

SN Computer Science

A High-Performance Computing Portal Applied to 3D Electron Microscopy

--Manuscript Draft--

Manuscript Number:	SNCS-D-25-06990R1	
Full Title:	A High-Performance Computing Portal Applied to 3D Electron Microscopy	
Article Type:	Original Research	
Funding Information:	Division of Biological Infrastructure (1707356)	Dr. Kristen Harris
	Division of Biological Infrastructure (2014862)	Dr. Kristen Harris
Abstract:	<p>Purpose</p> <p>The continued expansion in size and resolution of volumetric datasets generated by electron microscopy (3DEM) is making high-performance computing (HPC) an essential component. HPC provides the scalable, parallelized infrastructure required to process and analyze these increasingly large datasets.</p> <p>Methods</p> <p>The Texas Advanced Computing Center's (TACC) Core Experience Portal (CEP) is a science gateway specialized for leveraging HPC. The CEP can be adapted into customized versions based on a research community's needs. Here we have constructed 3dem.org as a gateway for the community exploring volumetric data generated by electron microscopy (3DEM).</p> <p>Results</p> <p>This gateway provides the 3DEM community with user-friendly browser-based access to raw and processed datasets including both private and shared data, image processing tools for alignment and segmentation, and simulation and analysis environments. All this is linked to the underlying HPC environment at TACC with the ability to connect to other data storage and compute systems.</p> <p>Conclusion</p> <p>3dem.org bridges advanced electron microscopy with HPC, providing the research community with scalable, accessible infrastructure for discovery.</p>	
Corresponding Author:	James Carson, Ph.D. The University of Texas at Austin UNITED STATES OF AMERICA	
Corresponding Author Secondary Information:		
Corresponding Author's Institution:	The University of Texas at Austin	
Corresponding Author's Secondary Institution:		
First Author:	James P Carson	
First Author Secondary Information:		
Order of Authors:	James P Carson	
	Tracy Brown	
	James Labyer	
	Thomas Bartol	
	Michael A Chirillo	

	John Mendenhall
	Andrea Nam
	Vijay Thiyagarajan
	Arthur Wetzel
	Weiling Yin
	Erik Ferlanti
	Jake Rosenberg
	Sal Tijerina
	Hedda Prochaska
	Jawon Song
	William J Allen
	Tiffany Huff
	Masaaki Kuawjima
	Thomas Middendorf
	Lyndsey Kirk
	Joe Stubbs
	Maytal Dahan
	Terrance Sejnowski
	Kristen Harris
Order of Authors Secondary Information:	
Author Comments:	
Response to Reviewers:	We thank the reviewers and the editors for their feedback. All revision requests have been directly addressed through the addition of relevant text in the Discussion section.

A High-Performance Computing Portal Applied to 3D Electron Microscopy

*James P. Carson, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

jcarson@tacc.utexas.edu <https://orcid.org/0000-0001-9009-5645>

Tracy Brown, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

<https://orcid.org/0009-0001-3486-9500>

James A. Labyer, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

<https://orcid.org/0009-0003-8222-2079>

Thomas M. Bartol, Salk Institute for Biological Studies, La Jolla, CA, USA. <https://orcid.org/0009-0002-2598-0052>

Michael A. Chirillo, The University of Texas at Austin, Center for Learning and Memory, Austin, TX, USA, and

Department of Biological Sciences, The University of Rhode Island, Kingston, RI USA. <https://orcid.org/0000-0003-0667-8050>

John Mendenhall, The University of Texas at Austin, Center for Learning and Memory, Austin, TX, USA.

Andrea J. Nam, The University of Texas at Austin, Center for Learning and Memory, Austin, TX, USA.

<https://orcid.org/0000-0001-5457-9167>

Vijay Venu Thiagarajan, The University of Texas at Austin, Center for Learning and Memory, Austin, TX, USA.

<https://orcid.org/0000-0002-8244-7150>

Arthur W. Wetzel, Carnegie Mellon University, Pittsburgh Supercomputing Center, Pittsburgh, PA, USA.

<https://orcid.org/0000-0003-1999-5688>

Weiling Yin, The University of Texas at Austin, Center for Learning and Memory, Austin, TX, USA.

<https://orcid.org/0009-0006-5574-587X>

Erik Ferlanti, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

<https://orcid.org/0000-0001-5128-1584>

Jake Rosenberg, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

<https://orcid.org/0000-0001-9454-8090>

Sal Tijerina, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

<https://orcid.org/0009-0001-9283-965X>

Hedda Prochaska, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

Jawon Song, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

<https://orcid.org/0000-0002-9916-3816>

William J. Allen, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

<https://orcid.org/0000-0003-1322-2764>

Tiffany Huff, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

<https://orcid.org/0000-0003-2849-8093>

Masaaki Kuwajima, The University of Texas at Austin, Center for Learning and Memory, Austin, TX, USA.

<https://orcid.org/0000-0002-1478-3726>

Thomas Middendorf, The University of Texas at Austin, Center for Learning and Memory, Austin, TX, USA.

<https://orcid.org/0000-0002-7179-291X>

Lyndsey Kirk, The University of Texas at Austin, Center for Learning and Memory, Austin, TX, USA.

<https://orcid.org/0000-0002-5125-652X>

Joe Stubbs, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

<https://orcid.org/0000-0002-8644-0300>

Maytal Dahan, The University of Texas at Austin, Texas Advanced Computing Center, Austin, TX, USA.

<https://orcid.org/0000-0003-3432-5789>

Terrance J. Sejnowski, Salk Institute for Biological Studies, La Jolla, CA, USA, and University of California, San Diego, La Jolla, CA <https://orcid.org/0000-0002-0622-7391>

Kristen M Harris, The University of Texas at Austin, Center for Learning and Memory, Austin, TX, USA.

<https://orcid.org/0000-0002-1943-4744>

Acknowledgements

The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing computational resources that have contributed to the research results reported within this paper. In addition, we thank all TACC staff as they provide the support and resources which make these projects possible. We

1
2
3
4 also thank all the users of the gateway for their feedback throughout the development process. This work was funded
5 by National Science Foundation grants DBI-1707356 and DBI-2014862.
6
7

8 **Abstract**

9
10 Purpose: The continued expansion in size and resolution of volumetric datasets generated by electron microscopy
11 (3DEM) is making high-performance computing (HPC) an essential component. HPC provides the scalable,
12 parallelized infrastructure required to process and analyze these increasingly large datasets.
13

14 Methods: The Texas Advanced Computing Center's (TACC) Core Experience Portal (CEP) is a science gateway
15 specialized for leveraging HPC. The CEP can be adapted into customized versions based on a research community's
16 needs. Here we have constructed 3dem.org as a gateway for the community exploring volumetric data generated by
17 electron microscopy (3DEM).
18

19 Results: This gateway provides the 3DEM community with user-friendly browser-based access to raw and processed
20 datasets including both private and shared data, image processing tools for alignment and segmentation, and
21 simulation and analysis environments. All this is linked to the underlying HPC environment at TACC with the
22 ability to connect to other data storage and compute systems.
23

24 Conclusion: 3dem.org bridges advanced electron microscopy with HPC, providing the research community with
25 scalable, accessible infrastructure for discovery.
26

27 **Keywords**

28 Science gateways, sustainable cyberinfrastructure, 3DEM, volumetric image processing, synapse modeling and
29 simulation
30
31

32 **Introduction**

33
34 3-dimensional electron microscopy (3DEM) is rapidly expanding the size of ultra-high-resolution datasets being
35 acquired to further our understanding of biological systems. For example, a cubic millimeter of imaged tissue can
36 now exceed one petabyte, necessitating the use of large-scale computational resources for storage and analysis. To
37 support broadening access to the many computationally demanding domains such as 3DEM, the Texas Advanced
38 Computing Center (TACC) Core Experience Portal (CEP) was developed as a template science gateway that can be
39 readily deployed in support of a wide variety of disciplines [1]. The design of the CEP reflects both the technical
40 demands of high-performance computing (HPC) and the collaborative priorities of the broader research ecosystem.
41 A key component of the CEP is the Application Programming Interface (API) that provides the link between the
42 web-browser and HPC. CEP uses the TACC API (Tapis) as the RESTful API layer that serves as the secure
43 connection between the web gateway and the data storage and computing resources, as well as managing user access
44 [2]. This approach forms a foundation of scientific community data sharing that supports the Findability,
45 Accessibility, Interoperability, and Reusability (FAIR) principles [3].
46
47

48 The 3dem.org web-based research platform was developed as part of the Next Generation Network for Neuroscience
49 (NeuroNex) program [4]. 3dem.org is focused on supporting and disseminating new technologies for enhanced
50 resolution 3DEM, with a specific emphasis on applications involving the ultrastructure of synapses [5]. By
51 integrating this new data collection technology with publicly funded high-performance computing (HPC), the
52 additional computational power is leveraged to process even larger datasets, and to promote reproducibility and
53 sharing of techniques. 3dem.org is designed to serve not only the project team, but all researchers interested in the
54 tools and capabilities. All visitors to 3dem.org may download the latest public datasets. Users with active TACC
55 accounts and allocations may request access to the interactive 3DEM Workbench. Within the Workbench, users can
56 access community datasets, workflows, and applications. Applications are focused on those that support processing
57 for enhanced resolution 3DEM. Users may also share their data and processes with other users.
58
59
60
61
62
63
64
65

3DEM is an appropriate domain for demonstrating the utility of the CEP. Image volumes are increasing in size as instrumentation improves, and the research is focused on seeking to further our functional understanding not only of cells but also the cell neighborhoods. These larger datasets increasingly depend on HPC for processing. Having a common portal allows shared access to the large datasets and processing applications. We describe here the application of the CEP and Tapis to this large-scale collaborative effort toward 3DEM of synapses.

Results

Public Front End

Visitors to 3dem.org can access a customized front end of the CEP that describes the domain and gives public access to project information, recorded workshops, and instructional materials (Figure 1). Information on how users can access gated resources and tutorials on using 3dem.org are provided publicly. In addition, publicly shared datasets can be downloaded.

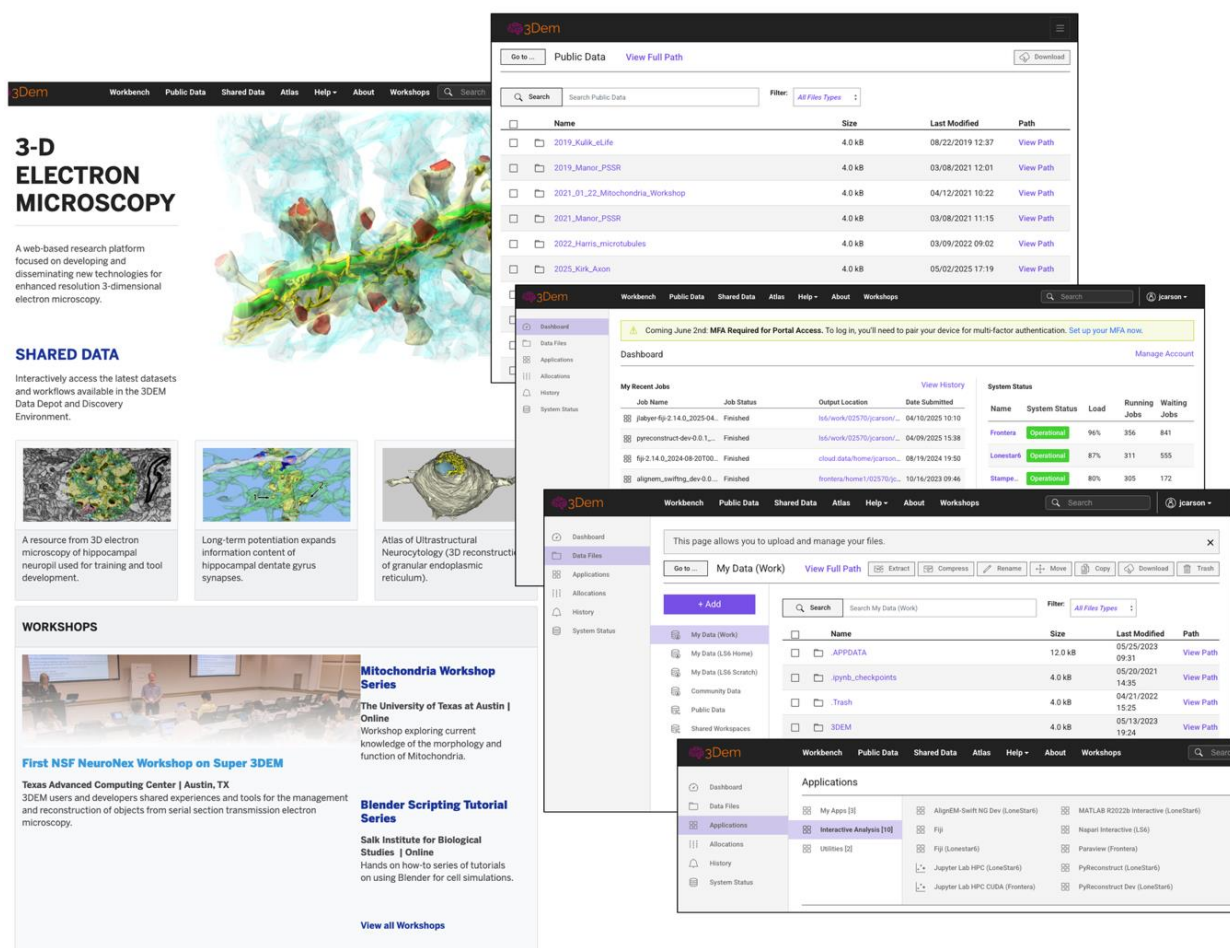


Fig. 1. The 3DEM Science Gateway provides a browser-based access point to advanced computing resources. Users can launch both interactive and batch applications on HPC systems, with both styles having access to the same set of queues. Users can also control their data sharing, from private, to selected collaborators via Shared Workspaces, to making their data Public.

Workbench Analytical Environment

After creating a TACC account, users can request and be granted access to the full features of 3dem.org. To facilitate access to data and analysis resources, the CEP provides a customized web-based graphical user interface (GUI) Workbench. Upon logging in via a laptop or even a smartphone, users first access their dashboard. The dashboard lists the user's most recently run HPC jobs, provides easy access to the user's support tickets, and lists the

operational status of the HPC systems. The Workbench also provides access to data and registered applications on HPC systems. Addressing project security needs, users can keep their data and applications fully private, share with specific users, or share with all members of the program. The infrastructure supports batch applications that submit jobs to HPC systems to run when resources are available, as well as interactive applications that open direct web-access to HPC nodes for more exploratory analysis.

Data Management

Integral to collaboration on 3dem.org is the capability to access and generate data and metadata from interactive environments and HPC resources. Within the GUI Workbench, 3dem.org has a “Data Files” tab that allows users to access various HPC filesystems. This includes TACC's Corral system which is used by 3dem.org as the central repository to facilitate sharing within custom workgroup projects as well as public data sharing, and TACC's Stockyard “\$WORK” filesystem within which each TACC user has 1TB available for their private data and active analysis. Users also have access to their home directories and unrestricted temporary storage on scratch filesystems on TACC's Lonestar6 (<https://docs.tacc.utexas.edu/hpc/lonestar6/>).

For custom workgroup projects, any user can also create these private directories and manage access with collaborators. This access uses both the fine-grained access control lists as well as TAPIS permissions. Together, this allows users to manage and share the data via the 3dem.org web-browser interface, while also making the data simultaneously available to both API-calls and direct command line access.

Featured Datasets

When research data is ready to be more broadly accessible than custom workgroups, 3dem.org provides two options, Community Data and Public Data.

Community Data

Community Data is a form of broad data access that allows anyone with a TACC account and 3dem.org access to also access Community Data. Currently, Community Data includes example datasets for trying out 3dem.org applications and testing application development. Also included in Community Data are datasets used in NeuroNex workshops, including 3DEM datasets from [5]. Community Datasets can be easily accessed by 3DEM applications as well as by command line. In addition, users can download the data if so desired.

Public Data

As with Community Data, 3dem.org users can readily access Public Data for use with 3DEM applications. However, Public Data is an even broader form of data access that allows anyone with a web-browser to download the data, even without a TACC account. For 3dem.org, the Public Data includes video recordings and transcripts from past NeuroNex workshop series presentations, as well as data made public as part of publications, which can serve as a complement to repositories that may have data size limitations (e.g., the Texas Data Repository at <https://dataverse.tdl.org/>). Publication data includes: deep learning-based pretrained models as well as all training and testing data for point-scanning super-resolution imaging that includes both volumetric EM and non-EM data [6]; serial EM images from hippocampal area CA1 from acute slices prepared from P15 rats along with trace, and series files analyzed for synaptogenesis [7] and for dendritic spine density and microtubules [8]; serial EM images acquired from hippocampal area CA1 from acute slices prepared from adult rats (P60-61), two hours following either control or theta burst stimulation [9]; serial-section EM images along with a corresponding PyReconstruct trace files collected from the middle and outer molecular layers of the hippocampal dentate gyrus of adult male rats (P121-185) [10]; and EM tomography images for analysis of presynaptic vesicles in long term potentiation [11].

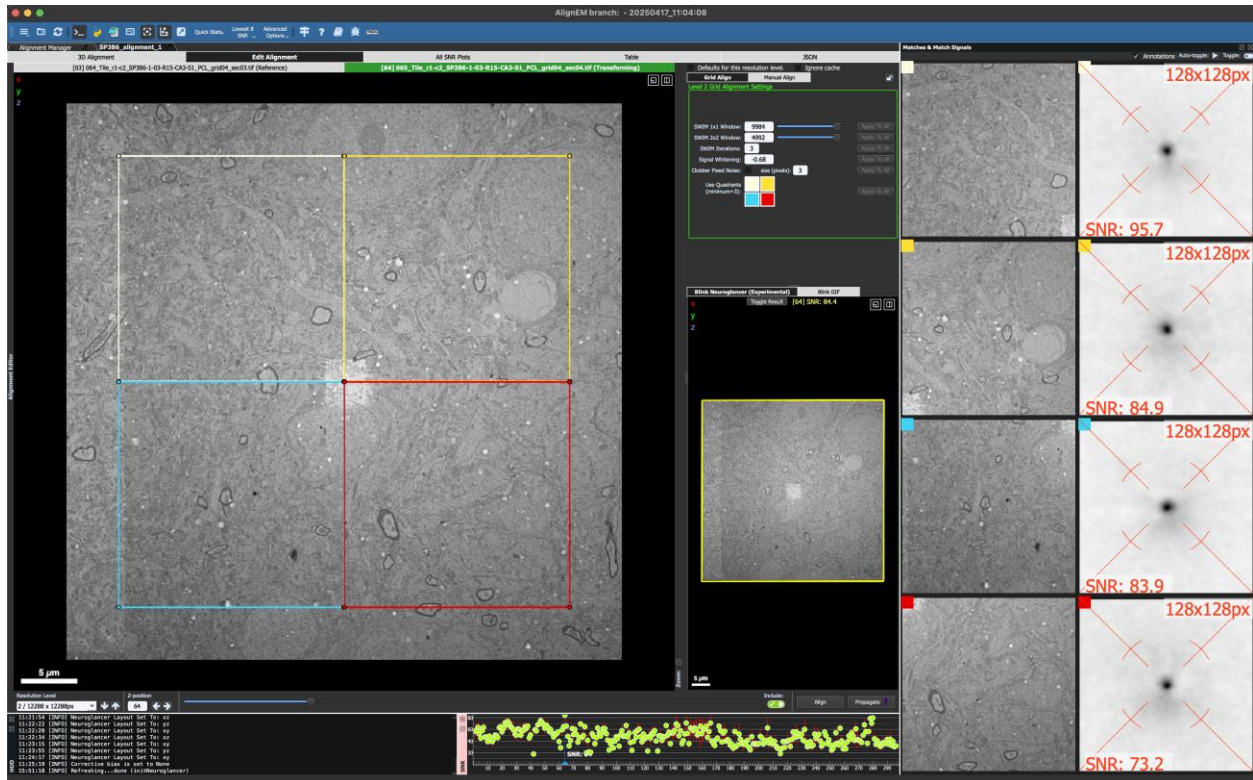
Applications

To support best practices in software development the use of software containers and public code repositories is encouraged for applications integrated into 3dem.org and other CEP gateways. Reproducibility is further supported by allowing users to individually or collaboratively develop and share their data processing workflows and analysis pipelines. Provenance of software versions is tracked by Tapis as well as inputs, outputs, and parameters for every job launched, allowing specific reproducibility and interactive development applicability. 3dem.org has a collection of cutting-edge applications, both pre-release and full-featured, designed to support the volumetric data community

1
2
3
4 that span the processing workflow, including generating a coherent 3D voxelated volume from acquired images,
5 segmenting morphological features, simulation, and analysis.
6

7 AlignEM-SWiFT

8 AlignEM-swift is an open-source software application under development for high-throughput alignment of serial
9 EM images (<https://github.com/mcellteam/swift-ir>). As such, this tool is fundamental to the creation and curation of
10 3DEM datasets. AlignEM-swift, written in Python, is based upon the powerful SWiFT-IR (Signal Whitening,
11 Fourier Transform Image Registration) toolkit [12], encapsulating the individual tools in the toolkit, which are
12 written in C. The graphical user interface of AlignEM-swift streamlines the workflow of 3DEM dataset curation
13 from intake of a raw image dataset, alignment of the dataset, analysis and evaluation of the quality of the alignment,
14 and remedies for problematic alignments. AlignEM-swift also encapsulates high-performance viewing of the
15 dataset, in zarr format (<https://zarr.dev/>), via Neuroglancer (<https://github.com/google/neuroglancer>).
16
17



18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43 Fig. 2. The GUI of AlignEM-swift under development shown here aligning a 3DEM image stack.

44
45 Figure 2 shows a screen capture of a typical use case: a stack of images from a 3DEM dataset is loaded into the tool
46 (297 images in this example). The images are then aligned to their previous neighbor in the stack in a pairwise
47 manner. Because the alignment process involves evaluation of pairwise images, it is highly parallelizable on HPC
48 platforms such as LoneStar6 at TACC. Alignment requires finding the affine matrix that maximizes the goodness of
49 fit between a given pair. This is accomplished by employing the signal whitening image matching (swim) function
50 of the SWiFT-IR toolkit. Swim measures the displacement of the signal whitening FFT match signals (dark spot
51 images displayed on the right panel in Figure 2) for several corresponding image patches in neighboring images in
52 the stack. Here 4 patches are used as indicated by the white, yellow, blue, and red rectangles shown on the left panel.
53 The set of displacements is used to compute the single affine matrix that brings the rectangles into correspondence
54 across the neighboring images. By iterating this process several times, the residual displacements measured by
55 swim decrease to sub-pixel magnitude and convergence to the best affine matrix is achieved. The goodness of fit of
56 the image alignment can be assessed from the signal-to-noise ratio (SNR) of the match signals, which is visually
57 apparent from the shape and compactness of the dark spot and its contrast with the light gray background of the
58 match signal. A graph showing the pairwise average SNR values obtained across the whole stack is shown in the
59 lower middle panel. Generally, SNR values greater than 8 indicate satisfactory alignment. In the example shown, the
60
61
62
63
64
65

1
2
3
4 SNR values across the stack ranged from 17 to 84. The graph can be used to identify quickly any problematic
5 alignments and the match signals for those cases can be assessed to determine appropriate remedies, available via
6 custom settings in the GUI.
7

8 The main output of AlignEM-swift is the curated project file which is a database in JSON format containing all
9 settings, steps performed, and results for a given project, including the final affine matrices aligning the stack.
10 PyReconstruct, introduced in the next section, uses the affine matrices in the project file as a fundamental part of its
11 non-destructive display of the 3DEM dataset.
12

13 **PyReconstruct**

14 PyReconstruct is a powerful, fully open-source application [13] for annotating and analyzing serial electron
15 microscopy data. PyReconstruct was designed as a modern successor to the widely-used legacy software
16 Reconstruct [14]. Created specifically to balance rapid segmentation with detailed structural analysis, PyReconstruct
17 empowers researchers to trace, segment, curate, and visualize 3D cellular structures across serial sections with
18 exceptional precision. The tool features an intuitive GUI with moveable widgets, sophisticated 3D rendering
19 capabilities, and built-in curation systems for team-based quality control. Its rich annotation environment supports
20 the application of tags, flags, and comments to traces and objects, creating a comprehensive documentation system
21 integrated with the image data. PyReconstruct's dynamic alignment strategy enables users to apply multiple
22 alignment profiles on-demand without permanently altering original images, making it ideal for correcting section
23 distortions discovered during annotation. The application excels at nanoscale quantification of structures like
24 synapses while maintaining compatibility with larger-scale connectomics workflows, making it valuable across
25 neuroscience as well as a variety of diverse fields [15–29].
26

27 PyReconstruct addresses key technical limitations through its multi-resolution zarr approach to image handling,
28 breaking through previous RAM restrictions to handle arbitrarily large datasets. PyReconstruct facilitates team-
29 based workflows through its robust import/export capabilities and trace conflict resolution system, which identifies
30 potential conflicts arising during multi-user segmentation efforts. The software accepts standard image formats as
31 input and produces multiple outputs including quantitative measurements, 3D meshes, and labelled zarrs for
32 integration with automated segmentation pipelines. PyReconstruct bridges manual and automated approaches,
33 enabling researchers to generate ground truth for machine learning while maintaining full control over annotation
34 quality. Interactive features like double-clicking on 3D objects to locate corresponding sections facilitate efficient
35 proofreading. Source code is fully open and available under a GNU General Public License (v3), with ongoing
36 development hosted on GitHub (<https://github.com/synapseweb/pyreconstruct>).
37

38 **Napari**

39 Napari is a Python-based application that supports n-dimensional image visualization, annotation, and analysis [30].
40 It can input n-dimensional images (2D images, 3D images, multi-channel images, Zarrs, tiffs) and output labels,
41 point annotations, line/curve/polygon annotations. In addition to enabling exploration, overlaying, and editing
42 complex images, Napari has a strong plugin community. Our installation of Napari [31] has several key plugins
43 important to 3DEM workflows. One of these is a Bootstrapper plugin for accessing and editing segmentations using
44 machine learning to generate segmentations from a sparse set of manual segmentation (Figure 3) [32]. Two other
45 featured Napari plugins are microSAM (Meta's Segment Anything Model finetuned for microscopy) [33] and
46 PSSR2, a deep learning-based point-scanning super-resolution imaging tool for improving image resolution [34].
47 Though Napari by itself does not have many parallel features, these plugins can have parallelized implementations
48 of functions. For this reason, Napari is best used for viewing, exploring, and annotating smaller cropped volumes.
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

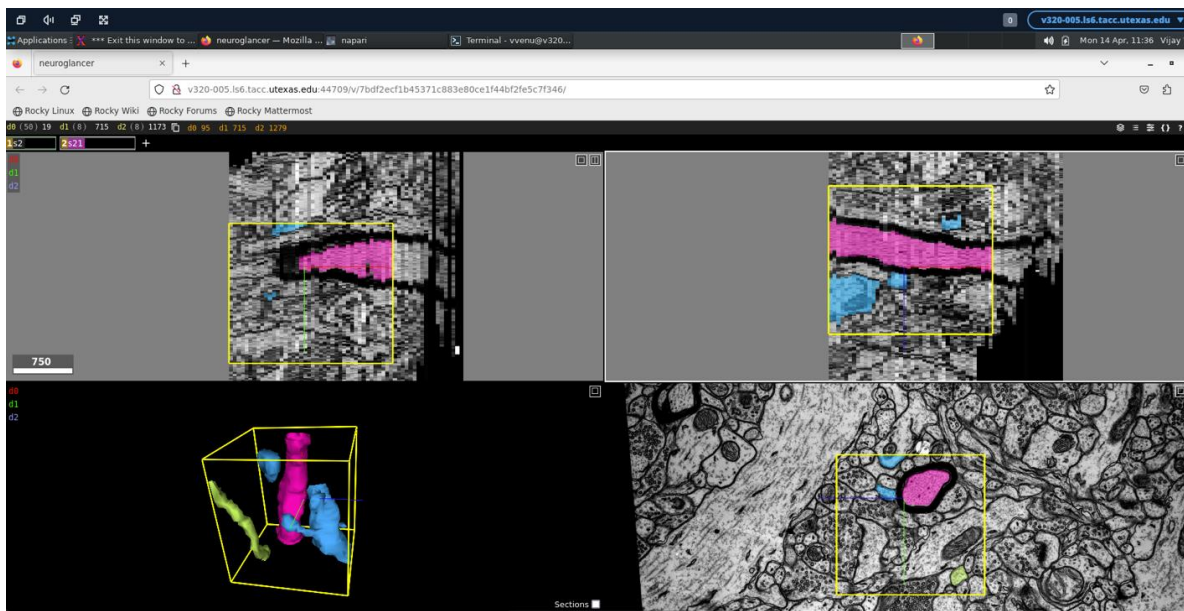
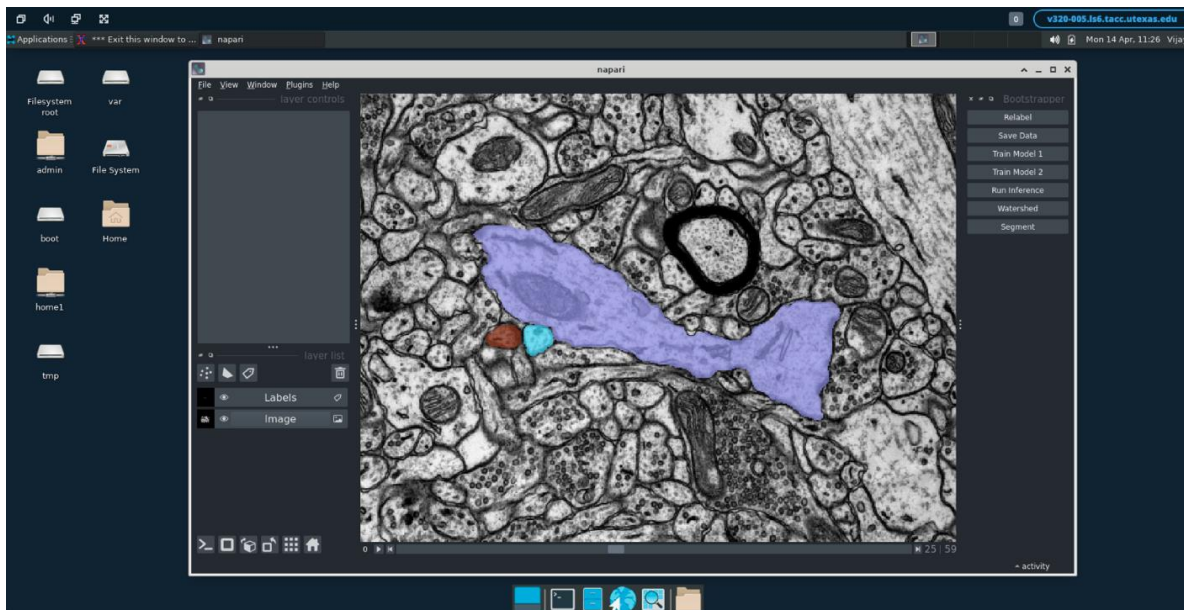


Fig. 3. The Napari application and Bootstraptrapper plugin. Top: Shown in use on 3dem.org. Sparse label annotations in color. Bottom: Screenshot of Neuroglancer viewer of same dataset illustrates ZY, ZX, and YX planes as well as 3D rendering of labels. <https://github.com/ucsdmanorlab/napari-bootstraptrapper/releases/tag/v0.2.0>

MCell-CellBlender-NeuropilTools

MCell-CellBlender-NeuropilTools is suite of tools designed for computer simulation and analysis of biochemical reaction-diffusion systems in 3D [35]. MCell is the particle/agent based Monte-Carlo simulation engine in the package. CellBlender is an add-on for Blender, a popular and powerful 3D CAD platform, and serves as the graphical frontend to the MCell physics engine. NeuropilTools works alongside CellBlender to provide computational quality mesh generation, annotation, and analysis of cellular and subcellular membranes (with emphasis on neuronal tissue) from their segmentation contours in PyReconstruct 3DEM datasets. Together this suite

of tools allows the researcher to bridge the gap from structure (given by the cellular ultrastructure in the 3DEM dataset) to dynamic cellular function, in a sense, bringing the tissue back to life *in-silico*.

Figure 4 illustrates an example of MCell-CellBlender-NeuropilTools used to simulate pre-synaptic release of neurotransmitter and post-synaptic activation at a spine synapse in hippocampal area CA1 [36, 37]. Here the cellular structure was hand-segmented in Reconstruct, meshed and annotated with NeuropilTools, and turned into a functional model of synaptic transmission with anatomically correct distributions of relevant molecular species, biochemical reaction pathways and kinetic rate constants with CellBlender, and then simulated with MCell. CellBlender displays the visualization of the simulation results within the Blender platform.

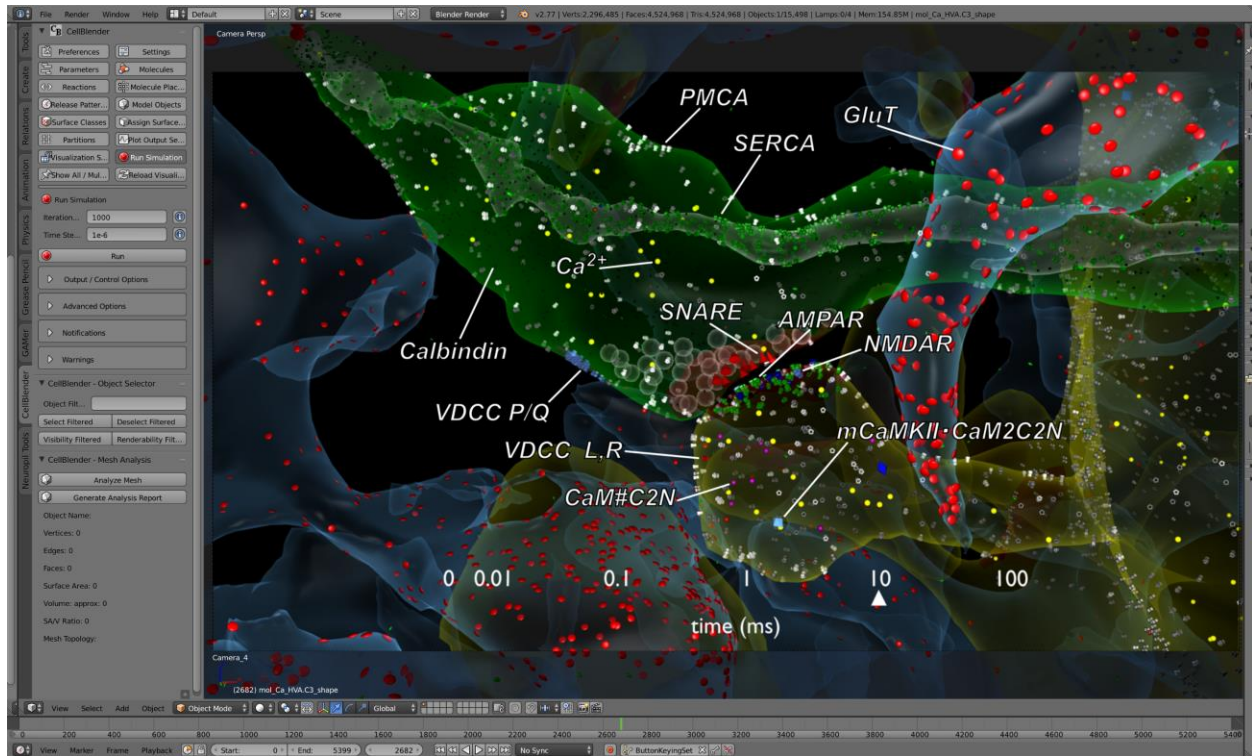


Fig. 4. Computational reconstitution and MCell simulation of glutamatergic synaptic transmission at dendritic spines of hippocampal area CA1 pyramidal neurons. The model encompasses many known components and molecular mechanisms of presynaptic neurotransmitter release (presynaptic axon shown in translucent green, molecular species as indicated) and postsynaptic activation on the dendritic spine and dendrite (shown in translucent yellow, molecular species as indicated). The width of field of view is ~ 3 micrometers. This model was constructed and simulated in MCell-CellBlender-NeuropilTools and is shown in the 3D view panel at $t=10$ ms into the simulation.

Additional Applications

In addition to these primary applications utilized by the NeuroNex project, several other software applications relevant to 3DEM are also currently available to all users. These include TrakEM2 [38] as part of Fiji [39] which is used for image registration. Notably, we have customized our Fiji container to allow specifying the number of processes as a parameter when launching the application (<https://github.com/eriksf/fiji-openjdk-8-ui>). This is to allow users to ensure that sufficient RAM is allocated to each process, which is important when aligning large images. Other applications include Paraview for 3D visualization [40], Matlab [41], Jupyter notebooks [42], and Neuroglancer (<https://github.com/google/neuroglancer/releases/tag/v2.40.1>), a fast web browser-based viewer for large zarr arrays which supports customization and workflow automation using Python. Running these applications on one or more HPC nodes provides a powerful approach for working with large 3DEM images and volumes. Beyond these applications, any user can create their own private application for their own use, share the application with collaborators, and if desired, request that the application be made public to all 3dem.org users.

Workflows

Figure 5 illustrates a typical workflow performed using 3dem.org. A raw transmission scanning EM (tSEM) image series collected for each experiment is transferred from the laboratory's local drive to 3dem.org My Data (temporary storage of LS6 Scratch). AlignEM-Swift in Interactive Analysis mode uses SNR to evaluate the quality of the series for issues such as poor focus or astigmatism, displaced region of interest (ROI), section defects, repeated ROIs, missing or out-of-sequence sections that may necessitate re-imaging. After passing the quality check, image series is coded for alignment using AlignEM-Swift in My Data. Aligned image series (.tiff, .swift) are transferred from My Data to project folder in Shared Workspaces for data storage and analysis. PyReconstruct, TrakEM2 (Fiji), Napari, and CellBlender are used for Interactive analysis to prepare data for research thesis, collaborative grant proposals, and publications. Datasets and associated metadata (.tiff, .swift, .jser, .zarr, .csv, .blend, .xlsx etc.) generated from experiments are saved and shared on the 3dem.org portal for easy retrieval and data sharing.

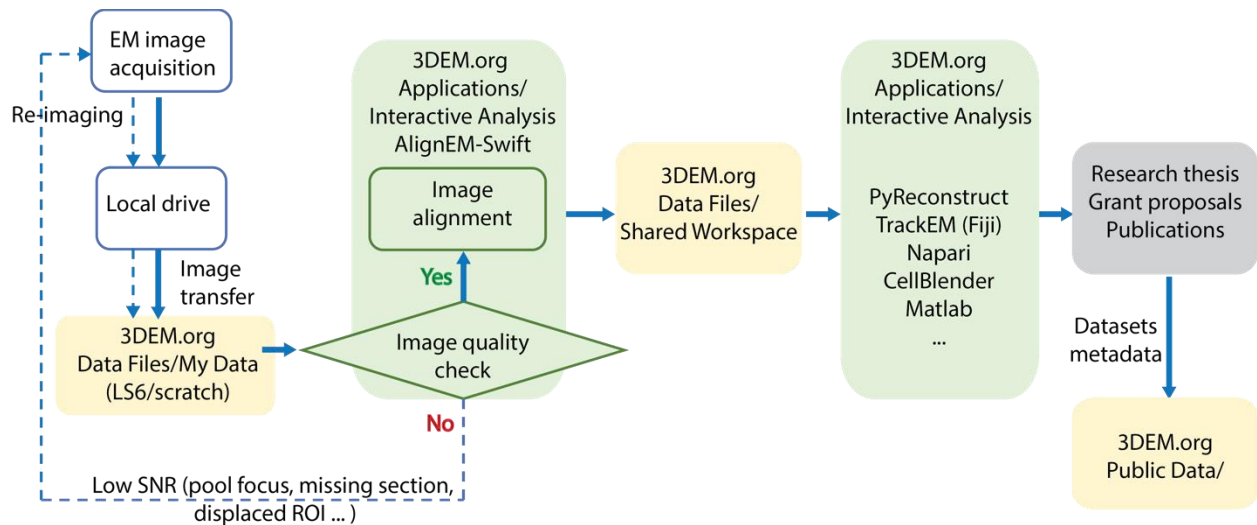


Fig. 5. Workflow of 3DEM image volumes, reconstructions, visualization, analysis and storage on 3dem.org.

Figure 6 illustrates how 3dem.org has been applied to publication workflows, beginning with ingesting the hippocampal dentate gyrus dataset into 3dem.org. For this dataset, in vivo long-term potentiation (LTP) and concurrent long-term depression (cLTD) were induced at granule cell synapses. Interactive applications available through 3dem.org facilitated sharing and analysis of the large datasets, accelerating research progress.

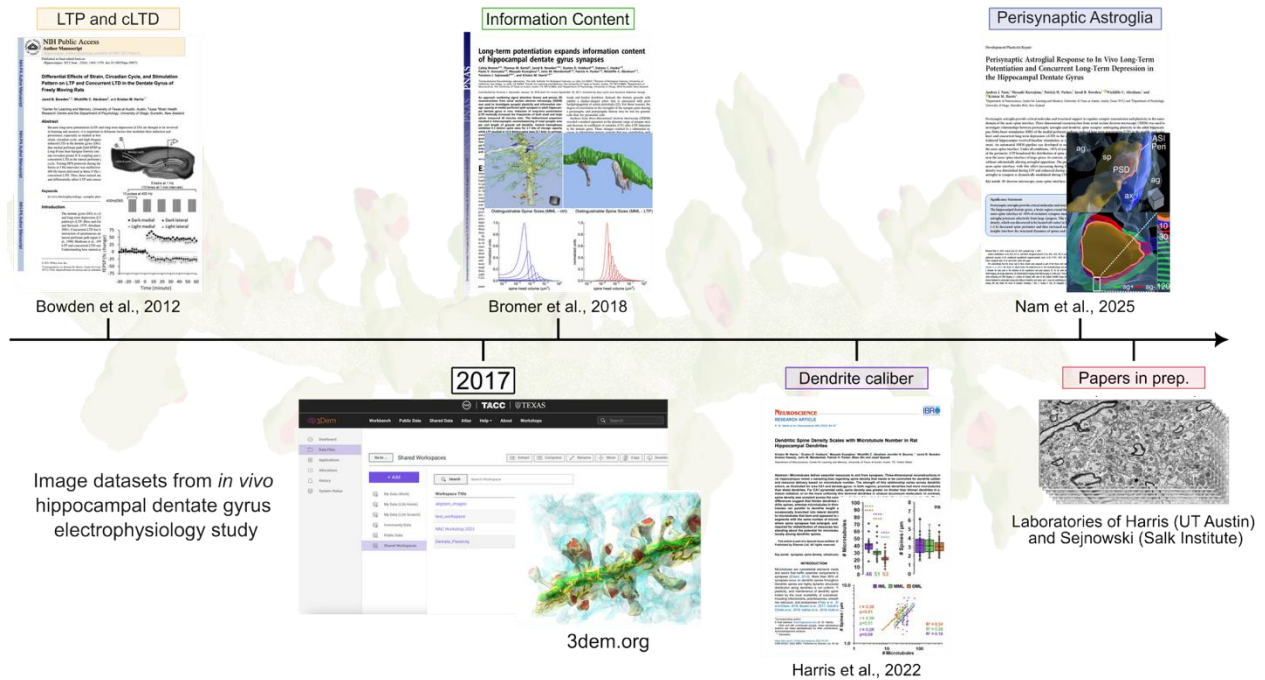


Fig. 6. Integrated data and analysis framework at 3dem.org supports data sharing and streamlined publication workflows.

Methods

Portal Deployment

The 3dem.org instance of the CEP is deployed on a virtual machine on Rodeo, TACC's in-house VMware cluster. The frontend is a web application using the React framework, providing users an interface for managing data and submitting HPC jobs. The backend is built with the Django web framework, providing API endpoints to primarily structure calls to Tapis apps, jobs, and systems and format the results for display on the front end. A PostgreSQL cluster is used for the database layer, which stores user profiles and project metadata. Searches for files, projects, and web content are managed by an Elasticsearch cluster running on a dedicated virtual machine. Search indexing and other long-running tasks are handled by an asynchronous Celery worker with RabbitMQ as the message broker. An Nginx server is placed in front of the rest of the services to route user requests and serve static web content.

The portal itself is completely Dockerized, with each service broken out into containers, with deployments managed by Jenkins, docker-compose, and an in-house customizable deploy-management suite called Camino (<https://github.com/TACC/Core-Portal>). The content management system (CMS) also exists as its own image and container, allowing for seamless upgrades and deployments targeting specific services, maximizing uptime.

API Tools

Agility and portability can be critical to achieving a rapid pace of iteration and discovery. Toward this end, 3dem.org uses Tapis version 3 as a common, unifying platform to harmonize the presentation of data, the publication of software applications (<https://github.com/james-labyer/3dem-apps>), and the provenance of analyses. As discussed, the portals provide a graphical, browser-based interface to the hardware infrastructure. These views were built on Tapis endpoints. Tapis managed the key functionality of the gateways, including managing user and data security.

Tapis is also accessible using powerful python libraries, Tapipy and TapisService (<https://tapis.readthedocs.io/en/latest/technical/pythondev.html>). These libraries enable the creation and maintenance of Tapis entities such as systems, applications, and jobs via a widely used language that is familiar to many scientists.

Application Deployment

Researchers can use the tooling and computing resources provided by TACC to make research software available via the 3dem.org workbench. The process begins with containerization: the desired software must be containerized before deployment via Tapis v3. Once the software has been containerized, it can be tested on TACC systems using an interactive development session or a DCV session that provides a GUI for applications with visualization requirements. Once the container is working, users can systematize the deployment of the container on TACC systems by creating a new Tapis application that stores the information needed to successfully deploy the container. The user can then run the new Tapis application from the 3dem.org workbench. Tapis applications are initially visible only to the user that created the application. Once the application has been tested, the user can request that an administrator make the application public, which makes the application available to all users via the 3dem.org workbench.

User Management

TACC has a robust, facile process for user enrollment into 3dem.org using the TACC User Portal combined with novel automation developed for these projects. Per-user quotas are not enforced, though user usage is tracked to ensure utilization efficiency.

Every user job, whether running on HPC or Cloud infrastructure, can be monitored for resource use efficiency. Metrics tracked include memory footprint, storage I/O, idle CPU, idle GPU, and network traffic. Historical records of each job are available for users to consult. In addition, TACC staff routinely survey the jobs database looking for usage that was less than optimally efficient. Staff members consult as needed with job owners to improve their computations, ensuring efficient use of federally-supported resources.

Hardware

TACC supports hierarchical, high-performance POSIX storage on the Stockyard and Corral filesystems, backed by TACC's Ranch tape archive for disaster recovery support. This storage is accessible to users on all TACC HPC and cloud computing platforms and used for data storage and analysis. TACC provides high-performance networking and data ingress services that demonstrably supports inbound data transfer rates of >1 TB/day. This is accomplished by operating a dedicated MinIO S3 server on an access node with 100GB/sec external networking to the public Internet. In addition, high-speed SSH, SFTP, and HTTPS access are provided.

TACC supports CPU and GPU compute capacity over the course of this program across Vista, Frontera, Stampede2, Stampede3, Lonestar5, Lonestar6, Jetstream, Jetstream2, and Maverick2 HPC systems [43–47]. This complement of systems allows users to run on a wide array of processor architectures including Intel, AMD, and NVIDIA. Cloud computing capacity is delivered via a combination of on-premises OpenStack and VMware clusters. Data storage, compute, network, and data transfer services operate at an availability of over 99%.

Discussion

We have described the capabilities of the NeuroNex 3dem.org portal, one of the many science gateways that have been deployed from the CEP and Tapis in recent years. The success of 3dem.org in serving its research community demonstrates the value of the template gateway approach. Due to the use of the CEP, the initial deployment of gateways can be rapid, with initial customization made available in less than a month. This efficiency highlights the robustness and adaptability of the template gateway model, which can be tailored swiftly to meet specific research needs. Collaboration was significantly enhanced through 3dem.org, with the impact and utility of the gateway in facilitating research and knowledge dissemination. The gateway's design makes it easy for researchers to locate and access relevant data, integrate it with other systems, and reuse it for various purposes.

Importantly, using a common gateway template is a key part of a long-term sustainability plan. This approach ensures that future CEP and Tapis updates can be shared quickly across all gateways, maintaining consistency and up-to-date functionality across the research community. Future functionality enhancements to 3dem.org can leverage these updates, for example, integrating automated workflows into 3dem.org using Tapis v3 workflow capabilities. Tapis v3 also serves as an API interface to external community resources, including software containers, distributed data stores, and external computing environments, enabling scalable and interoperable execution across platforms.

1
2
3
4 While Tapis supports external system integration, 3dem.org currently operates solely on TACC-hosted resources,
5 with the architecture designed to enable future extensions if needed. Users considering external integrations should
6 account for authentication configuration and data transfer times. There are no data transfer fees incurred on the
7 TACC side.
8

9 Providing the web-based GUI allows for execution of workflows without requiring command-line interaction or
10 knowledge. Users do not need prior experience with schedulers, batch scripts, or container technologies. More
11 advanced HPC familiarity becomes beneficial when users wish to optimize application parameters for large-scale
12 runs, develop and deploy new containerized applications, or integrate custom workflows via Tapis. To support
13 onboarding, 3dem.org provides public tutorials, recorded workshops, example datasets, and responsive user support,
14 allowing users to progressively deepen their engagement with HPC concepts as needed.
15

16 Leveraging HPC allows for scaling in both speed and size. In practice, application performance is influenced
17 primarily by hardware memory constraints and also by data movement. To mitigate this, workflows on 3dem.org are
18 structured to limit parallelism in order to optimize memory per process and to use in-node temporary storage when
19 practical. As datasets grow, bottlenecks most commonly emerge when per-process memory requirements or shared
20 filesystem I/O exceed the capabilities of a single node, rather than from insufficient aggregate compute capacity.
21 Applications designed to scale across multiple nodes can overcome these limitations.
22

23 Adhering to the FAIR principles further supports sustainability by promoting practices that enhance the longevity
24 and usability of the gateway. Long-term stewardship of public datasets on 3dem.org is supported through a
25 combination of curated project spaces, persistent identifiers within the portal, and explicit versioning of both data
26 and associated software artifacts. When public datasets evolve due to, for example, re-alignment or improved
27 segmentation, prior versions are retained, with clear provenance linking derived outputs to their source inputs and
28 software versions. The use of open and widely adopted file formats further reduces the risk of obsolescence as tools
29 and platforms evolve, while containerized applications and shared workflows ensure that historical analyses remain
30 reproducible over time. Community engagement is also central to our roadmap for long-term viability, with ongoing
31 efforts to solicit user feedback and align development priorities with emerging use cases. These strategies aim to
32 reinforce 3dem.org's role as a durable, community-driven gateway for volumetric imaging research.
33
34
35

36 Statements and Declarations

37
38 The authors have no financial or non-financial competing interests to declare that are relevant to the content of this
39 article. This work was funded by National Science Foundation grants DBI-1707356 and DBI-2014862 as part of the
40 NeuroNex program.
41

42 No human or animal research was performed to produce the work detailed in the Methods of this manuscript.
43 Datasets referenced and figures that include brain images rely on animal procedures approved by the Institutional
44 Animal Care and Use Committees (IACUC) at the respective institutions where live animals were handled.
45

46 No data was generated directly for this manuscript. Software described is openly available via GitHub at URLs
47 listed throughout the manuscript.
48

49 All co-authors contributed to the development of this manuscript. AJN, AWW, JAL, JM, JPC, JR, MAC, ST, TB,
50 TM, TMB, VVT, WY contributed text and figures. HP, JPC, JR, JSt, MD, ST, TMB implemented and maintained the
51 CEP instance, Tapis instance, and landing pages for 3dem.org. EF, JAL, JPC, JSo, TH, WJA integrated applications
52 into 3dem.org using Tapis. AJN, JAL, JM, JPC, LK, MAC, MK, TMB, VVT, WY tested 3dem.org and provided
53 feedback. KMH, JPC, TS led the funded research in support of 3dem.org.
54
55

56 References

- 57
58 1. Carson J, Fonner J, Vaughn M, Niall G, Brown T, Stubbs J, et al. Facilitating collaborative science through
59 portals connected to high-performance computing. Science Gateways 2023. Pittsburgh, PA: Zenodo; 2023. p. 4.
60 <https://doi.org/10.5281/ZENODO.10034840>.
61
62
63
64
65

2. Stubbs J, Cardone R, Packard M, Jamthe A, Padhy S, Terry S, et al. Tapis: An API Platform for Reproducible, Distributed Computational Research. In: Arai K, editor. *Advances in Information and Communication*, vol. 1363. Cham: Springer International Publishing; 2021. pp. 878–900. https://doi.org/10.1007/978-3-030-73100-7_61.
3. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018. <https://doi.org/10.1038/sdata.2016.18>.
4. Litvina E, Adams A, Barth A, Bruchez M, Carson J, Chung JE, et al. BRAIN Initiative: Cutting-Edge Tools and Resources for the Community. *J Neurosci*. 2019;39:8275–84. <https://doi.org/10.1523/JNEUROSCI.1169-19.2019>.
5. Harris KM, Spacek J, Bell ME, Parker PH, Lindsey LF, Baden AD, et al. A resource from 3D electron microscopy of hippocampal neuropil for user training and tool development. *Sci Data*. 2015;2:150046. <https://doi.org/10.1038/sdata.2015.46>.
6. Fang L, Monroe F, Novak SW, Kirk L, Schiavon CR, Yu SB, et al. Deep learning-based point-scanning super-resolution imaging. *Nat Methods*. 2021;18:406–16. <https://doi.org/10.1038/s41592-021-01080-z>.
7. Kulik YD, Watson DJ, Cao G, Kuwajima M, Harris KM. Structural plasticity of dendritic secretory compartments during LTP-induced synaptogenesis. *eLife*. 2019;8:e46356. <https://doi.org/10.7554/eLife.46356>.
8. Harris KM, Hubbard DD, Kuwajima M, Abraham WC, Bourne JN, Bowden JB, et al. Dendritic Spine Density Scales with Microtubule Number in Rat Hippocampal Dendrites. *Neuroscience*. 2022;489:84–97. <https://doi.org/10.1016/j.neuroscience.2022.02.021>.
9. Kirk L, Garcia G, Hanka D, Zatyko K, Bartol T, Sejnowski T, et al. Presynaptic vesicles supply membrane for axonal bouton enlargement during LTP. 2025. <https://doi.org/10.1101/2025.04.29.651313>.
10. Nam AJ, Kuwajima M, Parker PH, Bowden JB, Abraham WC, Harris KM. Perisynaptic astroglial response to *in vivo* long-term potentiation and concurrent long-term depression in the hippocampal dentate gyrus. 2025. <https://doi.org/10.1101/2025.05.13.653827>.
11. Jung JH, Kirk LM, Bourne JN, Harris KM. Shortened tethering filaments stabilize presynaptic vesicles in support of elevated release probability during LTP in rat hippocampus. *Proc Natl Acad Sci USA*. 2021;118:e2018653118. <https://doi.org/10.1073/pnas.2018653118>.
12. Wetzel AW, Bakal J, Dittrich M, Hildebrand DGC, Morgan JL, Lichtman JW. Registering large volume serial-section electron microscopy image sets for neural circuit reconstruction using FFT signal whitening. 2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR). Washington, DC, USA: IEEE; 2016. pp. 1–10. <https://doi.org/10.1109/AIPR.2016.8010595>.
13. Chirillo MA, Falco JN, Musslewhite MD, Lindsey LF, Harris KM. PyReconstruct: A fully open-source, collaborative successor to Reconstruct. *Proc Natl Acad Sci USA*. 2025;122:e2505822122. <https://doi.org/10.1073/pnas.2505822122>.
14. Fiala JC. *Reconstruct*: a free editor for serial section microscopy. *Journal of Microscopy*. 2005;218:52–61. <https://doi.org/10.1111/j.1365-2818.2005.01466.x>.
15. Castro RW, Lopes MC, De Biase LM, Valdez G. Aging spinal cord microglia become phenotypically heterogeneous and preferentially target motor neurons and their synapses. *Glia*. 2024;72:206–21. <https://doi.org/10.1002/glia.24470>.
16. Djama D, Zirpel F, Ye Z, Moore G, Chue C, Edge C, et al. The type of inhibition provided by thalamic interneurons alters the input selectivity of thalamocortical neurons. *Current Research in Neurobiology*. 2024;6:100130. <https://doi.org/10.1016/j.crneur.2024.100130>.
17. Haruwaka K, Ying Y, Liang Y, Umpierre AD, Yi M-H, Kremen V, et al. Microglia enhance post-anesthesia neuronal activity by shielding inhibitory synapses. *Nat Neurosci*. 2024;27:449–61. <https://doi.org/10.1038/s41593-023-01537-8>.
18. Holl D, Hau WF, Julien A, Banitalebi S, Kalkitsas J, Savant S, et al. Distinct origin and region-dependent contribution of stromal fibroblasts to fibrosis following traumatic injury in mice. *Nat Neurosci*. 2024;27:1285–98. <https://doi.org/10.1038/s41593-024-01678-4>.
19. Joyce MKP, Wang J, Barbas H. Subgenual and Hippocampal Pathways in Amygdala Are Set to Balance Affect and Context Processing. *J Neurosci*. 2023;43:3061–80. <https://doi.org/10.1523/JNEUROSCI.2066-22.2023>.
20. Morozov YM, Rakic P. Disorder of Golgi Apparatus Precedes Anoxia-Induced Pathology of Mitochondria. *IJMS*. 2023;24:4432. <https://doi.org/10.3390/ijms24054432>.
21. Ziółkowska M, Borczyk M, Cały A, Tomaszewski KF, Nowacka A, Nalberczak-Skóra M, et al. Phosphorylation of PSD-95 at serine 73 in dCA1 is required for extinction of contextual fear. *PLoS Biol*. 2023;21:e3002106. <https://doi.org/10.1371/journal.pbio.3002106>.

- 1
2
3
4 22. Amemiya S, Omori A, Tsurugaya T, Hibino T, Yamaguchi M, Kuraishi R, et al. Early stalked stages in ontogeny
5 of the living isocrinid sea lily *M etacrinus rotundus*. *Acta Zoologica*. 2016;97:102–16.
6 <https://doi.org/10.1111/azo.12109>.
- 7 23. Parke JL, Oh E, Voelker S, Hansen EM, Buckles G, Lachenbruch B. *Phytophthora ramorum* Colonizes Tanoak
8 Xylem and Is Associated with Reduced Stem Water Transport. *Phytopathology®*. 2007;97:1558–67.
9 <https://doi.org/10.1094/PHYTO-97-12-1558>.
- 10 24. Groh C, Lu Z, Meinertzhagen IA, Rössler W. Age-related plasticity in the synaptic ultrastructure of neurons in
11 the mushroom body calyx of the adult honeybee *Apis mellifera*. *J of Comparative Neurology*. 2012;520:3509–27.
12 <https://doi.org/10.1002/cne.23102>.
- 13 25. Seid MA, Junge E. Social isolation and brain development in the ant *Camponotus floridanus*. *Sci Nat*.
14 2016;103:42. <https://doi.org/10.1007/s00114-016-1364-1>.
- 15 26. Stökl J, Herzner G. Morphology and ultrastructure of the allomone and sex-pheromone producing mandibular
16 gland of the parasitoid wasp *Leptopilina heterotoma* (Hymenoptera: Figitidae). *Arthropod Structure &*
17 *Development*. 2016;45:333–40. <https://doi.org/10.1016/j.asd.2016.06.003>.
- 18 27. Moore BC, Mathavan K, Guillette LJ. Morphology and Histochemistry of Juvenile Male American Alligator (*Alligator mississippiensis*) Phallus. *The Anatomical Record*. 2012;295:328–37. <https://doi.org/10.1002/ar.21521>.
- 19 28. Baumgart M, Wiśniewski M, Grzonkowska M, Badura M, Dombek M, Małkowski B, et al. Morphometric study
20 of the two fused primary ossification centers of the clavicle in the human fetus. *Surg Radiol Anat*. 2016;38:937–
21 45. <https://doi.org/10.1007/s00276-016-1640-y>.
- 22 29. Dinnis CM, Dahle AK, Taylor JA. Three-dimensional analysis of eutectic grains in hypoeutectic Al–Si alloys.
23 *Materials Science and Engineering: A*. 2005;392:440–8. <https://doi.org/10.1016/j.msea.2004.10.037>.
- 24 30. Perkel JM. Python power-up: new image tool visualizes complex data. *Nature*. 2021;600:347–8.
25 <https://doi.org/10.1038/d41586-021-03628-7>.
- 26 31. Sofroniew N, Lambert T, Bokota G, Nunez-Iglesias J, Sobolewski P, Sweet A, et al. napari: a multi-dimensional
27 image viewer for Python. 2025. <https://doi.org/10.5281/ZENODO.15465370>.
- 28 32. Manor U, Thiyagarajan V, Harris K, Sheridan A. Sparse Annotation is Sufficient for Bootstrapping Dense Neuron
29 Segmentation. 2024. <https://doi.org/10.21203/rs.3.rs-5339143/v1>.
- 30 33. Archit A, Freckmann L, Nair S, Khalid N, Hilt P, Rajashekar V, et al. Segment Anything for Microscopy. *Nat*
31 *Methods*. 2025;22:579–91. <https://doi.org/10.1038/s41592-024-02580-4>.
- 32 34. Stites HC, Manor U. PSSR2: a user-friendly Python package for democratizing deep learning-based point-
33 scanning super-resolution microscopy. *BMC Methods*. 2025;2:1. <https://doi.org/10.1186/s44330-024-00020-5>.
- 34 35. Husar A, Ordyan M, Garcia GC, Yancey JG, Saglam AS, Faeder JR, et al. MCell4 with BioNetGen: A Monte
35 Carlo simulator of rule-based reaction-diffusion systems with Python interface. *PLoS Comput Biol*.
36 2024;20:e1011800. <https://doi.org/10.1371/journal.pcbi.1011800>.
- 37 36. Bartol TM, Keller DX, Kinney JP, Bajaj CL, Harris KM, Sejnowski TJ, et al. Computational reconstitution of
38 spine calcium transients from individual proteins. *Front Synaptic Neurosci*. 2015;7:17.
39 <https://doi.org/10.3389/fnsyn.2015.00017>.
- 40 37. Nadkarni S, Bartol TM, Sejnowski TJ, Levine H. Modelling vesicular release at hippocampal synapses. *PLoS*
41 *Comput Biol*. 2010;6:e1000983. <https://doi.org/10.1371/journal.pcbi.1000983>.
- 42 38. Cardona A, Saalfeld S, Schindelin J, Arganda-Carreras I, Preibisch S, Longair M, et al. TrakEM2 Software for
43 Neural Circuit Reconstruction. *PLoS ONE*. 2012;7:e38011. <https://doi.org/10.1371/journal.pone.0038011>.
- 44 39. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform
45 for biological-image analysis. *Nat Methods*. 2012;9:676–82. <https://doi.org/10.1038/nmeth.2019>.
- 46 40. Ahrens J, Geveci B, Law C. Paraview: An end-user tool for large data visualization. *The visualization handbook*.
47 2005.
- 48 41. The MathWorks Inc. MATLAB. 2024.
- 49 42. Kluyver Thomas, Ragan-Kelley Benjamin, Perez Fernando, Granger Brian, Bussonnier Matthias, Frederic
50 Jonathan, et al. Jupyter Notebooks - a publishing format for reproducible computational workflows. *Positioning*
51 *and Power in Academic Publishing: Players, Agents and Agendas*. IOS Press; 2016. [https://doi.org/10.3233/978-](https://doi.org/10.3233/978-1-61499-649-1-87)
52 [1-61499-649-1-87](https://doi.org/10.3233/978-1-61499-649-1-87).
- 53 43. Hancock DY, Fischer J, Lowe JM, Snapp-Childs W, Pierce M, Marru S, et al. Jetstream2: Accelerating cloud
54 computing via Jetstream. *Practice and Experience in Advanced Research Computing*. Boston MA USA: ACM;
55 2021. pp. 1–8. <https://doi.org/10.1145/3437359.3465565>.
- 56 44. Stanzione D, Barth B, Gaffney N, Gaither K, Hempel C, Minyard T, et al. Stampede 2: The Evolution of an
57 XSEDE Supercomputer. *Practice and Experience in Advanced Research Computing 2017: Sustainability,*
58 *Success and Impact*. New Orleans LA USA: ACM; 2017. pp. 1–8. <https://doi.org/10.1145/3093338.3093385>.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

45. Stanzione D, West J, Evans RT, Minyard T, Ghattas O, Panda DK. Frontera: The Evolution of Leadership Computing at the National Science Foundation. *Practice and Experience in Advanced Research Computing*. Portland OR USA: ACM; 2020. pp. 106–11. <https://doi.org/10.1145/3311790.3396656>.

46. Stewart CA, Turner G, Vaughn M, Gaffney NI, Cockerill TM, Foster I, et al. Jetstream: a self-provisioned, scalable science and engineering cloud environment. *Proceedings of the 2015 XSEDE Conference on Scientific Advancements Enabled by Enhanced Cyberinfrastructure - XSEDE '15*. St. Louis, Missouri: ACM Press; 2015. pp. 1–8. <https://doi.org/10.1145/2792745.2792774>.

47. Ruhela A, Cazes J, McCalpin J, Del-Castillo-Negrete C, Li J, Liu H, et al. Performance Analysis of Scientific Applications on an NVIDIA Grace System. *SC24-W: Workshops of the International Conference for High Performance Computing, Networking, Storage and Analysis*. Atlanta, GA, USA: IEEE; 2024. pp. 558–66. <https://doi.org/10.1109/SCW63240.2024.00078>.