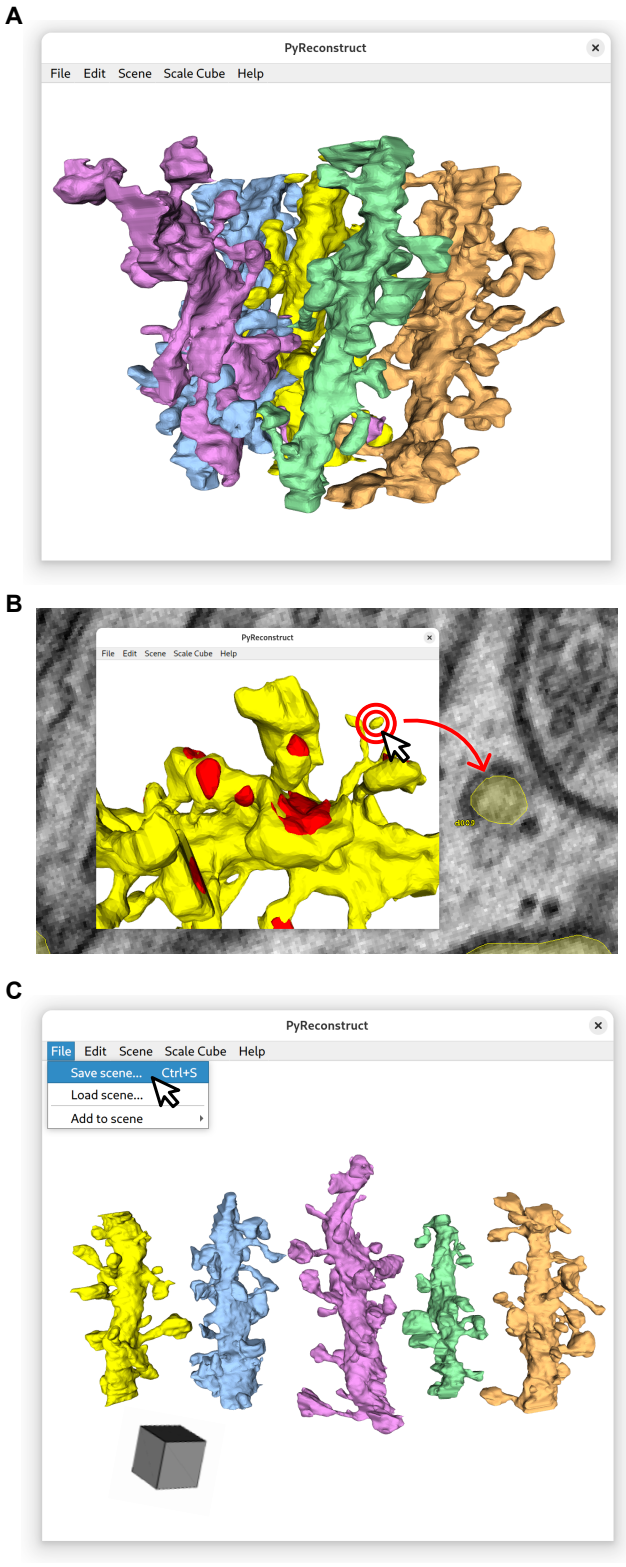


Figure 9. An enhanced 3D scene facilitates detailed curation and stores publication-ready visualizations. (**A**) 3D renderings of objects can be generated from the object list and are displayed in a featureful 3D scene. (**B**) Double-clicking in the 3D scene focuses on the corresponding 2D location in the main window, allowing users to navigate to sections for curation. Here, a disconnected part of an object identified in the 3D scene is rapidly uncovered on the 2D section, allowing users to curate more finely the object's segmentations. (**C**) The updated 3D scene offers users features that organize objects in an automated fashion. Here, all objects in the 3D scene have been organized along a single axis, providing a visual overview of all meshes. Objects can be moved and rotated interactively and their attributes (color, opacity, etc.) can be altered. Scenes can be exported, saved, and re-loaded at a later time, allowing users to produce publication-ready figures and store scenes with series data.

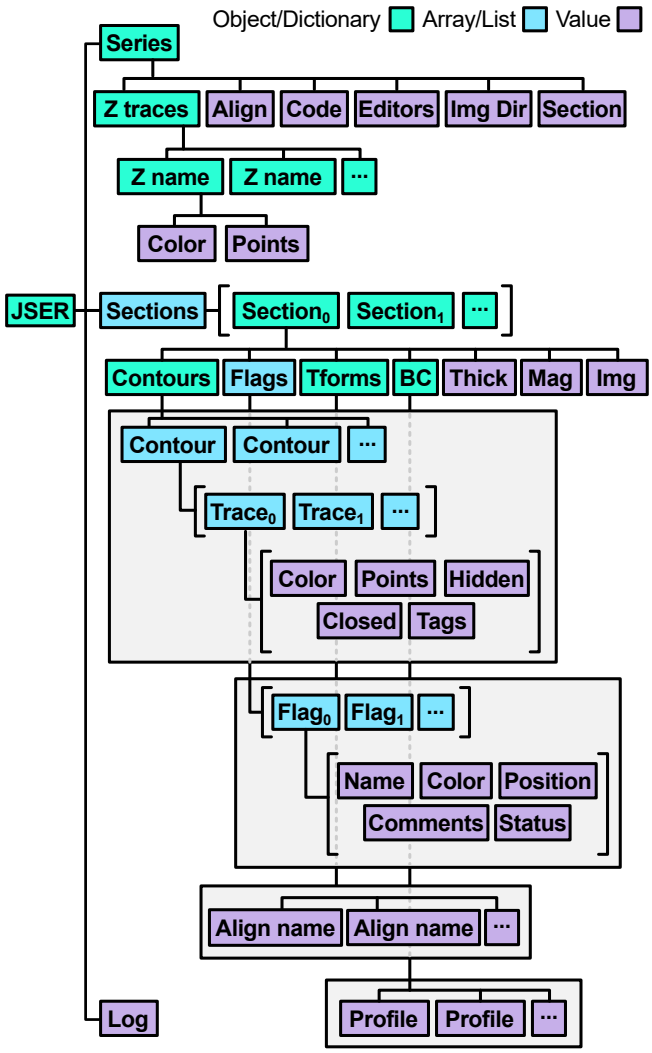


Figure 10. Schematic of PyReconstruct annotation data. Whereas legacy Reconstruct stored annotation data as multiple, non-canonical XML files, PyReconstruct stores annotation data locally as a single JSON-structured file (suffixed with .jser) loaded as a Python dictionary and edited using built-in modules or through PyReconstruct's API. Series-wide data is held in a "Series" key, section data in a "Section key", and series history in a "Log" key. The top-level key "Sections" maps to a list of section dictionaries with keys representing contour, flag, transformation data, etc. "Contours" is associated with a list of contours, which hold trace information. Traces are lists of trace attributes, such as trace color, points, visibility, etc. (Abbreviations: *Align* alignment, *Code* series code, *Img Dir* image directory, *Tforms* transforms, *BC* brightness/contrast profiles, *Thick* thickness, *Mag* pixel magnification, *Img* image file path.)

References

1. J. C. Fiala, Reconstruct: A free editor for serial section microscopy. *J microsc* **218**, 52–61 (2005).

2. J. C. Fiala, K. M. Harris, Extending Unbiased Stereology of Brain Ultrastructure to Three-dimensional Volumes. *Journal of the american medical informatics association* **8**, 1–16 (2001).

3. Harris Lab, SynapseWeb Software. (2025). Available at: https://synapseweb.clm.utexas.edu/software [Accessed 11 March 2025].

4. S. Saalfeld, A. Cardona, V. Hartenstein, P. Tomančák, CATMAID: Collaborative annotation toolkit for massive amounts of image data. *Bioinformatics* **25**, 1984–1986 (2009).

5. C. Sommer, C. Straehle, U. Köthe, F. A. Hamprecht, Ilastik: Interactive learning and segmentation toolkit in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, (2011), pp. 230–233.

6. A. Cardona, *et al.*, TrakEM2 Software for Neural Circuit Reconstruction. *Plos one* **7**, e38011 (2012).

7. C. M. Schneider-Mizell, *et al.*, Quantitative neuroanatomy for connectomics in Drosophila. *Elife* **5**, e12059 (2016).

8. P. A. Yushkevich, Y. Gao, G. Gerig, ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. *Conf proc ieee eng med biol soc* **2016**, 3342–3345 (2016).

9. K. M. Boergens, *et al.*, webKnossos: Efficient online 3D data annotation for connectomics. *Nat methods* **14**, 691–694 (2017).

10. D. R. Berger, H. S. Seung, J. W. Lichtman, VAST (Volume Annotation and Segmentation Tool): Efficient Manual and Semi-Automatic Labeling of Large 3D Image Stacks. *Frontiers in neural circuits* **12** (2018).

11. P. Hanslovsky, *et al.*, Paintera. Zenodo. https://doi.org/10.5281/zenodo.14454929. Deposited December 2024.

12. N. Sofroniew, *et al.*, Napari: A multi-dimensional image viewer for Python. Zenodo. https://doi.org/10.5281/zenodo.14719463. Deposited 22 January 2025.

13. K. L. Briggman, D. D. Bock, Volume electron microscopy for neuronal circuit reconstruction. *Current opinion in neurobiology* **22**, 154–161 (2012).

14. J. Kornfeld, W. Denk, Progress and remaining challenges in high-throughput volume electron microscopy. *Current opinion in neurobiology* **50**, 261–267 (2018).

15. C. S. Xu, S. Pang, K. J. Hayworth, H. F. Hess, "Transforming FIB-SEM Systems for Large-Volume Connectomics and Cell Biology" in *Volume Microscopy*, (Humana, New York, NY, 2020), pp. 221–243.

16. C. J. Peddie, *et al.*, Volume electron microscopy. *Nat rev methods primers* **2**, 1–23 (2022).

17. R. W. Castro, M. C. Lopes, L. M. De Biase, G. Valdez, Aging spinal cord microglia become phenotypically heterogeneous and preferentially target motor neurons and their synapses. *Glia* **72**, 206–221 (2024).

18. D. Djama, *et al.*, The type of inhibition provided by thalamic interneurons alters the input selectivity of thalamocortical neurons. *Curr res neurobiol* **6**, 100130 (2024).

19. K. Haruwaka, *et al.*, Microglia enhance post-anesthesia neuronal activity by shielding inhibitory synapses. *Nat neurosci* **27**, 449–461 (2024).

20. D. Holl, *et al.*, Distinct origin and region-dependent contribution of stromal fibroblasts to fibrosis following traumatic injury in mice. *Nat neurosci* **27**, 1285–1298 (2024).

21. M. K. P. Joyce, J. Wang, H. Barbas, Subgenual and Hippocampal Pathways in Amygdala Are Set to Balance Affect and Context Processing. *J. neurosci.* **43**, 3061–3080 (2023).

22. Y. M. Morozov, P. Rakic, Disorder of Golgi Apparatus Precedes Anoxia-Induced Pathology of Mitochondria. *International journal of molecular sciences* **24**, 4432 (2023).

23. M. Ziółkowska, *et al.*, Phosphorylation of PSD-95 at serine 73 in dCA1 is required for extinction of contextual fear. *Plos biology* **21**, e3002106 (2023).

24. S. Amemiya, *et al.*, Early stalked stages in ontogeny of the living isocrinid sea lily Metacrinus rotundus. *Acta zoologica* **97**, 102–116 (2016).

25. J. L. Parke, *et al.*, Phytophthora ramorum Colonizes Tanoak Xylem and Is Associated with Reduced Stem Water Transport. *Phytopathology®* **97**, 1558–1567 (2007).

26. C. Groh, Z. Lu, I. A. Meinertzhagen, W. Rössler, Age-related plasticity in the synaptic ultrastructure of neurons in the mushroom body calyx of the adult honeybee Apis mellifera. *J comp neurol* **520**, 3509–3527 (2012).

27. M. A. Seid, E. Junge, Social isolation and brain development in the ant Camponotus floridanus. *Sci nat* **103**, 1–6 (2016).

28. J. Stökl, G. Herzner, Morphology and ultrastructure of the allomone and sex-pheromone producing mandibular gland of the parasitoid wasp Leptopilina Heterotoma (Hymenoptera: Figitidae). *Arthropod structure & development* **45**, 333–340 (2016).

29. M. Baumgart, *et al.*, Morphometric study of the two fused primary ossification centers of the clavicle in the human fetus. *Surg radiol anat* **38**, 937–945 (2016).

30. B. C. Moore, K. Mathavan, L. J. Guillette, Morphology and histochemistry of juvenile male American alligator (Alligator mississippiensis) phallus. *Anat rec (hoboken)* **295**, 328–337 (2012).

31. C. M. Dinnis, A. K. Dahle, J. A. Taylor, Three-dimensional analysis of eutectic grains in hypoeutectic Al–Si alloys. *Materials science and engineering: A* **392**, 440–448 (2005).

32. J. Schindelin, *et al.*, Fiji: An open-source platform for biological-image analysis. *Nat methods* **9**, 676–682 (2012).

33. MCell Team, SWiFT-IR. GitHub. Deposited 2023.

34. J. Stiles, T. Bartol, "Monte Carlo methods for simulating realistic synaptic microphysiology using MCell" in *Computational Neuroscience: Realistic Modeling for Experimentalists*, (CRC Press, 2001), pp. 87–127.

35. R. A. Kerr, *et al.*, Fast Monte Carlo Simulation Methods for Biological Reaction-Diffusion Systems in Solution and on Surfaces. *Siam journal on scientific computing* **30**, 24 (2008).

36. MCell Team, CellBlender. GitHub. Deposited 2022.

37. T. Tasdizen, *et al.*, Automatic mosaicking and volume assembly for high-throughput serial-section transmission electron microscopy. *Journal of neuroscience methods* **193**, 132–144 (2010).

38. K. J. Hayworth, *et al.*, Ultrastructurally smooth thick partitioning and volume stitching for large-scale connectomics. *Nat methods* **12**, 319–322 (2015).

39. I. Lobato, T. Friedrich, S. Van Aert, Deep convolutional neural networks to restore single-shot electron microscopy images. *Npj comput mater* **10**, 1–19 (2024).

40. H. Hillman, K. Deutsch, Area changes in slices of rat brain during preparation for histology or electron microscopy. *Journal of microscopy* **114**, 77–84 (1978).

41. B. Cragg, Preservation of extracellular space during fixation of the brain for electron microscopy. *Tissue and cell* **12**, 63–72 (1980).

42. A. Schüz, G. Palm, Density of neurons and synapses in the cerebral cortex of the mouse. *Journal of comparative neurology* **286**, 442–455 (1989).

43. K. M. Harris, *et al.*, Uniform Serial Sectioning for Transmission Electron Microscopy. *J. neurosci.* **26**, 12101–12103 (2006).

44. M. Kuwajima, J. M. Mendenhall, L. F. Lindsey, K. M. Harris, Automated Transmission-Mode Scanning Electron Microscopy (tSEM) for Large Volume Analysis at Nanoscale Resolution. *Plos one* **8**, e59573 (2013).

45. M. Kuwajima, J. M. Mendenhall, K. M. Harris, Large-Volume Reconstruction of Brain Tissue from High-Resolution Serial Section Images Acquired by SEM-Based Scanning Transmission Electron Microscopy. *Nanoimaging: Methods and protocols* 253–273 (2013).

46. G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, A. V. Dalca, VoxelMorph: A Learning Framework for Deformable Medical Image Registration. *Ieee transactions on medical imaging* **38**, 1788–1800 (2019).

47. S. Hamzehei, *et al.*, 3D Biological/Biomedical Image Registration with enhanced Feature Extraction and Outlier Detection. *Acm bcb* **2023**, 1 (2023).

48. A. Hoopes, M. Hoffmann, B. Fischl, J. Guttag, A. V. Dalca, HyperMorph: Amortized Hyperparameter Learning for Image Registration in *Information Processing in Medical Imaging*, A. Feragen, S. Sommer, J. Schnabel, M. Nielsen, Eds. (Springer International Publishing, 2021), pp. 3–17.

49. S. Klein, M. Staring, K. Murphy, M. A. Viergever, J. P. W. Pluim, Elastix: A Toolbox for Intensity-Based Medical Image Registration. *Ieee transactions on medical imaging* **29**, 196–205 (2010).

50. L. Liu, *et al.*, Learning by Analogy: Reliable Supervision From Transformations for Unsupervised Optical Flow Estimation in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2020), pp. 6488–6497.

51. E. Mitchell, S. Keselj, S. Popovych, D. Buniatyan, H. S. Seung, Siamese Encoding and Alignment by Multiscale Learning with Self-Supervision. arXiv [Preprint] (2019). http://arxiv.org/abs/1904.02643 [Accessed 17 February 2025].

52. S. Popovych, *et al.*, Petascale pipeline for precise alignment of images from serial section electron microscopy. *Nat commun* **15**, 289 (2024).

53. S. Preibisch, S. Saalfeld, J. Schindelin, P. Tomancak, Software for bead-based registration of selective plane illumination microscopy data. *Nat methods* **7**, 418–419 (2010).

54. S. Saalfeld, R. Fetter, A. Cardona, P. Tomancak, Elastic volume reconstruction from series of ultra-thin microscopy sections. *Nat methods* **9**, 717–720 (2012).

55. J. Vargas, A.-L. Álvarez-Cabrera, R. Marabini, J. M. Carazo, C. O. S. Sorzano, Efficient initial volume determination from electron microscopy images of single particles. *Bioinformatics* **30**, 2891–2898 (2014).

56. A. W. Wetzel, *et al.*, Registering large volume serial-section electron microscopy image sets for neural circuit reconstruction using FFT signal whitening in *2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, (2016), pp. 1–10.

57. T. Xin, *et al.*, A novel registration method for long-serial section images of EM with a serial split technique based on unsupervised optical flow network. *Bioinformatics* **39**, btad436 (2023).

58. Dawson-Haggerty, Trimesh. (2023). Deposited 2023.

59. M. Desbrun, M. Meyer, P. Schröder, A. H. Barr, Implicit fairing of irregular meshes using diffusion and curvature flow in *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '99., (ACM Press/Addison-Wesley Publishing Co., 1999), pp. 317–324.

60. J. Vollmer, R. Mencl, H. Müller, Improved Laplacian Smoothing of Noisy Surface Meshes. *Computer graphics forum* **18**, 131–138 (1999).

61. M. Musy, *et al.*, Vedo. Zenodo. https://doi.org/10.5281/zenodo.8067437. Deposited 21 June 2023.

62. Q Project, Qt6. (2025). Deposited 2025.

63. A. Collette, *Python and HDF5: Unlocking Scientific Data* (O'Reilly, 2013).

64. P. D. Team, PyTables: Hierarchical datasets in Python. (2025). Deposited 2025.

65. C. Pape, *et al.*, Z5. Zenodo. https://doi.org/10.5281/zenodo.11671609. Deposited 15 June 2024.

66. A. Miles, *et al.*, Zarr-python. Zenodo. https://doi.org/10.5281/zenodo.14873428. Deposited 14 February 2025.

67. A. Shapson-Coe, *et al.*, A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. *Science* **384**, eadk4858 (2024).

68. A. Azevedo, *et al.*, Connectomic reconstruction of a female Drosophila ventral nerve cord. *Nature* 1–9 (2024). https://doi.org/10.1038/s41586-024-07389-x.

69. S. Dorkenwald, *et al.*, Neuronal wiring diagram of an adult brain. *Nature* **634**, 124–138 (2024).

70. J. A. Bae, *et al.*, Functional connectomics spanning multiple areas of mouse visual cortex. *Biorxiv* 2021.07.28.454025 (2023). https://doi.org/10.1101/2021.07.28.454025.

71. N. L. Turner, *et al.*, Reconstruction of neocortex: Organelles, compartments, cells, circuits, and activity. *Cell* **185**, 1082–1100.e24 (2022).

72. L. K. Scheffer, *et al.*, A connectome and analysis of the adult Drosophila central brain. *Elife* **9**, e57443 (2020).

73. J. A. W. Heymann, *et al.*, Site-specific 3D imaging of cells and tissues with a dual beam microscope. *Journal of structural biology* **155**, 63–73 (2006).

74. G. Knott, H. Marchman, D. Wall, B. Lich, Serial Section Scanning Electron Microscopy of Adult Brain Tissue Using Focused Ion Beam Milling. *J neurosci* **28**, 2959–2964 (2008).

75. S. B. Leighton, SEM images of block faces, cut by a miniature microtome within the SEM - a technical note. *Scan electron microsc* 73–76 (1981).

76. W. Denk, H. Horstmann, Serial Block-Face Scanning Electron Microscopy to Reconstruct Three-Dimensional Tissue Nanostructure. *Plos biology* **2**, e329 (2004).

77. K. D. Micheva, S. J. Smith, Array Tomography: A New Tool for Imaging the Molecular Architecture and Ultrastructure of Neural Circuits. *Neuron* **55**, 25–36 (2007).

78. R. Schalek, *et al.*, Development of High-Throughput, High-Resolution 3D Reconstruction of Large-Volume Biological Tissue Using Automated Tape Collection Ultramicrotomy and Scanning Electron Microscopy. *Microscopy and microanalysis* **17**, 966–967 (2011).

79. H. Horstmann, C. Körber, K. Sätzler, D. Aydin, T. Kuner, Serial Section Scanning Electron Microscopy (S3EM) on Silicon Wafers for Ultra-Structural Volume Imaging of Cells and Tissues. *Plos one* **7**, e35172 (2012).

80. K. J. Hayworth, *et al.*, Imaging ATUM ultrathin section libraries with WaferMapper: A multi-scale approach to EM reconstruction of neural circuits. *Front neural circuits* **8**, 68 (2014).

81. J. Maitin-Shepard, *et al.*, Neuroglancer: Web-based volumetric data visualization. (2021). https://doi.org/10.5281/zenodo.5573294. Deposited 16 October 2021.

82. K. M. Harris, *et al.*, A resource from 3D electron microscopy of hippocampal neuropil for user training and tool development. *Sci data* **2**, 150046 (2015).

83. N. Kasthuri, *et al.*, Saturated Reconstruction of a Volume of Neocortex. *Cell* **162**, 648–661 (2015).

84. J. T. Vogelstein, *et al.*, A community-developed open-source computational ecosystem for big neuro data. *Nat methods* **15**, 846–847 (2018).

85. T. Zhao, D. J. Olbris, Y. Yu, S. M. Plaza, NeuTu: Software for Collaborative, Large-Scale, Segmentation-Based Connectome Reconstruction. *Front. neural circuits* **12** (2018).

86. W. T. Katz, S. M. Plaza, DVID: Distributed Versioned Image-Oriented Dataservice. *Front. neural circuits* **13** (2019).

87. C. S. Xu, *et al.*, An open-access volume electron microscopy atlas of whole cells and tissues. *Nature* **599**, 147–151 (2021).

88. R. J. Hider, *et al.*, The Brain Observatory Storage Service and Database (BossDB): A Cloud-Native Approach for Petascale Neuroscience Discovery. *Frontiers in neuroinformatics* (2022). https://doi.org/10.3389/fninf.2022.828787.

89. W. Silversmith, *et al.*, Igneous: Distributed dense 3D segmentation meshing, neuron skeletonization, and hierarchical downsampling. *Front. neural circuits* **16** (2022).