



Introduction to the Brain Image Library:

Designated Repository for NIH BRAIN Initiative Microscopy Data



Marcel Bruchez (PI)
Greg Fisher (Microscope)



Alexander Ropelewski (Contact PI)
Kathy Benninger (Networking)
Greg Hood (Image Analysis+ HPC)
Derek Simmel (Systems+Data)
Arthur Wetzel (Image Analysis)
Luke Tuite (User Support+Web)



Simon Watkins (PI)
Alan Watson (Microscope)



www.brainimagelibrary.org

bil-support@psc.edu

Workshop Agenda

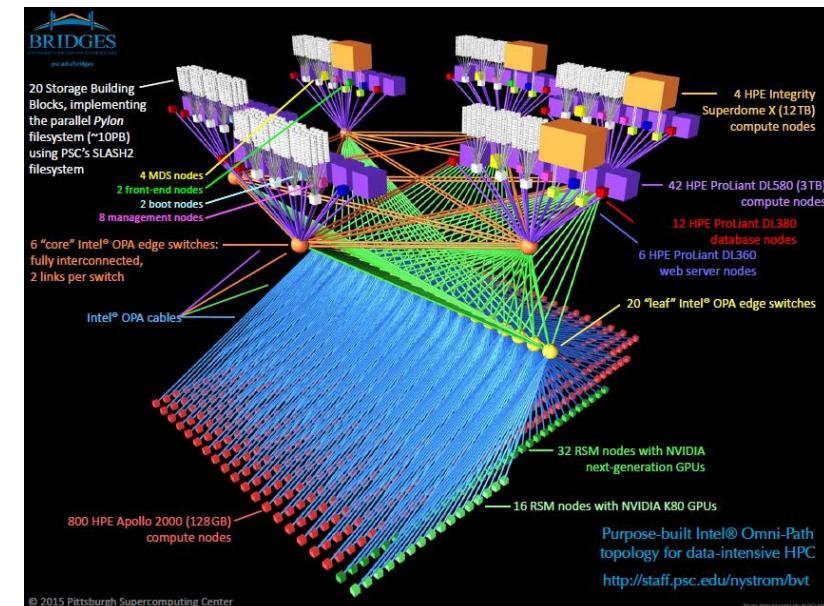
- What is the Brain Image Library:
 - The mission and scope of the archive
 - How to access/getting started
 - Who to contact/how to get additional information
 - Input on use cases/planned future features
- Exploring data in BIL:
 - Finding, accessing, downloading and using contributed data
 - Tools to interact with the data:
 - Input on use cases/planned future features
- Contributing data to BIL:
 - How to submit your image data to BIL
 - Input on data contribution/pre-submission processing
- Open time with BIL Staff

The Brain Image Library

Mission: National public resource enabling researchers to deposit, analyze, mine, share and interact with microscopy datasets of the brain.

Scope:

- Permanent repository for high-quality brain microscopy datasets
 - Whole brain images of mouse, rat, other mammals and model organisms
 - Targeted experiments Including connectivity between cells and spatial transcriptomics (*FISH)
 - Historical collections
- Provide HPC computing capability local to the data for pre-submission data processing and post-submission exploration
 - Enclave access to pre-release data
 - Research access to restricted-access, secured data
- Provide user access and support



Designated repository for NIH BRAIN Initiative Data

- Requirements for investigators outlined in NIH NOT-MH-19-010 **Notice of Data Sharing Policy for the BRAIN Initiative**
- BRAIN Initiative applications after March 1, 2020, are required to share the data they collect
- General expectation is that data will be submitted to the archives every 6 months
- Data will be shared with the research community no later than:
 - When papers using the data have been published
 - The end of the award period

Notice of Data Sharing Policy for the BRAIN Initiative

Notice Number: NOT-MH-19-010

Key Dates

Release Date: January 22, 2019

Related Announcements

[RFA-EY-18-001](#)

[RFA-MH-18-010](#)

[RFA-MH-18-135](#)

[RFA-MH-18-136](#)

[RFA-MH-18-147](#)

[RFA-MH-18-400](#)

[RFA-NS-18-018](#)

[RFA-NS-18-019](#)

[RFA-NS-18-020](#)

[RFA-NS-18-021](#)

[RFA-NS-18-022](#)

[RFA-NS-18-023](#)

[RFA-NS-18-036](#)

[RFA-NS-19-001](#)

[RFA-NS-19-002](#)

[RFA-NS-19-003](#)

Issued by

National Institute of Mental Health (NIMH)

National Eye Institute (NEI)

National Institute on Aging (NIA)

National Institute on Alcohol Abuse and Alcoholism (NIAAA)

National Institute of Biomedical Imaging and Bioengineering (NIBIB)

Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)

National Institute on Deafness and Other Communication Disorders (NIDCD)

National Institute on Drug Abuse (NIDA)

National Institute of Neurological Disorders and Stroke (NINDS)

National Center for Complementary and Integrative Health (NCCIH)

Purpose

The purpose of this notice is to inform prospective applicants and current awardees of a new policy concerning data collected with support from awards that are funded by the BRAIN Initiative. Specifically, this Notice clarifies the expectation that applicants to BRAIN Initiative funding opportunity announcements: 1) submit their data to one of the BRAIN data archives for sharing; 2) include specific required elements in the Resource Sharing Plan as further detailed below; and 3) include costs attributed to data preparation and submission to a data archive in grant applications.

BRAIN Initiative Overview

The Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative® is aimed at revolutionizing our understanding of the human brain. By accelerating the development and application of innovative technologies, researchers will be able to produce a new dynamic picture of the brain that, for the first time, will show how individual cells and complex neural circuits interact in both time and space. It is expected that the application of these new tools and technologies will ultimately lead to new ways to treat and prevent brain disorders.

NIH is one of several federal agencies involved in the BRAIN Initiative. Planning for the NIH component of the BRAIN initiative is guided by the long-term scientific plan "BRAIN 2025: A Scientific Vision," which details seven high-priority research areas and calls for a sustained federal commitment of \$4.5 billion over 12 years. This Notice and related FOAs issued as part of the BRAIN initiative are based on careful consideration by the NIH of the recommendations of the BRAIN 2025 Report, and input from the NIH BRAIN Multi-Council Working Group. Videocasts of the NIH BRAIN Multi-council Working Group are available at <http://www.braininitiative.nih.gov/about/mcwg.htm>.

To enable rapid progress in development of new technologies as well as in theory and data analysis, the BRAIN Initiative encourages collaborations between neurobiologists and scientists from statistics, physics, mathematics, engineering, and computer and information sciences.

BRAIN Initiative Data Sharing

The BRAIN 2025 report calls for establishing platforms for sharing data related to the BRAIN Initiative. In response, NIH has released RFAs related to data archives (RFA-MH-19-145 and RFA-MH-17-255), data standards (RFA-MH-19-146 and RFA-MH-17-256), and software to integrate and analyze data (RFA-MH-19-147 and RFA-MH-17-257). Each of those efforts is focused on data from a "sub-domain" which are usually defined by the data collection methodology.

Data archives that have been established include:

1) The Neuroscience Multi-omic Data Archive (<https://neuroarchive.org/about.php>, R24MH114788) to hold data from -omics experiments.

2) The Brain Image Library (<http://www.brainimagelibrary.org/index.html>, R24MH114793) to hold microscopy data.

3) Data Archive for the BRAIN Initiative (<https://dabi.loni.usc.edu>, R24MH114796) to hold data related to human electrophysiology experiments.

4) OpenNeuro (<https://openneuro.org/>, R24MH117179) to hold magnetic resonance imaging data.

5) Block and Object Storage Service (<https://bossdb.org/>, R24MH114785) to hold electron microscopy data.

A few additional data archives in distinct sub-domains are expected to be funded. The complete list of BRAIN Initiative infrastructure awards can be found [here](#). The awardees for the RFAs

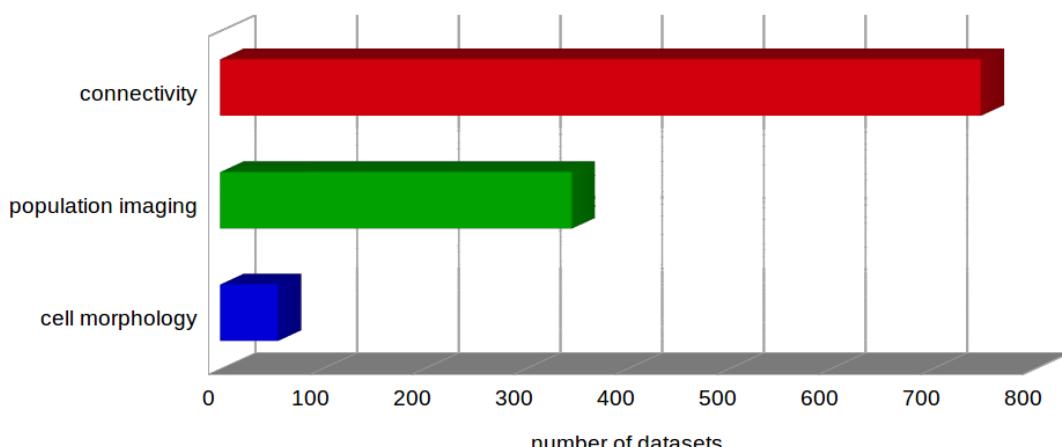


Publicly Accessible Data

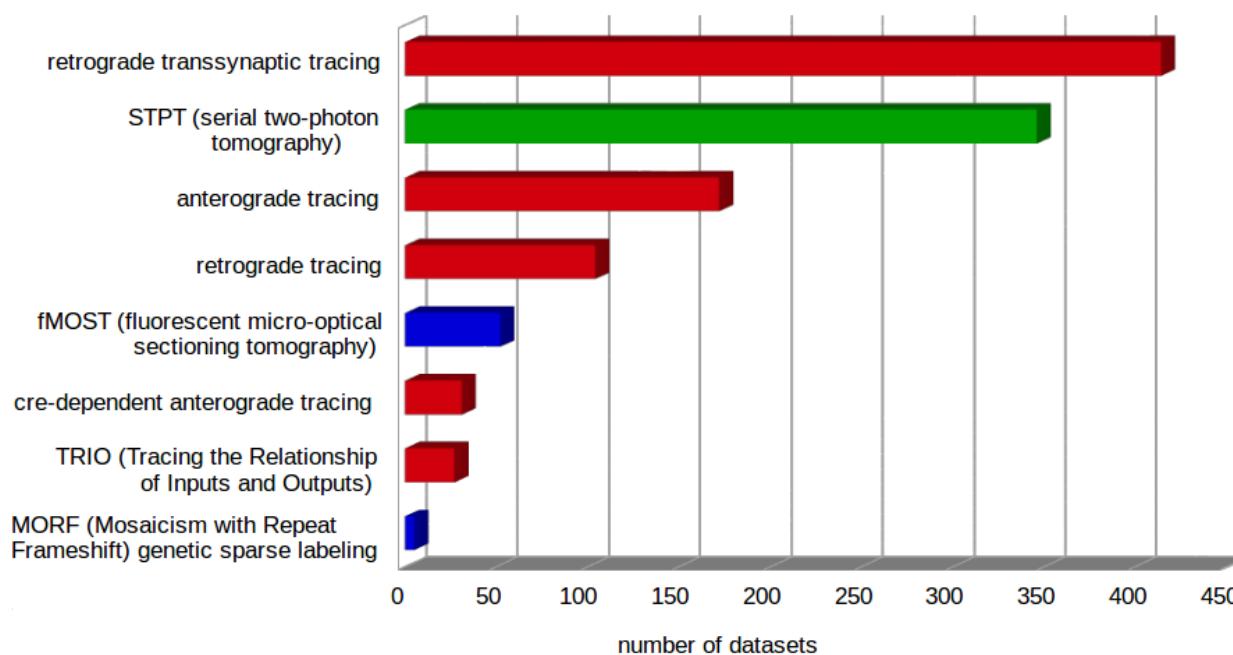
Over 1100 public mouse brain datasets are available at BIL



Modalities of datasets in BIL



Techniques used to generate datasets in BIL



Notable Data Contributors

Hongkui Zeng
Allen Institute for
Brain Science



Hongwei Dong
University of
California

Guoqiang Bi
University of
Science and
Technology
China



Z. Josh Huang
Cold Spring Harbor
Laboratory



Pavel Osten
Cold Spring Harbor
Laboratory



X. William Yang
University of
California Los
Angeles

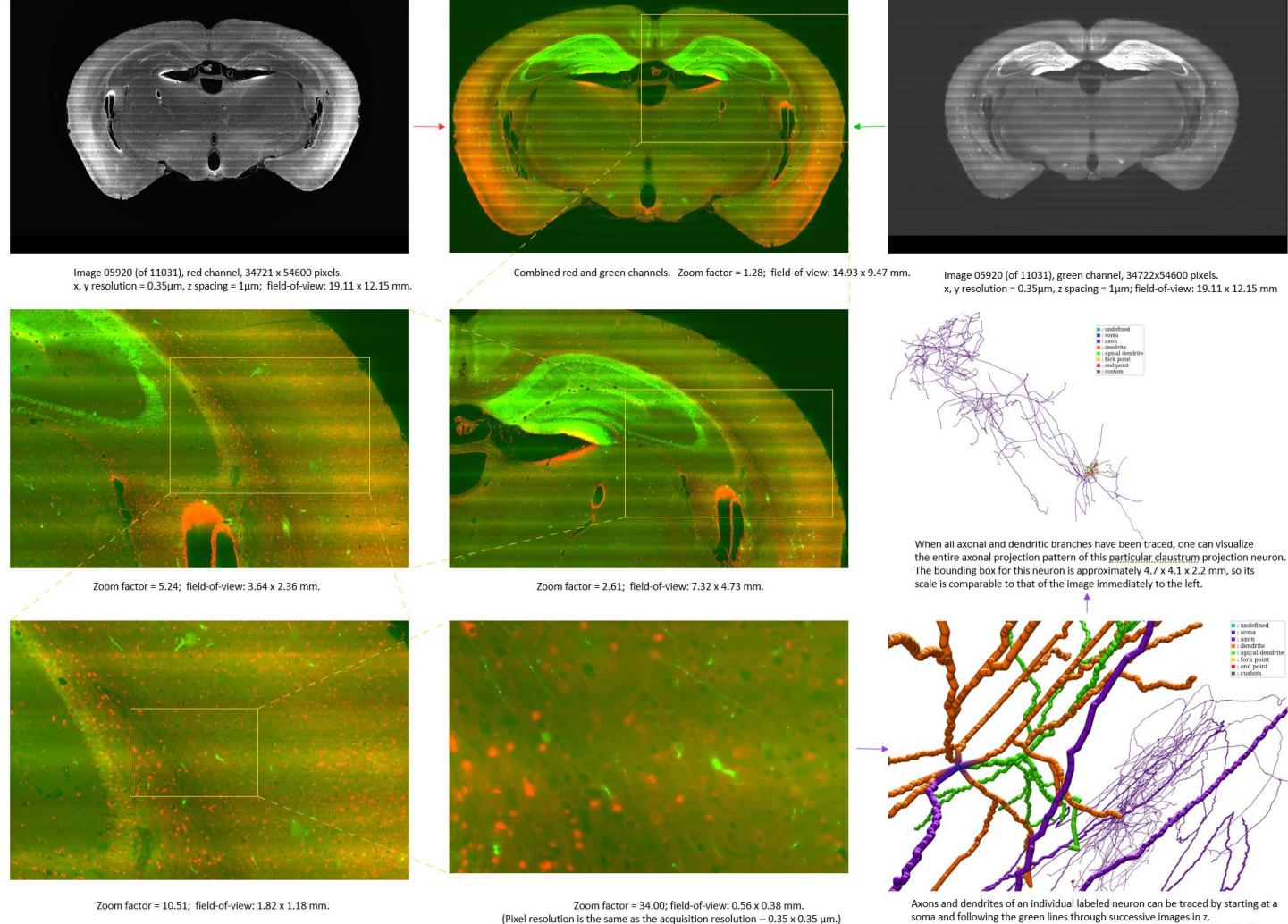
Exemplar Dataset



Contributor:
Hongkui Zeng
Allen Institute for Brain Science

fMOST (fluorescence micro-optical sectioning tomography) dataset showing one of several neuron morphologies reconstructed from this dataset.

We feature in this poster one of the 34 fMOST datasets received from the Zeng U19 project at the Allen Institute for Brain Science. Imaging was performed at Huazhong University of Science and Technology. This dataset represents whole-brain imaging of a transgenic mouse (Gnb4-IRES2-CreERT2/wt;Ai82(TIT2L-GFP)/Ai140(TIT2L-GFP-ICL-TA2)) in which a small sampling of neurons, including claustrum projection neurons, are fluorescently labeled [4]. The green channel shows where GFP is expressed, and can be used for tracing these neurons. It is important that only a sparse subset be labeled; otherwise, it would become much more difficult to trace individual neurons with high confidence. For this specimen BIL currently has available 4 neuron morphologies in .swc format, of which we illustrate one (17781_00001.swc). Potentially, thousands of other morphologies could be reconstructed from this dataset.



Axons and dendrites of an individual labeled neuron can be traced by starting at a soma and following the green lines through successive images in z. This figure was generated from the 17781_00001.swc file uploaded to BIL [7] and rendered with the SharkViewer tool [8].

Computational Services

- Computational Systems to process and explore BIL data in-place
available at no charge for open research and to support courses

BIL Hardware:

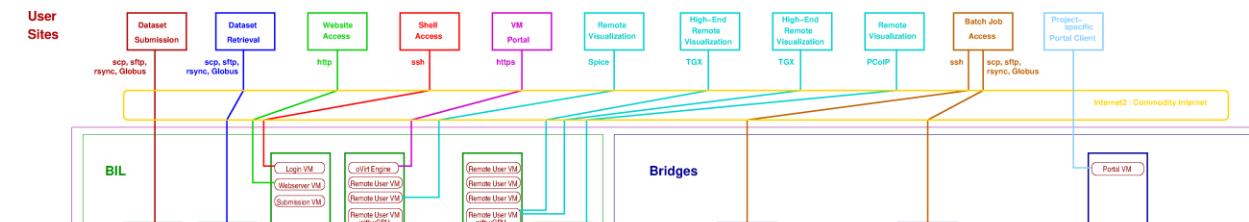
- VM computational system:
 - Remote Desktop Capable
 - NVIDIA V100 GPU's (shared)
- 4 PB storage (will grow to 10PB)
- Tape drives capable of receiving data on LTO 7/8 tapes
- "Brainball": 30TB portable drive systems
- Dedicated data transfer nodes

Bridges (HPC System):

- 28,628 Intel Xeon cores
- GPUS:
 - 64 NVIDIA K80
 - 64 NVIDIA P100
- 17 PB storage:
 - 10PB persistent
 - 7.3PB node-local

Nodes:

- 800x128GB RAM
- 42x3TB RAM
- 4x12TB RAM
- 88 NVIDIA Volta GPUs:
 - 9 Nodes**, with 8 NVIDIA V100
 - 1 NVIDIA DGX-2**, 16 Volta with 2.4 TB/s NVSwitch

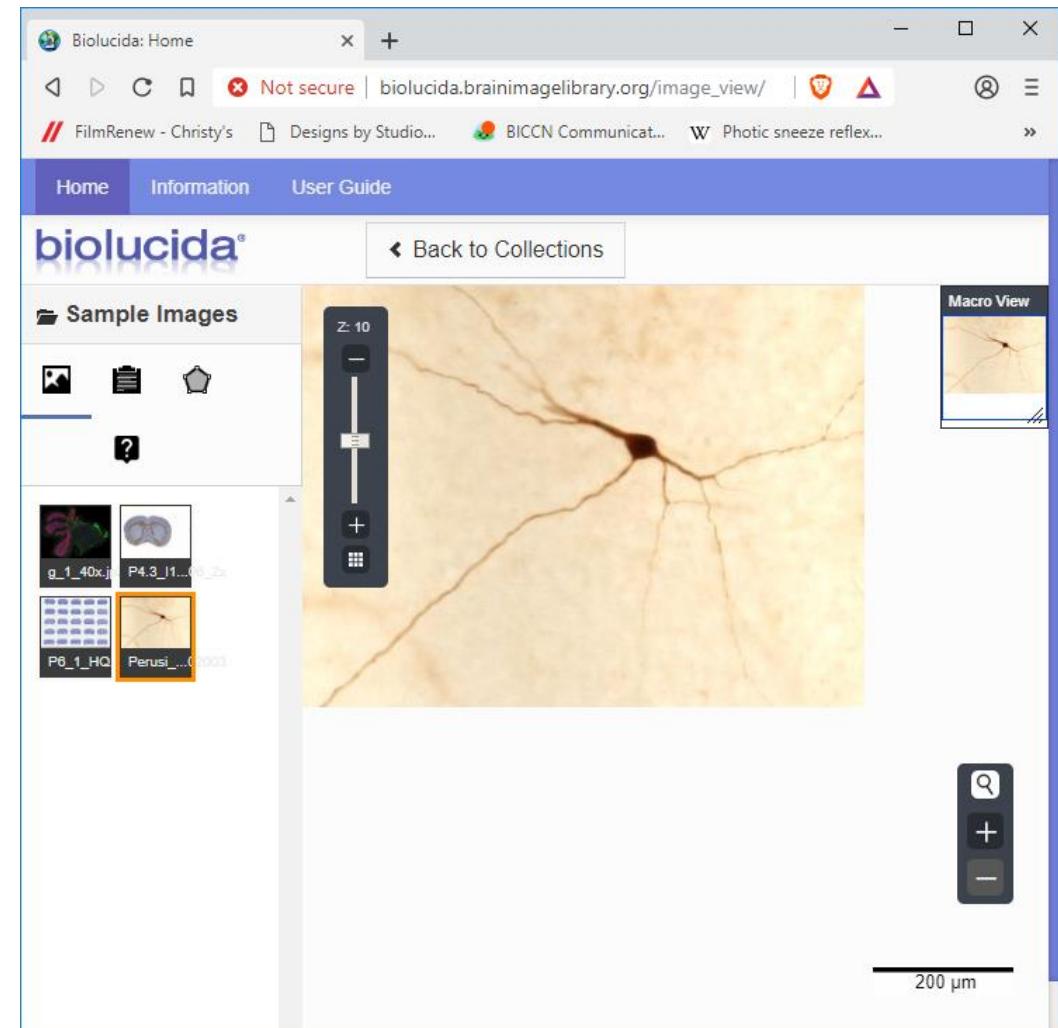


Coming Soon:

- Bridges-II: New \$10,000,000 HPC System arriving in June 2020

Web Portal Hosting

- Access data in unique ways
 - Access data in-place
 - Commercial software capable
 - Can take advantage of HPC to compute/deliver custom results on-the-fly



Help Desk Services

- Contact
 - Please email: bil-support@psc.edu.
 - Time-critical, call PSC hotline at: 412-268-6350.
 - Please note that we are located in the Eastern Time zone.
- Help Desk Scope:
 - Answer questions
 - Facilitate software installation to enable processing and visualization.
 - Networking and data transfer support
 - Diagnose and facilitate resolution of last mile and routing problems

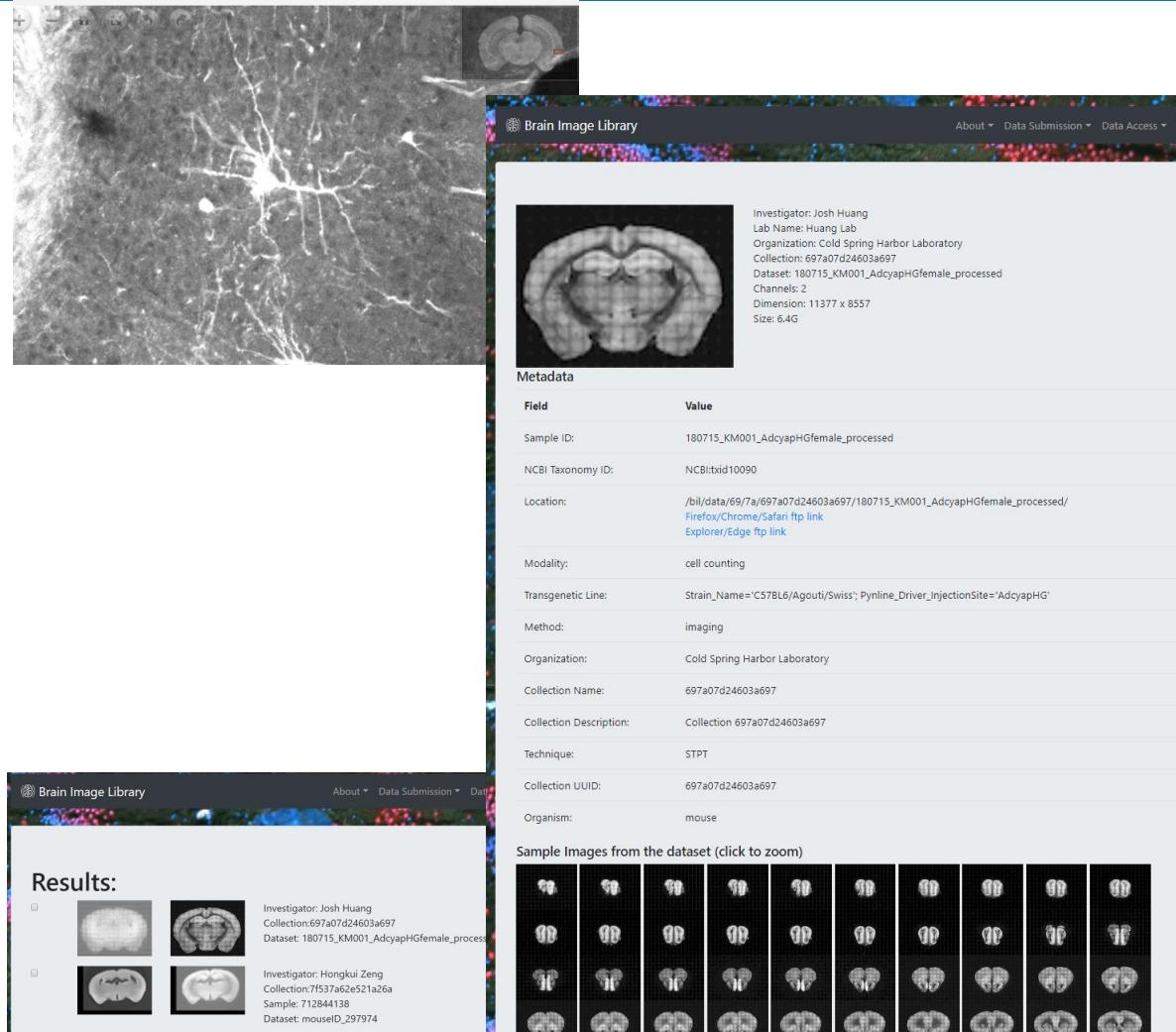
How to get started - Important Documentation

- BIL website: www.brainimagelibrary.org
- Locating public BIL data:
 - <http://www.brainimagelibrary.org/datasets.html>
- Downloading public BIL data:
 - <http://www.brainimagelibrary.org/download.html>
- Compute on the data without downloading:
 - <http://www.brainimagelibrary.org/compute.html>
- Contribute data to BIL:
 - <http://www.brainimagelibrary.org/submission.html>
- Help:
 - <http://www.brainimagelibrary.org/contact.html>
 - email: bil-support@psc.edu or call: 412-268-6350



Help Us Meet Your Needs

- How can we make BIL more accessible and usable by the research community?
 - What information would you expect to find within three mouse clicks?
 - What use cases should be supported?
 - What external resources should be linked to? Why?
 - What exploration & visualization software should be supported?
- Training/outreach needs and uses
- Feedback on future directions as well as general feedback



Data Contributors:

- Understanding your future computational and data storage needs including tools/software that are needed for your work



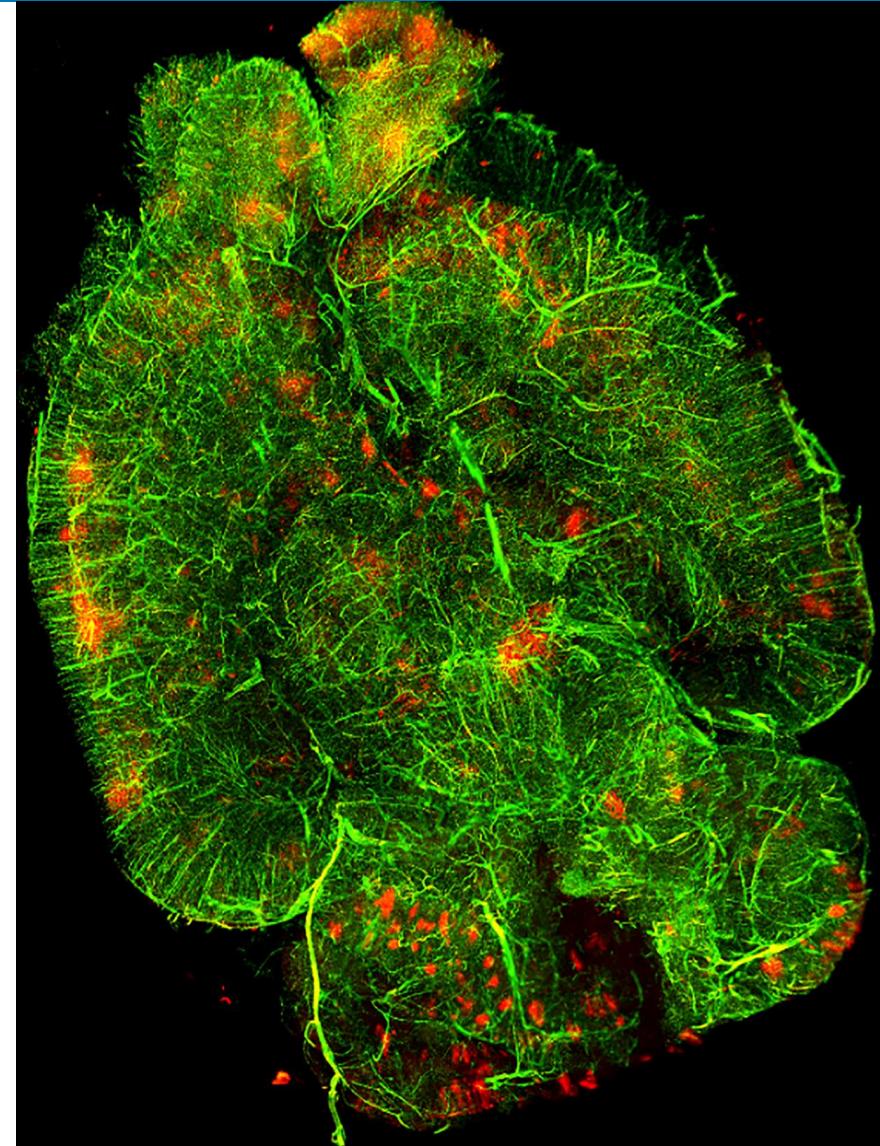
Exploring Data in the Brain Image Library

www.brainimagelibrary.org

bil-support@psc.edu

Wouldn't it be great if.....

- We had a collection of thousands of large, high-quality, high-resolution microscopy datasets of the brain...
- And this collection was attached to significant high performance computing resources...



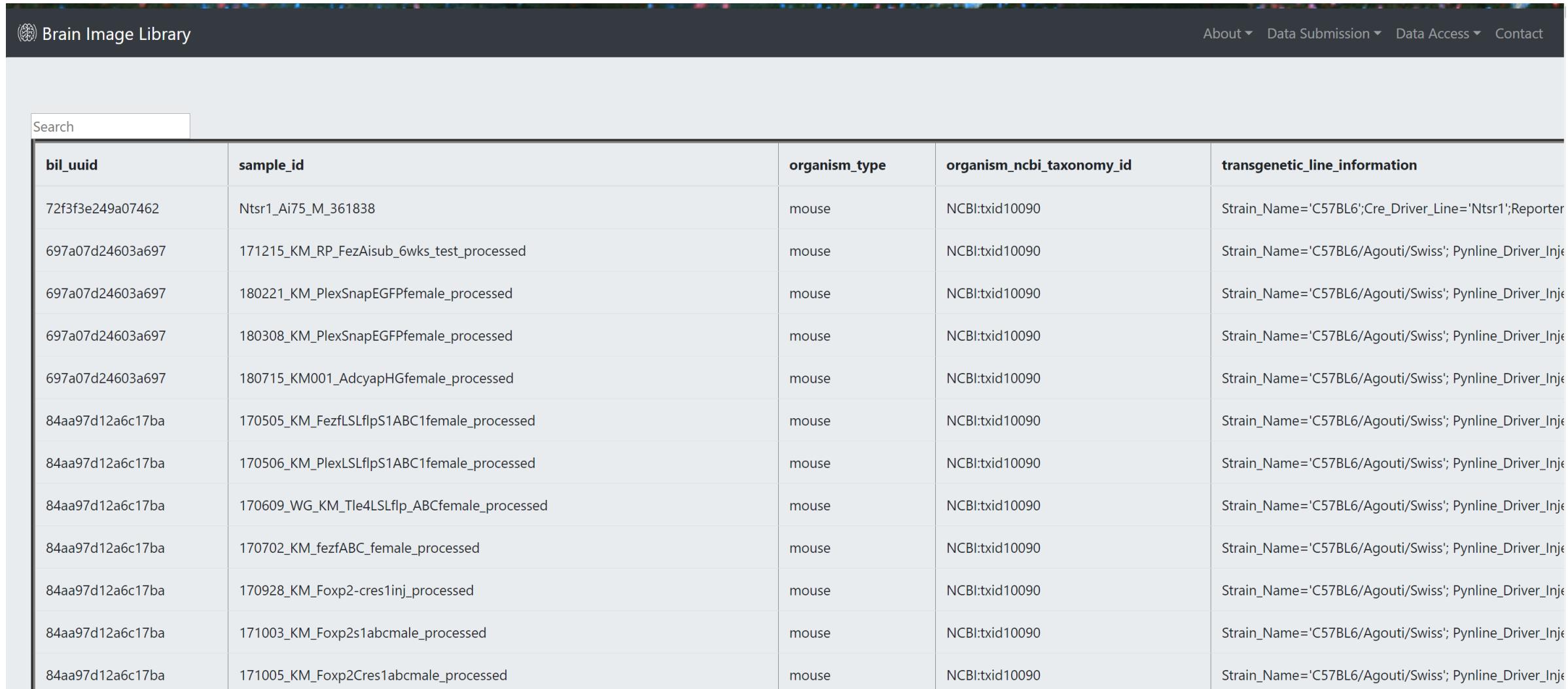
Watson et. al. <https://doi.org/10.1371/journal.pone.0180486>

Then we might be able to.....

- Confirm that neurons grown in tissue culture are “similar” to neurons found in whole brains. (Do they make the right connections with their neighbors?)
- Have sufficient training and validation sets to generate realistic models of key cellular-level brain structures
- Study variation across genetically similar (inbred) populations
- Develop special purpose maps joining the work from individual laboratories to elucidate broader connectivity patterns of localized circuits.
- Do several other interesting things.....

Finding Data: BIL Data Inventory

www.brainimagerlibrary.org Menu: Data Access -> Brain Inventory



The screenshot shows the Brain Image Library website with a search bar at the top. Below the search bar is a table with 12 rows of data. The columns are labeled: bil_uuid, sample_id, organism_type, organism_ncbi_taxonomy_id, and transgenic_line_information.

bil_uuid	sample_id	organism_type	organism_ncbi_taxonomy_id	transgenic_line_information
72f3f3e249a07462	Ntsr1_Ai75_M_361838	mouse	NCBI:txid10090	Strain_Name='C57BL6';Cre_Driver_Line='Ntsr1';Reporter_Gene='Ai75';
697a07d24603a697	171215_KM_RP_FezAisub_6wks_test_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Rear_Paws';
697a07d24603a697	180221_KM_PlexSnapEGFPfemale_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Brain';
697a07d24603a697	180308_KM_PlexSnapEGFPfemale_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Brain';
697a07d24603a697	180715_KM001_AdcyapHGfemale_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Brain';
84aa97d12a6c17ba	170505_KM_FezfLSLflpS1ABC1female_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Brain';
84aa97d12a6c17ba	170506_KM_PlexLSLflpS1ABC1female_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Brain';
84aa97d12a6c17ba	170609_WG_KM_Tle4LSLflp_ABCfemale_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Brain';
84aa97d12a6c17ba	170702_KM_fefzABC_female_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Brain';
84aa97d12a6c17ba	170928_KM_Foxp2-cres1inj_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Brain';
84aa97d12a6c17ba	171003_KM_Foxp2s1abcmale_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Brain';
84aa97d12a6c17ba	171005_KM_Foxp2Cres1abcmale_processed	mouse	NCBI:txid10090	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_Injection_Site='Brain';

Finding Data: Future

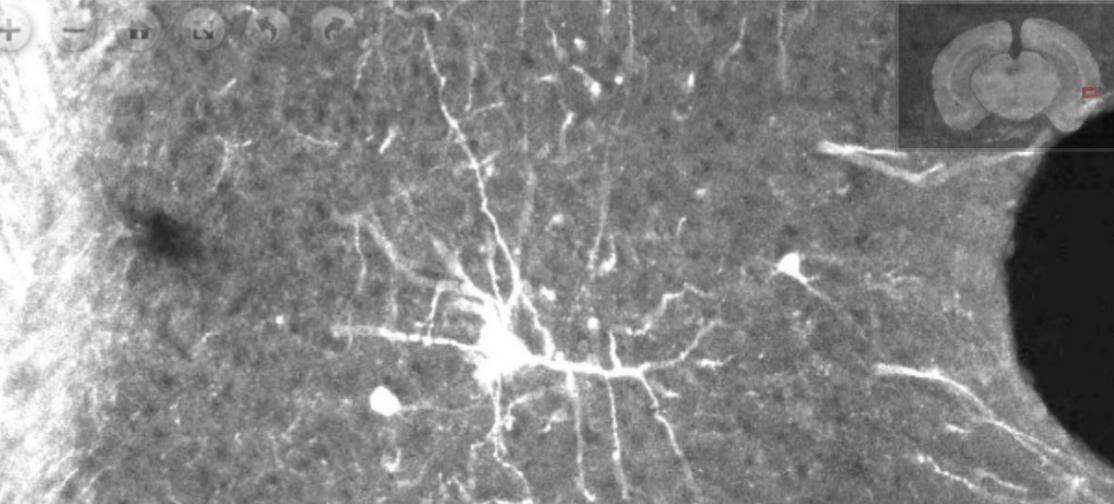
Brain Image Library

About ▾ Data Submission ▾ Data Access ▾ Contact

Results:

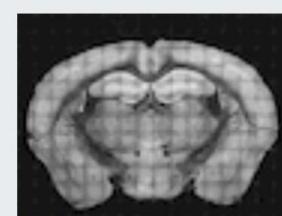
- Investigator: Josh Huang
Collection: 697a07d24603a697
Dataset: 180715_KM001_AdcyapHGfemale_processed
- Investigator: Hongkui Zeng
Collection: 7f537a62e521a26a
Sample: 712844138
Dataset: mouseID_297974
- Investigator: Hongwei Dong
Collection: 115bda6eadcd6507c
Sample: SW180709-02A
Dataset: SW180709-02A

[View Info](#)



Brain Image Library

About ▾ Data Submission ▾ Data Access ▾ Contact



Metadata

Field	Value
Sample ID:	180715_KM001_AdcyapHGfemale_processed
NCBI Taxonomy ID:	NCBItaxid10090
Location:	/bil/data/697a07d24603a697/180715_KM001_AdcyapHGfemale_processed/ Firefox/Chrome/Safari ftp link Explorer/Edge ftp link
Modality:	cell counting
Transgenic Line:	Strain_Name='C57BL6/Agouti/Swiss'; Pynline_Driver_InjectionSite='AdcyapHG'
Method:	imaging
Organization:	Cold Spring Harbor Laboratory
Collection Name:	697a07d24603a697
Collection Description:	Collection 697a07d24603a697
Technique:	STPT
Collection UUID:	697a07d24603a697
Organism:	mouse

Sample Images from the dataset (click to zoom)



What would you like to see?
How would you like to search?

Finding Data: File Systems

- First, check the data inventory for the file location
- On BIL computational resources, the file path to the BIL entry is:
 - /bil/data/<c1c2>/<c3c4>/<bil_uuid>/.../<Sample-Name>
 - where <c1c2> are replaced by the first and second characters of the unique identifier and <c3c4> are replaced by the third and fourth characters of the unique identifier.
 - Example bil_uuid: abc_{ded0123456789}, sample_name: example_dataset_01
 - /bil/data/_{ab}/_{cd}/abcdef0123456789/example_dataset_01/
- *In the future, DOI's will also be used to directly access datasets*



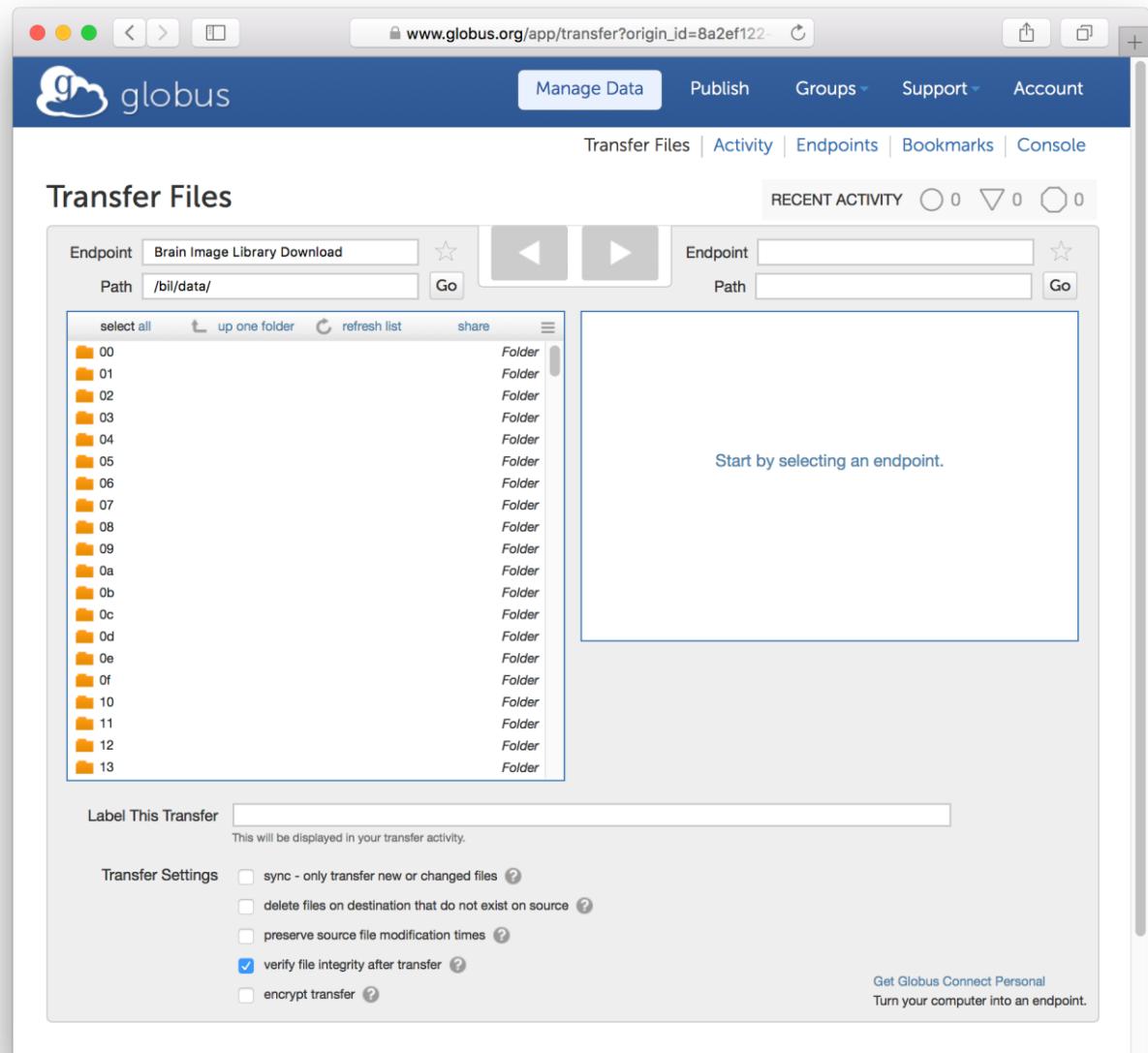
Downloading Data

- No Login Required:
 - Globus: bil#download
 - Preferred method, due to length of typical data transfers
 - Restart/resume capable
 - Anonymous FTP: download.brainimagelibrary.org
- Download Considerations:
 - Some BIL datasets can exceed 10TB and may be difficult to download over slow network connections.
 - 10 Terabyte mouse brain @ 1Gbit/sec ~ 1day (25hrs)
 - Same brain would take over 4 days @ 250 Mbits/sec
 - Consider compute/visualize in-place.



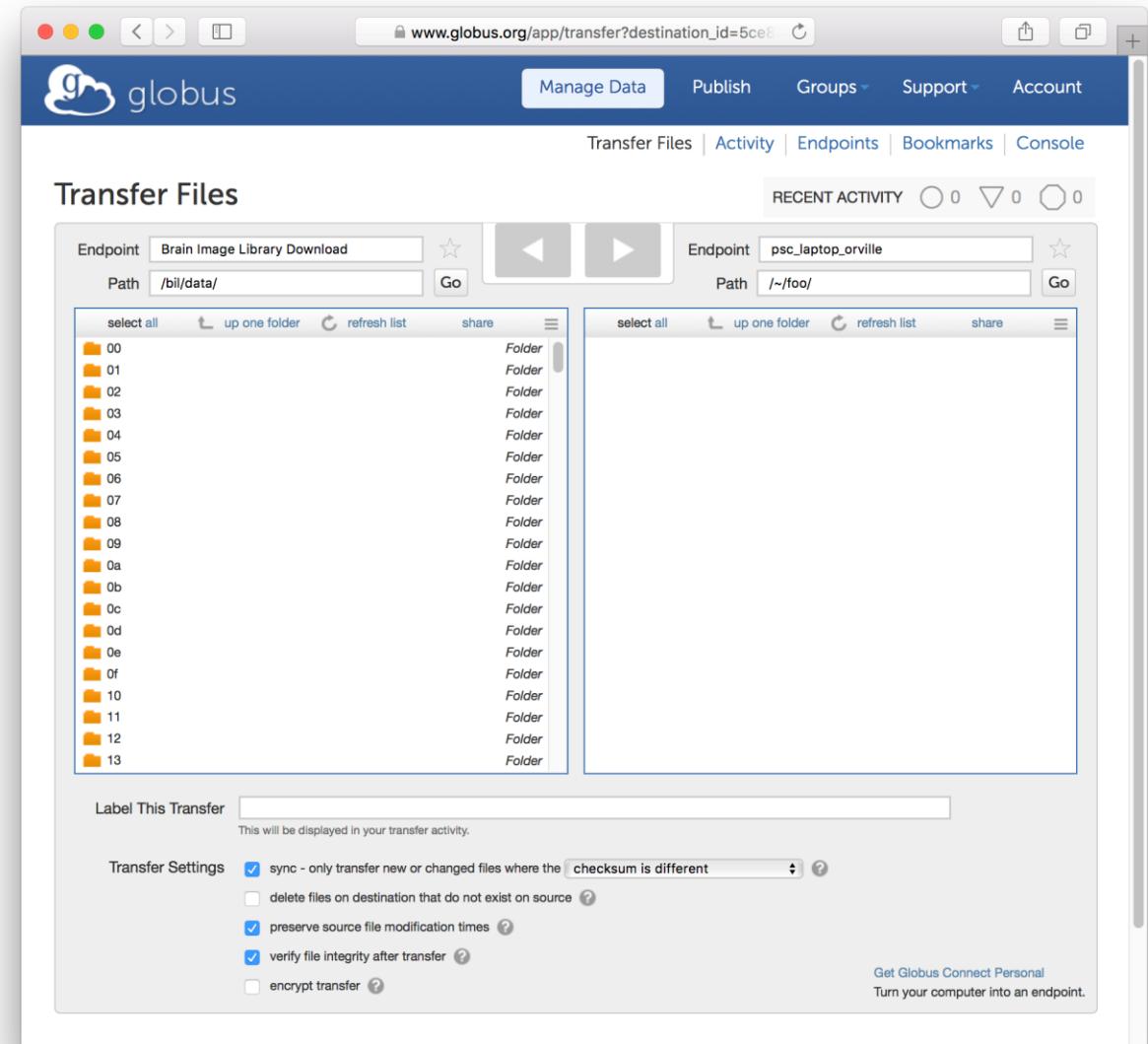
Downloading Data with Globus

- Login to the Globus transfer web service at: www.globus.org.
- Select an Endpoint field
 - Select 'Brain Image Library Download' (bil#download).
- Upon successful connection you will land in the /bil/data directory
- Navigate to the directory containing the dataset(s) that you wish to download.



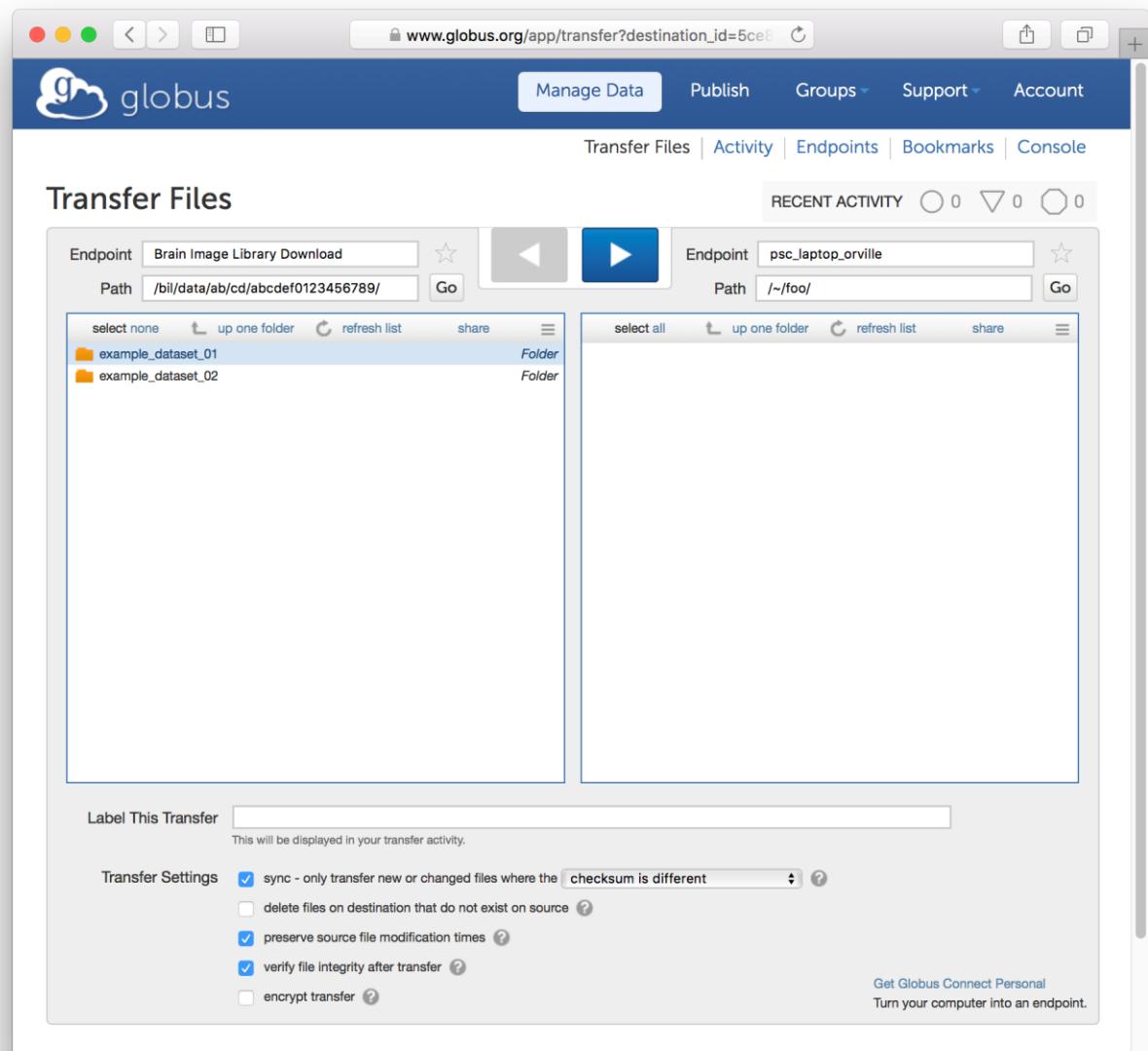
Downloading Data with Globus

- In the other Endpoint field, select your destination server.
- Select the appropriate transfer settings that apply to your download needs, e.g., "preserve source file modification times", "verify file integrity after transfer", etc.



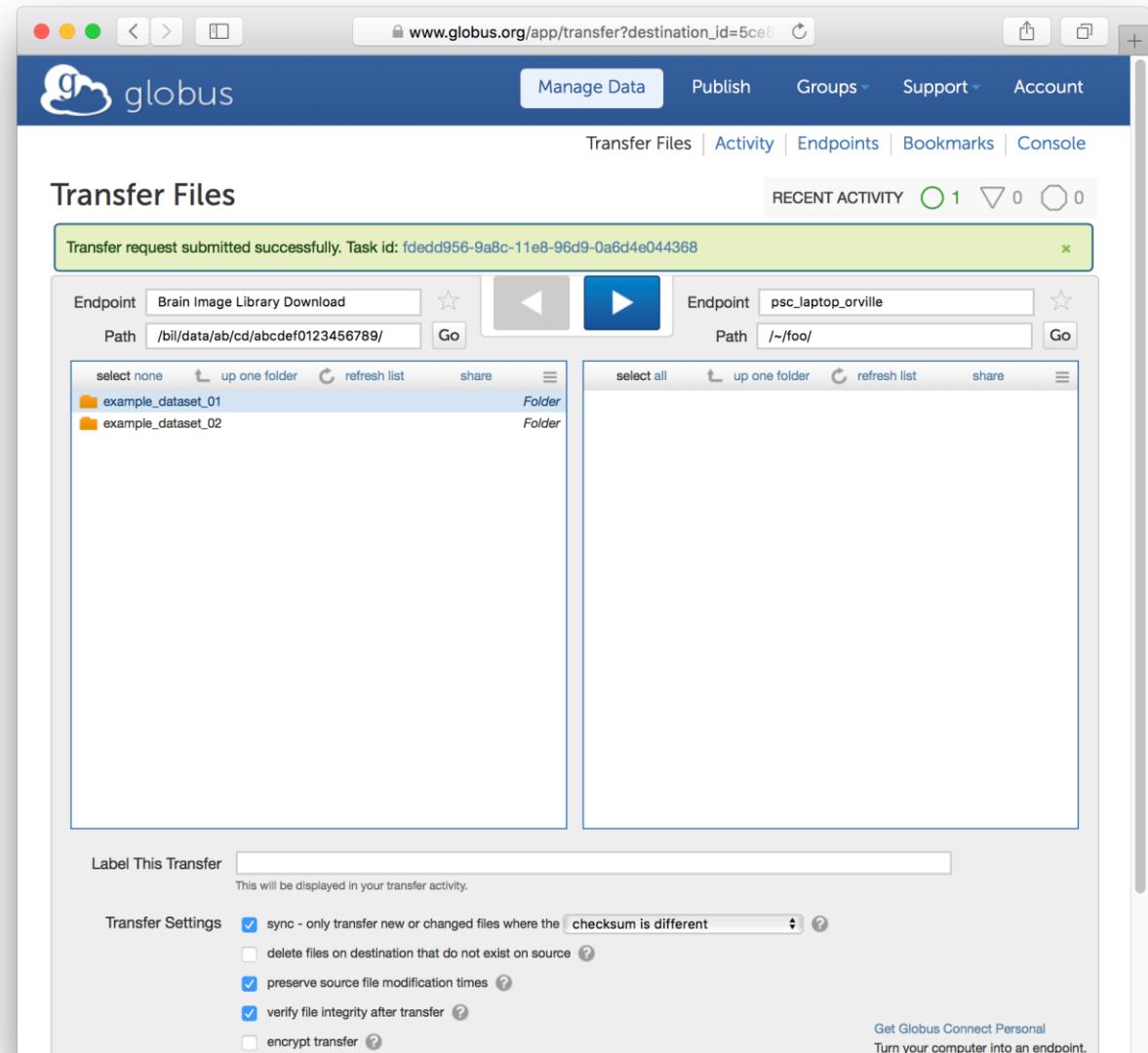
Downloading Data with Globus

- Select the dataset(s) and/or files that you wish to download by highlighting their directory and/or file names.



Downloading Data with Globus

- Click the arrow button in the direction of the transfer.
- This will schedule your transfer to begin.
- At this point you may log out of the Globus user interface and let the transfer proceed.



Downloading Data with Globus

- To check on the progress of your transfer, select the Activity tab
- Then select the job corresponding to your transfer.

The screenshot shows a web browser window for the Globus website (www.globus.org/app/activity/fdedd956-9a8c-11e8-96d9-0a6d4e044368). The page title is "Activity". The main content area displays a task card for a "Brain Image Library Download to psc_laptop_orville" task, which is currently "task queued". The card includes details such as Task ID, Owner (Derek Simmel), Source (Brain Image Library Download), Destination (psc_laptop_orville), Condition (ACTIVE), Requested Date (2018-08-07 05:58 pm), Deadline (2018-08-08 05:58 pm), and Transfer Settings (verify file integrity after transfer, transfer is not encrypted, preserve source file modification times, transfer new or changed files where the checksum is different (sync level 3)). To the right of the card, there is a summary of transfer statistics: Files 484, Directories 1, Bytes Transferred 0 B/s, Effective Speed 0 B/s, Pending 484, Succeeded 2, Cancelled 0, Expired 0, Failed 0, Retrying 0, and Skipped 0. At the bottom of the card, there is a link to "view debug data".

Downloading Data with Globus

- Upon successful completion of your transfer, the status page for your transfer will show 'Condition SUCCEEDED'
- You may also receive an e-mail to notify you if your Globus account settings are configured to do so.

The screenshot shows the Globus Activity page with a successful transfer listed. The transfer details are as follows:

- Task ID: fdedd956-9a8c-11e8-96d9-0a6d4e044368
- Owner: Derek Simmel (dsimmel@globusid.org)
- Source: Brain Image Library Download (owner: bil@globusid.org)
- Destination: psc_laptop_orville (owner: dsimmel@globusid.org)
- Condition: SUCCEEDED
- Requested: 2018-08-07 05:58 pm
- Completed: 2018-08-07 05:59 pm
- Transfer Settings:
 - verify file integrity after transfer
 - transfer is not encrypted
 - preserve source file modification times
 - transfer new or changed files where the checksum is different (sync level 3)

On the right side, there is a summary of transfer statistics:

Files	484
Directories	1
Bytes Transferred	282.80 MB
Effective Speed	4.91 MB/s
Pending	0
Succeeded	486
Cancelled	0
Expired	0
Failed	0
Retrying	0
Skipped	0

[view debug data](#)

Tools for Interacting with Data in-Place

- Rich suite of software available for use on BIL VMs and Bridges
- VMs capable to bring-your own commercial software
- Contact us @ bil-support@psc.edu



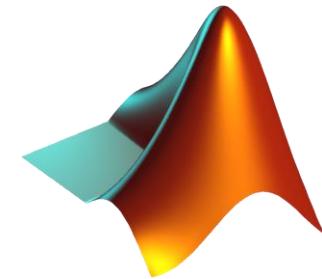
Singularity



TensorFlow



SPICE

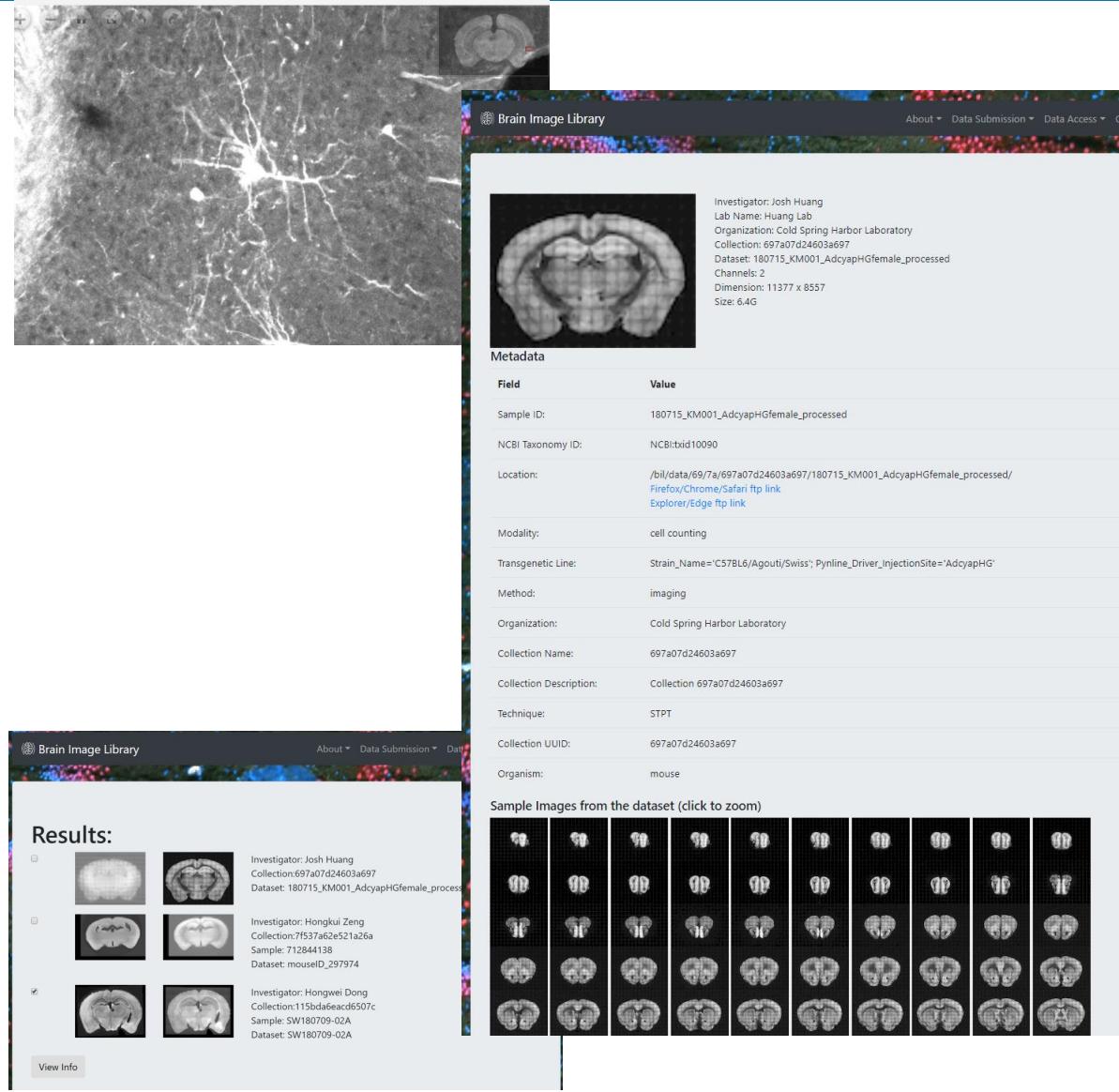


Require Significant Computational Resources?

- Apply for an XSEDE resource allocation on Bridges:
 - **There is NO CHARGE for accessing these computational resources.**
 - Normally a research starter allocation provides up to:
 - 50,000 Regular memory core-hours (50,000 Service Units)
 - 1000 Large Memory terabyte-hours. (1,000 LM Service Units)
 - 2,500 K80 (or 1000 P100) GPU hours (2,500 Service Units)
 - 1500 (Volta,DGX-2) AI GPU hours (1500 Service Units)
 - Research allocations can be for substantially more service units
 - Educational/instructional allocations also available.
 - We can help groups with the process – contact: bil-support@psc.edu
- Multiple users can be added to and share a single resource allocation:
 - PI's collaborators (including non-US) students

Help Us Meet Your Data Exploration Needs

- How can we make BIL more accessible and usable by the research community?
 - What information would you expect to find within three mouse clicks?
 - What use cases should be supported?
 - What external resources should be linked to? Why?
 - What exploration & visualization software should be supported?
- Training/outreach needs and uses
- Feedback on future directions as well as general feedback





Contributing Data to the Brain Image Library

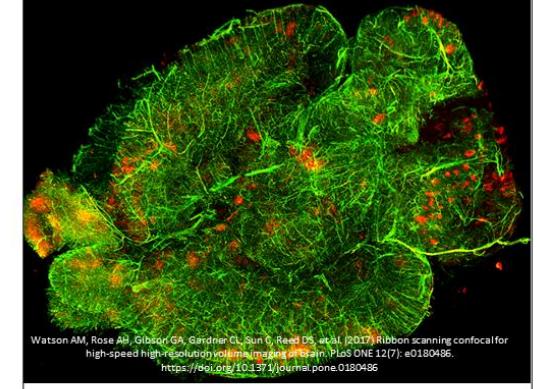
www.brainimagelibrary.org

bil-support@psc.edu

Contributing Data to BIL

- There is no-charge to contribute data to BIL or access BIL services.
- BIL can handle (and we encourage!) the submission of very large datasets
 - If you will be submitting more than 100TB of data, please let us know in advance.
- We can also provide letters of collaboration for proposals submitted to funding agencies
- Detailed data submission instructions are at:
 - www.brainimagelibrary.org/submission.html

WANTED
BY THE BRAIN IMAGE LIBRARY



Watson AM, Rose AP, Gibson GA, Gardner CL, Sun C, Reed DS, et al. (2017) Ribbon scanning confocal for high-speed high-resolution volume imaging of brain. PLoS ONE 12(7): e0180486. <https://doi.org/10.1371/journal.pone.0180486>

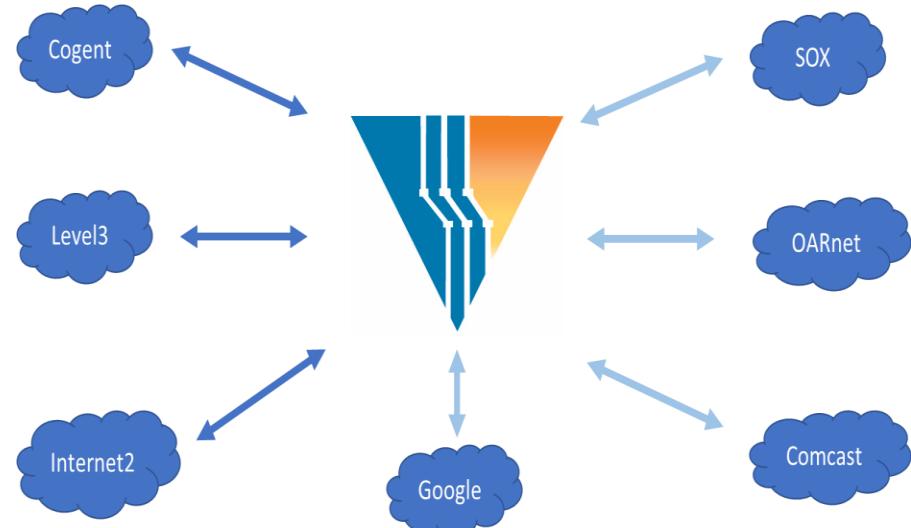
The Brain Image Library (BIL) is an NIH funded public resource (R24-MH-114793) to analyze, mine, share and interact with large high-resolution brain image datasets. BIL relies on contributions of brain image data from researchers like you. There is no charge for contributing data to the library and a help desk is available to provide data submission assistance.

CONTACT

bil-support@psc.edu
www.brainimagelibrary.org

Moving Data to BIL

- Preferred method is over the internet
 - BIL is connected to multiple internet providers, peering networks, and a high-speed link to the internet2 backbone.
 - We can help debug networking issues.
- Special arrangements can be made to receive data on disk or tape
 - BrainBall
 - Portable 30TB RAID system
 - Tape Cartridge:
 - Capability to receive data on LTO7 and LTO8 tapes.



Pre-Submission Processing

- Processing:
 - Remote desktop VMs
 - HPC Access to Bridges
- Individual and project storage space
- Rich suite of software available on BIL VMs and Bridges
- Contact us @ bil-support@psc.edu

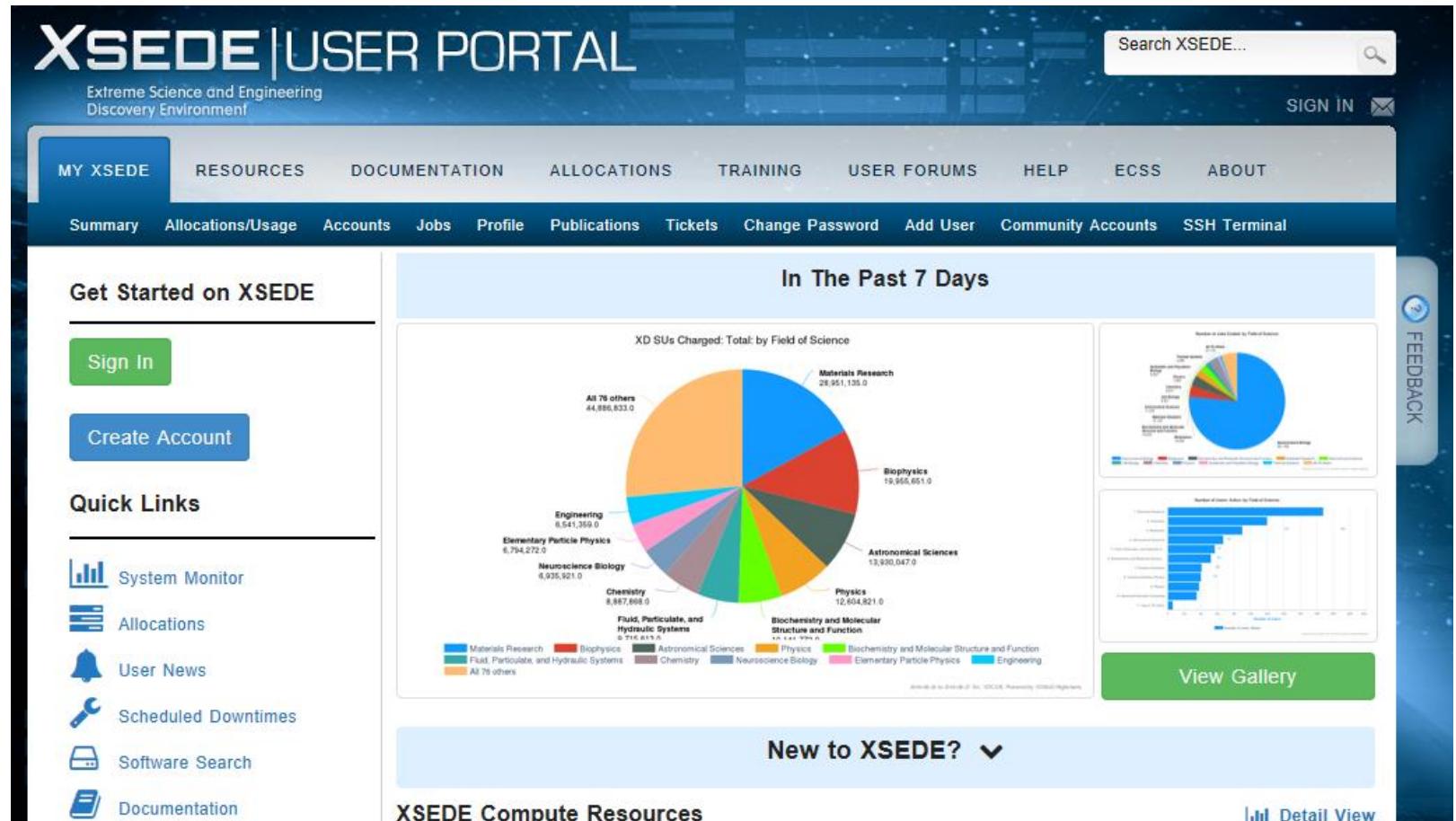


Contributing Data: Process

1. **Set up Submission Account:** Use the XSEDE portal to set up a submission account, sending that portal account name to bil-support@psc.edu
2. **Enter the Submission Portal:** Log into the data submission portal.
3. **Create a Collection:** A collection is a wrapper associating image data with metadata.
4. **Upload Image Data:** Because this is the time-limiting step, we suggest this step be started before uploading metadata for the collection.
5. **Upload Metadata:** Collection metadata can be uploaded in spreadsheet format.
6. **Validate Data:** Validate the data and correct any errors discovered.
7. **Review and Submit Data:** Correct all errors and check all warnings in the validation log prior to submission

Set Up Submission Account (Step 1)

- Visit portal.xsede.org
- Click on:
Create Account
- Fill in requested information



The screenshot shows the XSEDE User Portal homepage. At the top, the XSEDE logo and "USER PORTAL" are displayed, along with a search bar and sign-in links. Below the header, a navigation menu includes "MY XSEDE", "RESOURCES", "DOCUMENTATION", "ALLOCATIONS", "TRAINING", "USER FORUMS", "HELP", "ECSS", and "ABOUT". The main content area features a "Get Started on XSEDE" section with "Sign In" and "Create Account" buttons, and a "Quick Links" sidebar with links to "System Monitor", "Allocations", "User News", "Scheduled Downtimes", "Software Search", and "Documentation". To the right, a large chart titled "In The Past 7 Days" shows the distribution of XD SUs Charged by Field of Science. A legend indicates the colors for different fields: Materials Research (orange), Biophysics (red), Astronomical Sciences (dark blue), Physics (light blue), Biochemistry and Molecular Structure and Function (green), Neuroscience Biology (yellow), Fluid, Particulate, and Hydraulic Systems (purple), Elementary Particle Physics (pink), and Engineering (cyan). The chart data is summarized in the following table:

Field of Science	XD SUs Charged
Materials Research	29,951,135.0
Biophysics	19,865,651.0
Astronomical Sciences	13,920,047.0
Physics	12,604,821.0
Biochemistry and Molecular Structure and Function	8,887,668.0
Fluid, Particulate, and Hydraulic Systems	6,794,272.0
Chemistry	6,541,359.0
Neuroscience Biology	6,935,921.0
Engineering	44,886,833.0
All 79 others	44,886,833.0

Below the chart, there is a "View Gallery" button and a "New to XSEDE? ▾" link. The footer of the page includes "XSEDE Compute Resources" and "Detail View" links.



Set Up Submission Account (Step 1)

- Fill in personal information
- Choose a passkey (this is something you make up)
- System will send you email with a verification code
- You will use the verification code together with passkey in the next step to finish creating your account

The screenshot shows a web-based form titled "Create an XSEDE Portal Account". The form is divided into several sections: "NAME" (with fields for First Name, Middle Name, and Last Name), "CONTACT DETAILS" (with fields for Organization and Department), "POSITION" (a dropdown menu with "Choose one"), "ADDRESS (LINE 1)" and "ADDRESS (LINE 2)" (text input fields), "CITY" and "ZIP/POSTAL CODE" (text input fields), "COUNTRY" (a dropdown menu with "United States" selected) and "STATE/PROVINCE" (a dropdown menu with "Choose one"), "EMAIL ADDRESS" and "WORK PHONE" (text input fields), and "COUNTRY OF CITIZENSHIP" (a dropdown menu with "United States" selected). Below these fields is a section titled "CHOOSE A PASSKEY" with a note: "YOU WILL USE THIS PASSKEY FOR ACCOUNT VERIFICATION. MAXIMUM OF 6 CHARACTERS." There is also a link "Same as above".

Set Up Submission Account (Step 1)

- Verify your account using email address, passkey, and verification code
- Choose username
- Choose password (that complies with password rules)

Verify your XSEDE Portal Account

ENTER YOUR EMAIL ADDRESS

EMAIL ADDRESS

VERIFY YOUR ACCOUNT REQUEST

ENTER THE PASSKEY YOU CHOSE DURING REGISTRATION

ENTER THE VERIFICATION CODE SENT TO YOUR EMAIL

CHOOSE YOUR TERAGRID USER NAME AND PASSWORD

USER NAME

User names must be between 3 and 8 characters in length, must contain only lowercase letters and numbers, and must begin with a letter.

PASSWORD

Passwords must be a non dictionary word, at least 8 characters in length and must contain at least 3 of the following character classes:

- lowercase letters
- uppercase letters
- numbers
- symbols, excluding quotes ' or " and space(s)

CONFIRM PASSWORD

PROVE YOU ARE HUMAN





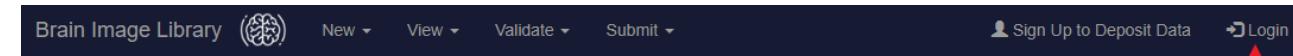
SUBMIT

Set Up Submission Account (Step 1)

- Email bil-support@psc.edu with the following:
 - Your XSEDE Portal Account
 - Information about the project that you are associated with (e.g. NIH or NSF Grant number & Title)
- You will receive automated email messages:
 - That you have been added to the system
 - One with your username + instructions to set your password

Enter Submission Portal (Step2)

- With a web browser, visit: submit.brainimagelibrary.org
- Log in with your
 - Username
 - Password
- Problems?
 - bil-support@psc.edu
 - 412-268-



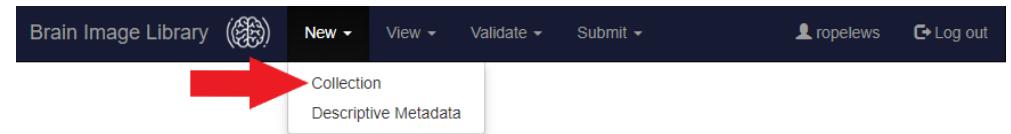
Data Submission Portal

The Brain Image Library (BIL) is a public resource enabling researchers to deposit, analyze, mine, share and interact with large brain image datasets. More information about the Brain Image Library can be found on our informational website: www.brainimagelibrary.org. The BIL is supported by the National Institute of Mental Health of the National Institutes of Health under award number R24MH114793.

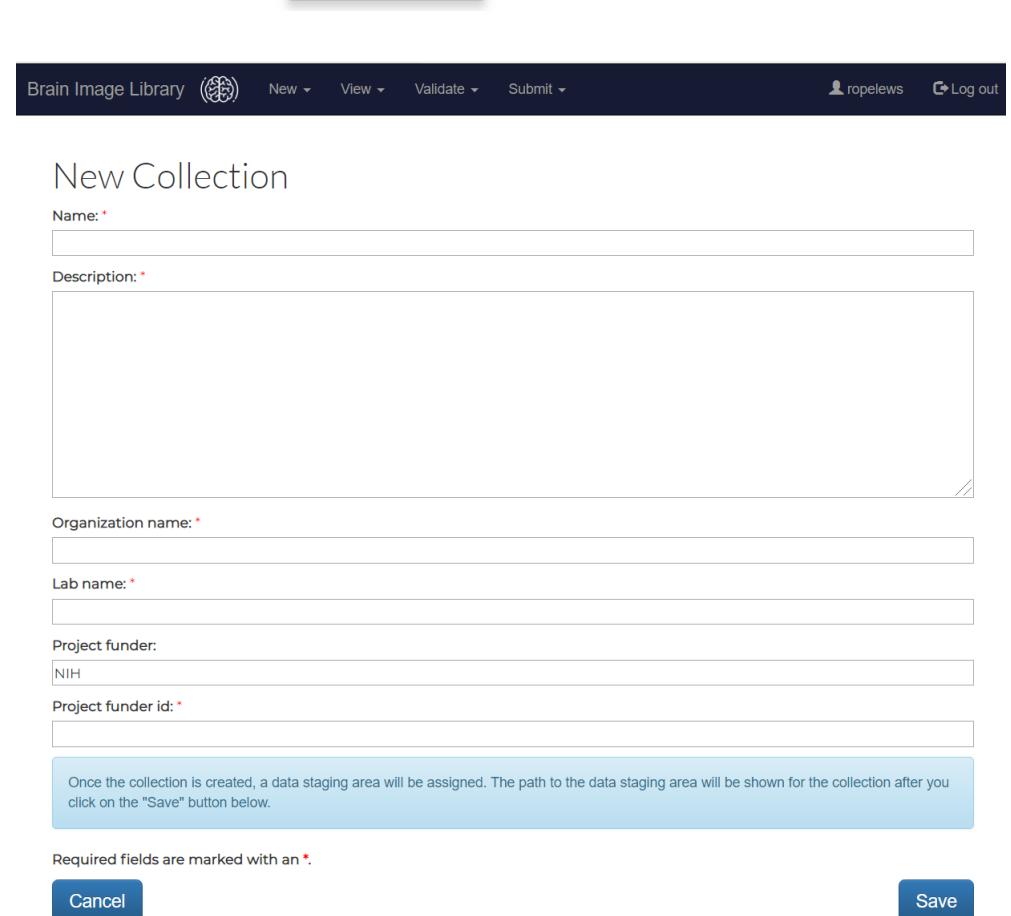
Please [log in](#) or [sign up](#) to begin creating collections for publication in the BIL.

Create a Collection (Step 3)

- Menu: **New -> Collection**
- A “Collection” is a flexible ingest wrapper to tie data and metadata together:
 - One per brain
 - One per experiment
- Creating a collection also creates a “Landing Zone” (a private space for you to upload this data to)
 - **/bil/lz/user/collection**



The screenshot shows the Brain Image Library interface. At the top, there is a navigation bar with the title "Brain Image Library" and icons for "New", "View", "Validate", and "Submit". On the right side of the bar, there is a user profile icon labeled "ropelews" and a "Log out" button. A red arrow points to the "New" dropdown menu, specifically highlighting the "Collection" option.



The screenshot shows the "New Collection" creation form. The form has several input fields:

- Name: * (empty field)
- Description: * (empty field)
- Organization name: * (empty field)
- Lab name: * (empty field)
- Project funder: (text input field containing "NIH")
- Project funder id: * (empty field)

A note at the bottom of the form states: "Once the collection is created, a data staging area will be assigned. The path to the data staging area will be shown for the collection after you click on the "Save" button below." At the bottom of the form, there is a note: "Required fields are marked with an *." and two buttons: "Cancel" and "Save".



Upload Image Data (Step 4)

- Upload endpoints support Globus, rsync, sftp, and scp:
 - upload.brainimagelibrary.org (rsync,sftp,scp)
 - #bil-upload (globus)
- Authentication is required
- Once authenticated, change to the landing-zone directory

```
$ sftp testuser@upload.brainimagelibrary.org
```

The authenticity of host 'upload.brainimagelibrary.org (128.182.108.164)' can't be established.

ECDSA key fingerprint is 32:cf:46:44:3d:9c:8e:b2:1d:14:03:66:45:0b:11:29.

Are you sure you want to continue connecting (yes/no)? yes

Warning: Permanently added 'upload.brainimagelibrary.org,128.182.108.164' (ECDSA) to the list of known hosts.

testuser@upload.brainimagelibrary.org's password:

Connected to upload.brainimagelibrary.org.

```
sftp> cd /bil/lz/testuser/1234567890abcdef
```

Upload Metadata (Step 5)

- Menu: New-> Descriptive Metadata
- Download and fill out the metadata spreadsheet
 - Be sure to enter in the spreadsheet the collection subdirectory where the data is located.
- Select the collection the metadata is for
- Upload filled-in metadata spreadsheet

Brain Image Library (brain) New View Validate Submit ropelew Log out

Upload Descriptive Metadata Spreadsheet

Download and fill out either the [Excel](#) or [LibreOffice Calc](#) template.

Choose a collection and upload your metadata.

Associated collection:

[Upload Metadata](#) [Cancel](#)

Open

This PC

3D Objects

Desktop

Documents

Downloads

Name

Date modified

Type

descriptive_metadata_template

7/10/2019 12:29 P... Microsoft

File name: All Files

[Open](#) [Cancel](#)

Brain Image Library (brain) New View Validate Submit ropelew Log out

Your Descriptive Metadata

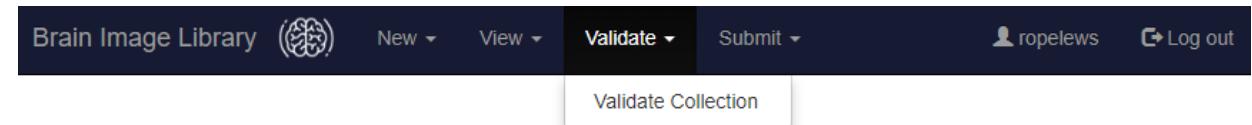
Collection	Date Created	Last Edited	Locked	Sample Id	Organism Type	Organism Ncbi Taxonomy Id	Transgenic Line Information	Modality	Method	Technique	Anatomical Structure
Tues	10/20/2019 4:20 p.m.	10/20/2019 4:20 p.m.	🔒	abcdef	mouse	NCBI:bxid10090	—	Anatomy/Morphology	cell type distribution	Ribbon	Whole brain
Whole Brain - September 2018	09/04/2018 5:21 p.m.	03/25/2019 3:29 p.m.	🔒	abcdef	mouse	NCBI:bxid10090	—	Anatomy/Morphology	cell type distribution	Ribbon	Whole brain
Test of new acls	09/11/2018 10:29 p.m.	03/25/2019 3:03 p.m.	🔒	abcdef	mouse	NCBI:bxid10090	—	Anatomy/Morphology	cell type distribution	Ribbon	Whole brain
Test of new acls	09/11/2018 10:33 p.m.	03/25/2019 3:03 p.m.	🔒	abcdef	mouse	NCBI:bxid10090	—	Anatomy/Morphology	cell type distribution	Ribbon	Whole brain

[+ Create Metadata](#) [Cancel](#)

Descriptive Metadata successfully uploaded



Validate Data (Step 6)

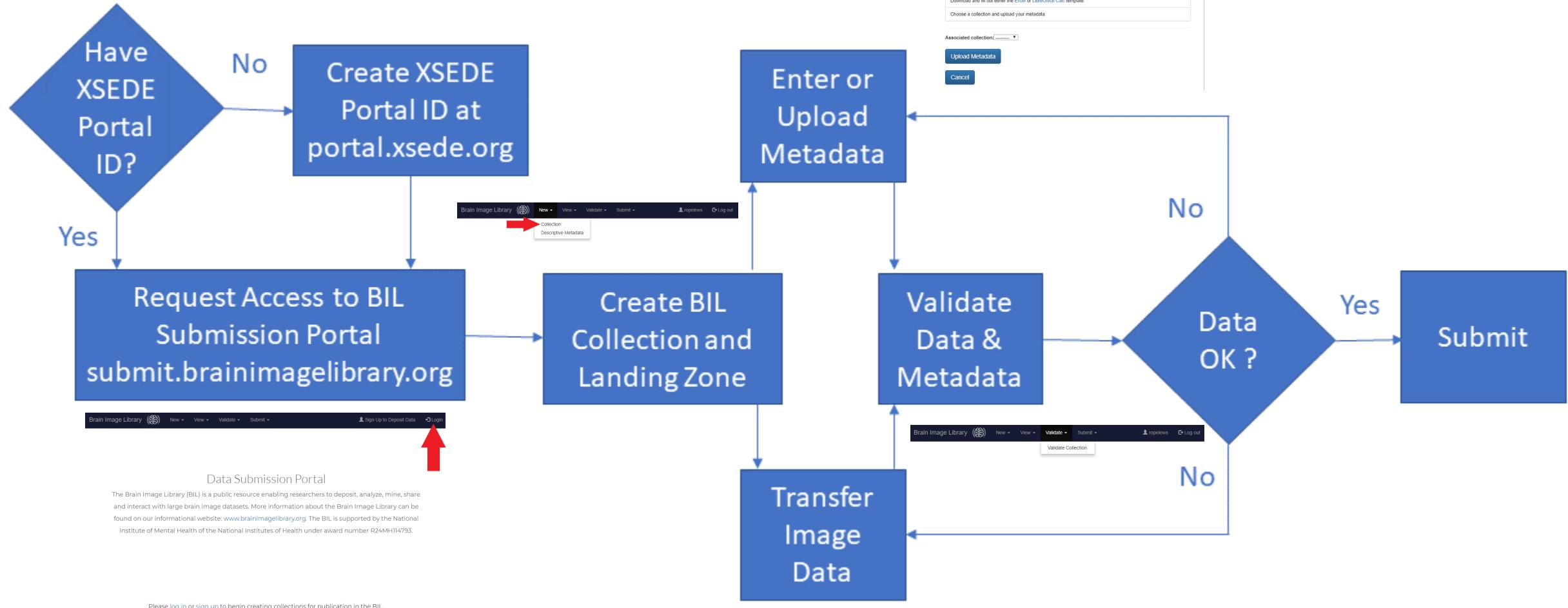


- Menu: **Validate -> Validate Collection**
- Checks performed on files in directories identified in the metadata spreadsheet (+ subdirectories of those directories):
 - Zero length file
 - For image files, bioformats can successfully read image file header information
 - Removal of (unix) problematic special characters from filenames (e.g. space, quotes asterisk, non-printable characters)
- Checks not performed:
 - Any directory that is not identified in the metadata spreadsheet

Submit Data (Step 6)

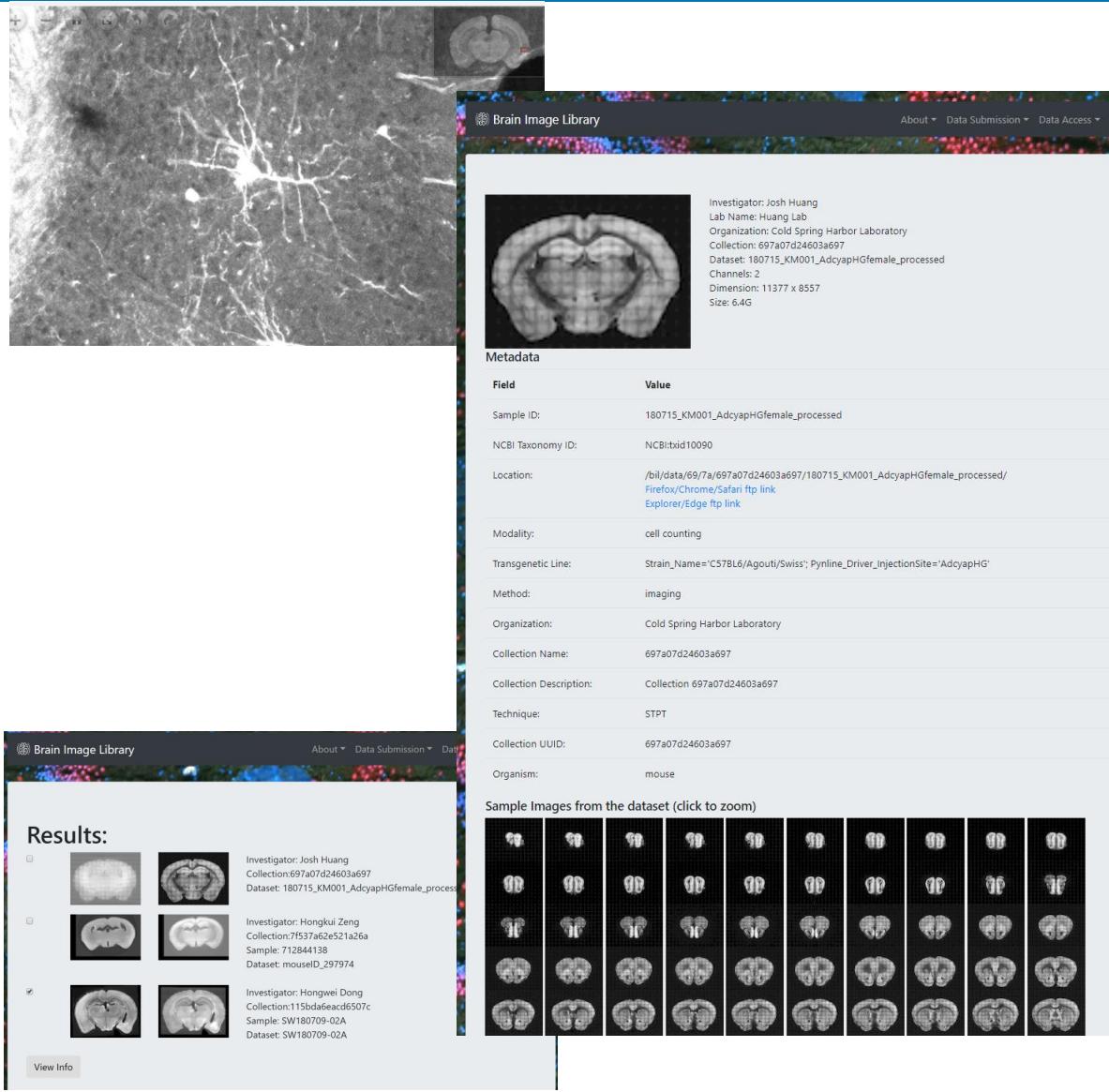
- Submitting data will make the data publicly available
- Current:
 - To submit data and make the data publicly available, send email to bil-support@psc.edu
 - Additional manual checks are performed
- Future:
 - Automated submit with embargo periods

Data Contribution Workflow



Help Us Meet Your Data Contribution Needs

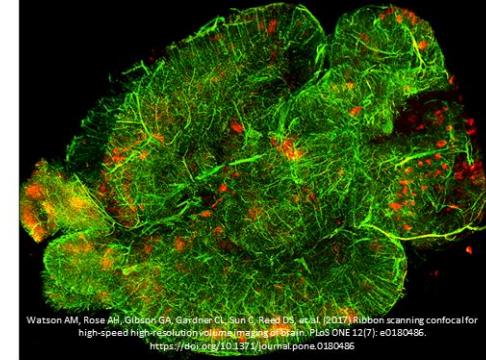
- To submit data that you have now, how much disk space would you need?
- What software and visualization tools are needed for your pre-analysis work?
- What validation/QA/QC feedback would you find useful for your datasets?
- What higher-level use cases might your data support if similar datasets were found in BIL?
- Training/outreach needs
- Feedback on future directions as well as general feedback



For More Information

- Contact us at:
 - Email: bil-support@psc.edu
 - Telephone : 412-268-6350
- Web Portals
 - Contributor Portal and Data Upload
 - submit.brainimagelibrary.org (Metadata)
 - upload.brainimagelibrary.org (Image Data)
 - Data Access
 - download.brainimagelibrary.org
 - Informational Website
 - www.brainimagelibrary.org

WANTED
BY THE BRAIN IMAGE LIBRARY



Watson AM, Rose AJ, Cifkova GA, Gardner CL, Sun C, Reed DS, et al. (2017) Ribbon scanning confocal for high-speed high-resolution volumetric imaging of brain. PLoS ONE 12(7): e0180486. <https://doi.org/10.1371/journal.pone.0180486>

The Brain Image Library (BIL) is an NIH funded public resource (R24 MH-114793) to analyze, mine, share and interact with large high-resolution brain image datasets. BIL relies on contributions of brain image data from researchers like you. There is no charge for contributing data to the library and a help desk is available to provide data submission assistance.

CONTACT

bil-support@psc.edu
www.brainimagelibrary.org

Acknowledgements: BIL is operated as a partnership between the Biomedical Applications Group at the PSC, the Center for Biological Imaging at the University of Pittsburgh and the Molecular Biosensor and Imaging Center at Carnegie Mellon University. BIL is supported by the National Institute of Mental Health under award number R24MH114793. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The Bridges supercomputing system at the PSC was acquired through NSF award ACI-1445606 [2]. User access to Bridges is supported through the Extreme Science and Engineering Discovery Environment (XSEDE) which is supported through NSF grant OCI-1053575 [5]. Any opinions findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. The authors wish to acknowledge the Brain Initiative Cell Census Network (BICCN) for their data contributions to BIL.