Open Neuroscience Research Data through EBRAINS

Delivering on the FAIR, TRUST and CARE principles
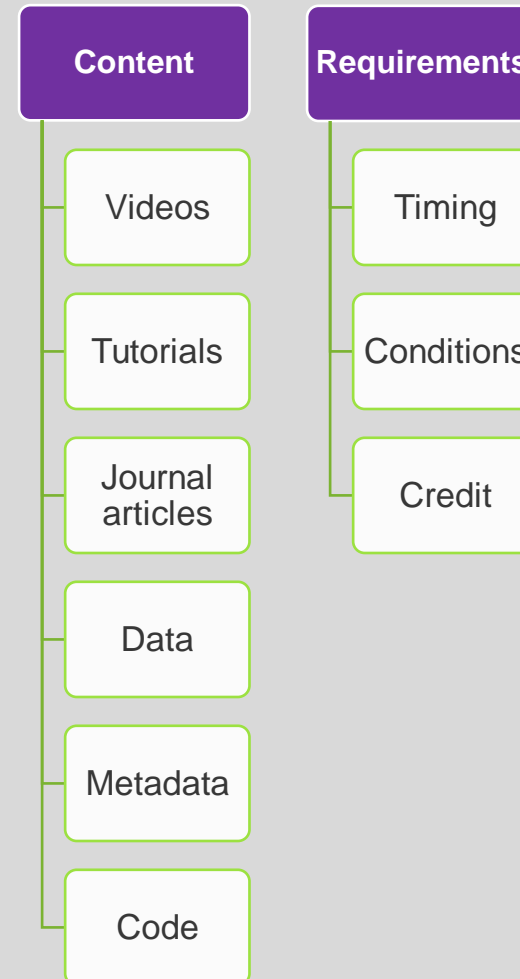
Jan G. Bjaalie, MD PhD

EBRAINS Data and Knowledge leader
EBRAINS Management Board Member
Former Infrastructure Director Human Brain Project 2019 -2023
Professor, Institute of Basic Medical Sciences, University of Oslo
Dean of Research and Innovation, Faculty of Medicine, University of Oslo

Open Science

What is Open Science?

- *"Open science is a research accelerator"* (Woelfle, M. et al, Nature Chemistry 2011)

- *"…. scientific knowledge of all kinds should be openly shared as early as is practical ….."* (openscienceasap.org)

Content
- Videos
- Tutorials
- Journal articles
- Data
- Metadata
- Code

Requirements
- Timing
- Conditions
- Credit

Open Science, EU initiatives

Open Science

An approach to the scientific process that focuses on spreading knowledge as soon as it is available using digital and collaborative technology. Expert groups, publications, news and events.

PAGE CONTENTS

The EU's open science policy

Ambitions:

- Open data
- EOSC
- New metrics
- Rewards
- Open access
- Research integrity & reproducibility
- Education and skills
- Citizen science

Aims for open science policy under Horizon Europe
- ensure that beneficiaries retain the intellectual property rights
- require research data to be FAIR and open by default (with exceptions notably for commercial purposes)
- promote the adoption of open science practices
- engage and involve citizens, civil society organisations and end-users
- ….

Drivers of science: Ideas - Tools - Data

- **Ideas – or "systems of ideas" (paradigms)**
  - Change: gradual or sudden – new ideas replacing old ones (paradigm shift)
  - Top down: "Ideas based on equations"
  - Bottum up: "Ideas based on experiments"

- **Tools**
  - New instruments producing digital output / data with vastly improved precision, or enabling new type of measurements
  - Software for digital data processing
  - Computing capacity

- **Data**
  - The outputs from the instruments
  - The outputs from the data processing

The importance of access to data: data integration

- Combining data from different sources into a single, unified view

- The most complex topics in neuroscience can only be studied by combining information from different levels of investigations (many methods, different granularity / spatial scales)

- Answering specific research questions / performing hypothesis driven research (topic or question determined before the research is performed)

- Data driven research (topic or question determined based on what is found in the data collections)

- Integration begins with access to interpretable and actionable data and ultimately enable data analysis aimed at acquiring knowledge

EBRAINS 2.0

EBRAINS

What is FAIR, TRUST, and CARE?

The principles of data sharing

Data repositories
T ransparency
R esponsibility
U ser focus
S ustainability
T echnology

Lin et al., Sci Data 2020

Data and metadata
F indable
A ccesible
I nteroperable
R eusable

Wilkinson et al., Sci Data 2016

Data governance
C ollective benefit
A uthority to control
R esponsibility
E thics

Carroll et al., Data Science Journal 2020

What are the contributions of EBRAINS to data sharing according to the FAIR, TRUST, and CARE principles?

- **Enable collaborative neuroscience through FAIR data.** Provide professionally maintained and user-friendly mechanisms for making data FAIR as well as reusable with **easily accessible tools and workflows**

- **Support multimodal data integration.** Systematic use of semantic and spatial (atlas) metadata

- **Foster data and metadata harmonisation.** Curation and harmonisation of data and metadata acquisition for a range of foundational data types; data integrity, comparability and reliability across studies

- **Accelerate the adoption of data sharing services** by strategically concentrating efforts on **prioritized data and workflow categories**

- **Promote a responsible open science culture** - transparency, reproducibility, and replicability - while ensuring ethical and lawful sharing through a **data governance structure** based upon People, Processes and Technology (PPT)

EBRAINS 2.0

EBRAINS

EBRAINS is providing a strong data sharing platform with 3 main user facing services

EBRAINS 2.0

Share data, models and software

Request curation ⬈

Deliver your data and metadata

Metadata is curated

Data published on EBRAINS

KG Automation
Automate your processes

KG Editor
Manage & publish your metadata

Find neuroscience data, models and tools

Explore data in the Knowledge Graph ⬈

KG Search
Search and use data

KG Query Builder
Collect metadata with queries

KG Statistics
Visualize the graph

KG Core API & SDKs
Programmatic access

Live Papers

Learn more ⬈

Structured and interactive supplementary documents to complement journal publications

An efficient analytical reduction of detailed nonlinear neuron models

Amsalem et al. (2020) [Live Paper]   DOI:10.25493/1D2M-CV0

The microcircuits of striatum in silico

Hjorth et al. (2020) [Live Paper]   DOI:10.25493/RR3S-54

EBRAINS

EBRAINS 2.0

EBRAINS
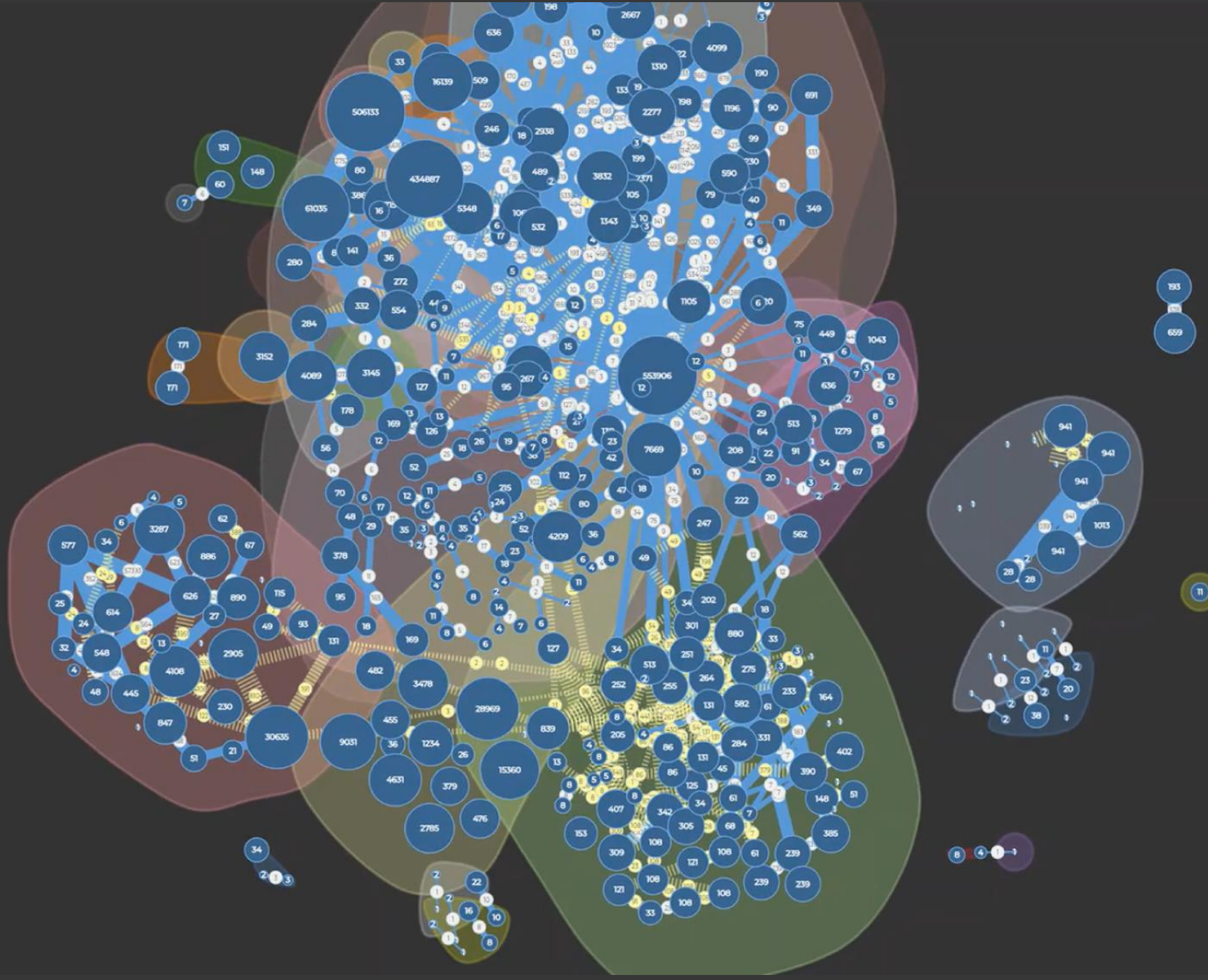
The foundational triangle

Knowledge Graph

Integrative development of metadata framework and graph database management system

Joint development of tools and workflows for data and metadata submission and quality control
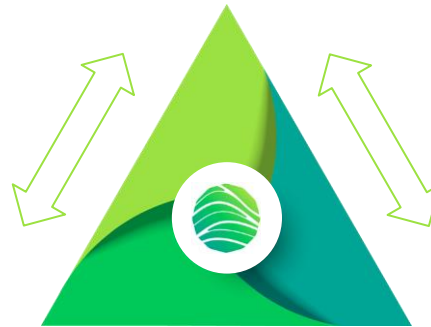
Metadata framework

Curation Service

Collaborative development of metadata framework and ontology-driven terminologies
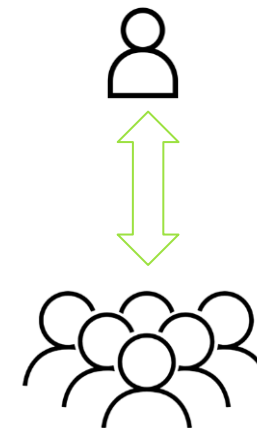
The 4ᵗʰ component: EBRAINS Compliance Management, Data Governance and Data Protection
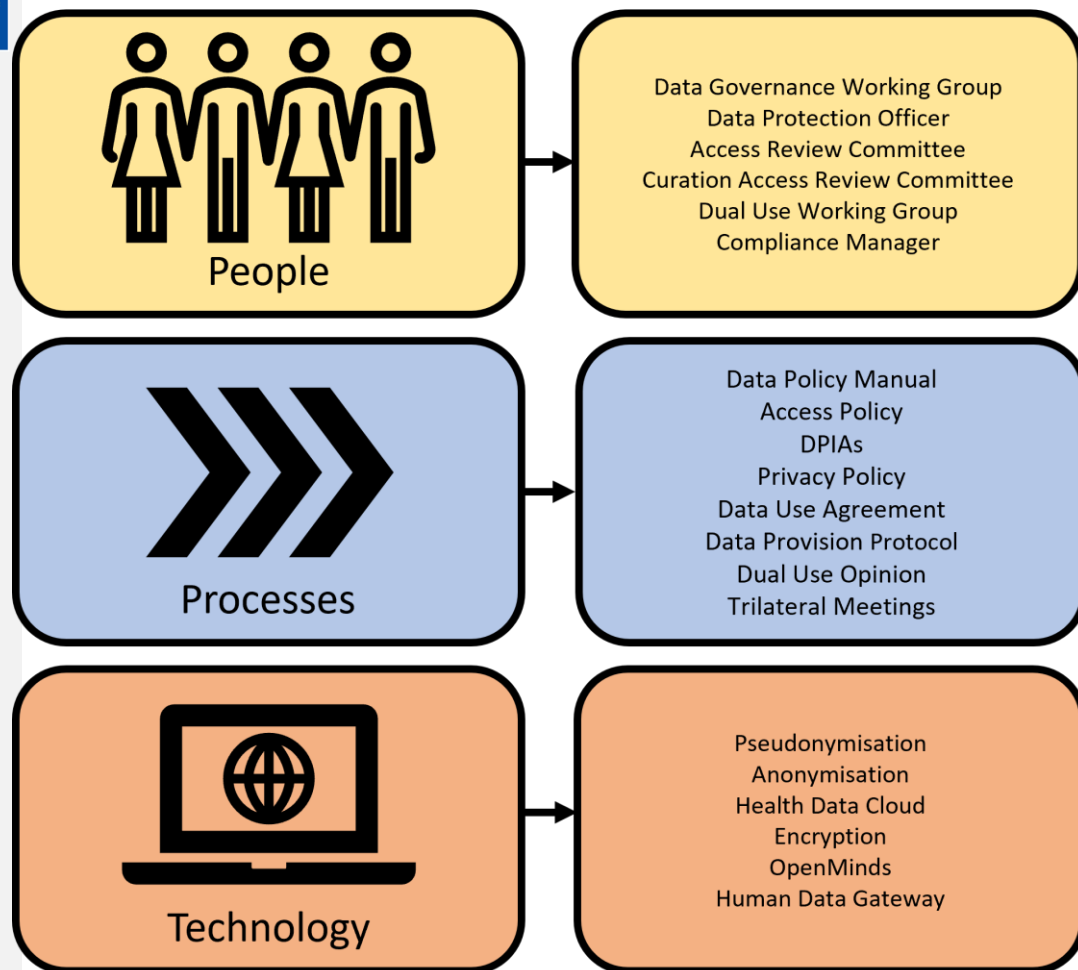
EBRAINS has developed and implemented a data governance structure based upon **People**, **Processes** and **Technology** (PPT).

- ***People*** – stakeholders involved in all EBRAINS data processing pipelines.

- ***Processes*** - policies, procedures and practical processes set up to ensure responsible data governance.

- ***Technology*** - technologies developed and applied in EBRAINS workflows.

Built upon these three pillars, EBRAINS complies with its ethical, legal and professional responsibilities, including compliance with GDPR.

The SHARE data service

Why metadata?

➔ Give context to dataset, models and software

➔ Transform unstructured information into structured information

➔ Create relations between datasets

➔ Increase the findability and interoperability

➔ ... when standardized ...

openMINDS: open Metadata **I**nitiative for **N**euroscience **D**ata **S**tructures

openminds@ebrains.eu

https://wiki.ebrains.eu/bin/view/Collabs/openminds
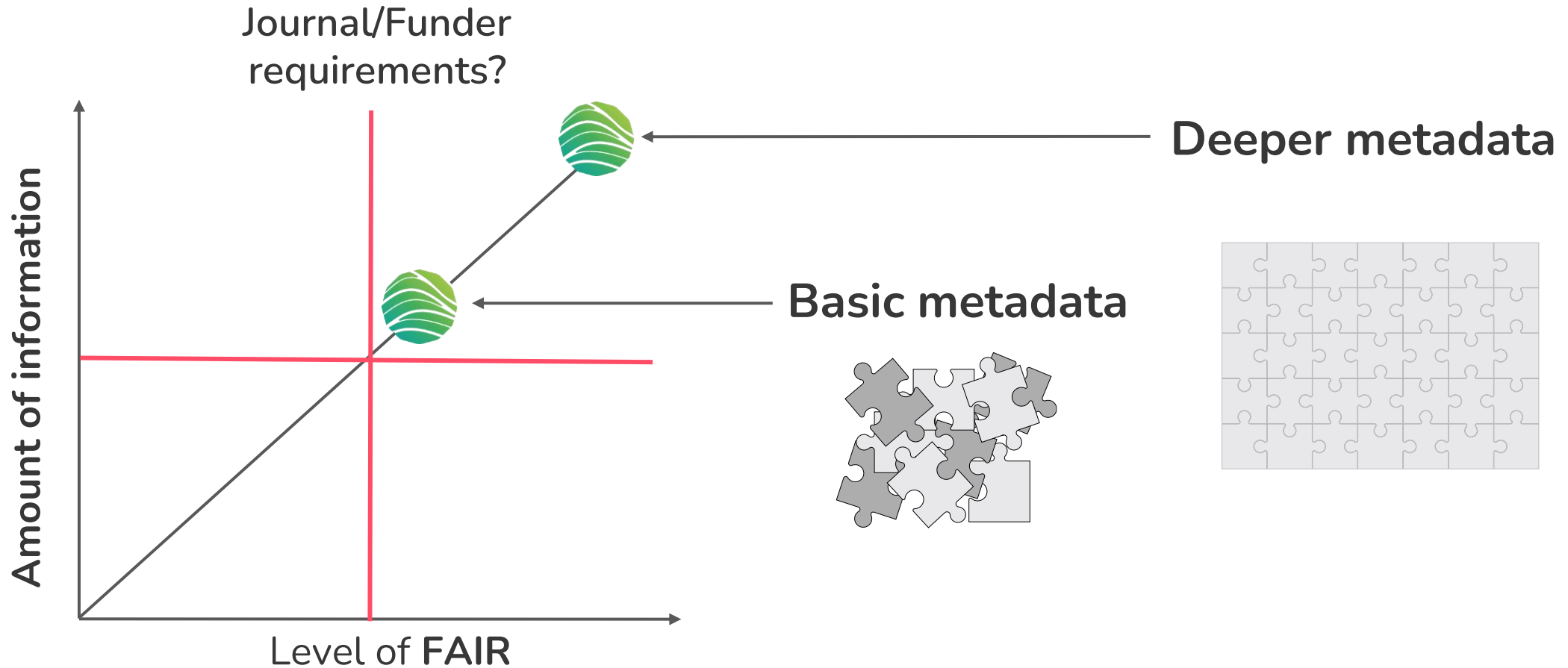
GitHub https://github.com/HumanBrainProject/openMINDS

How much metadata should be provided? (depth of curation)

Levels of sharing: non-sensitive data

Data freely accessible	Data under embargo	Data temporarily hidden
• DOI for your data • Data is discoverable • Data is accessible for download and reuse	• URL for your data • Metadata is discoverable • Data is inaccessible until publication	• Private URL for reviewers • Dataset is **not** discoverable • Data is inaccessible until publication

Publish a dataset in parallel to submitting a peer-reviewed journal publication

Levels of sharing: sensitive data

**Anonymous metadata findable:
<u>Controlled</u> access to the sensitive
data**

- User will have to be identified through an EBRAINS account
- Terms of use to be accepted
- User gets temporary access to data in EBRAINS storage for download
- All user activity is tracked

**Anonymous metadata findable:
<u>Restricted</u> access to the
sensitive data**

- User will have to be identified through an EBRAINS account
- User will be provided with instructions on how to apply for access
- Data not hosted by EBRAINS but by a trusted repository interacting with EBRAINS
- Metadata is discoverable

Publish anonymous metadata – make data discoverable

What motivates researchers to share data through EBRAINS?

- My funding agency requests that I share my research data. Can I get more out of this?

- Get more citations / credit (citable DOI)

- Communiate the conditions for use (license)

📥 Download Dataset   66 Cite dataset   🗐 Data-descriptor

DOI : ❶ 10.25493/D2CW-JR

License : Creative Commons Attribution-NonCommercial-ShareAlike

Project : Macaque Brain Atlas

Custodians : ❶ Palomero-Gallagher, Nicola

What motivates researchers to share data through EBRAINS?

- I would like to make my data available together with a journal publication. Can EBRAINS Data and Knowledge services deliver a solution?

- Track for publishing research data alongside a journal article

- Special track for journals publishing Data descriptors, including Nature Scientific Data

SHARE DATA - GUIDELINES

Publishing a Dataset that Accompanies a Peer-Reviewed Journal Publication

What motivates researchers to share data through EBRAINS?

- The data from my research is of a foundational nature. I want to easily find back to my own data and also make sure the data can be used in future research. Why should I use EBRAINS Data and Knowledge services?

- I am looking for research data to complement my own research. Where should I look?

- Curation of metadata to increase discoverability and opportunities for re-use

- Specialized service for neuroscience, metadata standard for neuroscience data

- Advanced discoverability: Search interface and programmatic access

EBRAINS 2.0

EBRAINS

What motivates researchers to share data through EBRAINS?

The FIND data service

maps

SEARCH

CATEGORIES

Project	27
Dataset	520
Model	4
Software	10
Web service	1
Contributor	58

Viewing **1-20** of **520** results.

Did you mean main, mann or map?

Top trending

Julich-Brain Atlas, cytoarchitectonic maps

Released : 2023-07-13

Accessibility : free access

Custodians : Amunts, K.

This dataset contains the Julich-Brain Atlas, Cytoarchitectonic *maps* in different coordinate spaces....

anatomical parcellation technique   brain   brain mapping   cytoarchitectonic mapping   magnetic resonance imaging

maximum probability projection   silver staining

FILTERS

Reset

ACCESSIBILITY

☐ free access	489
☐ under embargo	25
☐ controlled access	5
☐ restricted access	1

New

Contrast maps obtained from Individual Brain Charting

Badges

DATASET

Anterogradely labeled axonal projections from the insular cortex in rat (v1)

Versioning

Mathiasen, M. L.; Hansen, L.; Monterotti, B.; Laja, A.; Reiten, I.; Leergaard, T. B.; Witter, M. P.

Overview

Data descriptor

How to cite

Get data

Publications

Specimens

Related resources

How to use

DOI : ℹ 10.25493/WK4W-ZCQ

Released : ℹ 2020-04-02

Accessibility : free access

License : Creative Commons Attribution 4.0 International

Ethics assessment : EU-compliant

Project :
The efferent connections of the orbitofrontal, posterior parietal, and insular cortex of the rat brain

Custodians : ℹ Witter, M. P.

High-resolution fluorescence microscopy images of serial coronal brain sections showing anterogradely labeled axons originating from different parts of the insular cortex in adult female Sprague Dawley rats. The dataset includes image series from 8 brains in which one or both of the anterograde tracers biotinylated dextran amine (BDA) and *Phaseolus vulgaris Leucuagglutinin* (PHA-L) were injected into different subregions of the insular cortex, always in the right hemisphere. Discrete injections were placed in the agranular (AI), dysgranular (DI) or granular (GI) subdivisions of the insular cortex. Brain sections were 50 μm thick and sampled every 300 μm. The data provides an overview of the efferent connections of the rat insular cortex.

Version specification : This is the first version of this dataset.

View data in LocaliZoom :
* tissue sample collection (subject 12949_BDA)
* tissue sample collection (subject 13018_PHAL)

Service links
(e.g. atlas viewer)
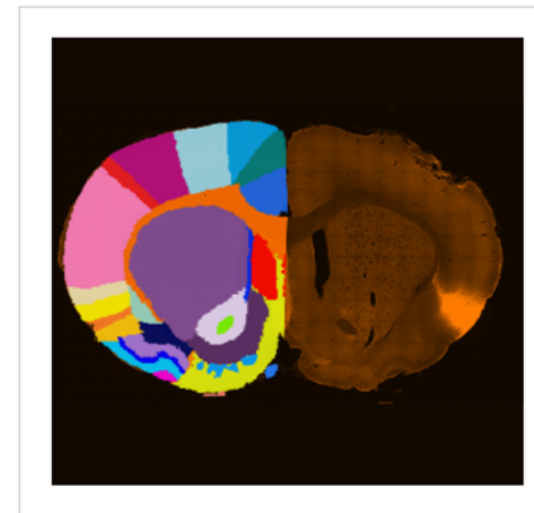
Preview image

Studied brain region :
insular region (Waxholm Space Rat Brain Atlas)

Anat. location of tissue samples : ℹ
brain (Waxholm Space Rat Brain Atlas)

Anatomical
localization of data

Preparation :
* ex vivo
* in vivo

Experimental approach :
* anatomy
* microscopy
* histology
* neural connectivity

Controlled
ontology driven
annotations

DATASET

Anterogradely labeled axonal projections from the insular cortex in rat (v1)

Mathiasen, M. L.; Hansen, L.; Monterotti, B.; Laja, A.; Reiten, I.; Leergaard, T. B.; Witter, M. P.

Overview

Data descriptor

How to cite

Get data

Publications

Specimens

Related resources

How to use

Programmatic access to metadata :
To make programmatic use of the (meta-)data of EBRAINS, you have different options to interact with the EBRAINS Knowledge Graph (KG):

KG Query Builder
With the KG Query Builder, you can design your own query to retrieve metadata for this instance and those of the same type conveniently via UI without the requirement of learning a graph query language (see the tutorial). You can also save the query and use it with the REST-API and the KG Core SDKs (see below).

Build a query for this instance

KG REST-API
You can use the KG REST-API to access and/or manipulate metadata on the EBRAINS KG as well as to run queries saved previously in the KG Query Builder.

KG Core SDKs
The KG Core SDKs provide convenient ways to authenticate and make use of the functionality of the KG REST-API with your favorite programming language (currently available for Python and JavaScript/TypeScript).

EBRAINS

Programmatic access to metadata in the Knowledge Graph

Can I get more fMRI data of stroke patients in BIDS?

Who else does research on this particular brain region?

Do they have these type of measurements in other species?

What is the typical age range of monkeys performing a specific type of behavioral training?

1. Research question

2. KG Query Builder *

3. Query from script

KG Core API & SDKs *

4. Review results

5. Download & reuse data

* New in M22-M42

- Build complex queries tailored to your research questions
- Receive easily reusable results in standard format
- Set up fast & reproducible computational workflows
- Automate data analysis using software specific to your analytical needs
- Perform large scale meta-analyses over hundreds of datasets
- Gather training data for your machine learning model
- Find EBRAINS software to visualize your results based on content type

EBRAINS 2.0

EBRAINS

A data journey in EBRAINS

DATASET

Excitability profile of CA1 pyramidal neurons in APPPS1 Alzheimer disease mice and control littermates (v1)

Salgueiro-Pereira, A. R.; Marie, H.

DOI: 10.25493/YJFW-HPY
License: CC BY 4.0
Study target: Alzheimer's disease
Species: Mus musculus (13 subjects)
Technique:
- whole cell patch clamp
- current clamp

WEB SERVICE

NeuroFeatureExtract (v2.1.1)

Bologna, L. L.; Smiriglia, R.; Curreri, D.; Migliore, M.

Accessibility: free access

SOFTWARE

NEURON (7.7.2)

Moore, J. W.; Hines, M.; Carnevale, T.

License: The 3-Clause BSD License

MODEL

Age-dependent excitability of CA1 pyramidal neurons in APPPS1 Alzheimer's model (v1)

Migliore, R.; Vitale, P.

Published in: https://modeldb.science/266848
License: CC BY-NC 4.0
Model scope: single cell
Abstraction level: spiking neurons: biophysical
Study target:
- Mus musculus
- hippocampus CA1 pyramidal neuron
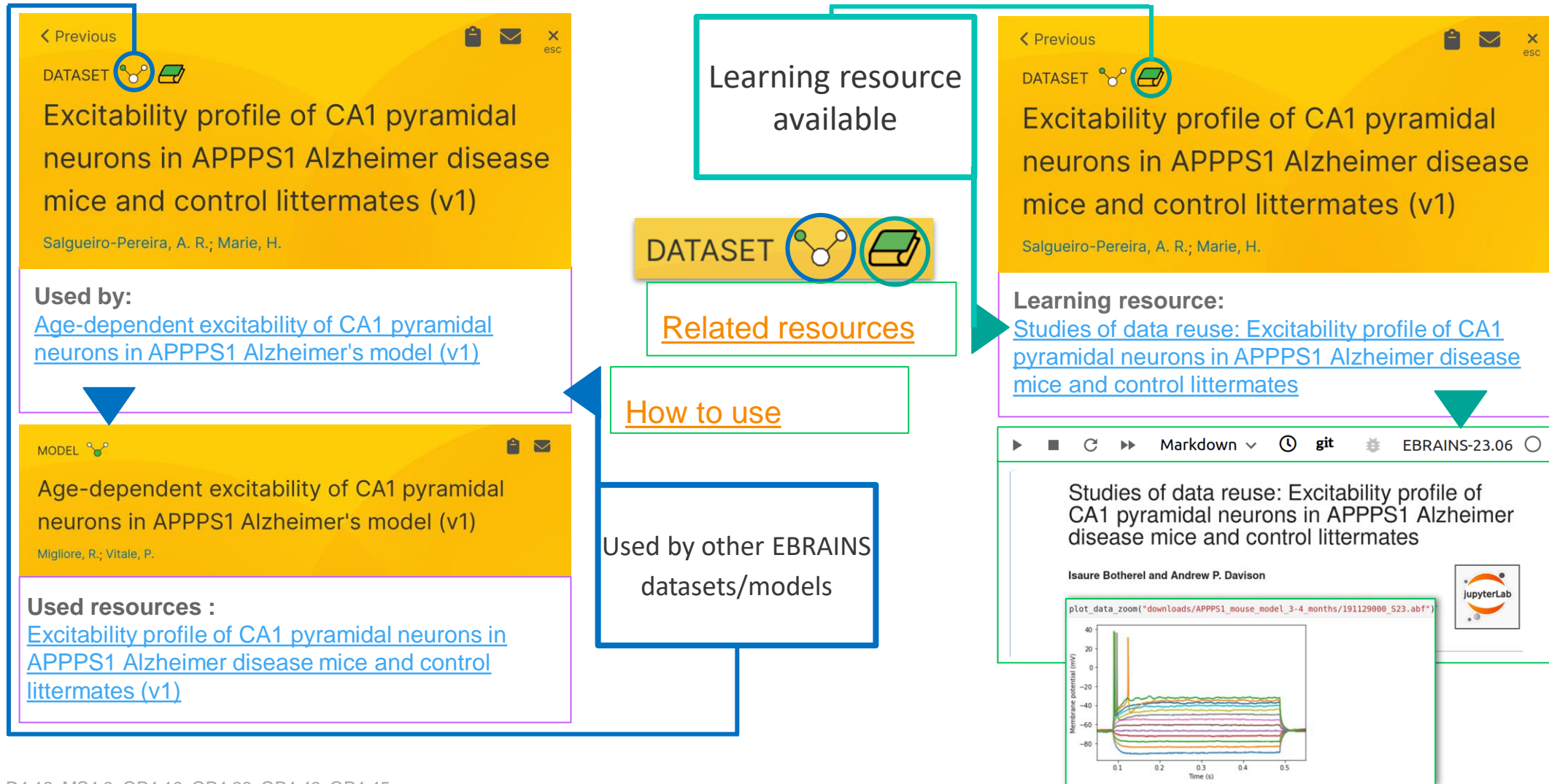- Alzheimer's disease

frontiers

Analysis of Age-Dependent Alterations in Excitability Properties of CA1 Pyramidal Neurons in an APPPS1 Model of Alzheimer's Disease

Vitale et al. (2021) [Article]   **DOI:** 10.3389/fnagi.2021.668948

STATIC PUBLICATION

EBRAINS

Analysis of Age-Dependent Alterations in Excitability Properties of CA1 Pyramidal Neurons in an APPPS1 Model of Alzheimer's Disease

Vitale et al. (2021) [Live Paper]   **DOI:** 10.25493/D4PT-QNB

INTERACTIVE PUBLICATION

Slide courtesy of L. Zehl

EBRAINS 2.0

EBRAINS

Co-funded by the European Union

KG-enabled integration across EBRAINS services

DATASET

Excitability profile of CA1 pyramidal neurons in APPPS1 Alzheimer disease mice and control littermates (v1)

Salgueiro-Pereira, A. R.; Marie, H.

Used by:
Age-dependent excitability of CA1 pyramidal neurons in APPPS1 Alzheimer's model (v1)

MODEL

Age-dependent excitability of CA1 pyramidal neurons in APPPS1 Alzheimer's model (v1)

Migliore, R.; Vitale, P.

Used resources :
Excitability profile of CA1 pyramidal neurons in APPPS1 Alzheimer disease mice and control littermates (v1)

Learning resource available

DATASET

Related resources

How to use

Used by other EBRAINS datasets/models

DATASET

Excitability profile of CA1 pyramidal neurons in APPPS1 Alzheimer disease mice and control littermates (v1)

Salgueiro-Pereira, A. R.; Marie, H.

Learning resource:
Studies of data reuse: Excitability profile of CA1 pyramidal neurons in APPPS1 Alzheimer disease mice and control littermates

Markdown   git   EBRAINS-23.06

Studies of data reuse: Excitability profile of CA1 pyramidal neurons in APPPS1 Alzheimer disease mice and control littermates

Isaure Botherel and Andrew P. Davison

jupyterLab

plot_data_zoom("downloads/APPPS1_mouse_model_3-4_months/191129000_S23.abf")

Slide courtesy of L. Zehl

EBRAINS 2.0

EBRAINS

Semantic data integration within and beyond EBRAINS through tags with ontology-driven controlled terminology

Dataset cards are annotated with controlled terms

Controlled terms linked to external resources ^{openMINDS service}

Knowledge Graph

DATASET

The Swedish National Facility for Magnetoencephalography Parkinson's Disease Dataset (v1.0)

DOI: 10.25493/NMD2-2FW

Study targets: Parkinson's disease **1**

1

DISEASE

Parkinson's disease

openMINDS

Interlex **2**   Knowledge Space **3**

EBRAINS metadata standard developed in partnership with representatives from EBRAINS, NIF and INCF

openMINDS

Controlled terms are linked to detailed ontologies

InterLeX   **FDI Lab** FAIR Data Informatics **2**

Parkinson disease

A progressive degenerative disorder of …

Synonym(s)	Ontology identifiers
paralysis agitans; Parkinson's disease	MONDO:0005180 ILX:0764151 DOID:14330

Ontology terms create linkages across databases

Knowledge Space   incf **3**

EBRAINS
16 results

OpenNEURO
29 results

SPARC **SPARC**
3 results

CONP Portal
12 results

EBRAINS

T4.1, T4.2, T4.3, T4.4 , D4.12, MS4.3, OP4.19, OP4.28, OP4.29, OP4.43, OP4.45

Slide courtesy of L. Zehl

EBRAINS Data and Knowledge solutions developed by teams across Europe

UiO
Norway

FZ Jülich
Germany

CNRS
France

KI
Sweden

CSCS
Switzerland

DMU
UK

EBRAINS AISBL
Belgium

The SHARE data service

The FIND data service

Multiple services for USING data and analysing new combinations of data

Example: Atlas services, in close interaction with the Knowledge Graph

Overview: EBRAINS atlas services
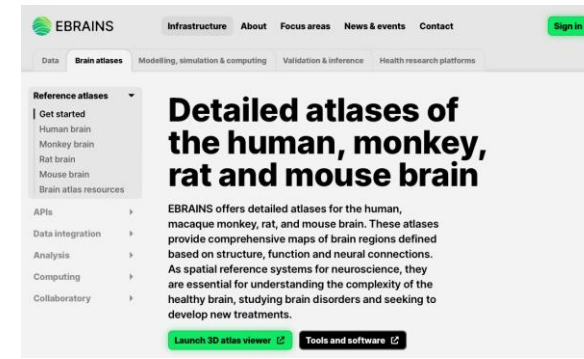
Atlases of the human, macaque, rat & mouse brain at unmatched detail

A software toolsuite that makes atlases accessible and interoperable

A set of tools for spatial anchoring to atlases that fascilitate extensibility

Atlas-guided analysis tools that demonstrate usefulness

Highly detailed reference atlases

Human multilevel atlas
- Volumetric, surface & microscopic templates
- >200 cytoarchitectonic structures
- Deep & superficial fibre bundles
- Functional maps
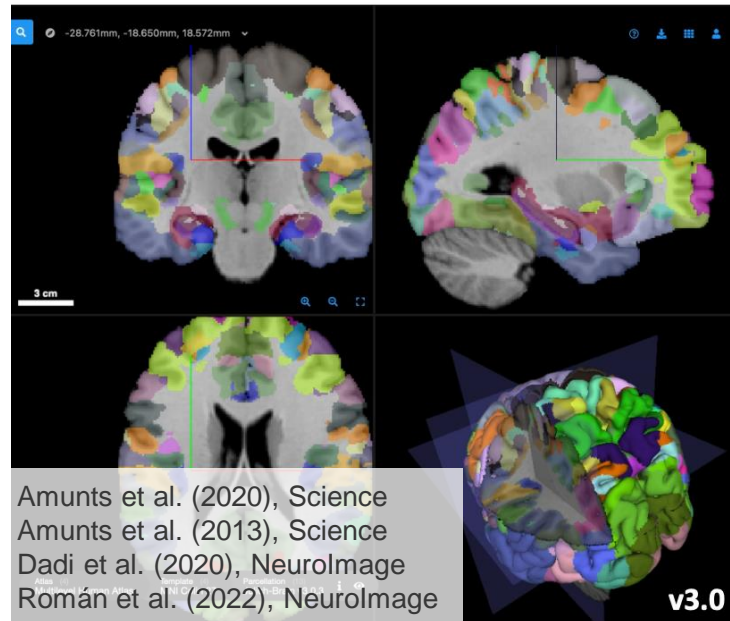
Waxholm Space rat atlas

- 222 regions delineated in V4
- Commercial uptake by Mbf Bioscience and Gubra A/S
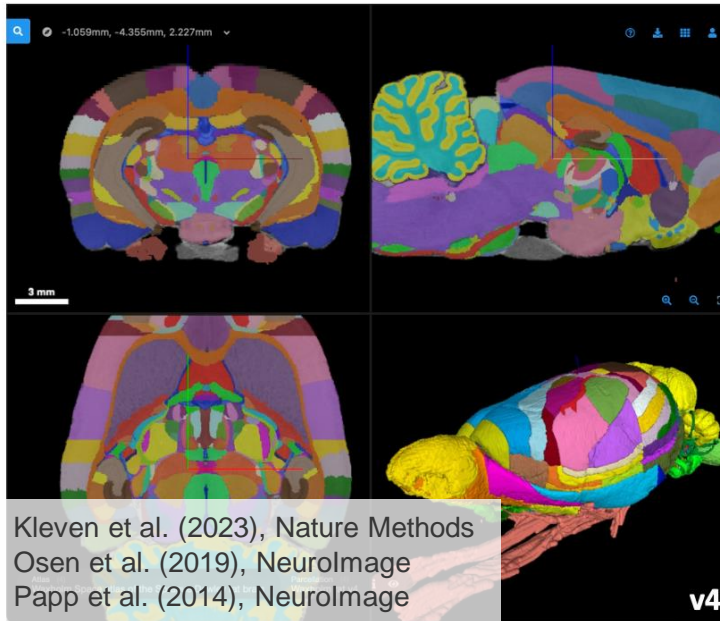
Macaque atlas

- High-resolution template
- Cytoarchitectonic maps

Amunts et al. (2020), Science
Amunts et al. (2013), Science
Dadi et al. (2020), NeuroImage
Román et al. (2022), NeuroImage
v3.0

Kleven et al. (2023), Nature Methods
Osen et al. (2019), NeuroImage
Papp et al. (2014), NeuroImage
v4

Balan et al. (2023), preprint

atlases.ebrains.eu/viewer/go/human

atlases.ebrains.eu/viewer/go/rat
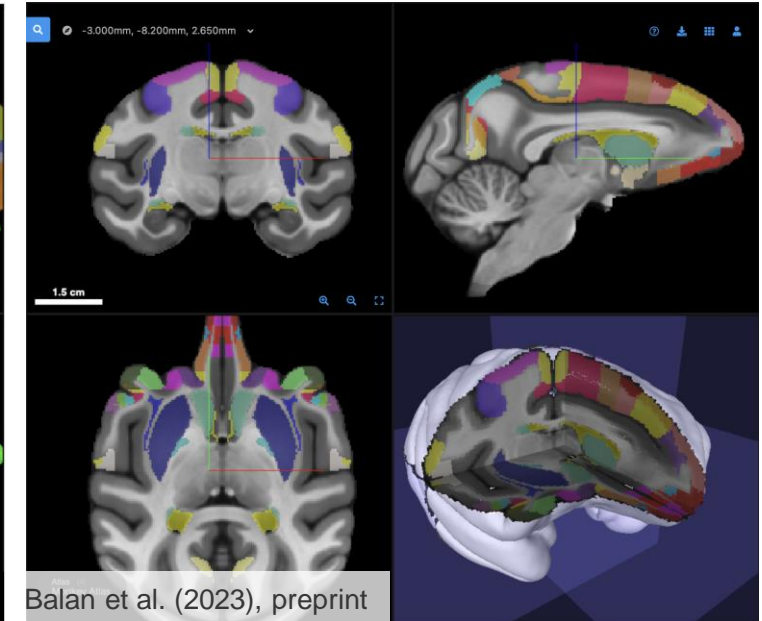
atlases.ebrains.eu/viewer/go/monkey

Slide courtesy of T. Dickscheid

Detailed atlases of the human, monkey, rat and mouse brain

EBRAINS offers detailed atlases for the human, macaque monkey, rat, and mouse brain. These atlases provide comprehensive maps of brain regions defined based on structure, function and neural connections. As spatial reference systems for neuroscience, they are essential for understanding the complexity of the healthy brain, studying brain disorders and seeking to develop new treatments.

A unique concept of connecting scales and modalities

The atlas framework is well accessible

https://atlases.ebrains.eu/viewer/go/human
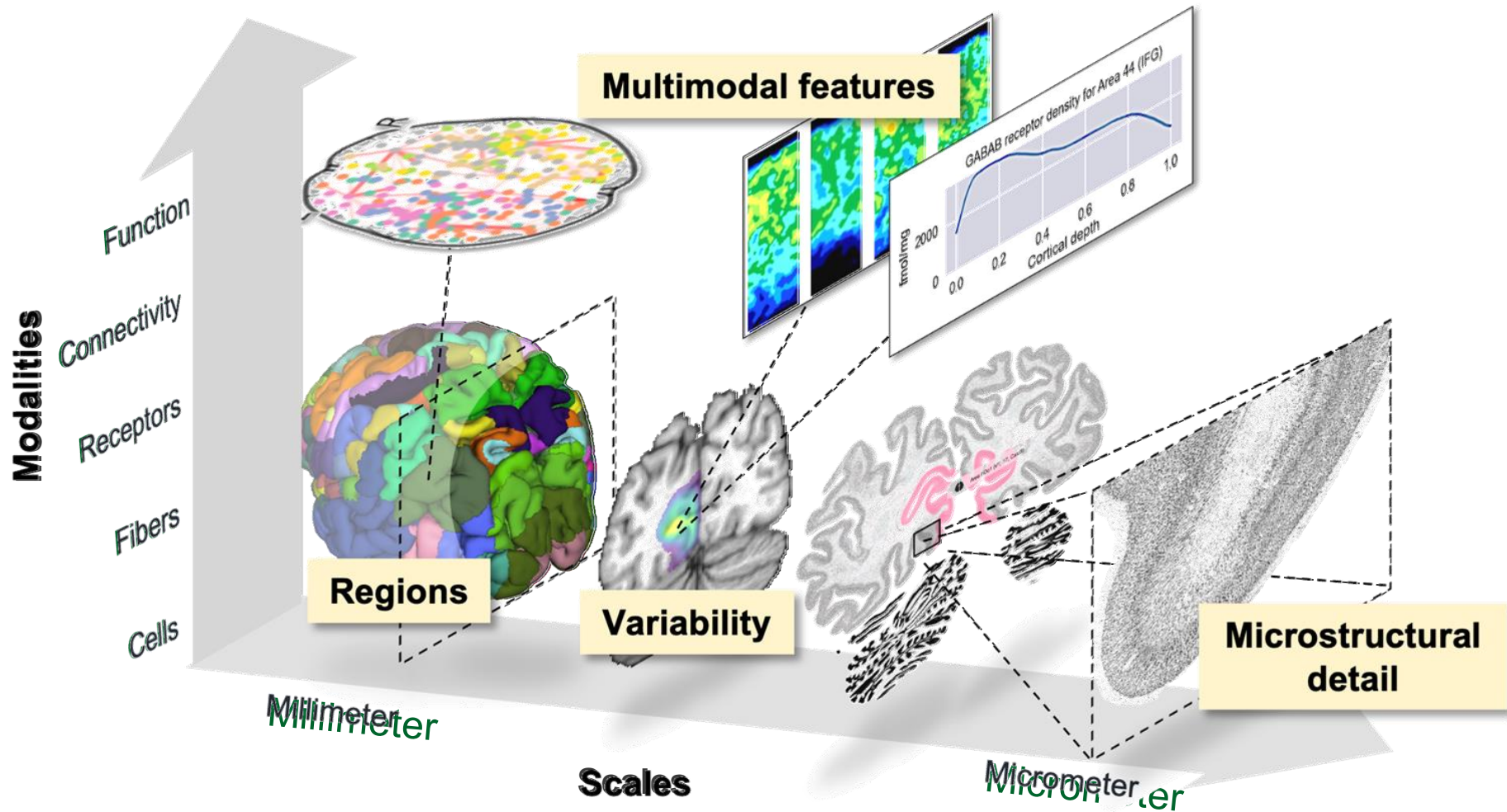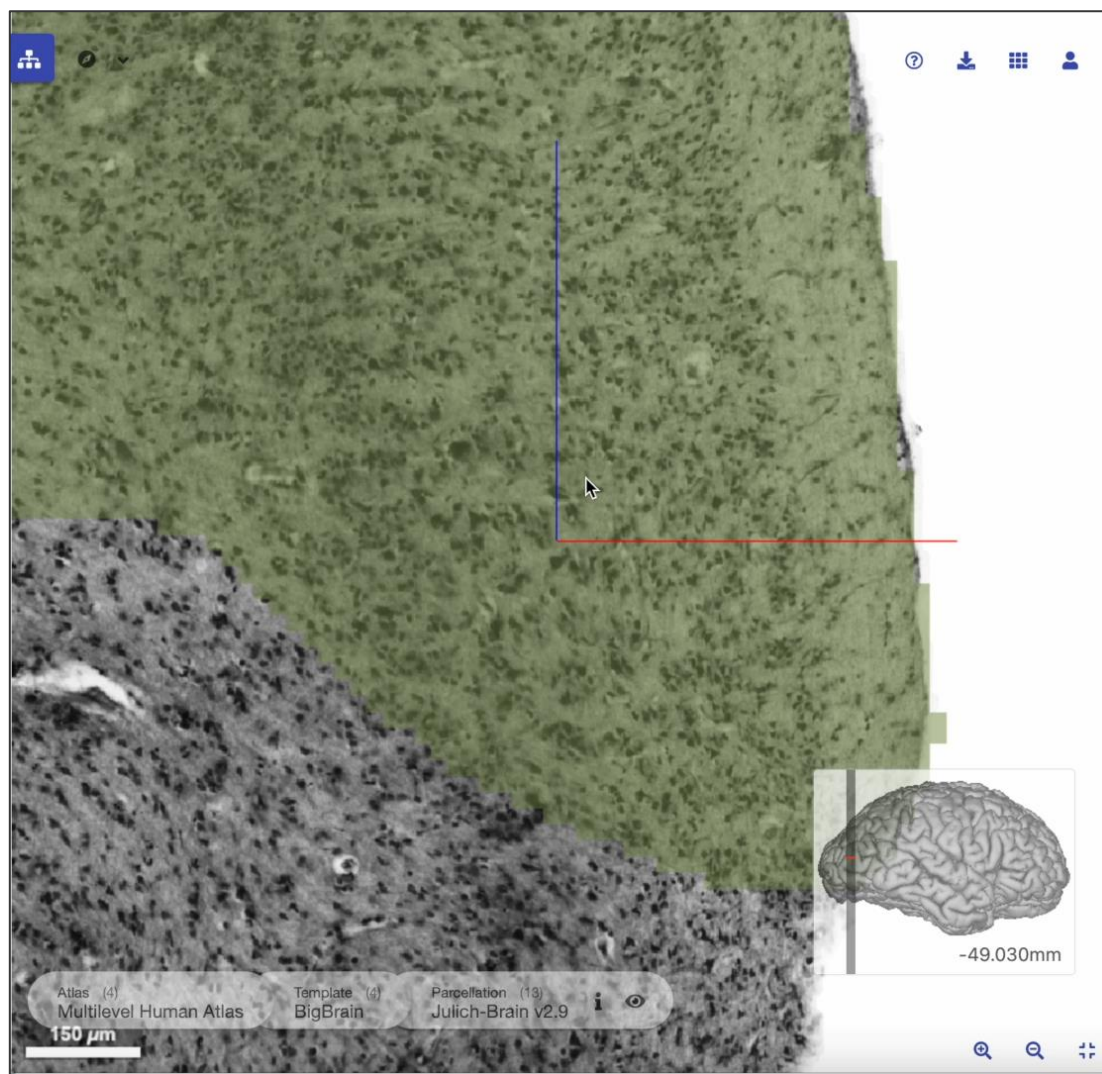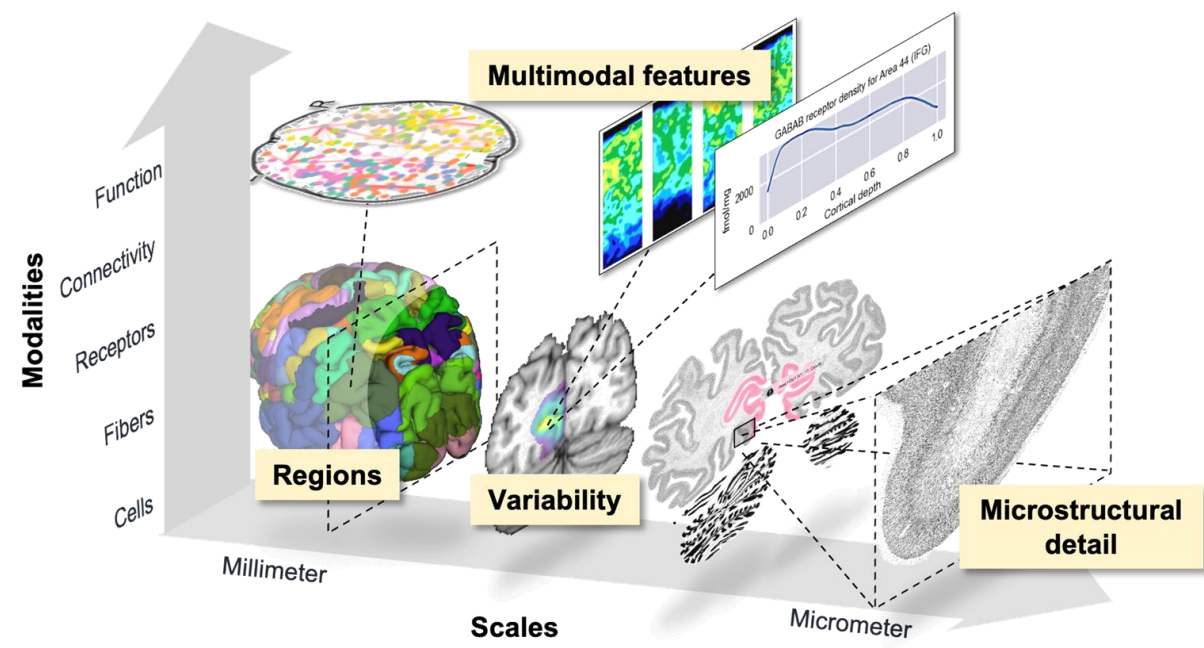
Slide courtesy of T. Dickscheid

The atlas framework is well accessible

https://siibra-python.readthedocs.io

Can run on EBRAINS collab, HIP, HDC and local computers
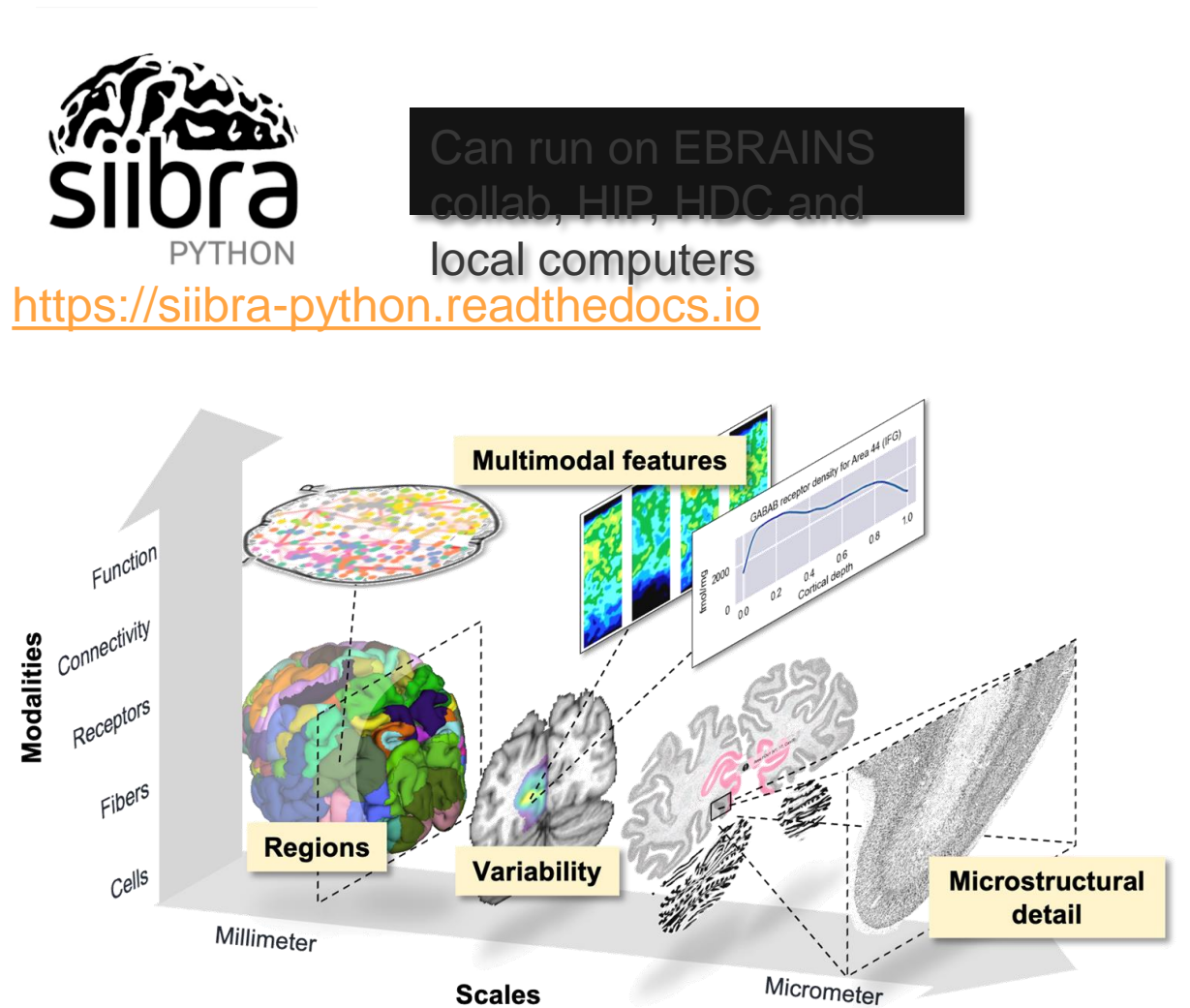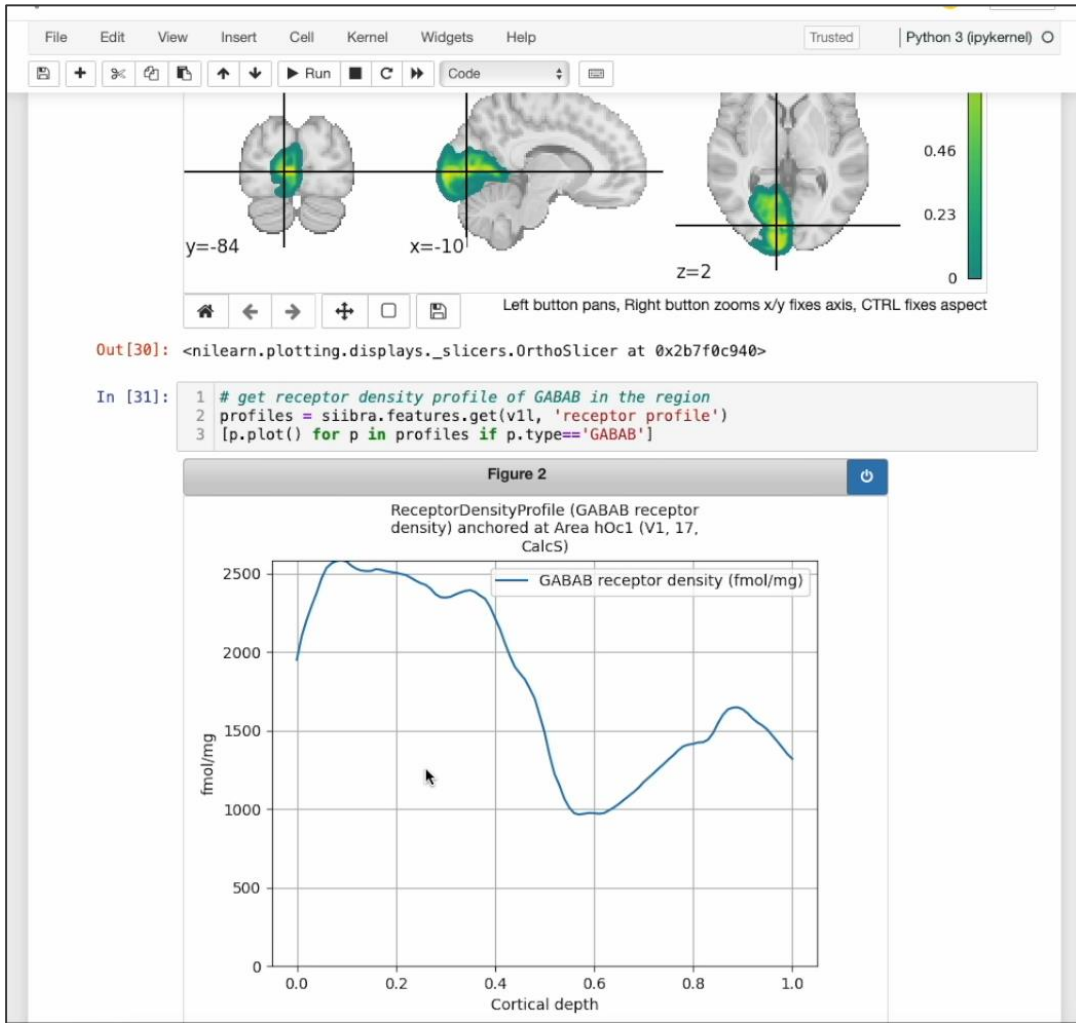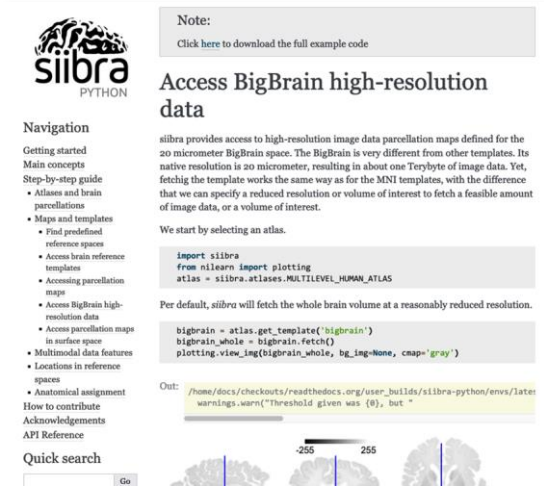
The atlas is computable and interoperable

siibra-python
Fully functional Python client
siibra-python.readthedocs.io
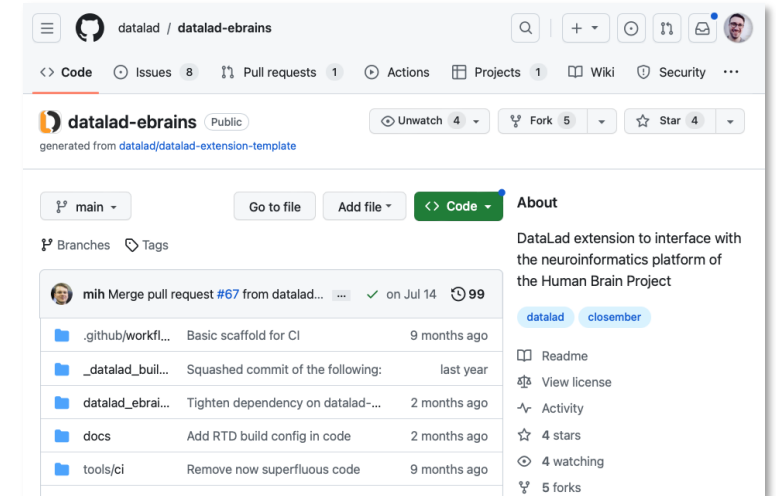
siibra-api
HTTP interface for external apps
siibra-api-stable.apps.hbp.eu/v3_0

datalad-ebrains
„git for EBRAINS datasets"
facilitates off-site processing
https://github.com/datalad/datalad-ebrains

All atlas content is modeled as FAIR datasets
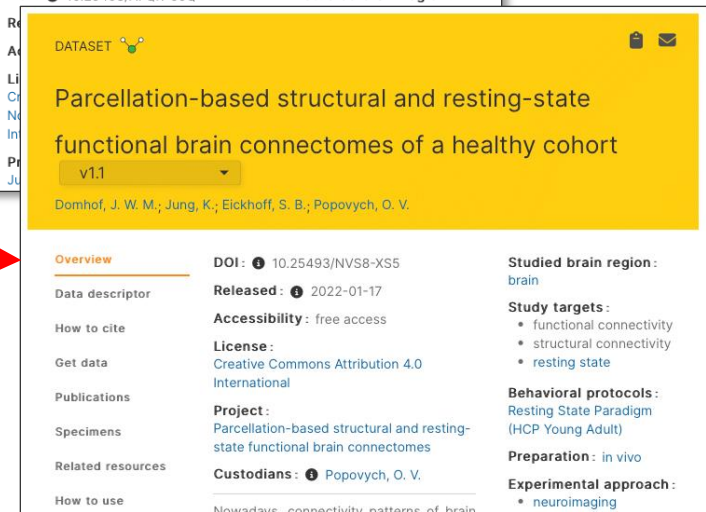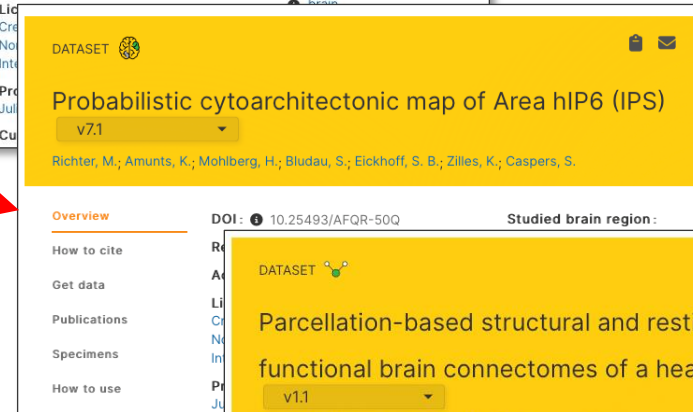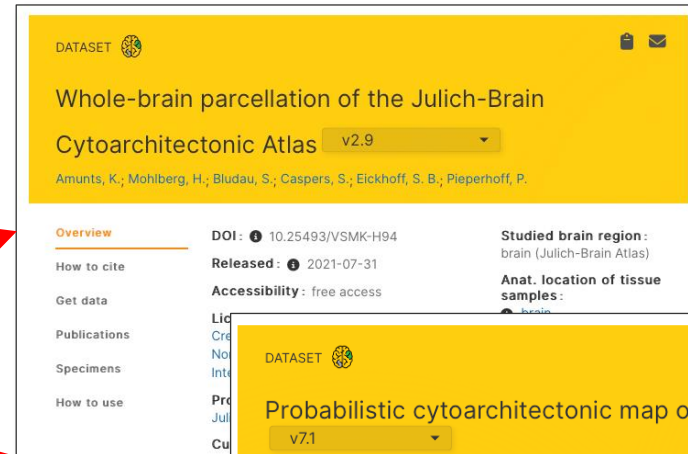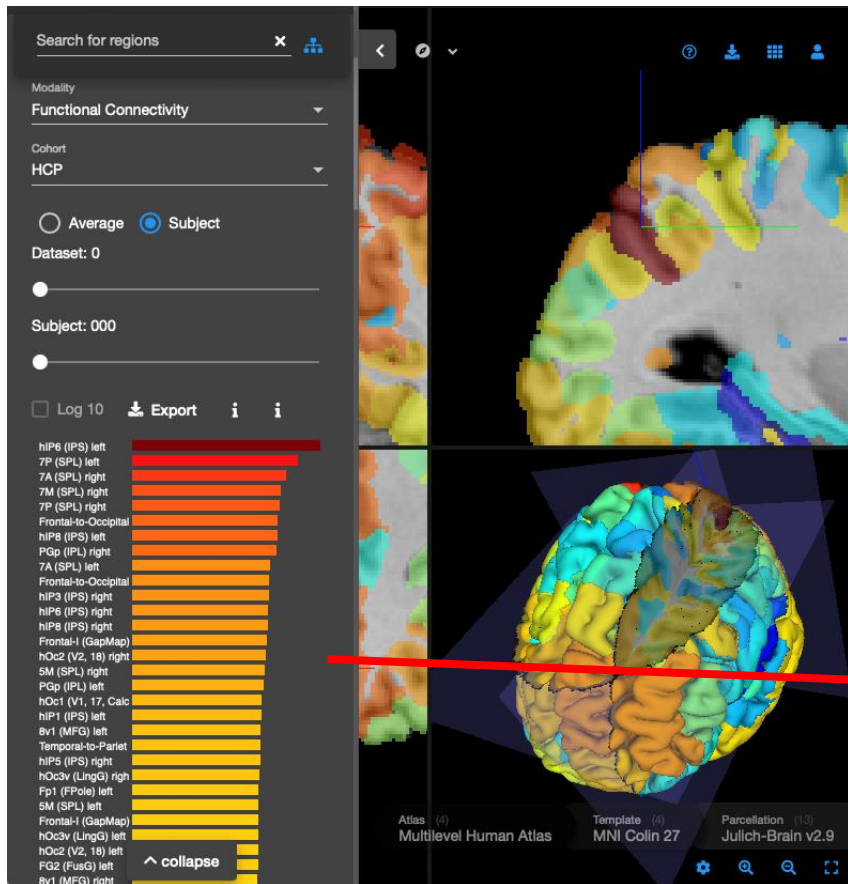
AtOM: Atlas **O**ntology **M**odel
(Kleven et al., Scientific Data;
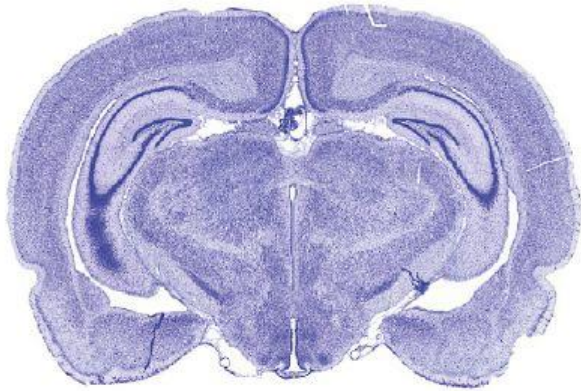https://doi.org/10.1038/s415
97-023-02389-4)

Tools are in place to integrate more data into atlases

2D sections in rodent brains

QuickNII & VisuAlign

https://www.ebrains.eu/tools/visualign
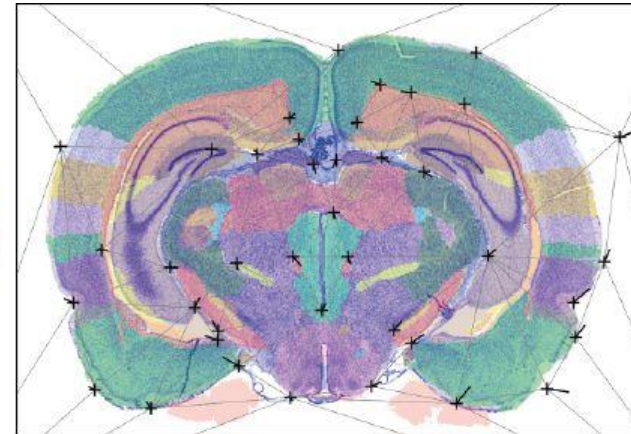https://www.ebrains.eu/tools/quicknii
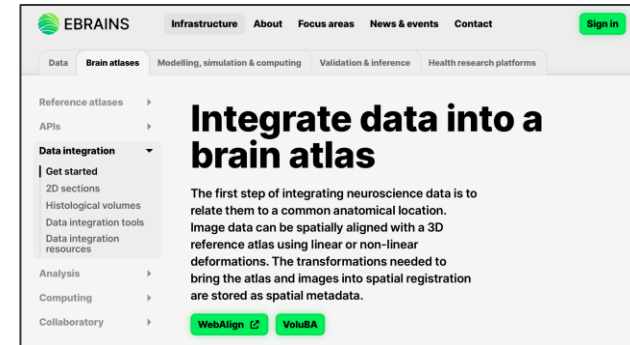
Section image
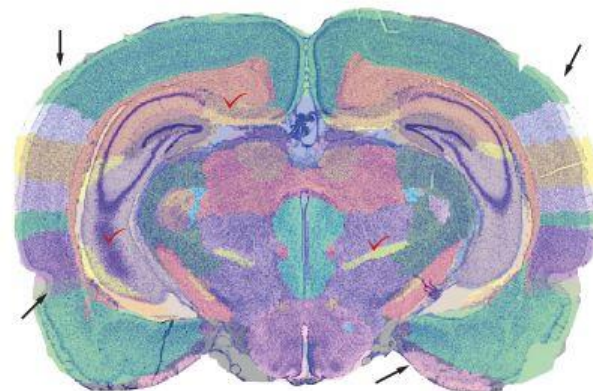
QuickNII registration:
find deviation angles

VisuAlign registration:
non-linear adjustment

ebrains.eu/brain-atlases/data-integration

Tools are in place to integrate more data into atlases

2D sections in rodent brains

QuickNII & VisuAlign

https://www.ebrains.eu/tools/visualign
https://www.ebrains.eu/tools/quicknii

Histological volumes of interest

voluba

https://ebrains.eu/tools/voluba

ebrains.eu/brain-atlases/data-integration

Thank you