



CNR-INO

ISTITUTO NAZIONALE DI OTTICA
CONSIGLIO NAZIONALE DELLE RICERCHE



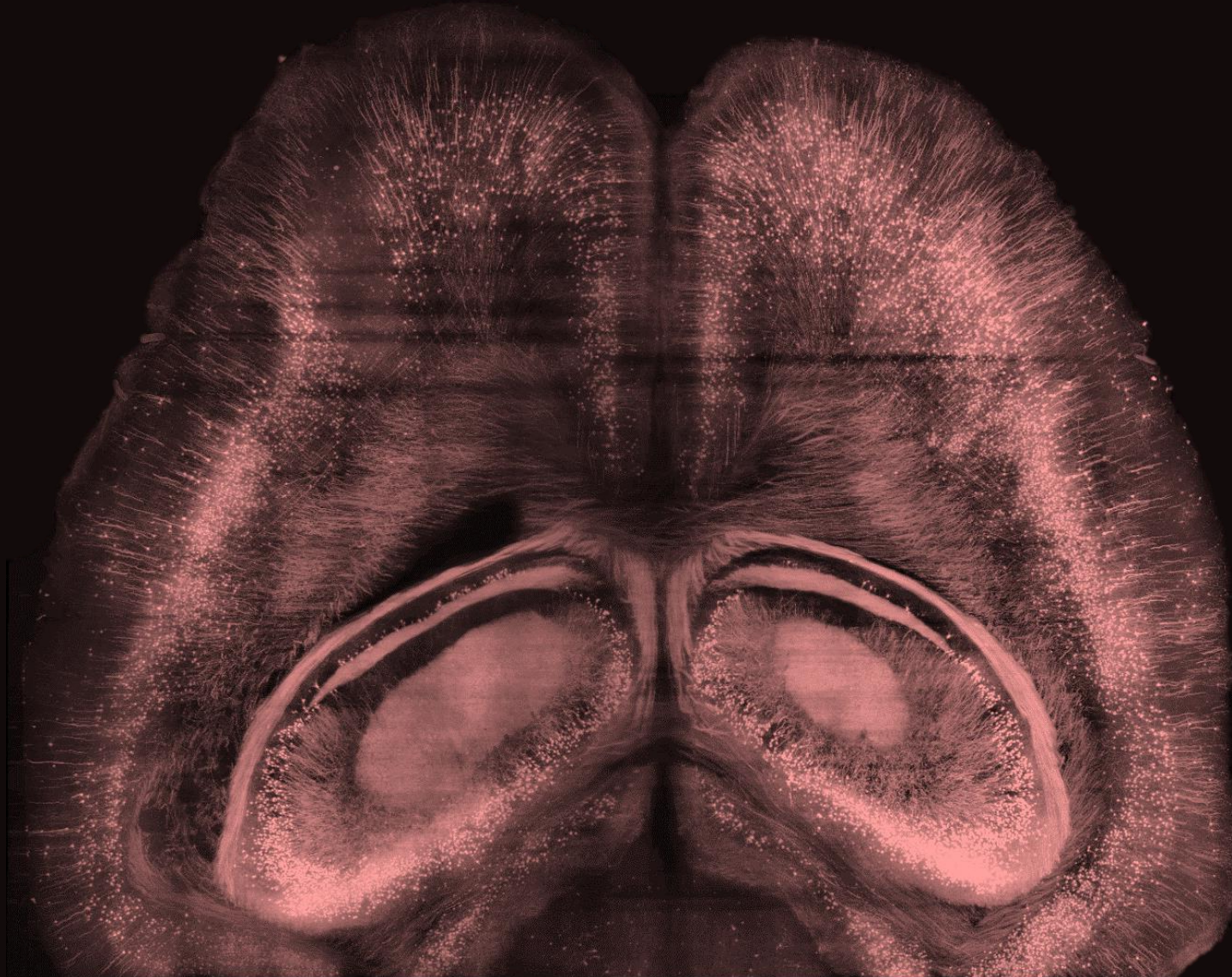
Image processing and management of large datasets in Light-Sheet Microscopy

Giacomo Mazzamuto

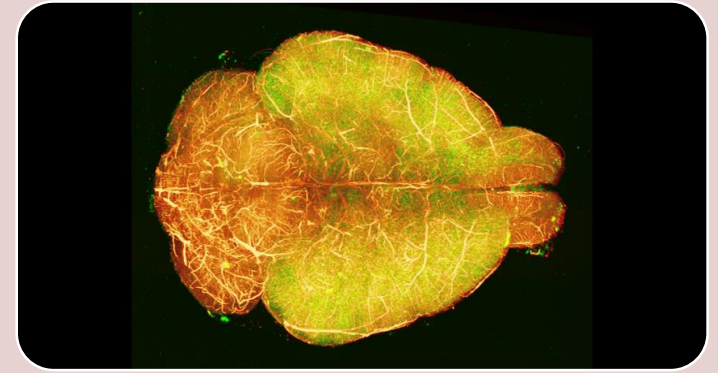
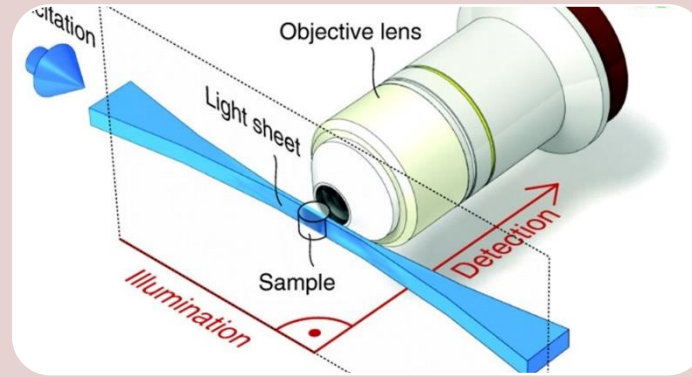
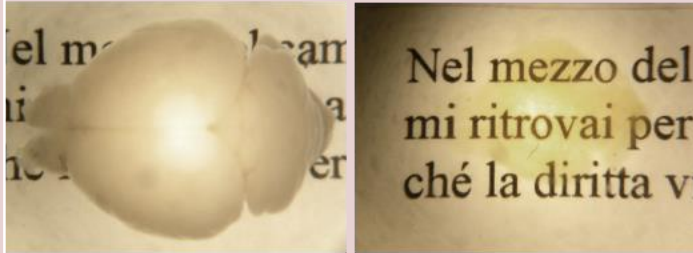
Laserlab-Europe / ELI/
CASUS Workshop

Better data for better
science

29/10/2021



The three pillars of selective illumination microscopy

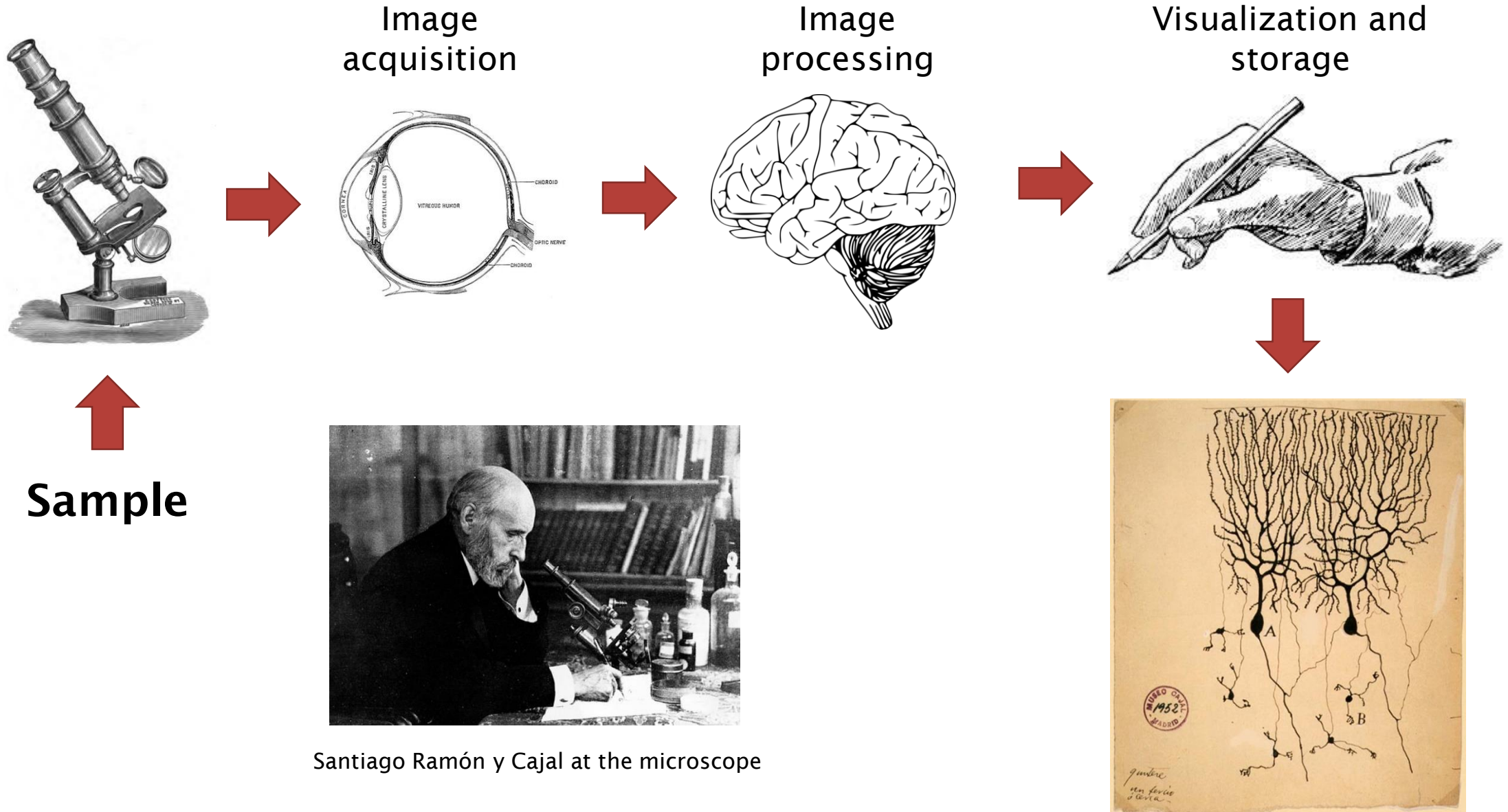


Optical
clearing

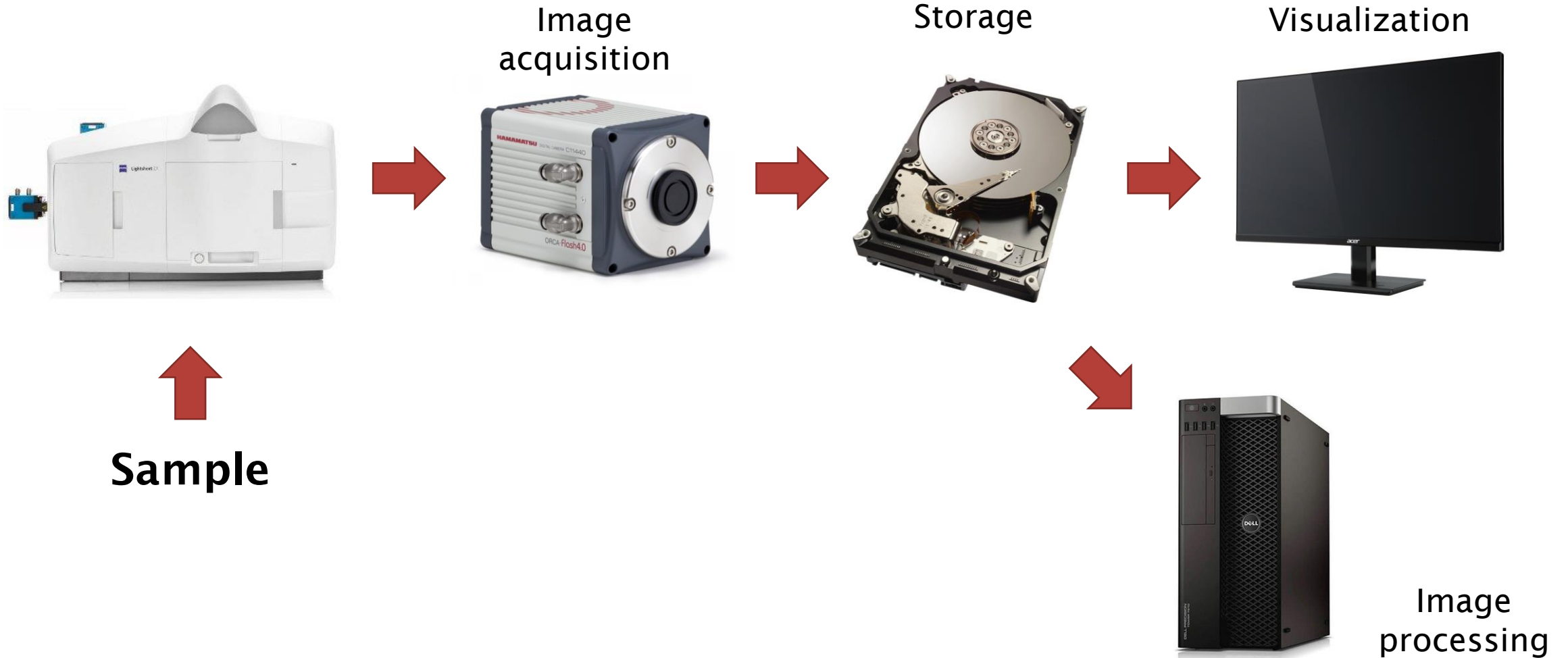
Instruments
Light-Sheet
Two-Photon

Image
processing

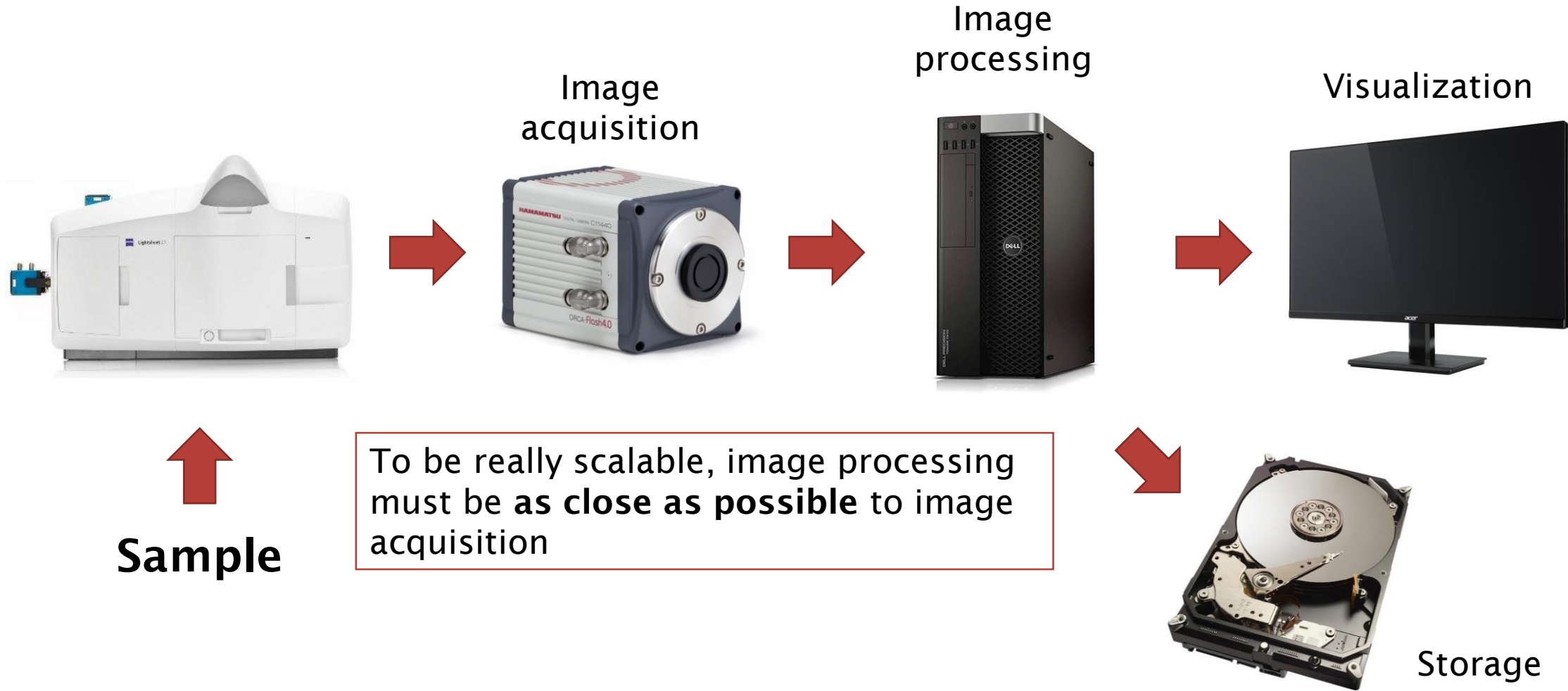
Microscopy Imaging: acquisition and processing (old days)



Microscopy Imaging: acquisition and processing (today)



Microscopy Imaging: acquisition and processing (tomorrow)

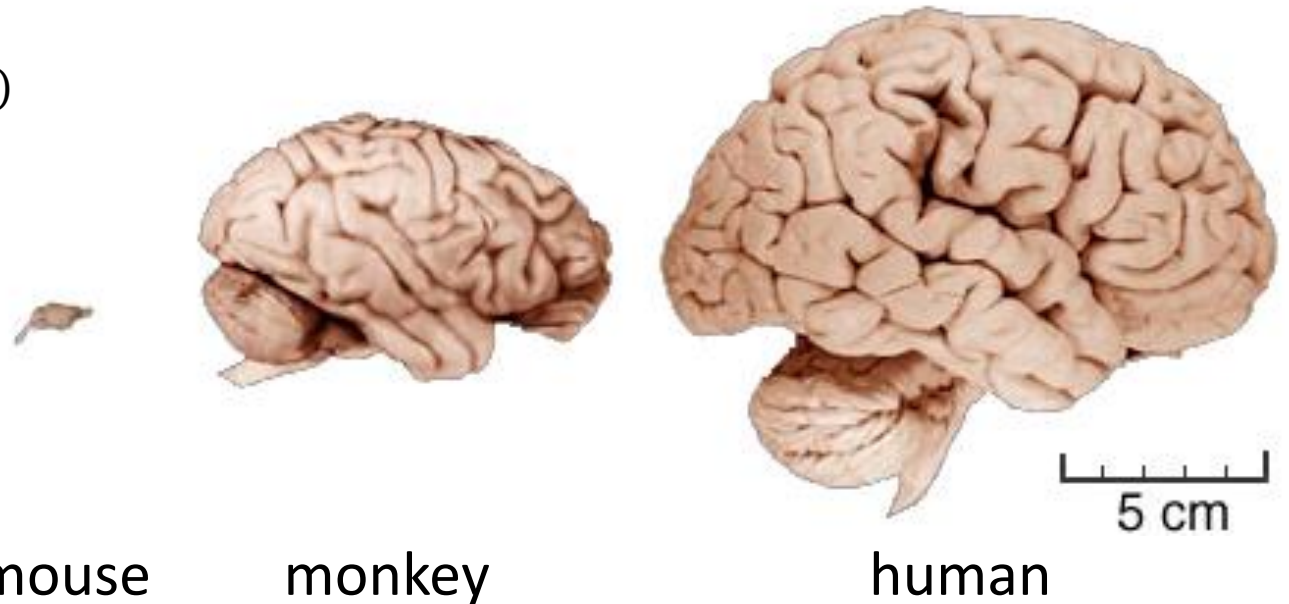


Big data challenges in whole-brain imaging



		Mouse	Monkey (Rhesus)	Human
Approx volume		1 cm ³	100 cm ³	1500 cm ³
Raw data size	1 channel, 10 μm res.	2 GB	200 GB	3 TB
	3 channels, 10 μm res.	6 GB	600 GB	9 TB
	1 channel, 1 μm res.	2 TB	200 TB	3000 TB
	3 channels, 1 μm res.	6 TB	600 TB	9000 TB

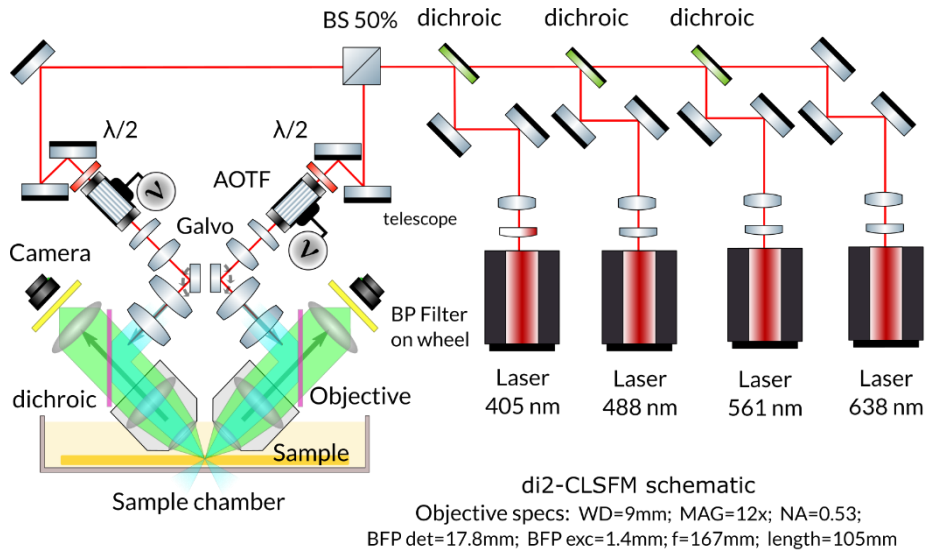
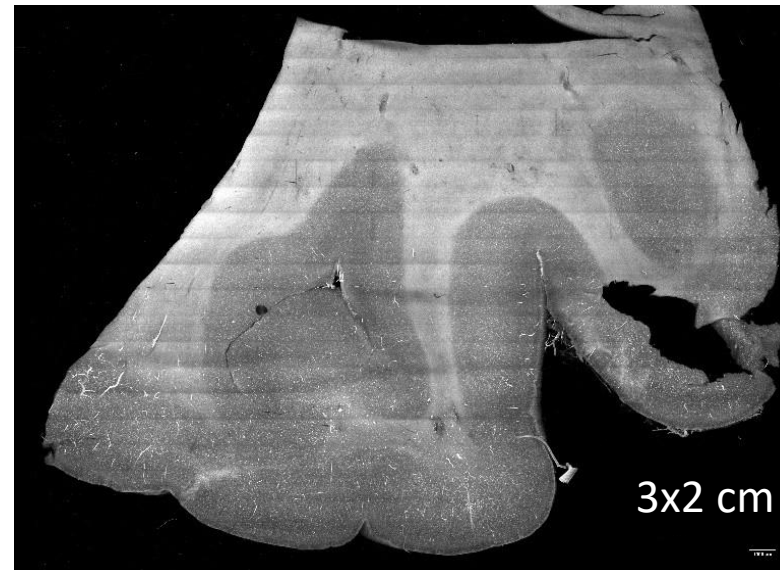
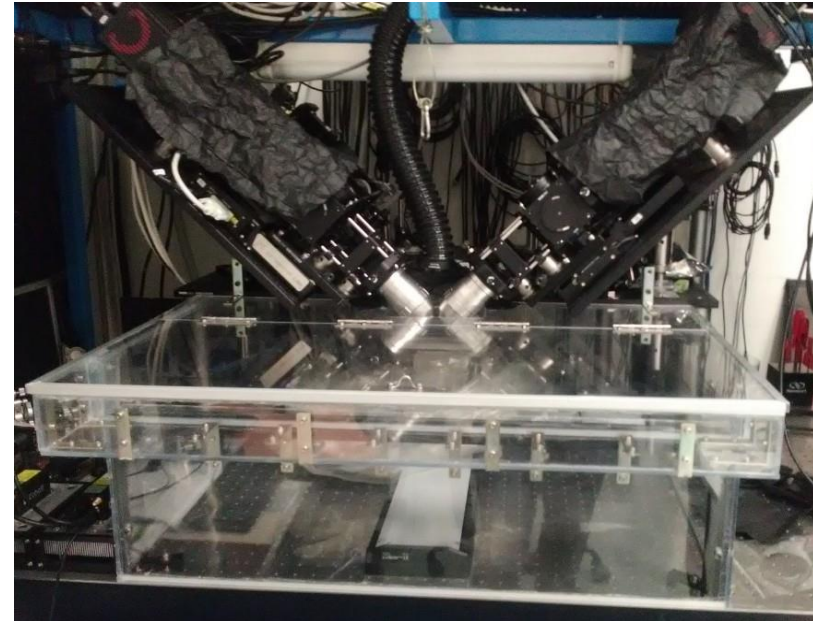
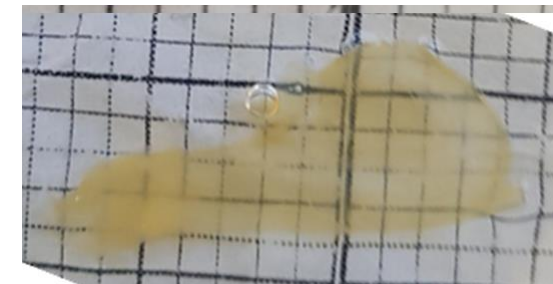
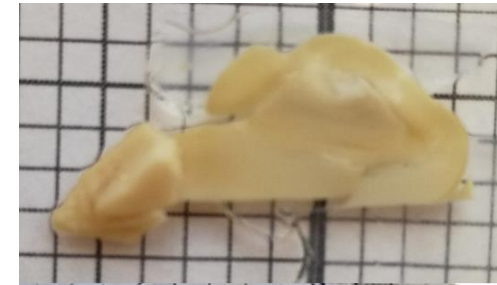
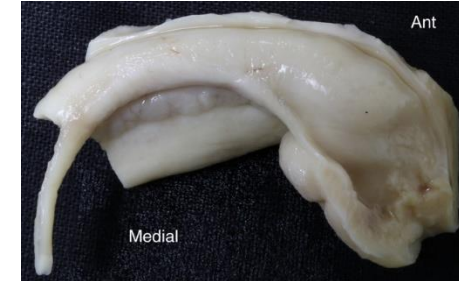
$$\frac{1 \text{ cm}^3}{(10 \text{ } \mu\text{m})^3} = 10^9 \text{ voxels} = 2\text{GB (@ 16bits/pixel)}$$



Dual view inverted Light Sheet Fluorescence Microscope



Human Brain Project



- dual-view inverted SPIM apparatus

- Isotropic 1 μm resolution
- Up to 4 channels
- Designed to image 1 mm thick sample
- High data throughput (up to 1 GB/s, one channel)
- Volumetric rate: 0.1 mm³/s.

Hippocampus volume:

20 slices of 3cm x 1cm x 0.05 cm = 0.15cm³ = 0.15 x 10¹² μm³

- Voxel size ≈ 1 μm³ => 0.15 Tvoxels = 0.15 TB per channel = 0.6 TB per slice, Grand total: 12TB

- Data throughput: 24TB per sample (dual view)

HBP data repository: EBRAINS



DATASET

Layer-specific excitatory and inhibitory neuronal maps of hippocampus (v1.1)

Costantini, I.; Mazzamuto, G.; Pesce, L.; Gavryusev, V.; Laurino, A.; Scardigli, M.; Pavone, F.

Download Dataset

Cite dataset

Data-descriptor

DOI: [10.25493/MD0E-BW5](https://doi.org/10.25493/MD0E-BW5)

License: [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Custodians: [Pavone, Francesco S.](#)

The aim of this work is to reconstruct the 3D organization of neurons in a left hippocampus from sub-cellular resolution images obtained by a custom-made dual view inverted light sheet fluorescence microscope. We analyze the left hippocampus of a 99-year old woman, not affected by Alzheimer's disease, nor hypertension, but presenting a cognitive decline, obtained from the University of Tours. The sample was cut into 72 slices of 500 μm thickness before imaging. Each slice was cleared with the SWITCH/TDE method and stained with NeuN (all neurons) and GAD67 (inhibitory neurons). Imaging was performed at a resolution of $1.1 \times 1.1 \times 3.8 \mu\text{m}^3$.

A previous data version of "Layer-specific excitatory and inhibitory neuronal maps of hippocampus" can be found here:

[Pavone et al. \(2020\) \[Data set, v1.0\] DOI: 10.25493/1GZV-ZU](#)

Modality:

- microscopy
- histological approach
- molecular expression characterization
- cell population imaging

Brain region: Hippocampus

Preparation: Ex vivo

Methods:

- SWITCH/TDE clearing method
- Dual View Inverted Light Sheet Fluorescence Microscope

Keywords:

- clearing method
- imaging

www.ebrains.eu

“EBRAINS is a new digital research infrastructure, created by the EU-funded Human Brain Project, that gathers an extensive range of data and tools for brain-related research.”

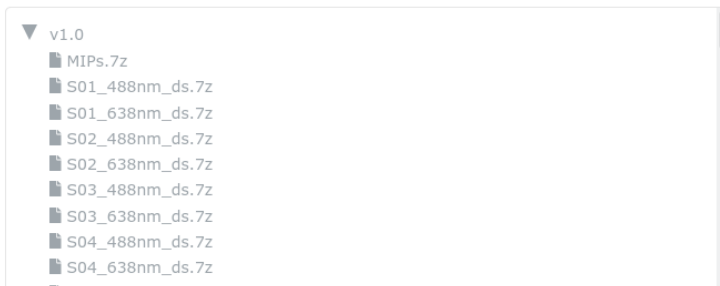
Lower-resolution dataset: $(4.4 \times 4.4 \times 3.3) \mu\text{m}^3$

Total size: 116 GiB

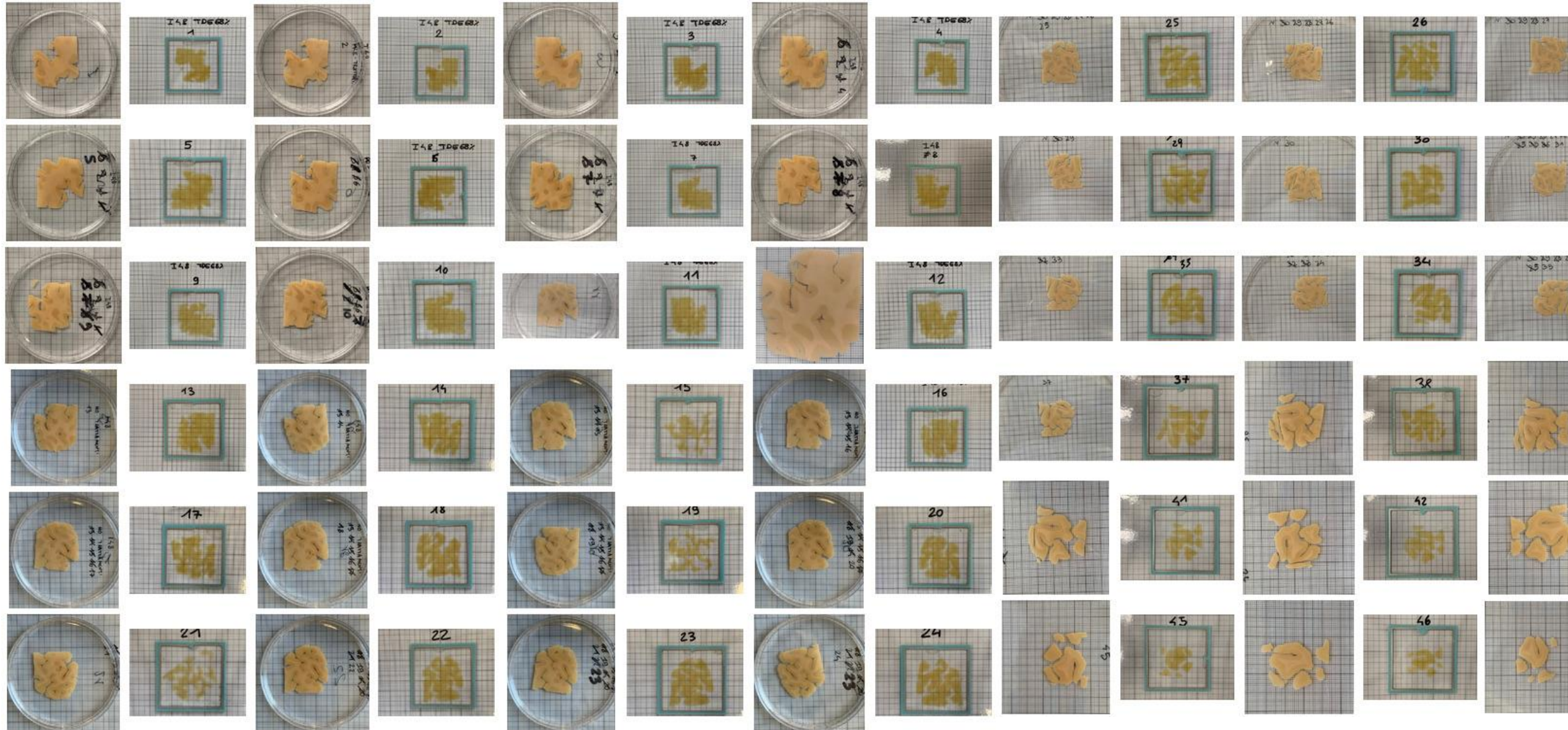
DOI: <https://doi.org/10.25493/MD0E-BW5>

Files (145)

Subjects (1)



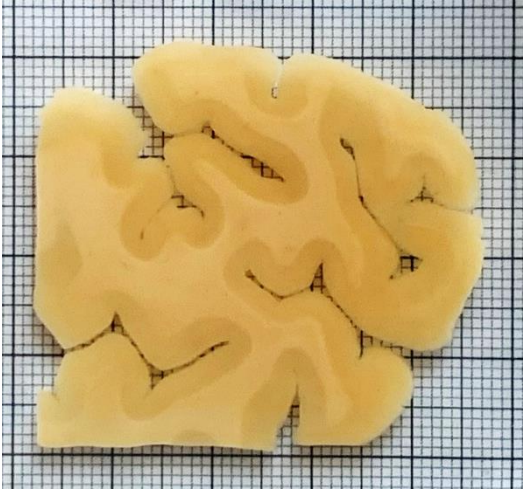
A bigger challenge: human Broca's area (4x4x2cm³, 48 slices)



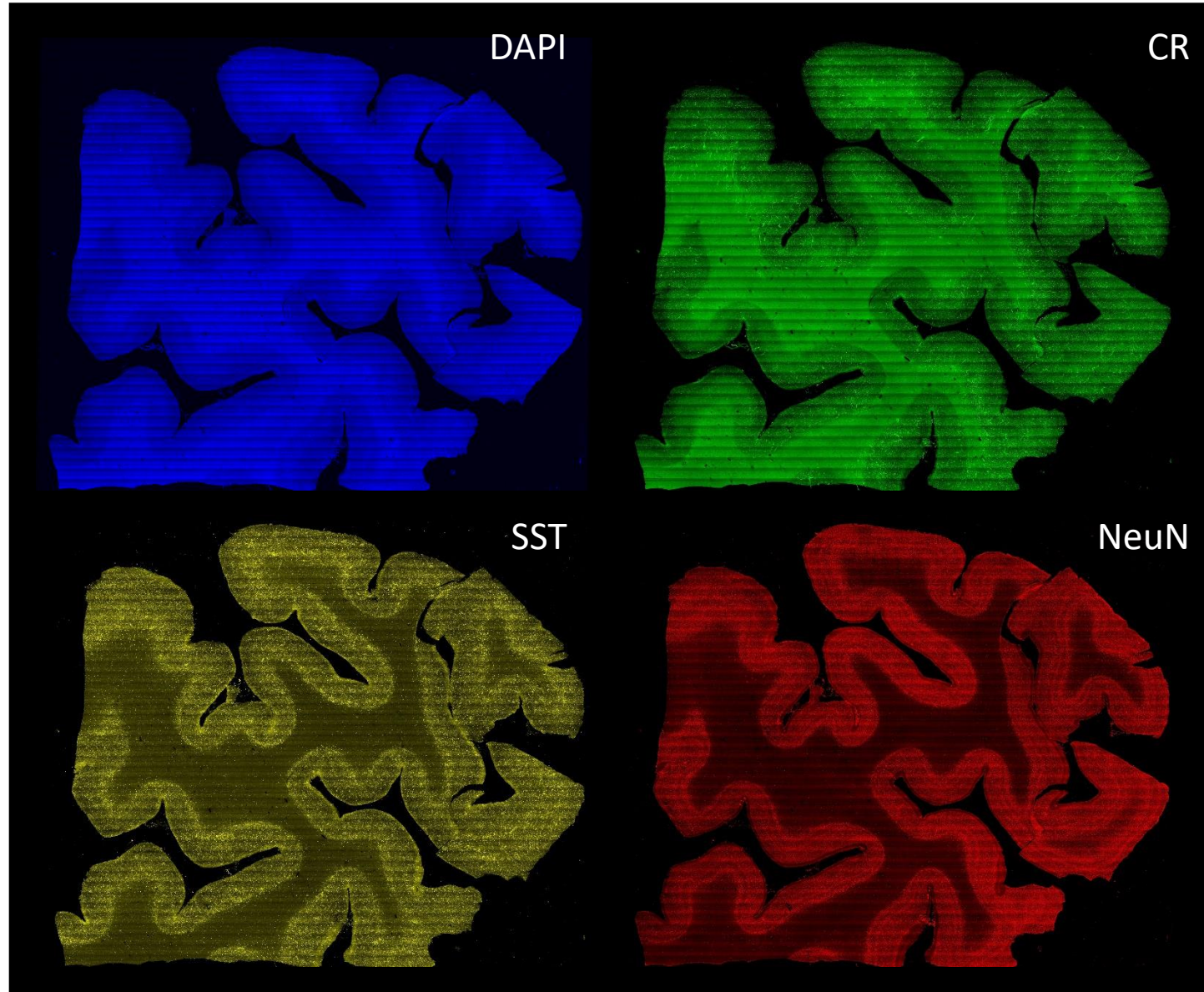
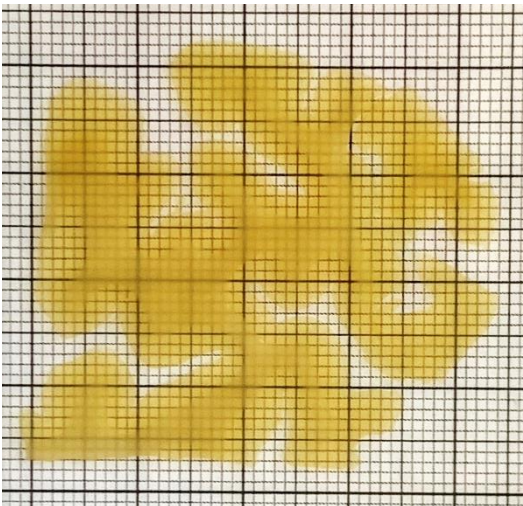
A bigger challenge: human Broca's area ($4 \times 4 \times 2 \text{cm}^3$, 48 slices)



Fixed tissue



Cleared tissue



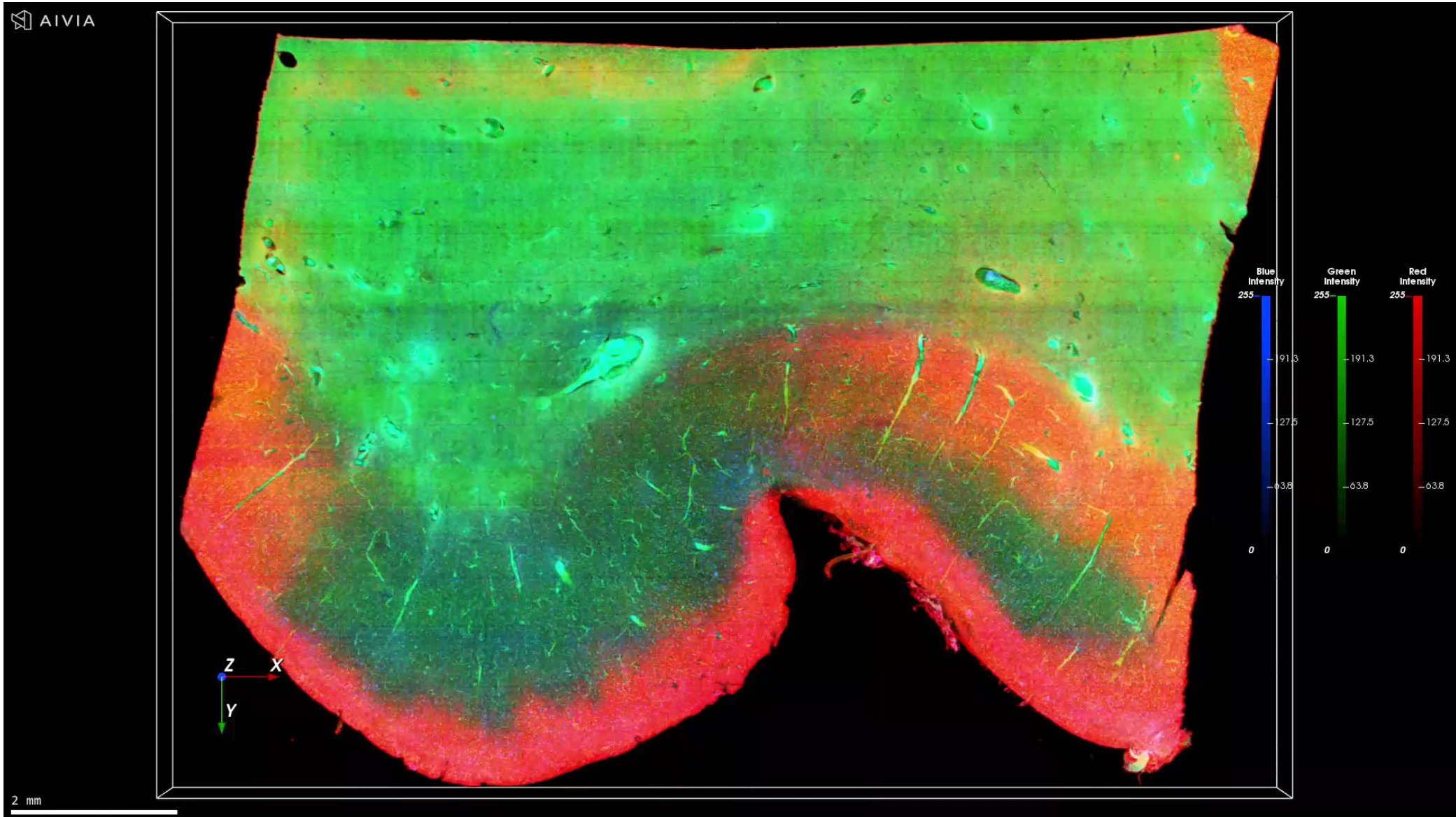
We successfully applied the SHORT clearing method to a whole Broca's Area sliced in 48 slabs of $400 \mu\text{m}$ of thickness.

Brain Initiative Cell Census Network (BICCN)

Supported by The General Hospital Corporation Center of the National Institutes of Health, award number 1U01MH117023-01



3D slice



Some numbers



August 2021

4 channels, 48 slices

Number of files	8602
Raw data	507 TiB
Lossy Compression, JPEG2000, original resolution (0.5x0.5x3.5 μ m)	25 TiB
Resliced (lower res, 3.5 μ m isotropic)	10.4 TiB
Resliced (lossless compression, 3.5 μ m, delivered)	4.2 TiB

BRAIN initiative data repository: DANDI



Human brain cell census for BA 44/45



ID: 000026 **DRAFT** Contact **Morgan, Leah** File Count **18530** File Size **6.2 TB**

Created **June 24, 2020**

Last update **October 19, 2021**

[Mazzamuto, Giacomo](#); [Costantini, Irene](#); [Gavryusev, Vladislav](#); [Castelli, Filippo Maria](#); [Pesce, Luca](#); [Scardigli, Marina](#); [Pavone, Francesco Saverio](#); [Roffilli, Matteo](#); [Silvestri, Ludovico](#); [Hof, Patrick R.](#); [Boas, David A.](#); [Fischl, Bruce](#); [Morgan, Leah](#); [Yang, Jiarui](#); [Chang, Shuaibin](#); [Laffey, Jessie](#); [Magnain, Caroline](#); [Varadarajan, Divya](#); [Wang, Hui](#); [Frost, Robert](#); [Kouwe, Andre van der](#); [Player, Allison Stevens](#); [Atzeni, Alessia](#); [Gonzalez, Juan Eugenio Iglesias](#); [Balbastre, Yael](#); [Vera, Matthew](#); [Cordero, Devani](#); [Nestor, Kimberly](#); [Ammon, William](#); [Nolan, Jackson](#); [Mora, Jocelyn](#); [Pallares, Erendira Garcia](#); [Augustinack, Jean](#); [Diamond, Bram](#); [Fogarty, Morgan](#); [Boyd, Emma](#); [Varghese, Merina](#); [Dalca, Adrian V.](#); [Edlow, Brian](#); [Frosche, Matthew](#); [Chen, I-Chun Anderson](#); [Wicinski, Bridget](#)

Magnetic resonance imaging (MRI) is used to establish a macroscopic reference coordinate system of laminar and cytoarchitectural boundaries. Cell counting is obtained with both traditional immunohistochemistry, to provide a stereological gold standard, and with a custom-made inverted confocal light-sheet fluorescence microscope (LSM) for 3D imaging at cellular resolution. Finally, polarization-sensitive optical coherence tomography (PSOCT) enables registration of the distorted histological cell typing obtained with LSM to the MRI-based atlas coordinate system. [- see less]

Keywords: [multi-modal imaging](#) [MRI](#) [OCT](#) [SPIM](#) [human cortex](#) [Broca's area](#) [Motor cortex](#) [Stereology](#)

Licenses: [spdx:CC-BY-4.0](#)

Subject matter of the dataset

- Broca's Area
- Motor Cortex

Access information

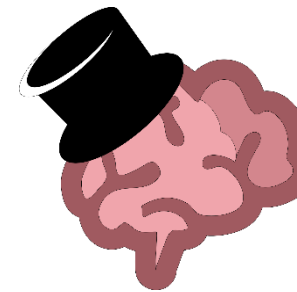
[dandi:OpenAccess](#)

Assets Summary

DANDI: Distributed Archives for Neurophysiology Data Integration

www.dandiarchive.org

“The BRAIN Initiative archive for publishing and sharing neurophysiology data including electrophysiology, optophysiology, and behavioral time-series, and images from immunostaining experiments.”



DANDI

```
pip-install dandi
```

BIDS extension proposal 31 (BEP31)
(The Brain Imaging Data Structure)

Acquisition and control software (SPIMlab + QtLab)



SPIMlab

File ?

Cameras Lasers Settings Messages

2,000 1,500 1,000 500 0

0 500 1,000 1,500 2,000

2,000 1,500 1,000 800 600 400 200 0

0 500 1,000 1,500 2,000

0 500 1,000 1,500 2,000

0 500 1,000 1,500 2,000

Translational stages

	Curr. Pos.	Set Pos.	Step down	Step up	Step size	Velocity
X	140.1000	140.1000	-	+	0.1000	1.0000
Y	126.0000	126.0000	-	+	1.0000	5.0000
Z	15.4000	15.4000	-	+	1.0000	0.5000
Z L	5.4100	5.4100	-	+	0.0100	0.1000
Z R	5.4400	5.5000	-	+	0.0100	0.1000

Acquisition

From To Step

X axis 150.0000 mm 180.0000 mm 0.1000 mm

Y axis 260.0000 mm 260.0000 mm 0.0000 mm

Path /mnt/ramdisk

Exposure Time 0.100 ms Set

Galvo Ramp

Offset	Amplitude	Delay	Fraction
-0.110 V	4.150 V	-0.450 ms	97.5 %
0.000 V	-4.150 V	-0.400 ms	97.5 %

Progress

Time / ETA:

Controls

Initialize

Start free run

Start acquisition

Stop capture

Free run

- developed in C++ using Qt
- 7000 SLOC (Single Lines of Code)
- multi threaded
- data rate 1 GB/s with two cameras
- flexible and modular architecture

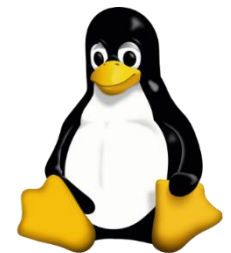
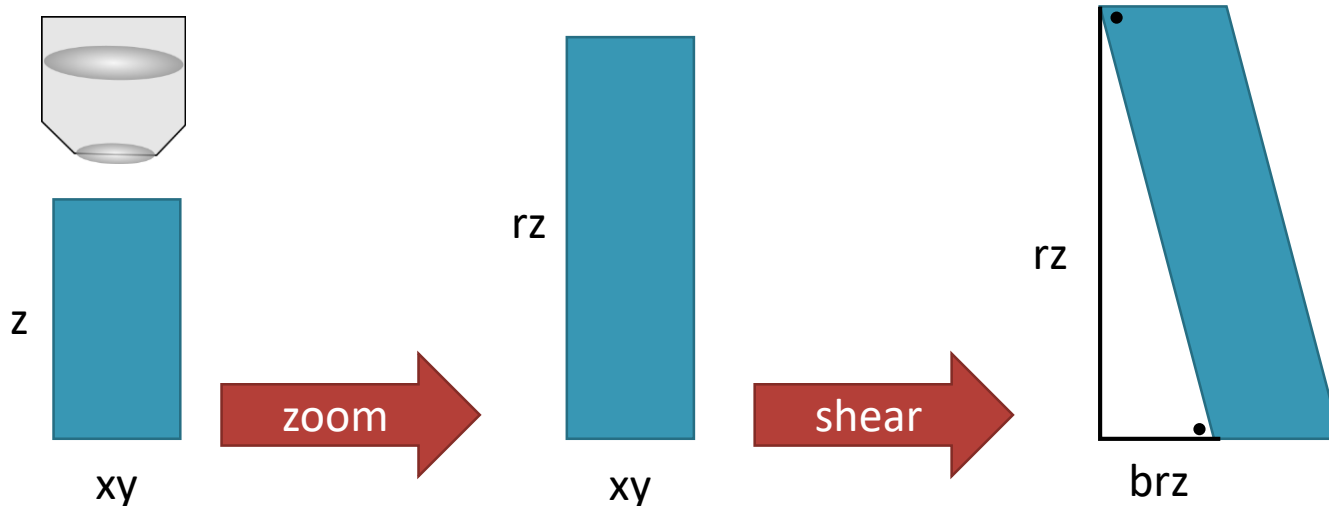


Image reslicing



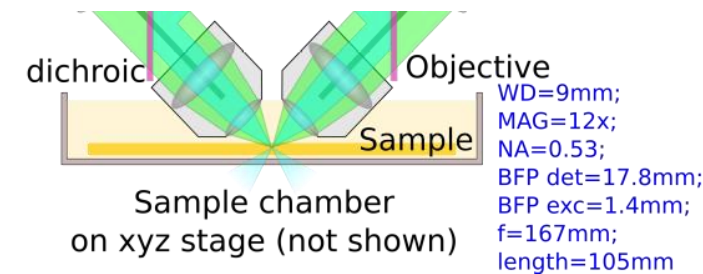
Objective reference frame



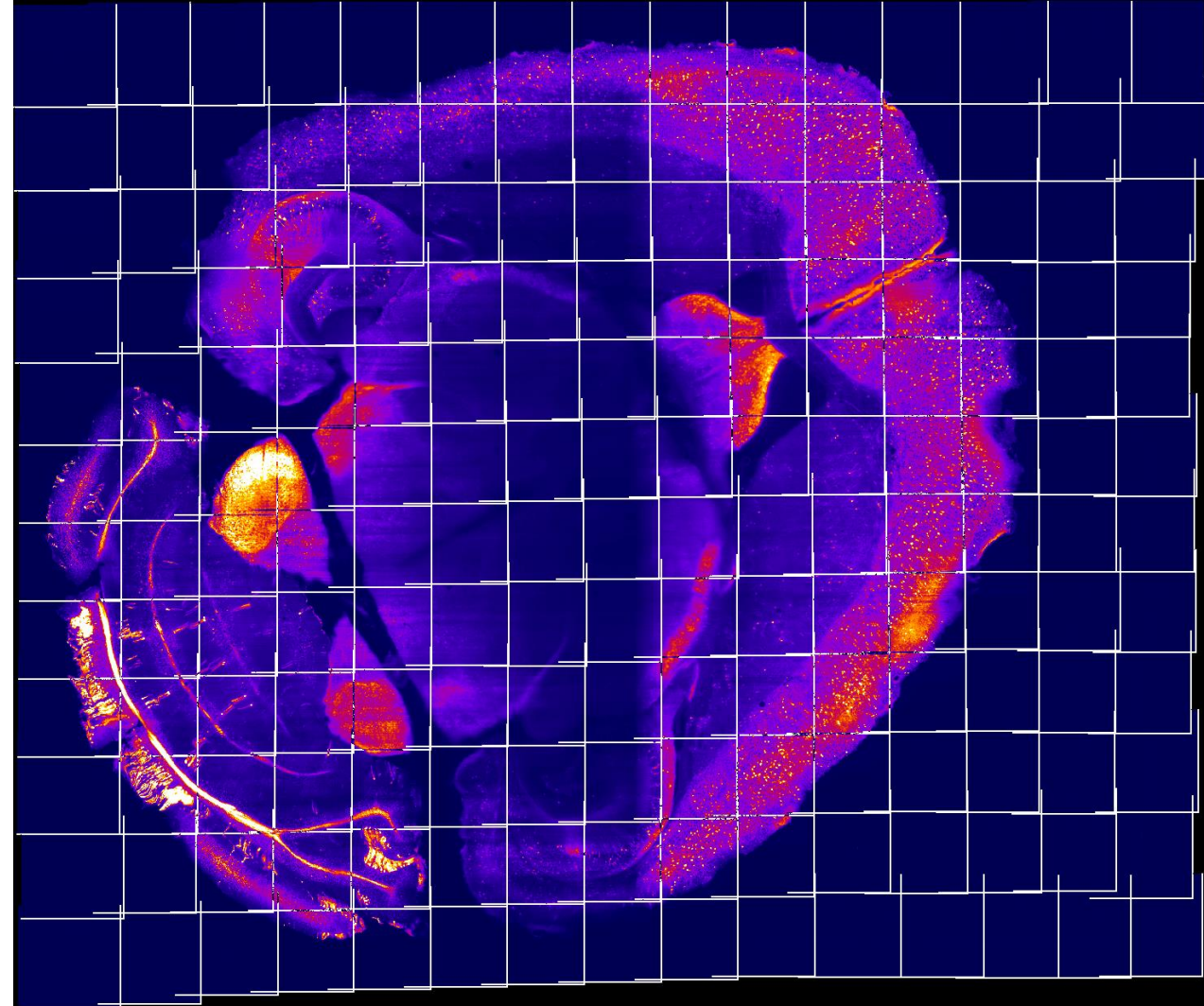
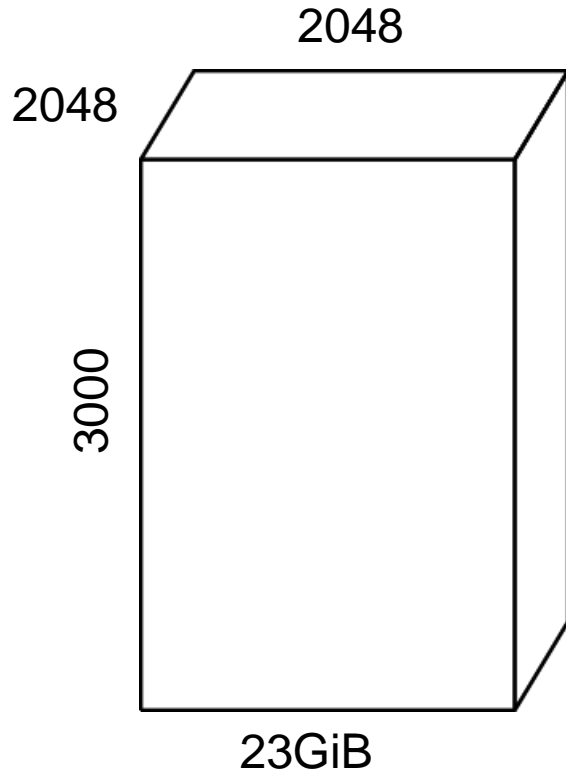
Affine transformation using zooms, rotations and shears to go from the objective's reference frame to the lab's reference frame.

$$A = T R S Z$$

Lab reference frame

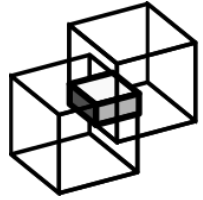


Volumetric stitching



- Sample size (half mouse brain): $0.6 \times 1.2 \times 1.2 \text{ cm}^3$ (ZYX)
- Pixel size: $2 \times 0.65 \times 0.65 \text{ }\mu\text{m}^3$ (ZYX), 16 bit
- Mosaic: 15 x 12 stacks, 23 GiB each
- Whole dataset: 4.2 TiB

Volumetric stitching: ZetaStitcher



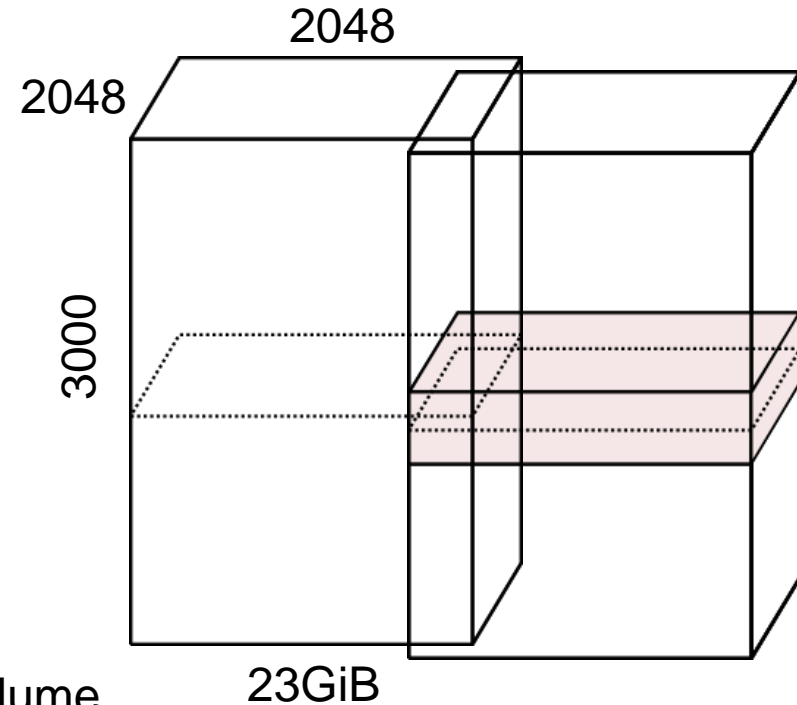
ZetaStitcher



GitHub

<https://github.com/lens-biophotonics/ZetaStitcher>

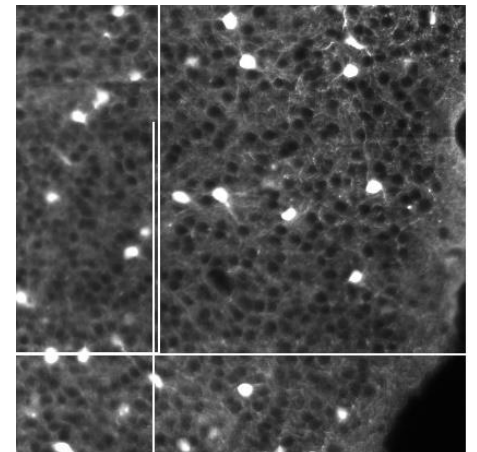
- Developed entirely in Python
- Open source (GPLv3)
- High throughput



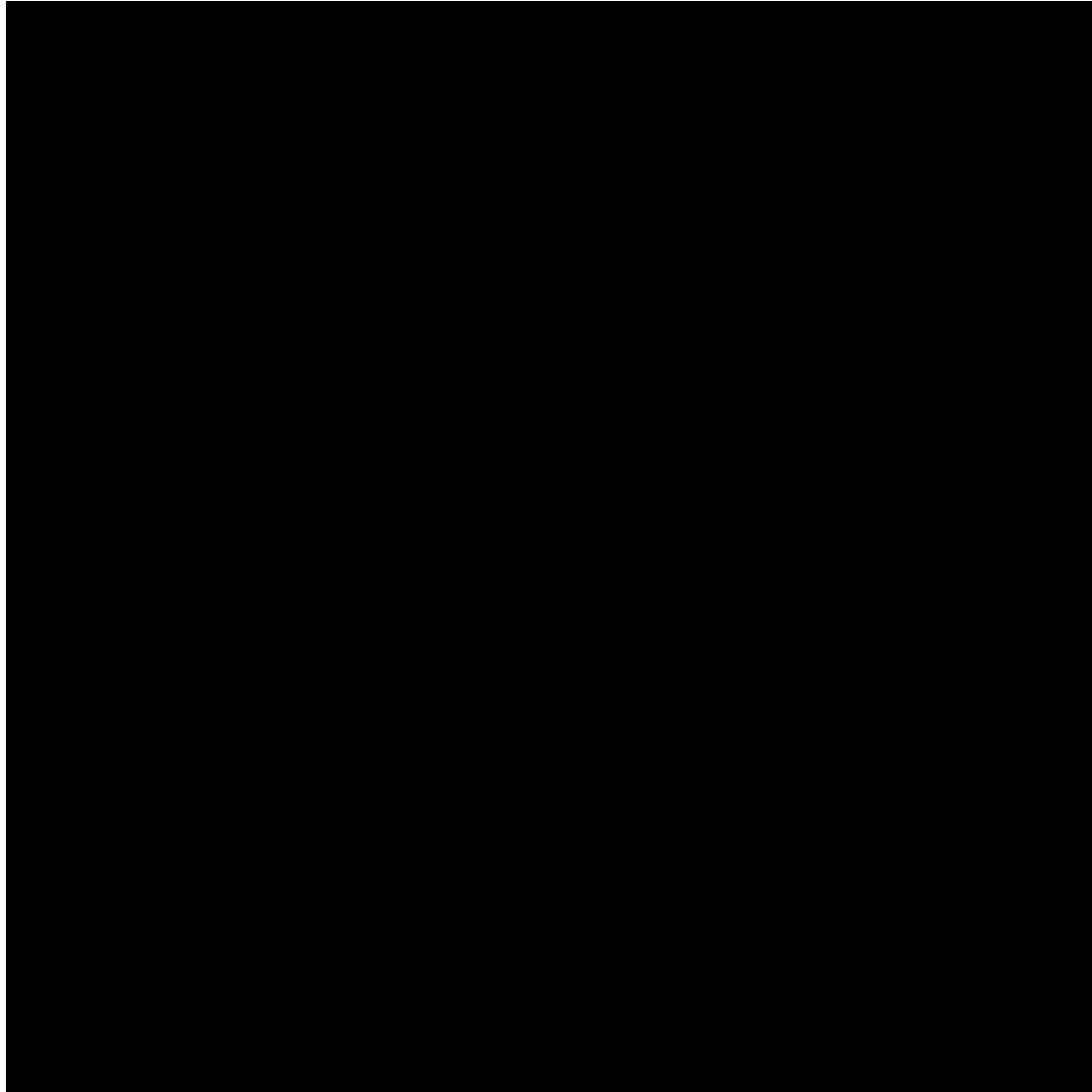
A powerful and simple Python API to query arbitrary regions within the fused volume.

```
>>> from zetastitcher import VirtualFusedVolume  
>>> vfv = VirtualFusedVolume('stitch.yml')  
>>> vfv.shape  
(2985, 18924, 23486)
```

```
a = vfv[2000:2500, 12000:13000, 15500:16500]
```



Fused volume (half mouse brain)



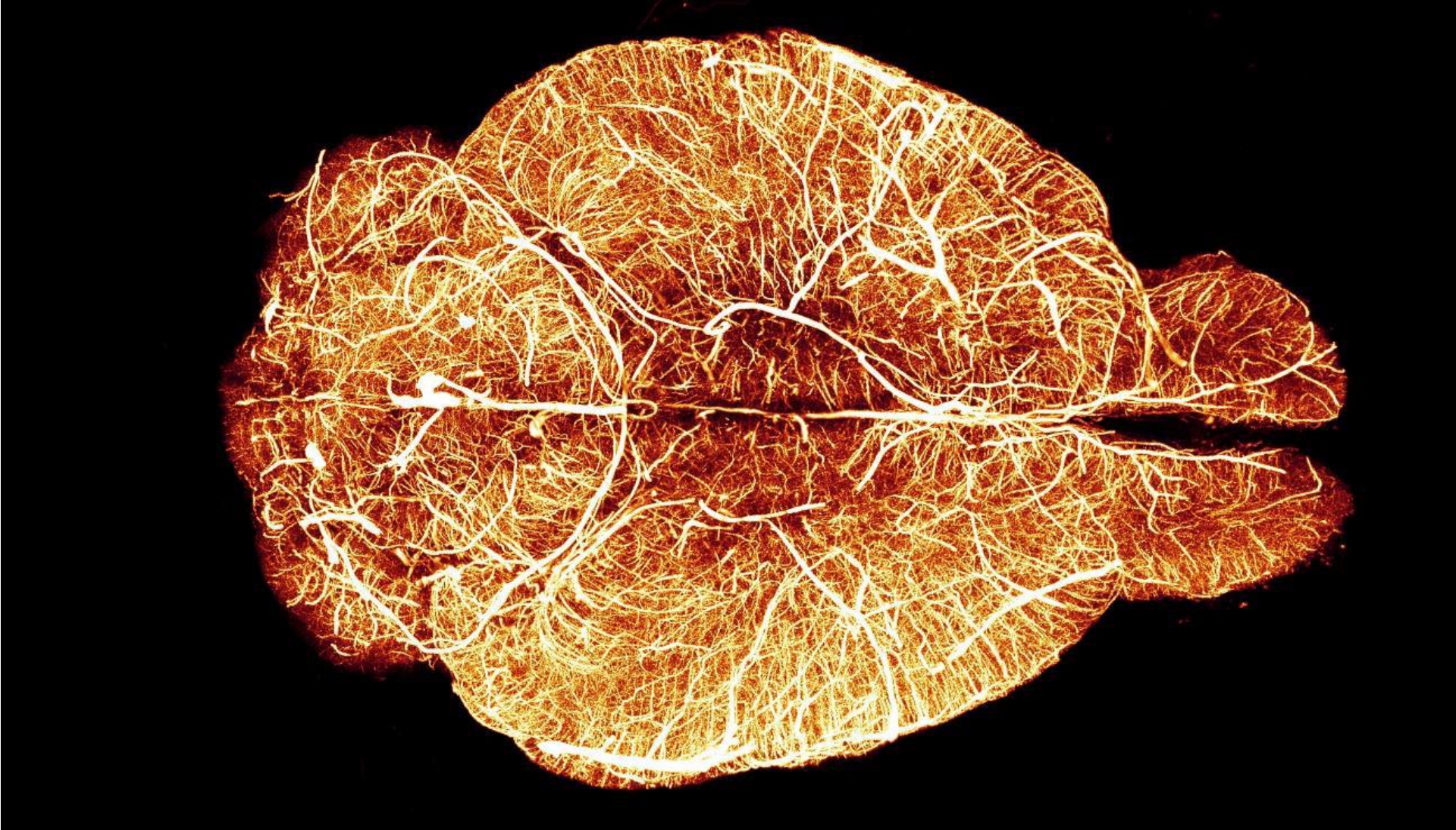
Downscaled volume:

px size: $10 \times 10.4 \times 10.4 \mu\text{m}^3$ (ZYX)

Original volume:

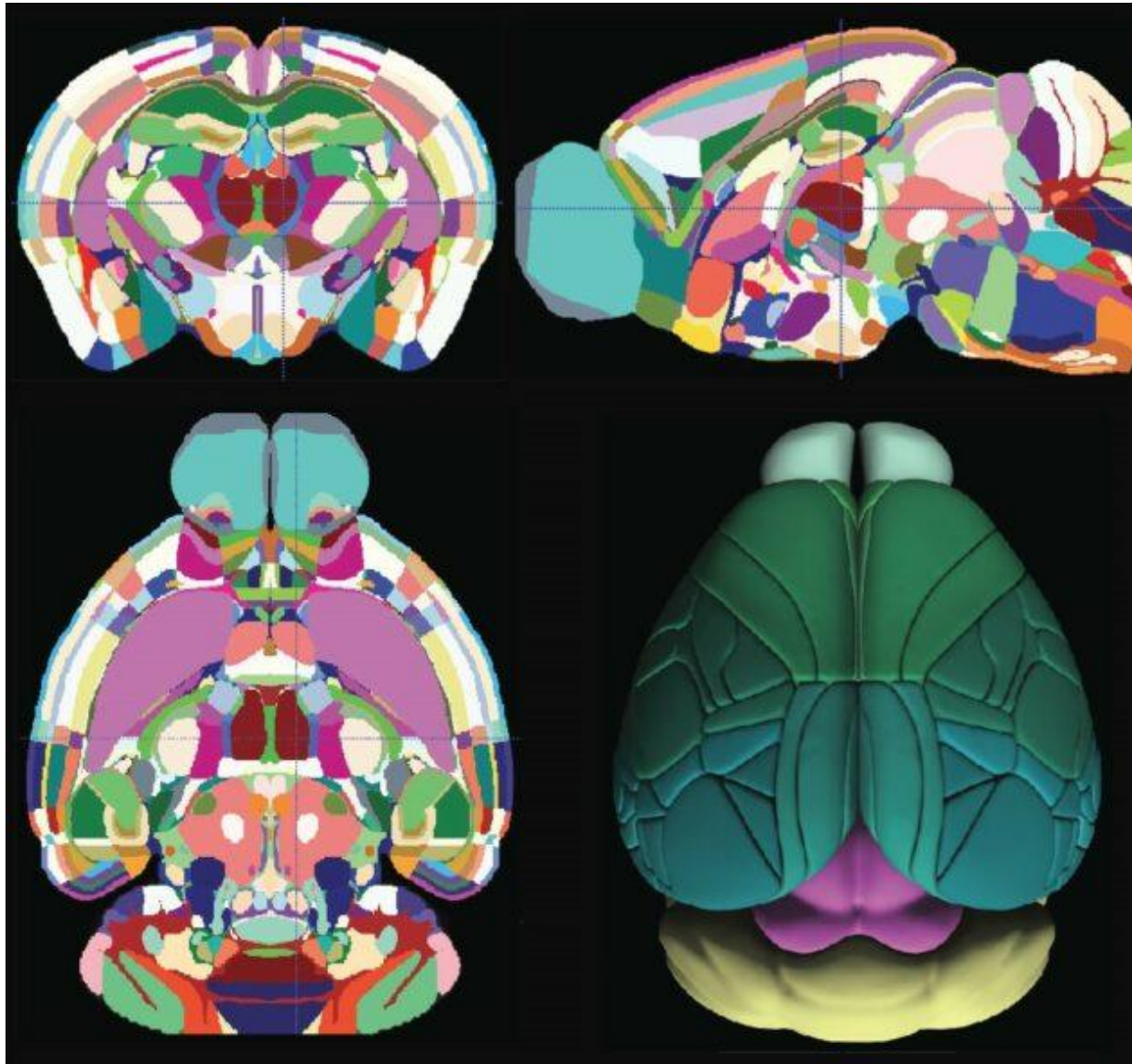
$2985 \times 18924 \times 23486$ px (Z \times Y \times X) $\approx 1.3 \cdot 10^{12}$ voxels

Whole mouse brain tomography with LSFM: vasculature



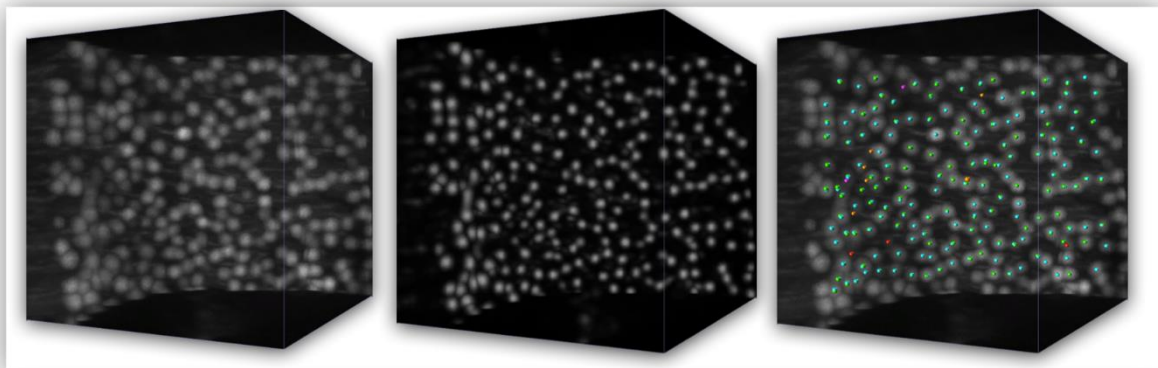
Di Giovanna, Antonino Paolo, et al. "Whole-brain vasculature reconstruction at the single capillary level." *Scientific reports* 8.1 (2018): 1-11.

Alignment to the Allen Mouse Brain Reference Atlas



Atlas warping using ANTs on downsampled version of the dataset (Avants et al. IEEE transactions on medical imaging, <http://stnava.github.io/>)

Neuron localization in the whole mouse brain

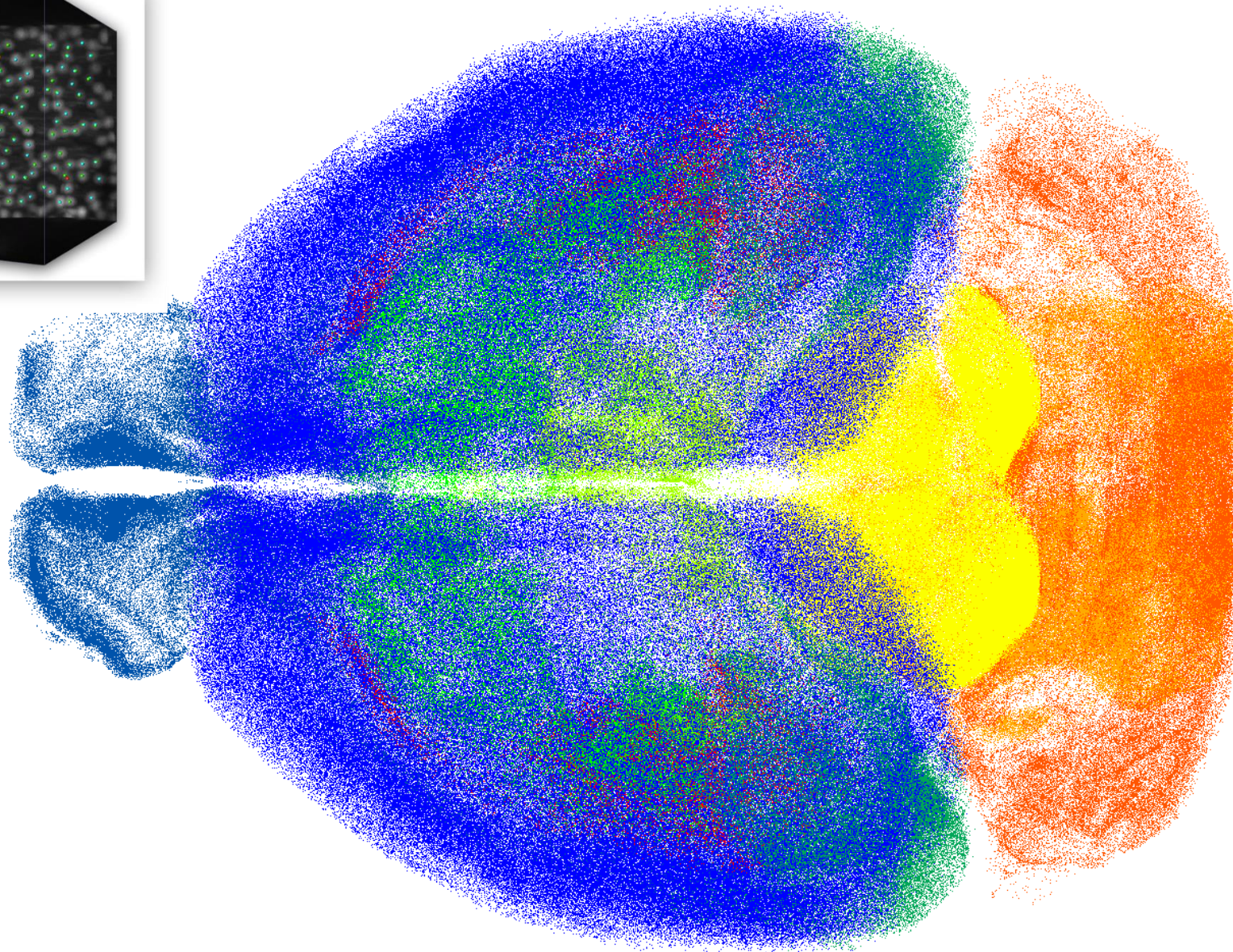


Semantic deconvolution

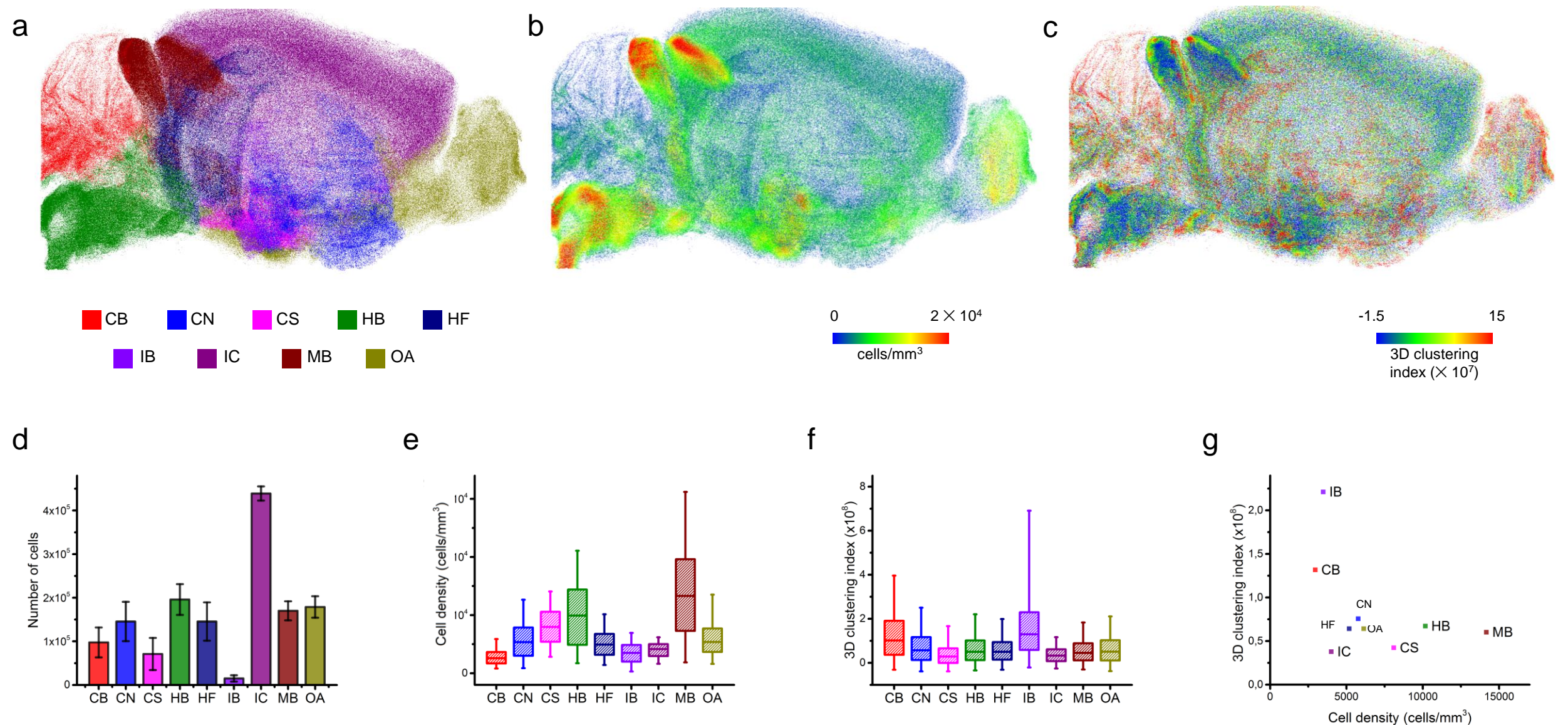
SST-positive cells

About 1.5×10^6 neurons

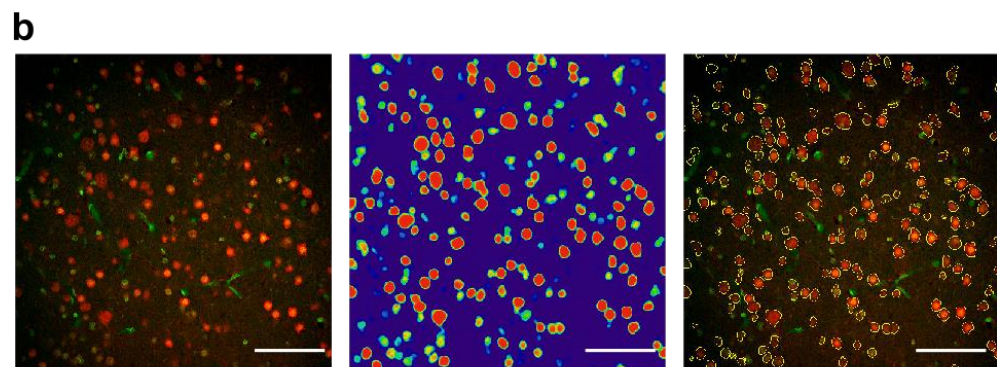
University of Florence, Department of
Information Engineering (DINFO)
Paolo Frasconi



Quantitative data analysis



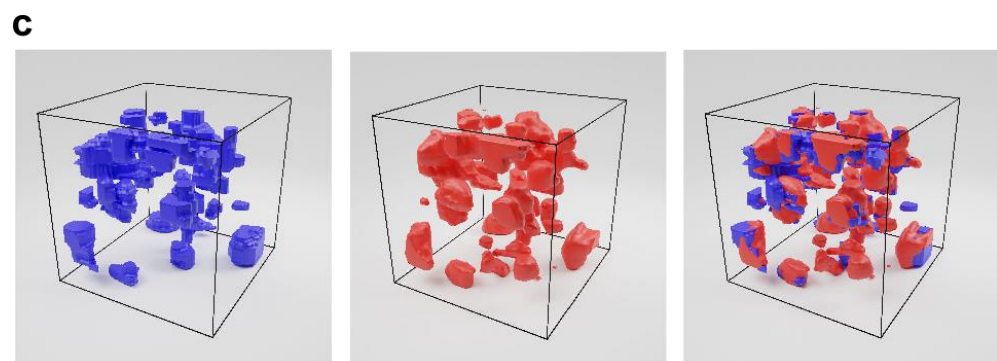
Digital histology



TPFM image

2D probabilistic heatmap

2D segmentation

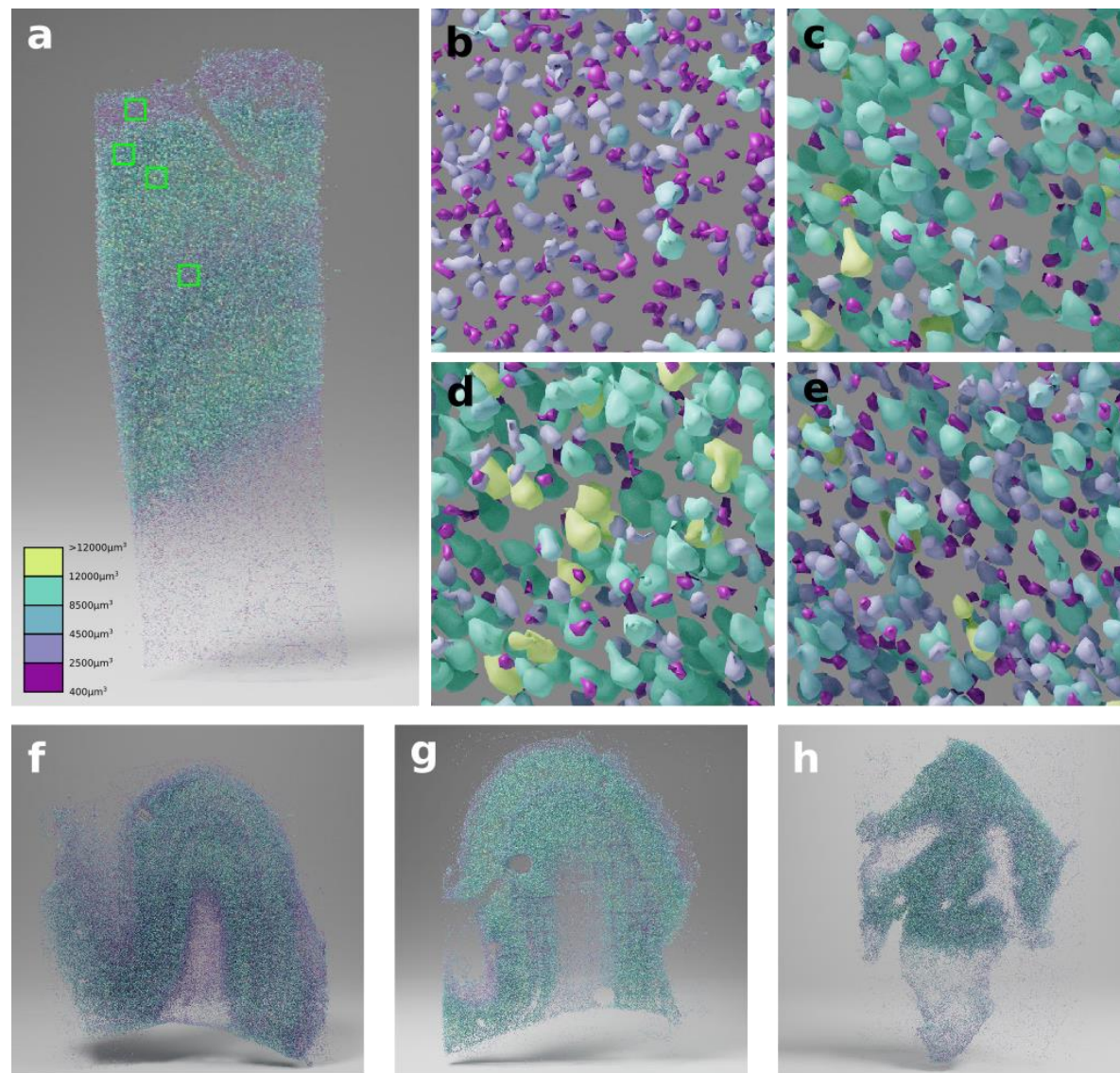


GT

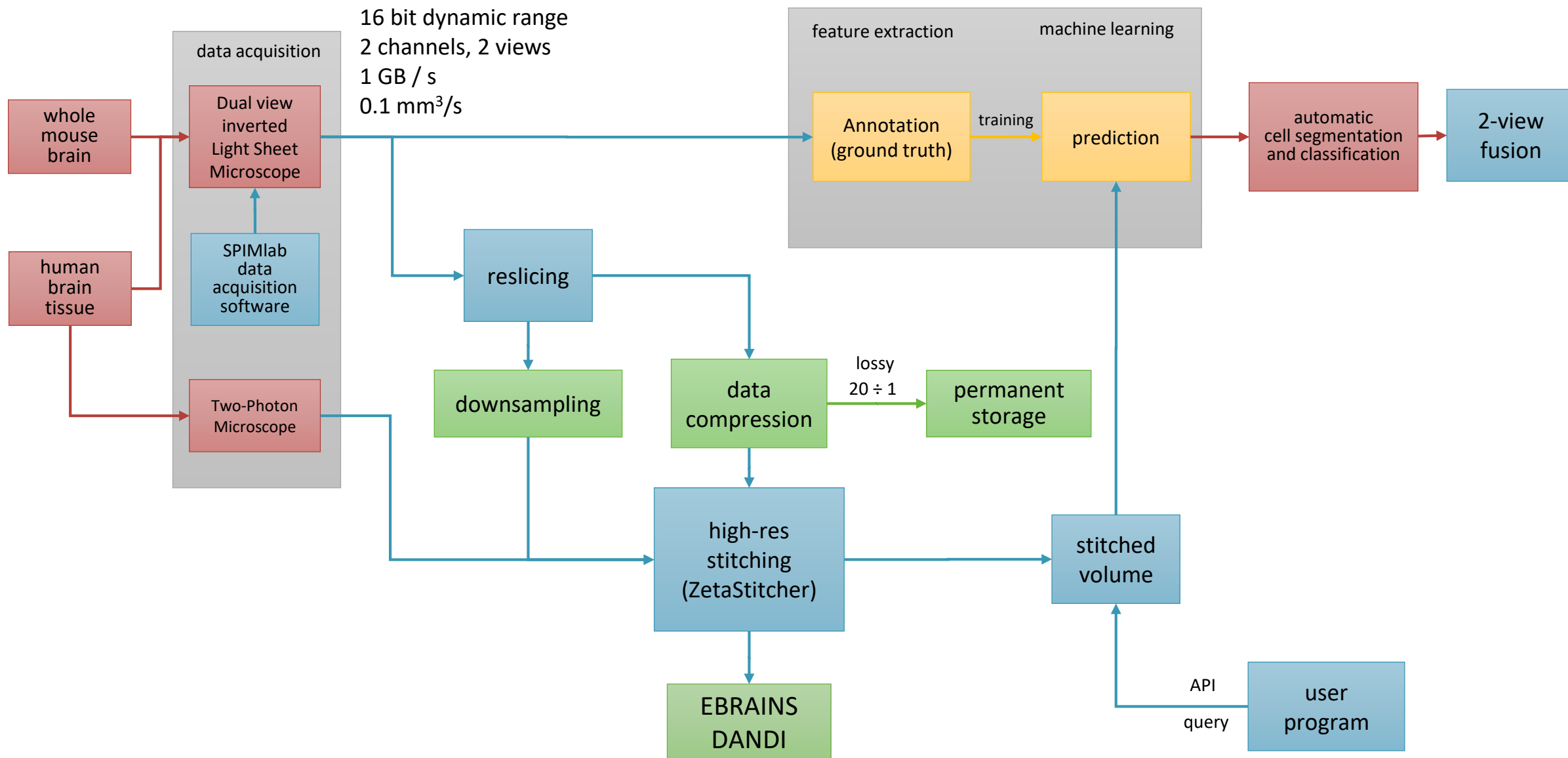
2.5D

Both

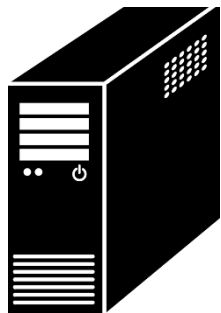
I. Costantini, G. Mazzamuto et al. **Biomedical Optics Express**, Vol. 12, [Issue 6](#), pp. 3684-3699 (2021)



Data management workflow



Computing cluster



1 GPU server with:
1x GeForce 3090 24GB
256GB RAM
AMD Ryzen Threadripper 32 core



2 GPU servers each with:
4x GeForce 2080Ti 12GB
192GB RAM
Intel Xeon 24 core



1 GPU server with:
1x Pascal P100 12GB
128GB RAM
Intel Xeon 16 core



NAS storage
500 TB



Central manager node

- HTCondor
- Docker registry
- LDAP



CINECA (Bologna)



Internet

1 Gbit/s

10 Gbit/s

Acknowledgements



European Laboratory for Non-linear Spectroscopy (LENS)

Ludovico Silvestri, Irene Costantini, Vladislav Gavryusev, Giuseppe Sancataldo, Marina Scardigli, Luca Pesce, Filippo Maria Castelli, Francesco S. Pavone



University of Florence, Department of Information Engineering (DINFO)

Francesco Orsini, Paolo Frasconi



Bioretics Srl, Cesena

Mattia Neri, Matteo Roffilli

Human Brain Project



A grayscale, high-magnification microscopic image of a biological structure, possibly a cross-section of a lens or a similar tissue. The image shows a complex, layered structure with numerous small, bright, granular spots scattered throughout. A central, darker region is visible, and the overall appearance is highly textured and detailed. A semi-transparent black rectangular box is overlaid in the center, containing white text.

Thank you!
www.ino.it
bio.lens.unifi.it
github.com/lens-biophotonics