FET FLAGSHIPS

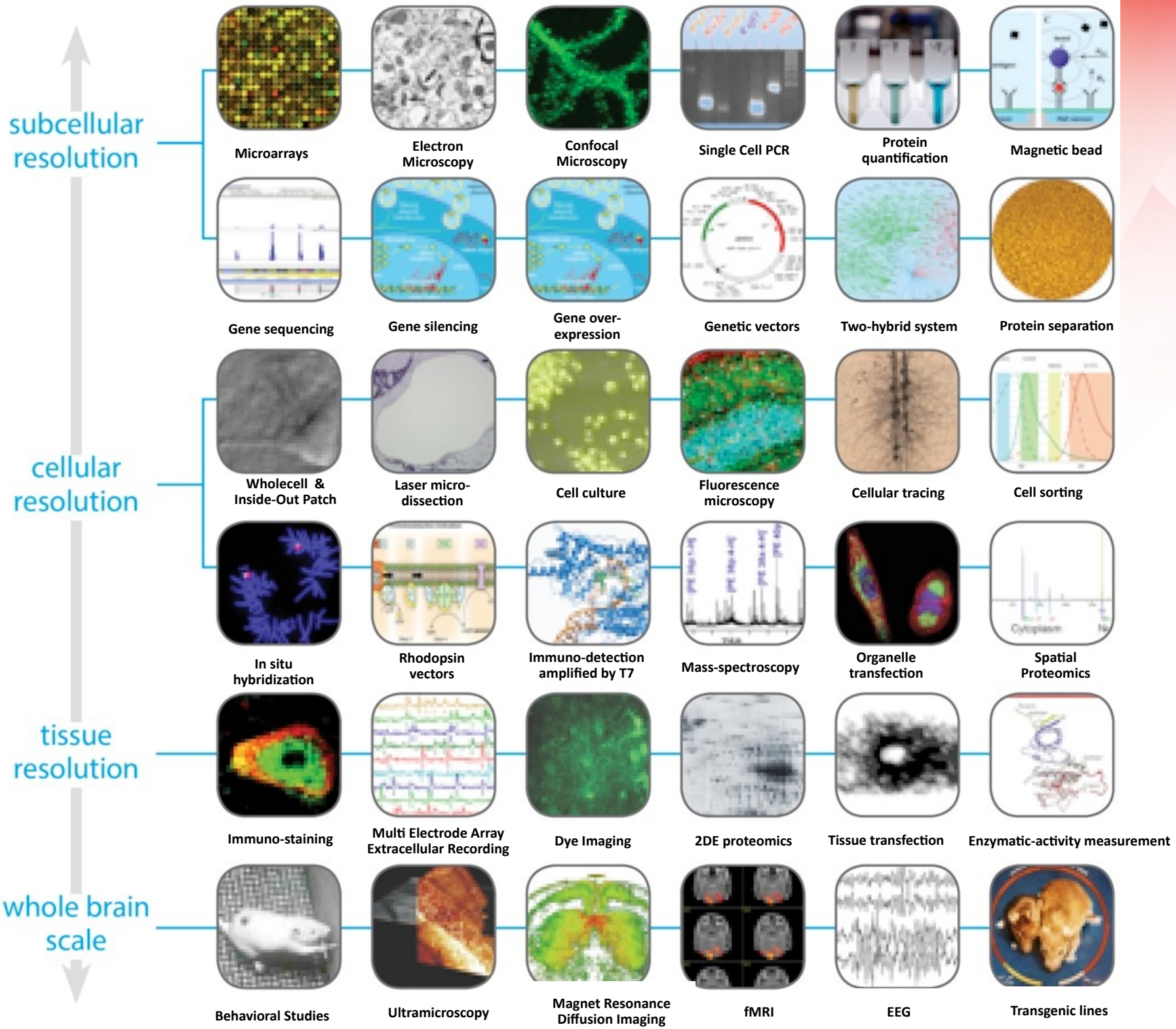
Human Brain Project
Unifying our understanding of the human brain.

European Commission

Open Data and the Human Brain Project

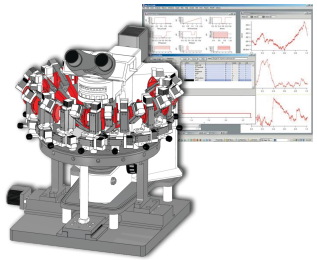
Sean Hill

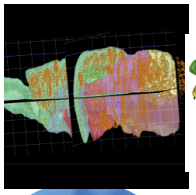
EGI 2015, Lisbon

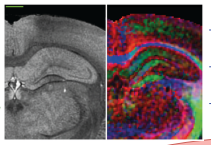

subcellular resolution
cellular resolution
tissue resolution
whole brain scale
Microarrays
Electron Microscopy
Confocal Microscopy
Single Cell PCR
Protein quantification
Magnetic bead
Gene sequencing
Gene silencing
Gene over-expression
Genetic vectors
Two-hybrid system
Protein separation
Wholecell & Inside-Out Patch
Laser micro-dissection
Cell culture
Fluorescence microscopy
Cellular tracing
Cell sorting
In situ hybridization
Rhodopsin vectors
Immuno-detection amplified by T7
Mass-spectroscopy
Organelle transfection
Spatial Proteomics
Immuno-staining
Multi Electrode Array Extracellular Recording
Dye Imaging
2DE proteomics
Tissue transfection
Enzymatic-activity measurement
Behavioral Studies
Ultramicroscopy
Magnet Resonance Diffusion Imaging
fMRI
EEG
Transgenic lines

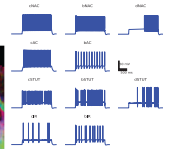
Build, Simulate and Validate Unifying Brain Models

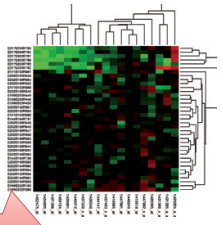
Experimental Data Gathering



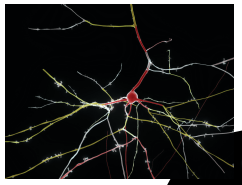


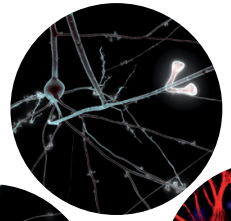


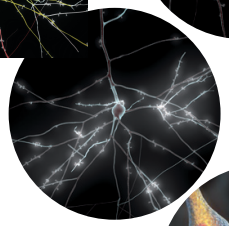


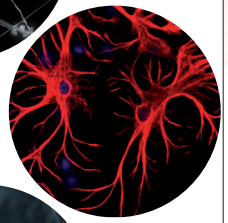


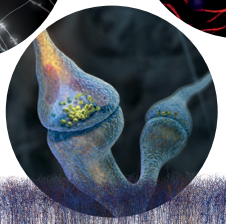
Model Building





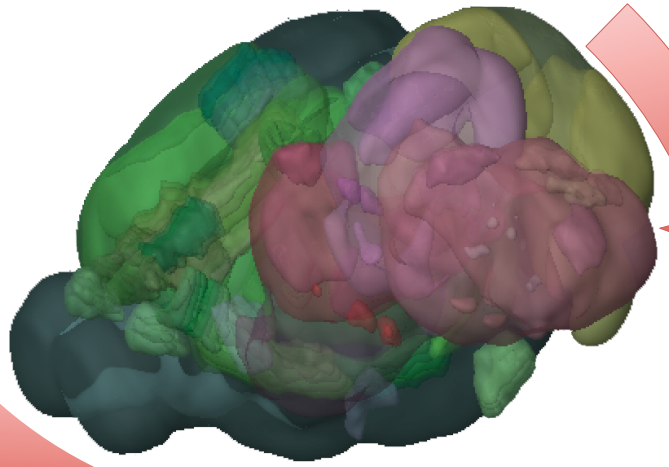


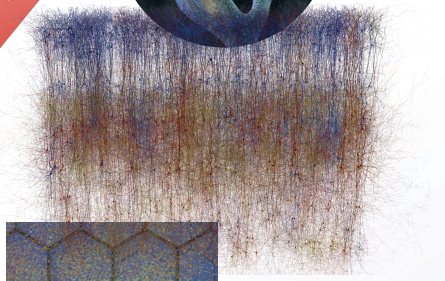


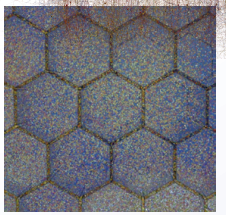


Worldwide Published Data, Models and Literature

Refinement of Models and Experiments

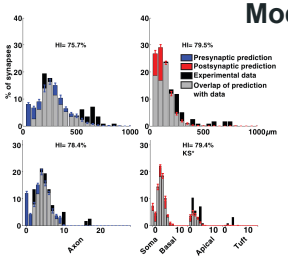


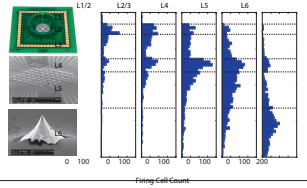


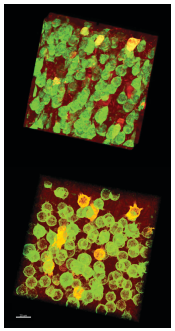


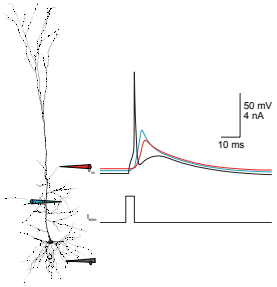
Simulation

Model Validation

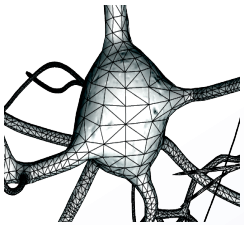








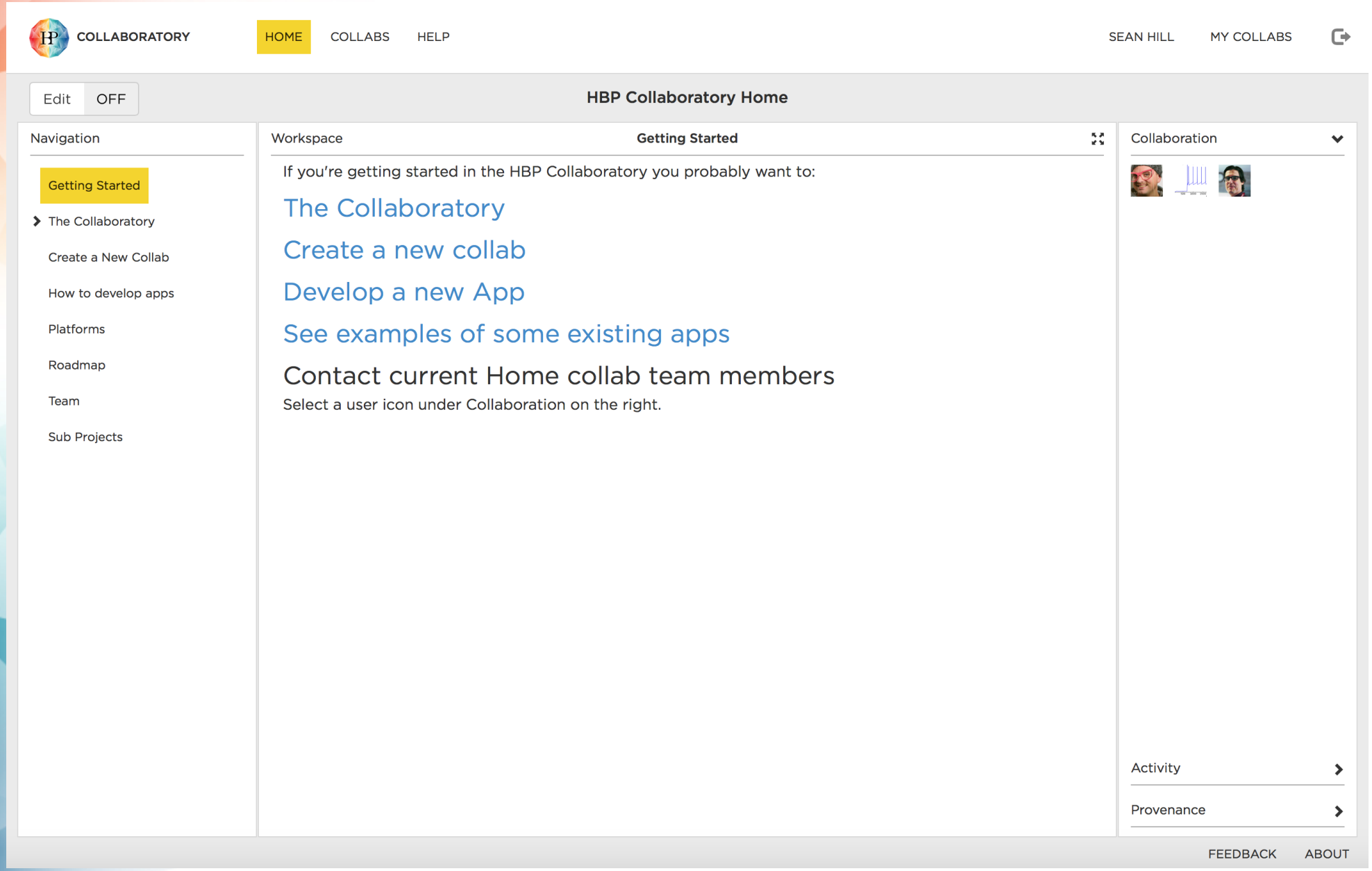
Analysis and Visualization



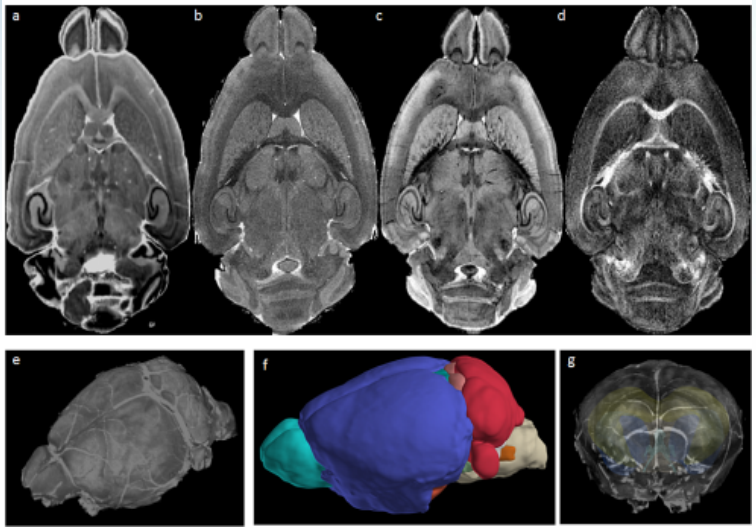
Supercomputing

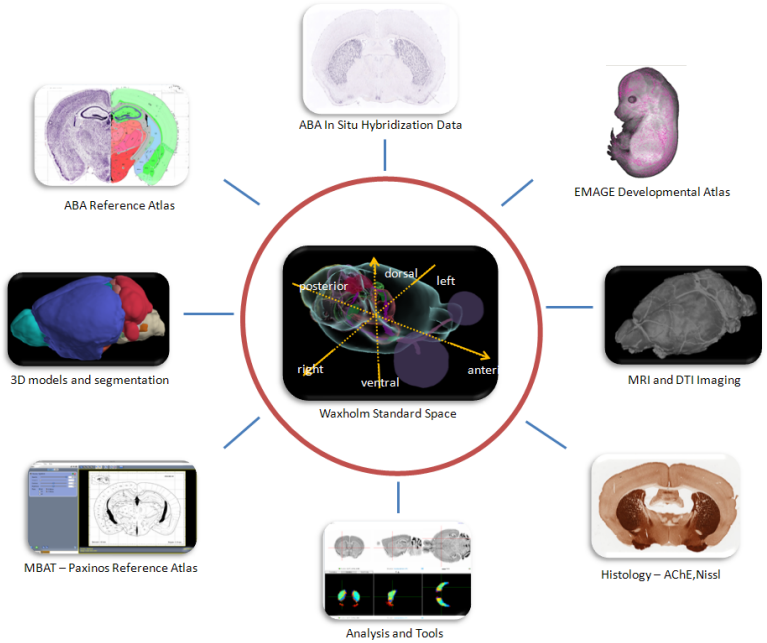
Mouse atlas data courtesy of Allen Institute for Brain Science

HBP Collaboratory



Neuroinformatics Platform





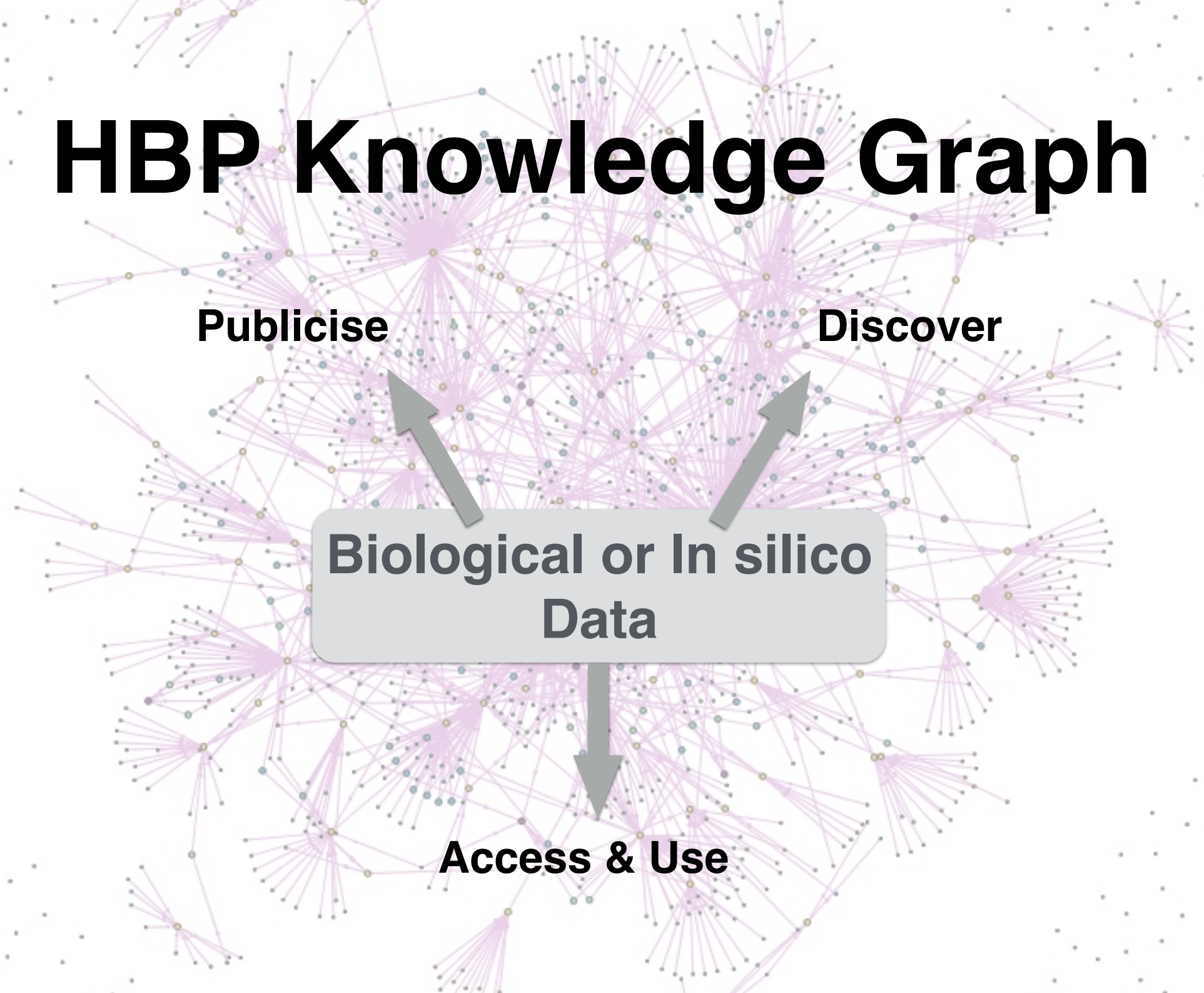
Provide technical capabilities to **federate** neuroscience data, **analyze** structural and functional brain data and to **build and navigate multi-level brain atlases.** This involves:

- **spatial and temporal data registration**
- **ontology development and semantic annotation**
- **predictive neuroscience**
- **machine learning, data mining**
- **track provenance, build workflows.**

Goal: enable an integrated view of the neuroscience data. Prepare data for modeling pipelines

Builds on standards and infrastructure of International Neuroinformatics Coordinating Facility (INCF)

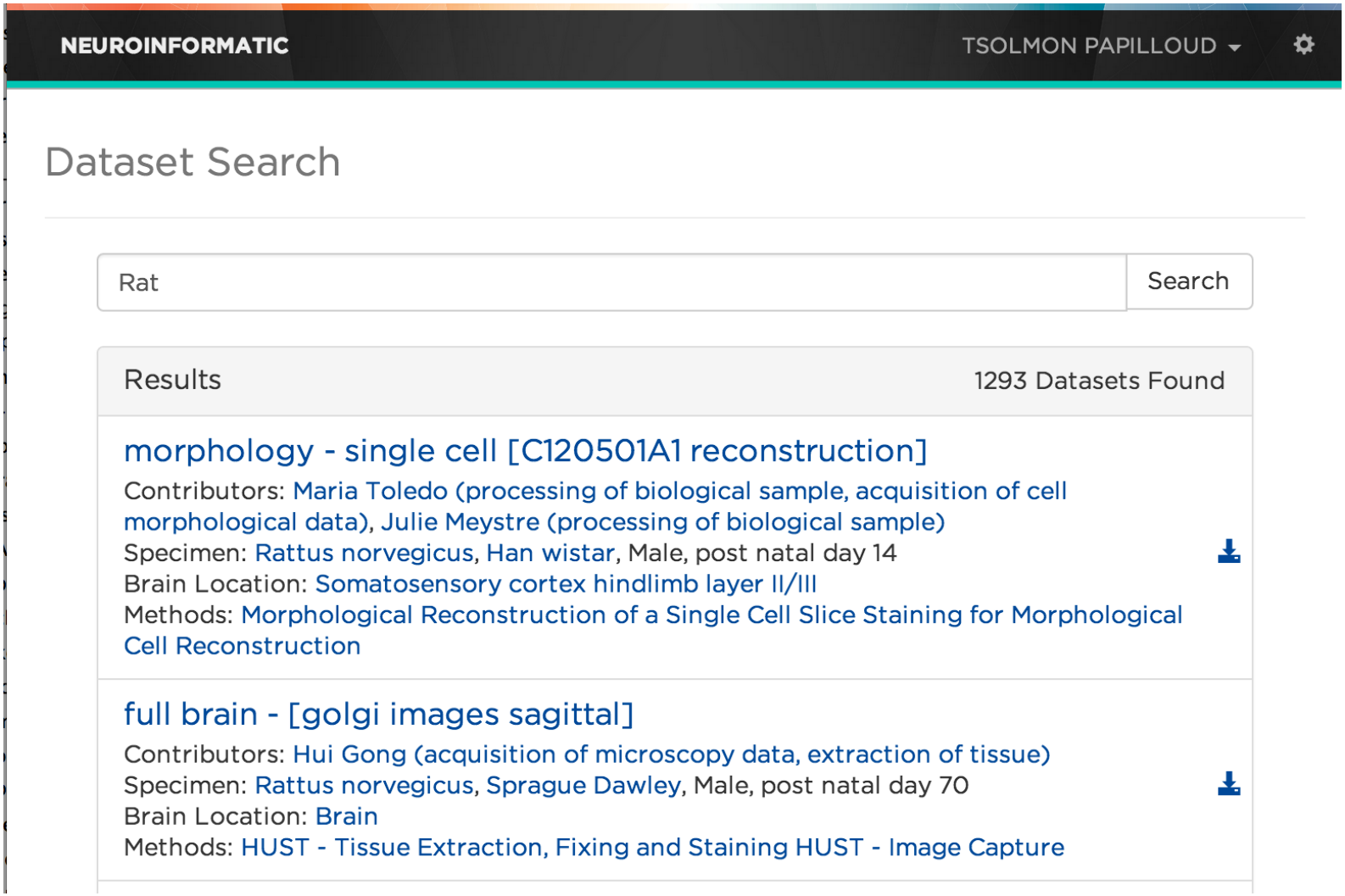
HBP Knowledge Graph



HBP Core Data Model

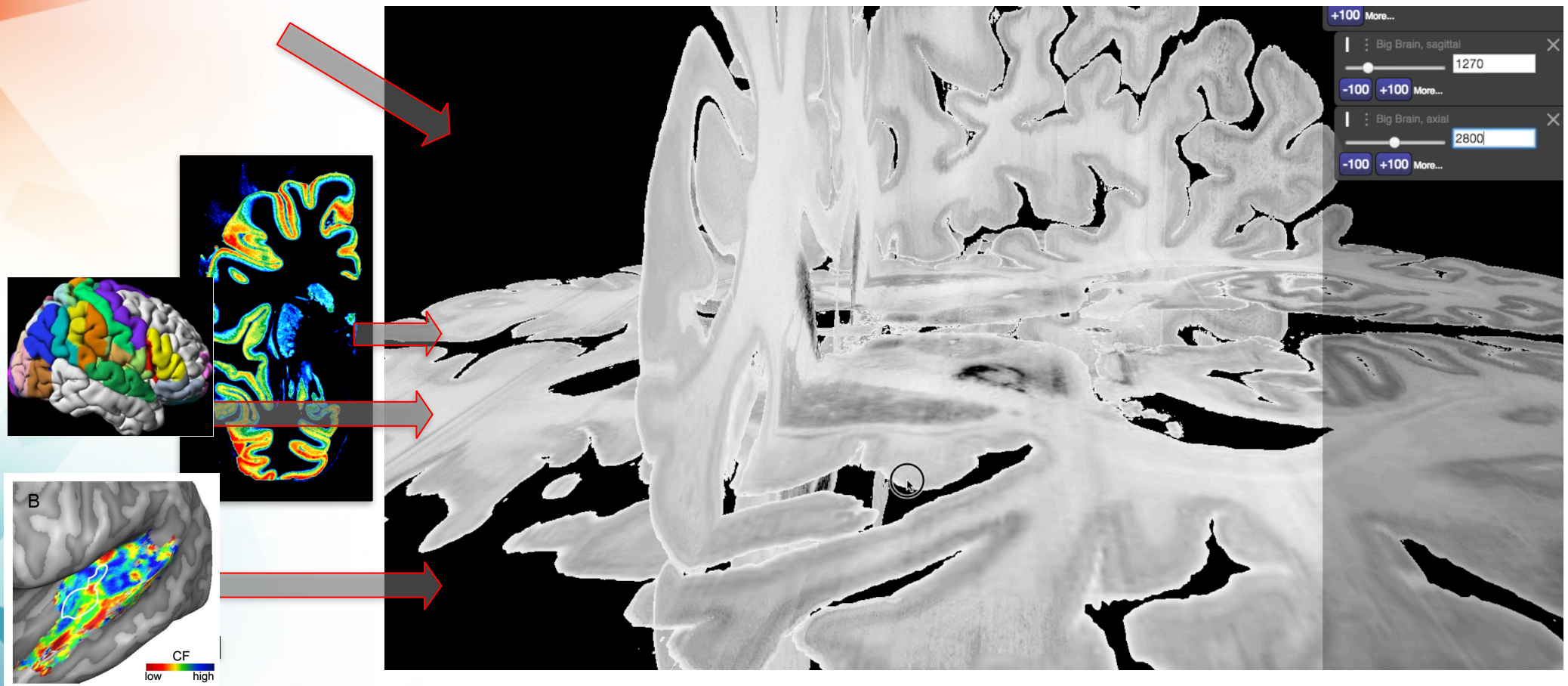
- Built on **W3C PROV-O**
- Represents in vivo, in vitro and in silico **entities**
- Represents **observations**
- Describes properties using **ontologies**
- Records **where** an entity or observation is located
- Tracks **how** data is produced
- Tracks **who** performed experiments/manipulations

Search


NEUROINFORMATIC
TSOLMON PAPILLOUD
Dataset Search
Rat
Search
Results
1293 Datasets Found
morphology - single cell [C120501A1 reconstruction]
Contributors: Maria Toledo (processing of biological sample, acquisition of cell morphological data), Julie Meystre (processing of biological sample)
Specimen: Rattus norvegicus, Han wistar, Male, post natal day 14
Brain Location: Somatosensory cortex hindlimb layer II/III
Methods: Morphological Reconstruction of a Single Cell Slice Staining for Morphological Cell Reconstruction
full brain - [golgi images sagittal]
Contributors: Hui Gong (acquisition of microscopy data, extraction of tissue)
Specimen: Rattus norvegicus, Sprague Dawley, Male, post natal day 70
Brain Location: Brain
Methods: HUST - Tissue Extraction, Fixing and Staining HUST - Image Capture

BigBrain(s): cellular architecture

In the HBP Collaboratory

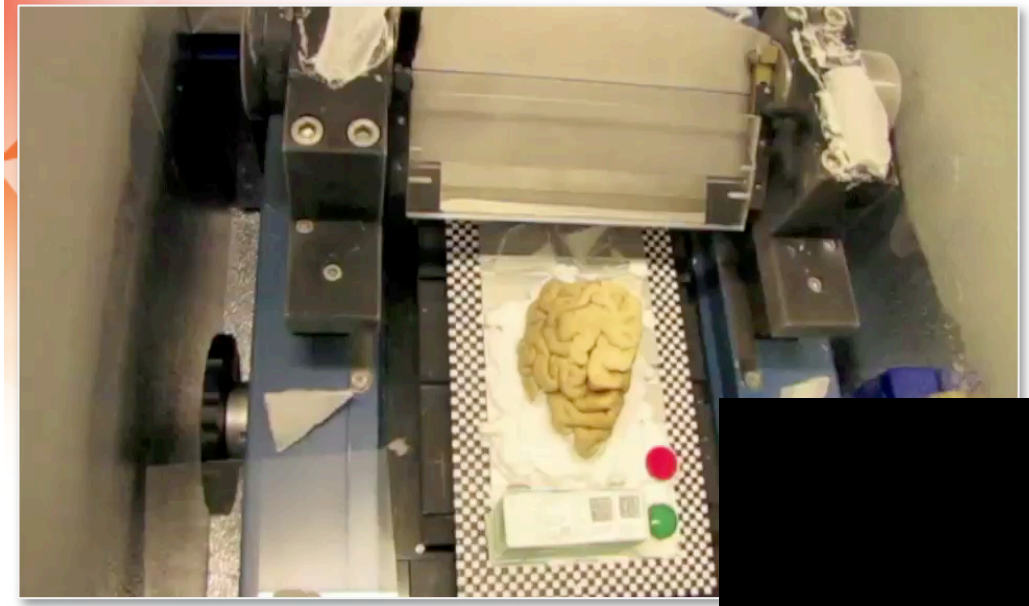


MIT Technology Review 10 Breakthrough Technologies 2014

Amunts et al. *Science*, 340(6139): 1472-1475, 2013

unts et al., in *Brain Inspired Computing*, Grandinetti, Lippert, Petkov, (Ed.), 2014, *Lecture Notes in Computer Science* 8603: 3-14, Springer.

BigBrain(s): fiber architecture



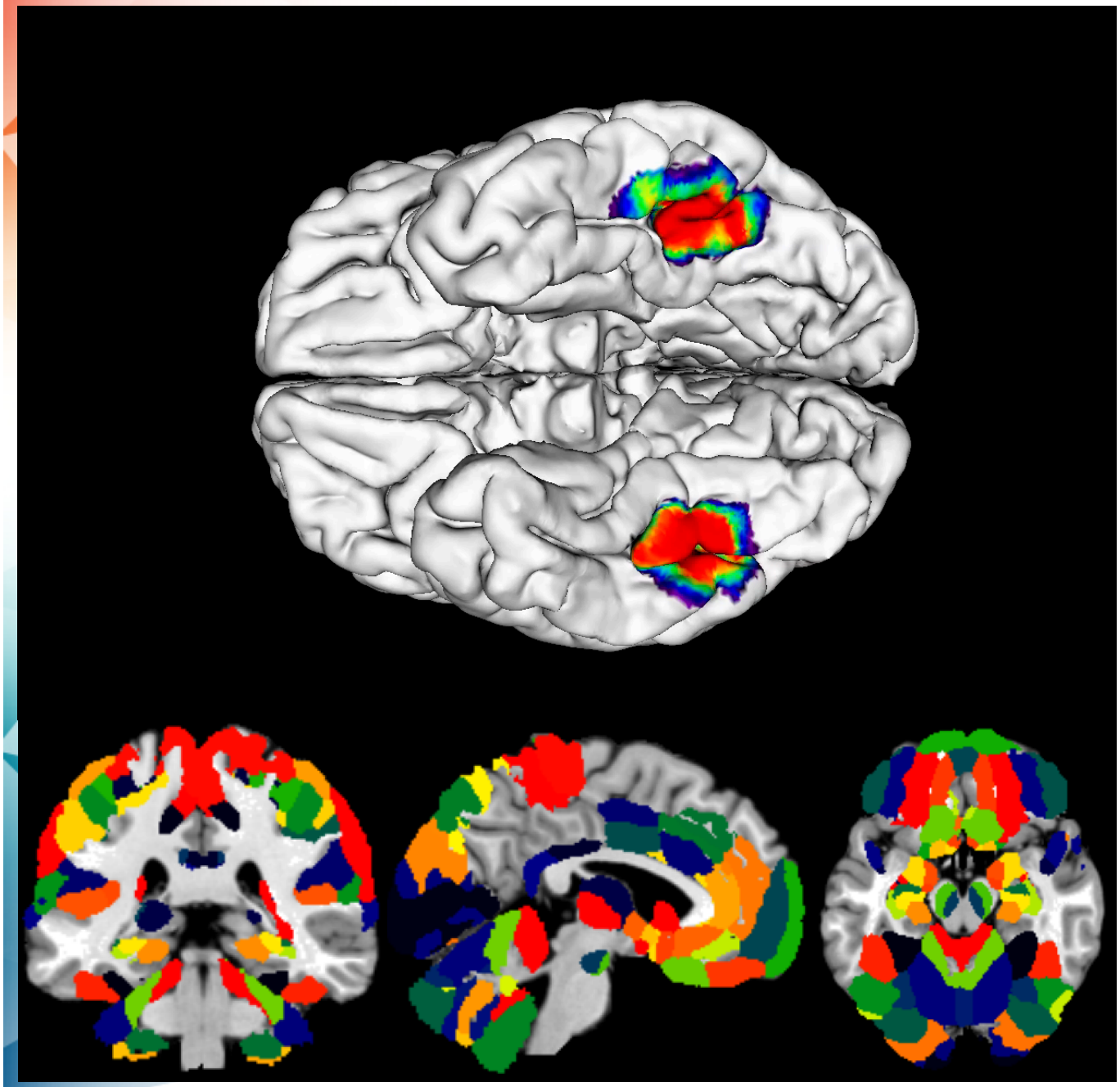
3D-PLI

Polarized Light Imaging as a tool to analyze the fibres at microscopical resolution in 3D

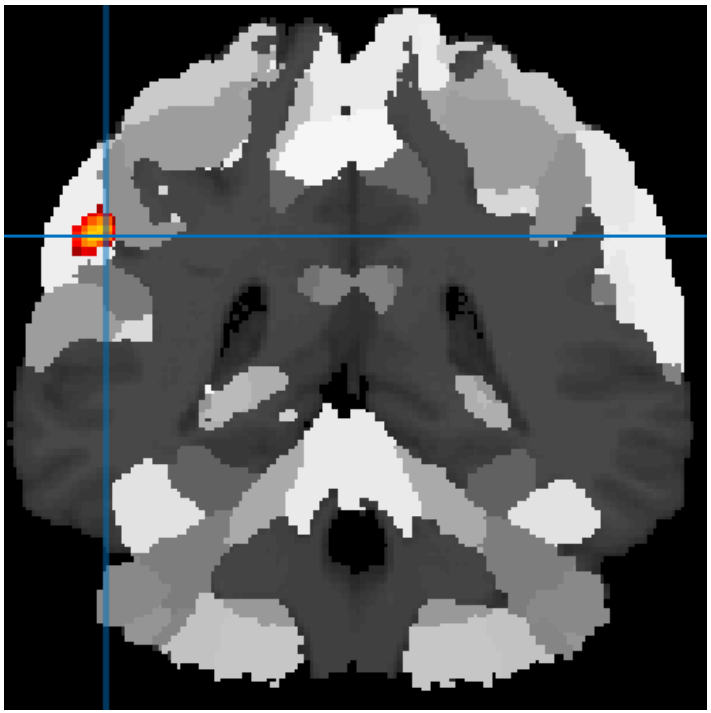
N: 1.500 – 3.000 per brain
Thickness: 100 – 50 µm
Resolution: 1.3 µm in-plane

Axer et al. (2011) *NeuroImage* 54: 1091-1101 & (2011) *Frontiers Neuroinformatics* 5:34; Westhoff (2014) *Proceedings: Architecture of Computing Systems, ARCS2014;* Wiese et al., (2014) in: Chenault, Goldstein (Eds.) *SPIE Proceedings, Polarization: Measurement, Analysis, and Remote Sensing XI*

The MNI space as a common template for different data modalities

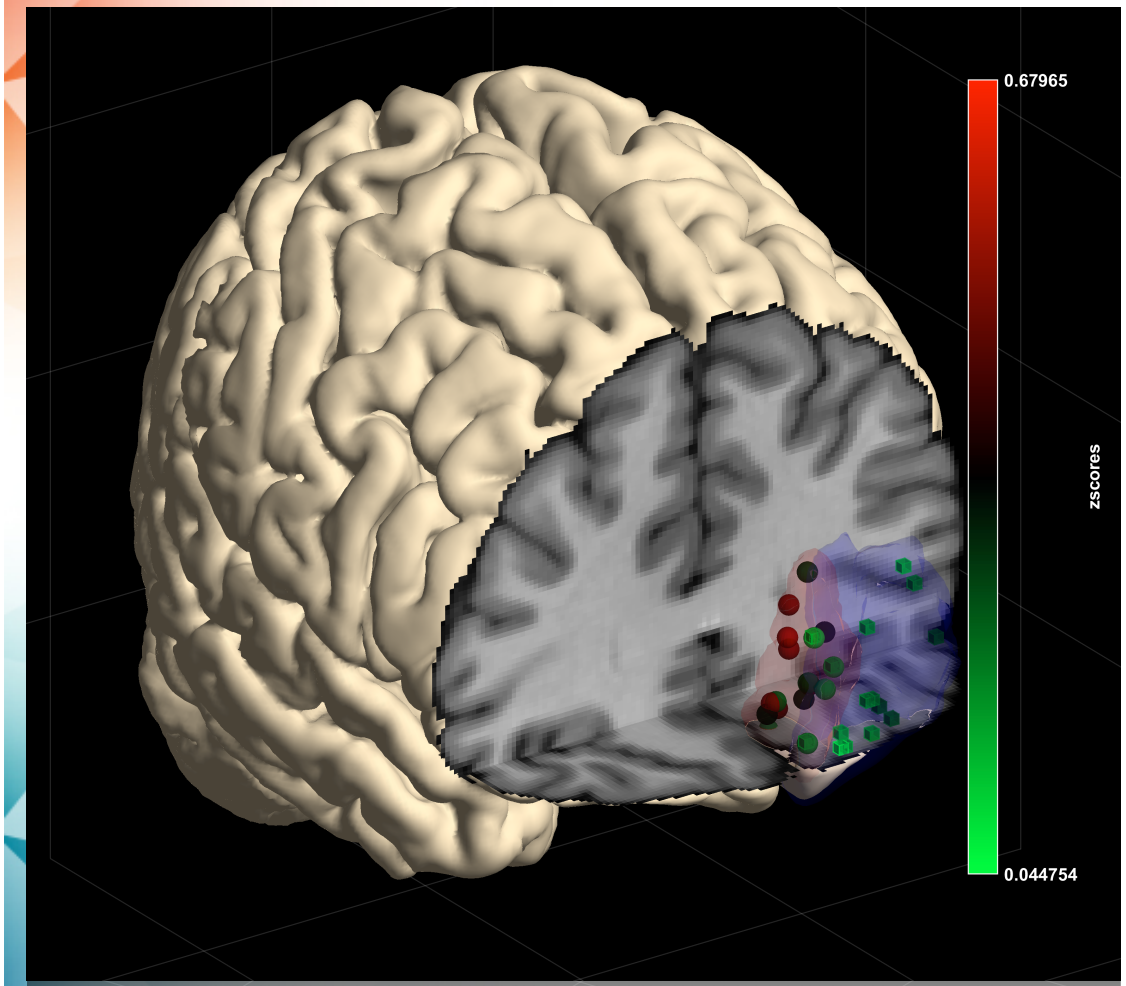


JuBrain atlas for interpreting fMRI data on visuo-motor coordination

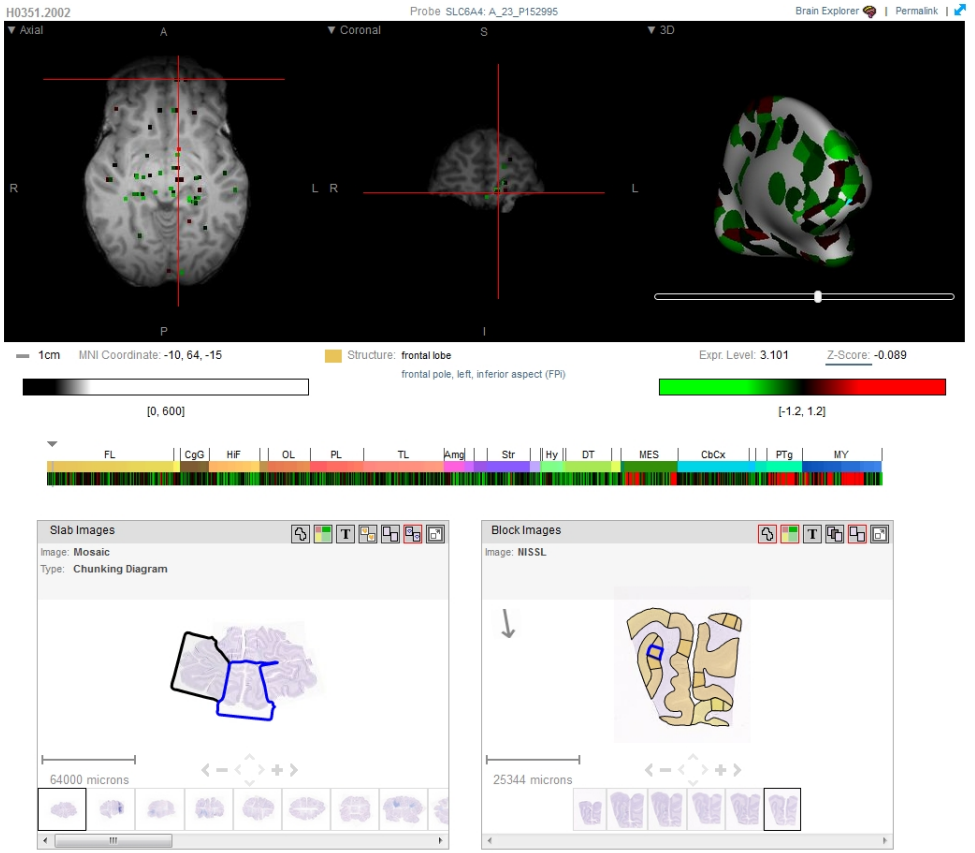


80.5% in left Area PF (IPL) (11.7% activated)
14.9% in left Area PFt (IPL) (1.7% activated)
3.5% in left Area hIP2 (IPS) (1.0% activated)

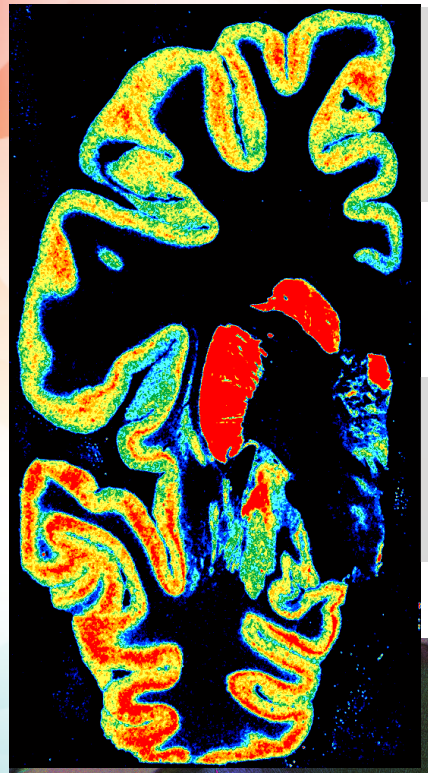
JuBrain connects to *Allen Brain*



ALLEN BRAIN ATLAS
DATA PORTAL

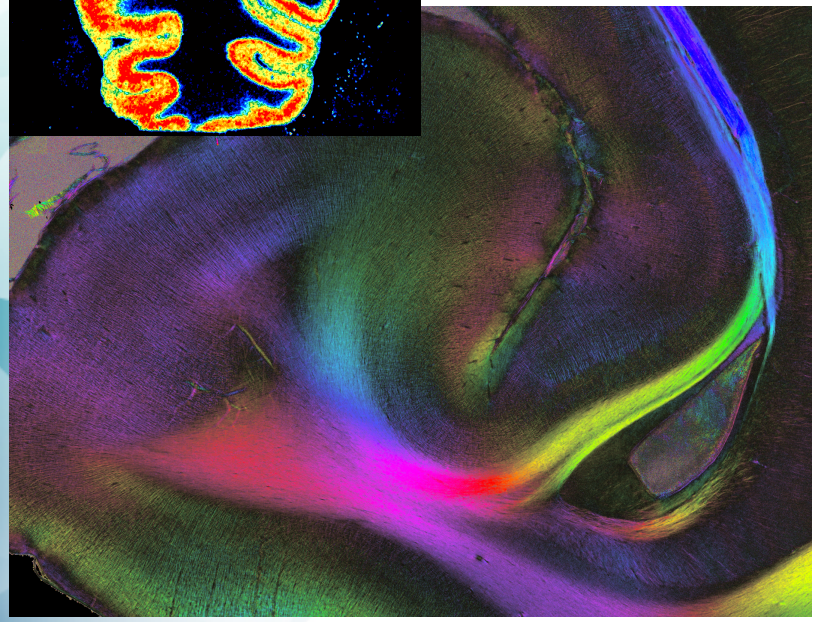


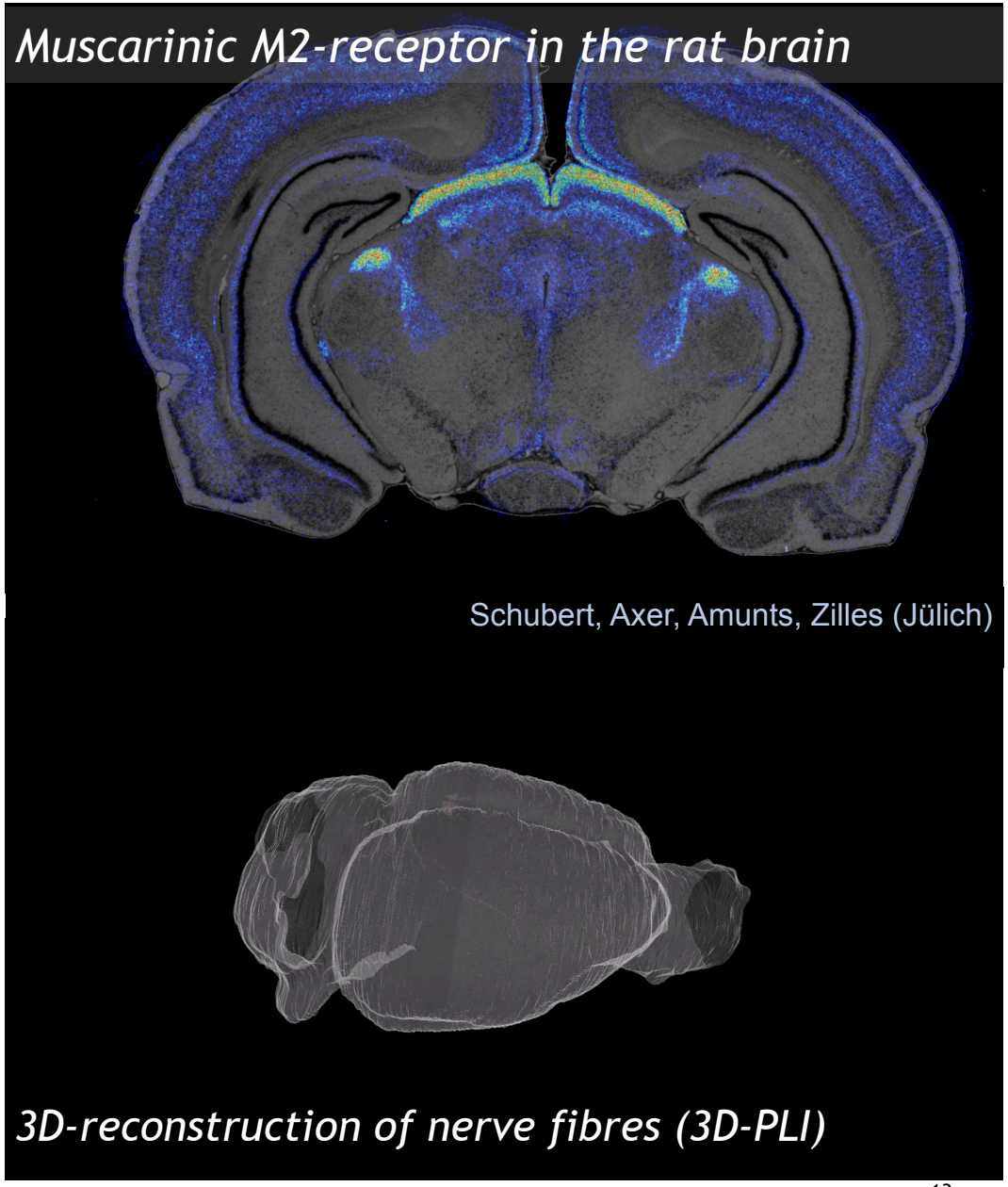
Human and rodent brain atlases



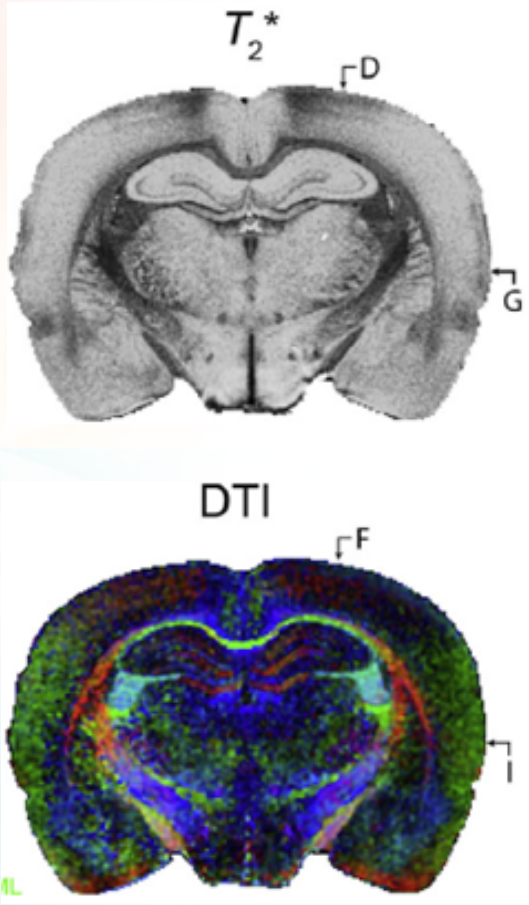
Muscarinic M2-receptor in the human brain

3D-reconstruction of nerve fibres (3D-PLI)





Rodent brain atlases

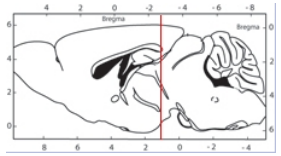


Papp et al., Neuroimage, 2014
Kjonigsen et al., Neuroimage, 2015
Boccara et al., Hippocampus, 2015, in press

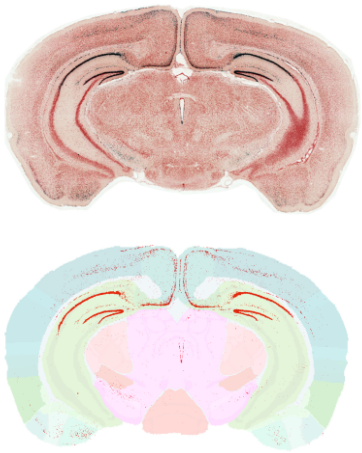
HBP Waxholm Space rat brain atlas

Waxholm Space rat brain atlas v2.0: updated and detailed delineations of the hippocampus

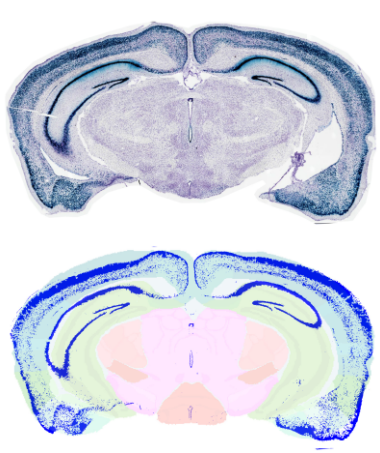
Data integration in atlas space



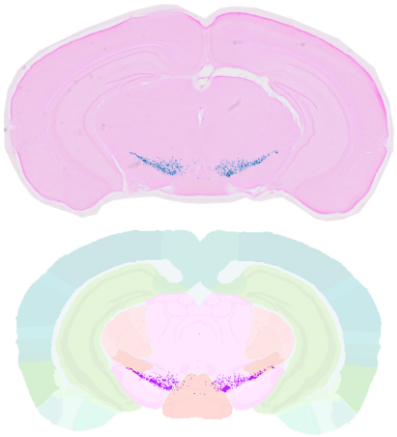
PrP-tTA



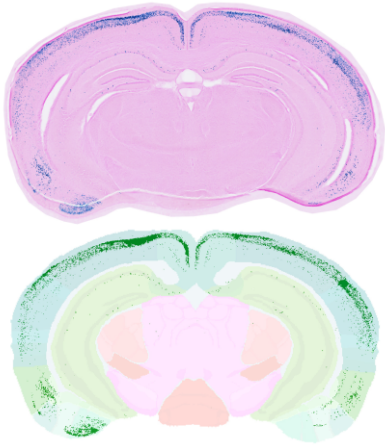
CamKII-tTA



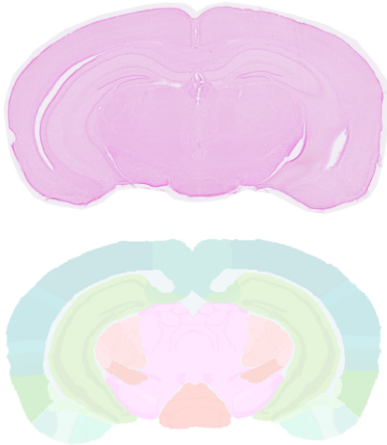
PitX3-tTA



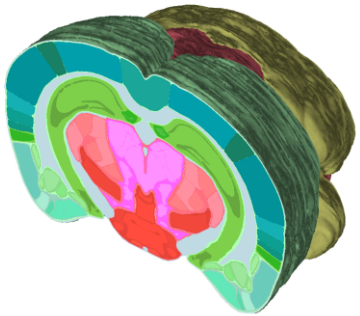
NOP(EC)-tTA

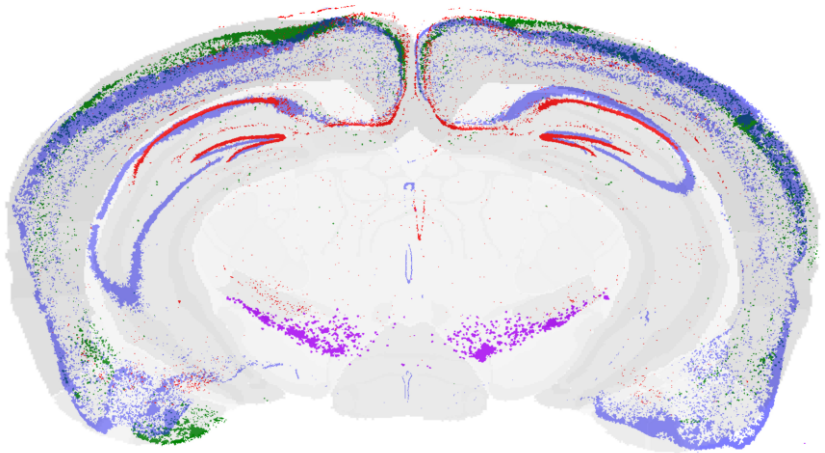


PcP2-tTA



3D Allen Mouse Brain Atlas





PrP

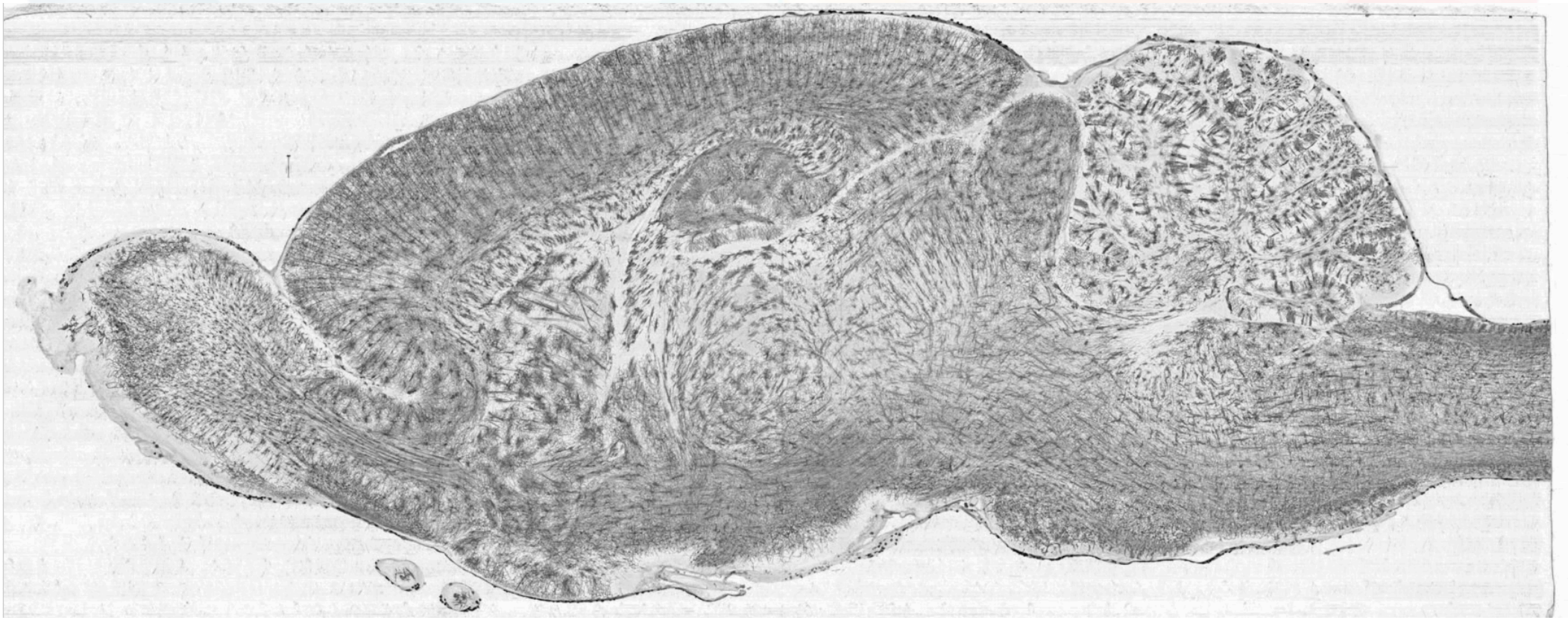
CamKII

PitX3

NOP

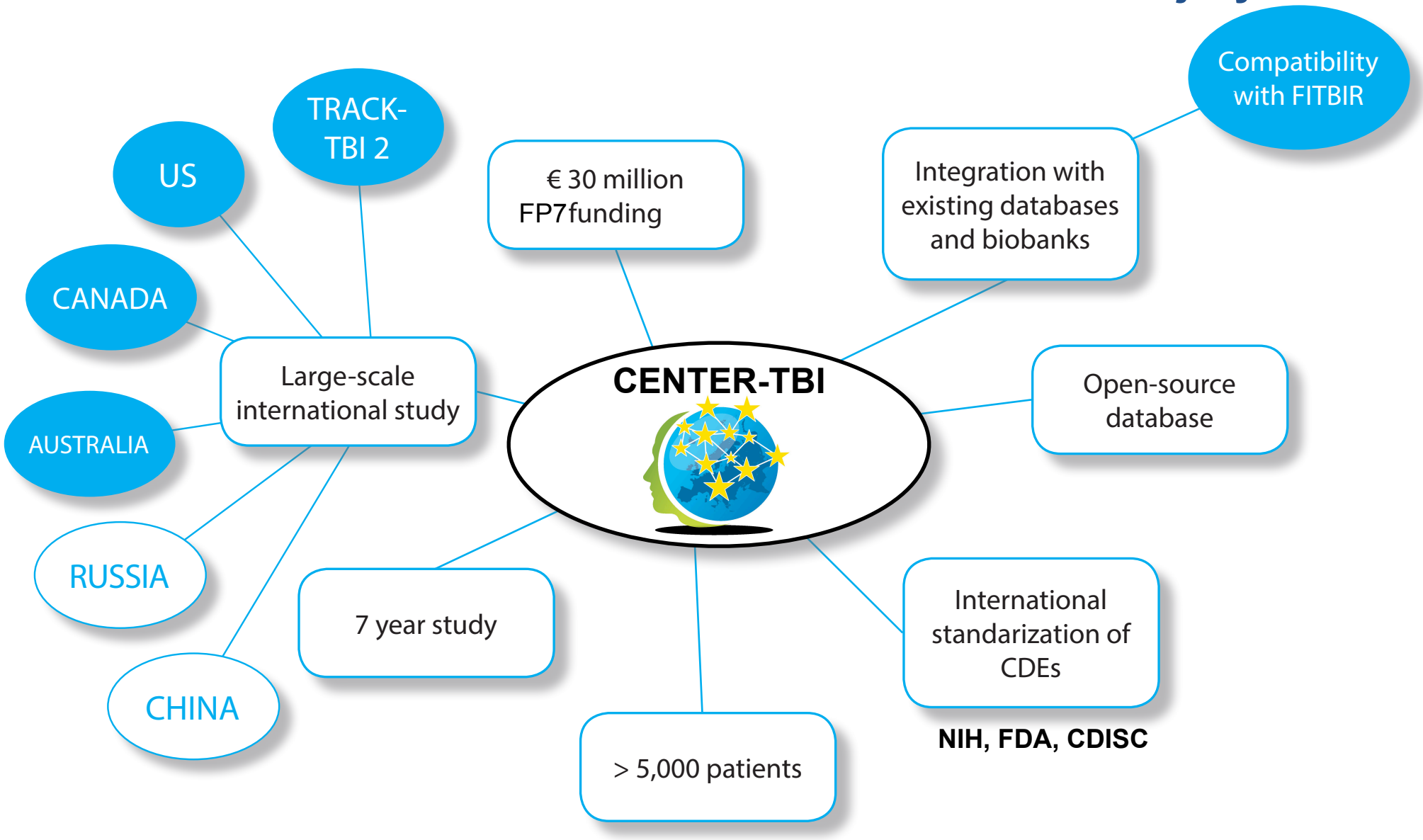
PcP2

Transgene expression in five tTA driver mouse lines integrated in Allen mouse brain atlas space



Funding open, collaborative research

International collaboration on traumatic brain injury



Andrew Maas, David Menon

Open Data Challenges

- Not black and white:
  - Not just OPEN or CLOSED
  - Need granularity to be explicit about what is open, when and for what purpose
  - Then gradually develop the culture of loosening these restrictions

Open Data Challenges

- Willing to share data but:
  - Expensive to produce (intellectual capital - experimental design, acquisition cost, time)
  - Many possible uses (multiple research questions) for large data sets
  - Currently, reward currencies are intellectual advances, publications and citations
  - No clear reward or motivation for providing data completely free of any constraint

Data use agreements

- Data use agreement
  - Sets conditions for data use
    - Don't abuse privileges (e.g. deidentify human data)
    - Don't redistribute, go to approved repository for registered access (maintain data integrity, tracking accesses)
    - Agree to acceptable use policies (e.g. investigate non-embargoed questions only)
  - Embargo duration
  - Share and share-alike - if data combined with others that should be shared, result should be shared
  - Commercial use?
  - Stakeholders
  - Points to research registry for dataset
    - Owners can reserve (embargo) data access for specific research questions for limited time period
    - Others may access for non-embargoed use

Open Data - tiered access

- Tier 0 - Unrestricted
  - All metadata and/or data freely available (includes contributor, specimen details, methods/ protocols, data type, access URL)
  - **Reward: Potential citation, collaboration**
- Tier 1 - Restricted use
  - Data available for restricted use, developing analysis algorithms
  - **Reward: Data citation**
- Tier 2 - Restricted Use
  - Data available for restricted use, nonconflicting research questions
  - **Reward: Co-authorship**
- Tier 3 - Restricted use
  - Full data available for collaborative investigation, joint research questions
  - **Reward: Collaboration, co-authorship**