

KARAIB

<http://project.inria.fr/karaib>

Bertrand Thirion, bertrand.thirion@inria.fr



Motivation

- Cognitive concepts and brain pathologies are ill-defined [Yarkoni Poldrack 2016]
 - Psychological constructs
 - Pathologies or symptoms ?
- Leverage existing data w. representation learning
 - Mostly Image repositories & publications
 - Identify latent factors

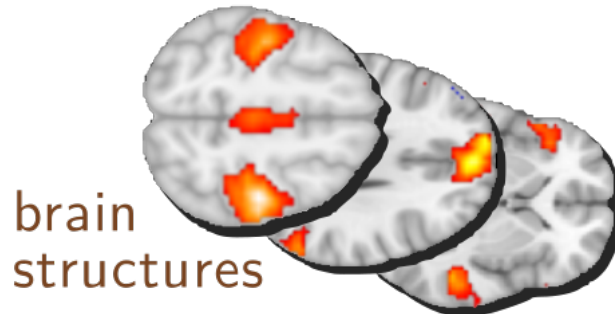
Existing resources

Data source	Data type	Volume	Notes
OpenNeuro	Raw fMRI data, mostly task-related activation	247 studies (as of June 2019), > 6000 subjects	Wide variety of cognitive tasks
NeuroVault	Statistical maps, mostly group-level activation maps	80k maps from 2032 different collections (June 2019)	idem, noisy annotations.
NeuroSynth & Neuroquery	Coordinates from the literature	500 000 from 14 000 publications (June 2019)	Neuroquery developed by PARIETAL
Human Connectome Project	Raw fMRI and diffusion data, mostly rest fMRI	1 200 subjects, with 11 GB per subject	Homogeneous, high-quality data
Individual Brain Charting (IBC)	Task fMRI	12 subjects, 50 acquisitions covering many cognitive domains	Wide cognitive coverage, acquisition and processing by PI team.
UK Biobank Image Data	Rest and task fMRI	100 000 subjects (2021)	A lot of individual behavioral information, aging effects, diseases.

Aim 1: Build coordinated representations bridging psychology and brain maps

Unsupervised learning from large image sets

$$\begin{matrix} \text{studies} \\ \left[\begin{matrix} \text{voxels} \\ X \end{matrix} \right] = \left[\begin{matrix} D \\ A \end{matrix} \right] \end{matrix}$$



cognitive loadings

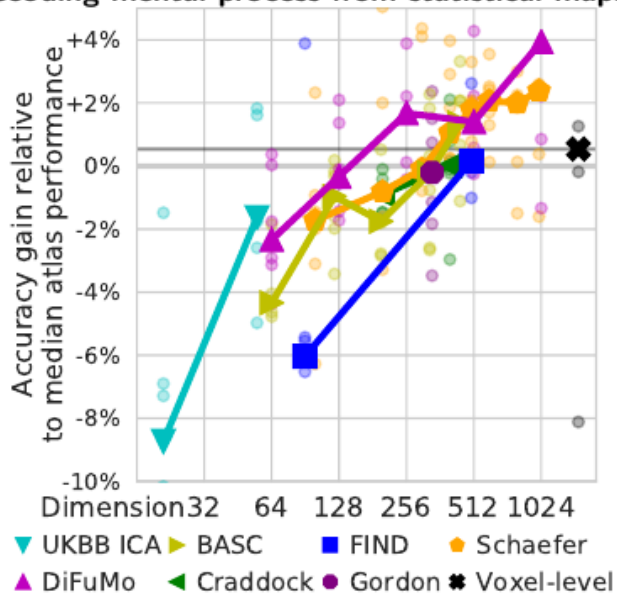
Beyond plain dictionary learning

- Nonlinear versions
- Connectivity model
- multi-resolution

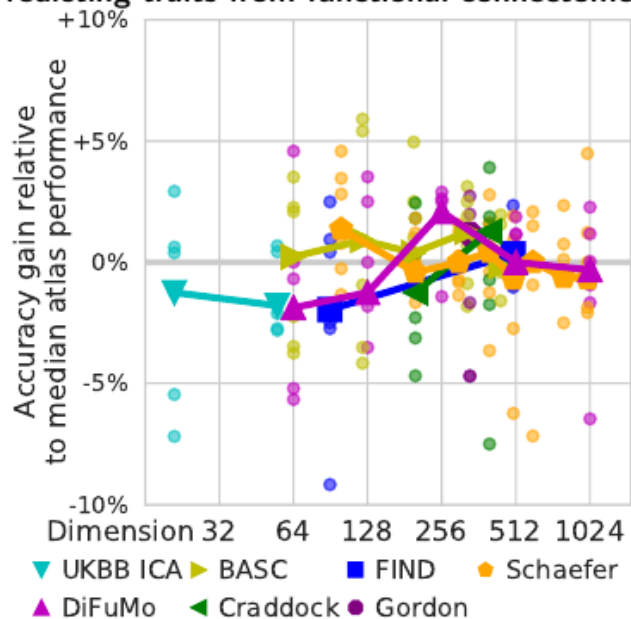
Preliminary results: DiFuMo

The most powerful brain atlas for functional data analysis !

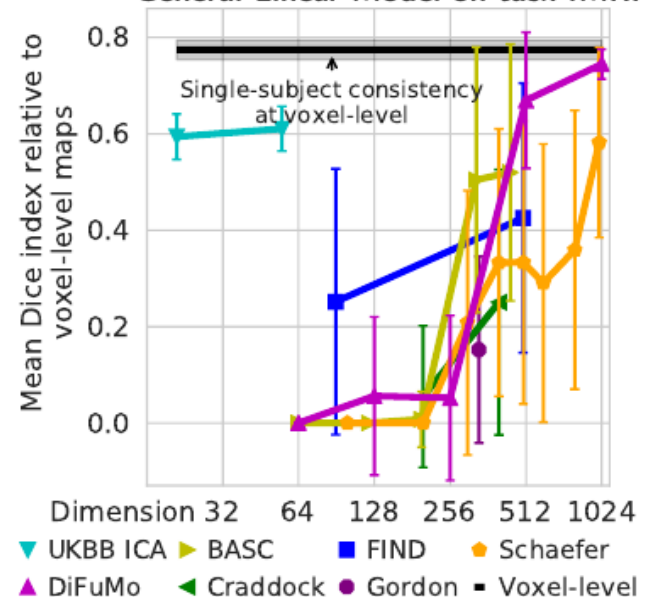
Decoding mental process from statistical maps



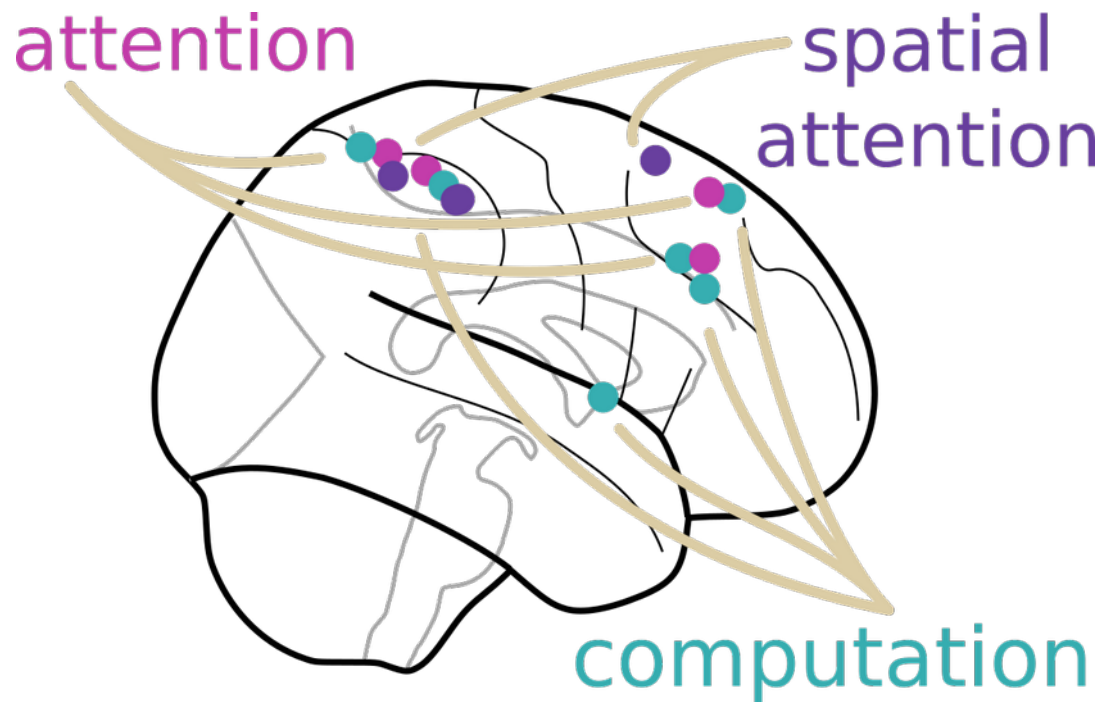
Predicting traits from functional connectomes



General Linear Model on task fMRI

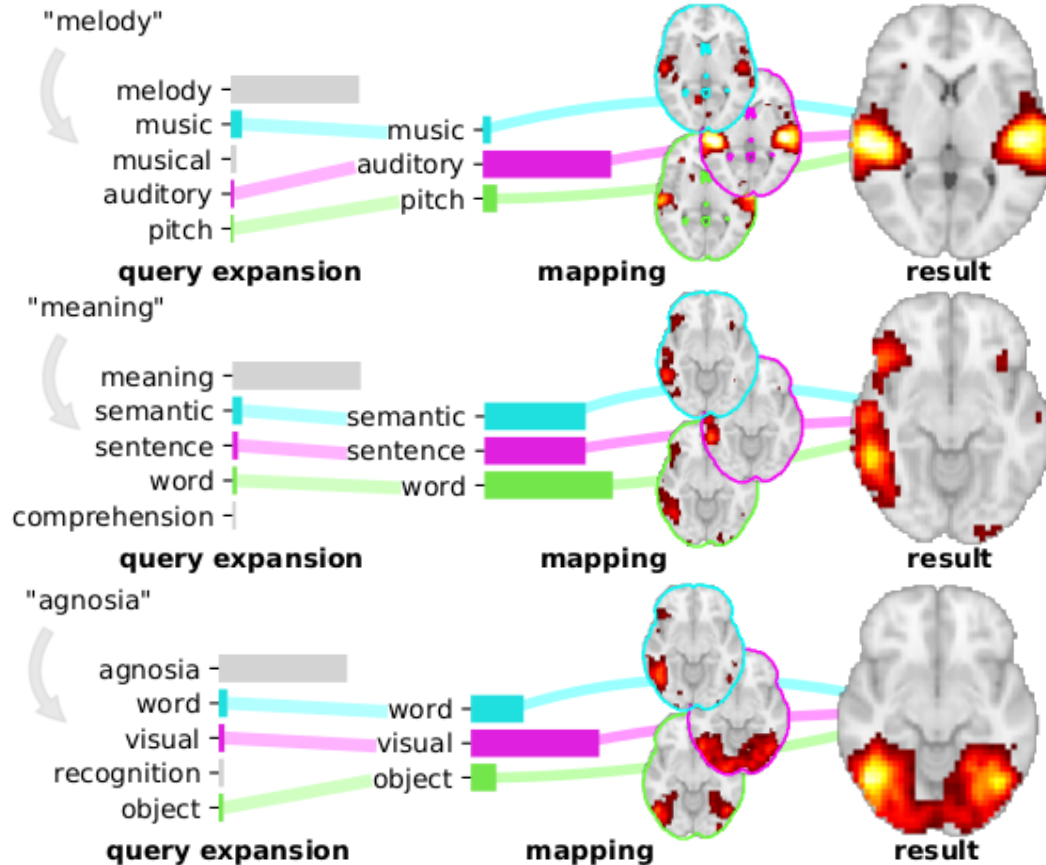


Leveraging Modern language models of psychological literature



More powerful (pre-trained) language models in Neuroquery

Preliminary work: Neuroquery



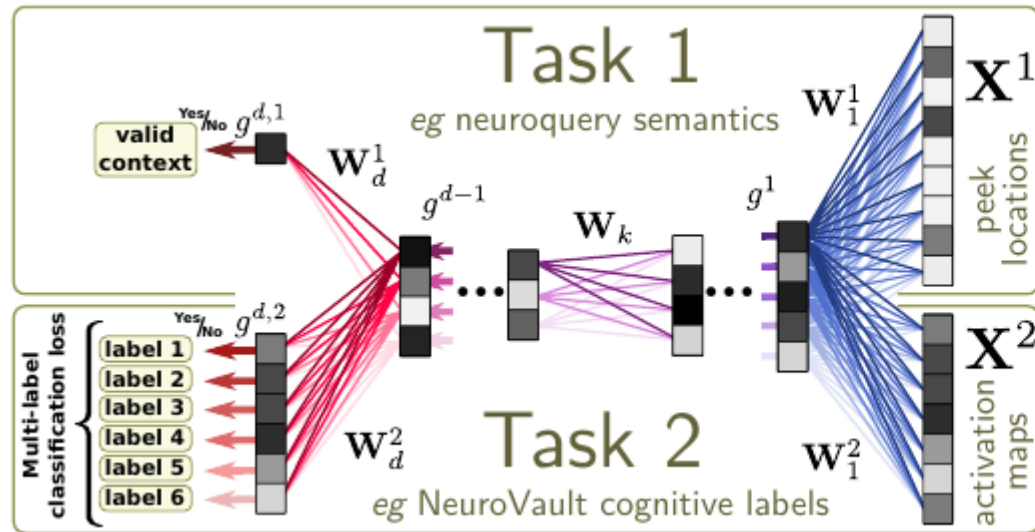
Semantic structure
→ map concepts
with few/no data

[Dockes et al. Elife 2020]

Deep learning for coordinated representations

Multi-task architecture:

- models for Tasks 1 and 2 share parameters.
- Joint Minimization → common parameters benefit from both tasks and data sources.



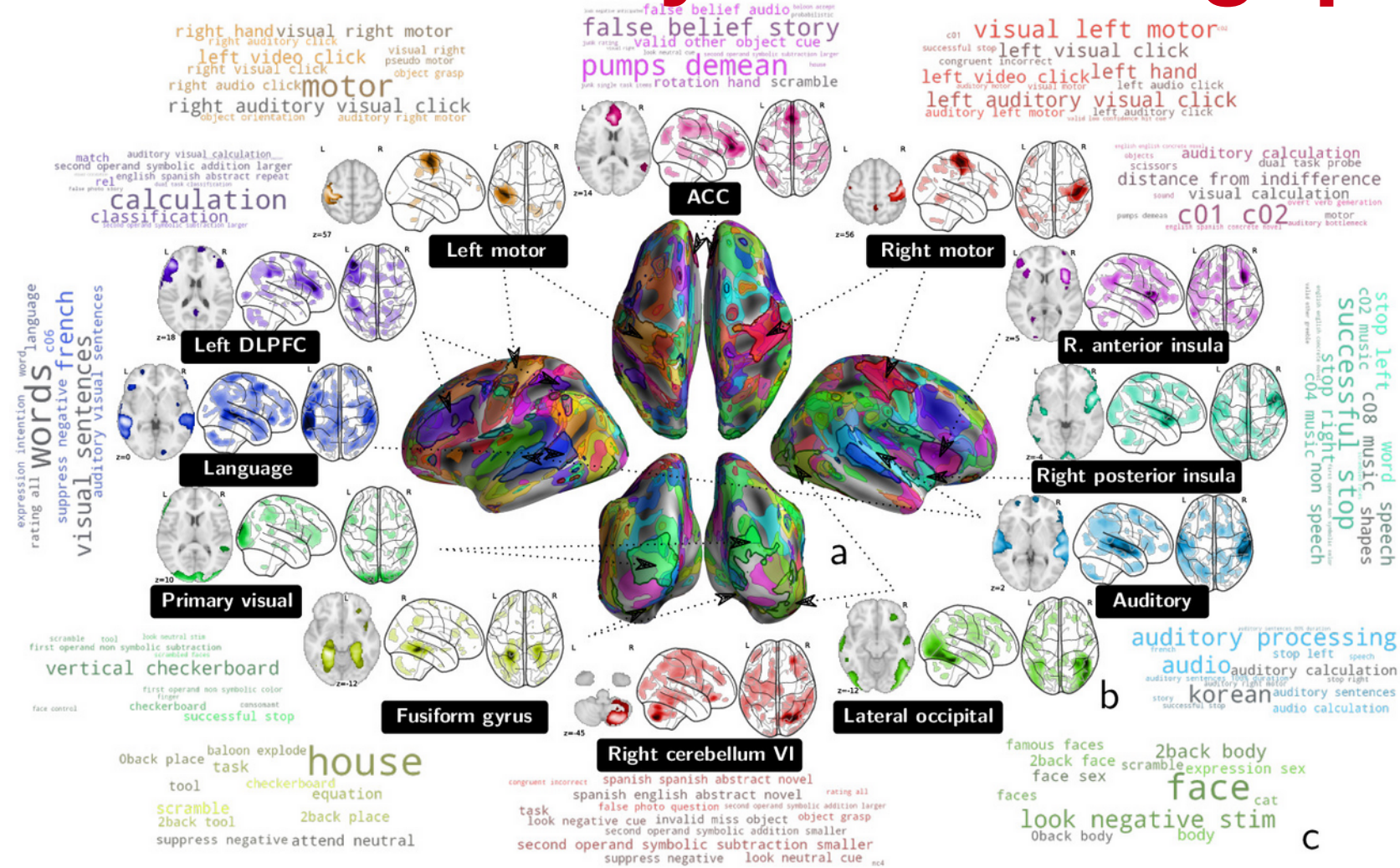
Model for task 1:

$$y^1 = g^{d,1} \left(\left(W_d^2 \dots g^k \left(W_k \dots g^1 \left((W_1^1 x^1)^T \right) \right) \right) \right)^T$$

Model for task 2:

$$y^2 = g^{d,2} \left(\left(W_d^2 \dots g^k \left(W_k \dots g^1 \left((W_1^2 x^2)^T \right) \right) \right) \right)^T$$

Preliminary result: Cogspaces

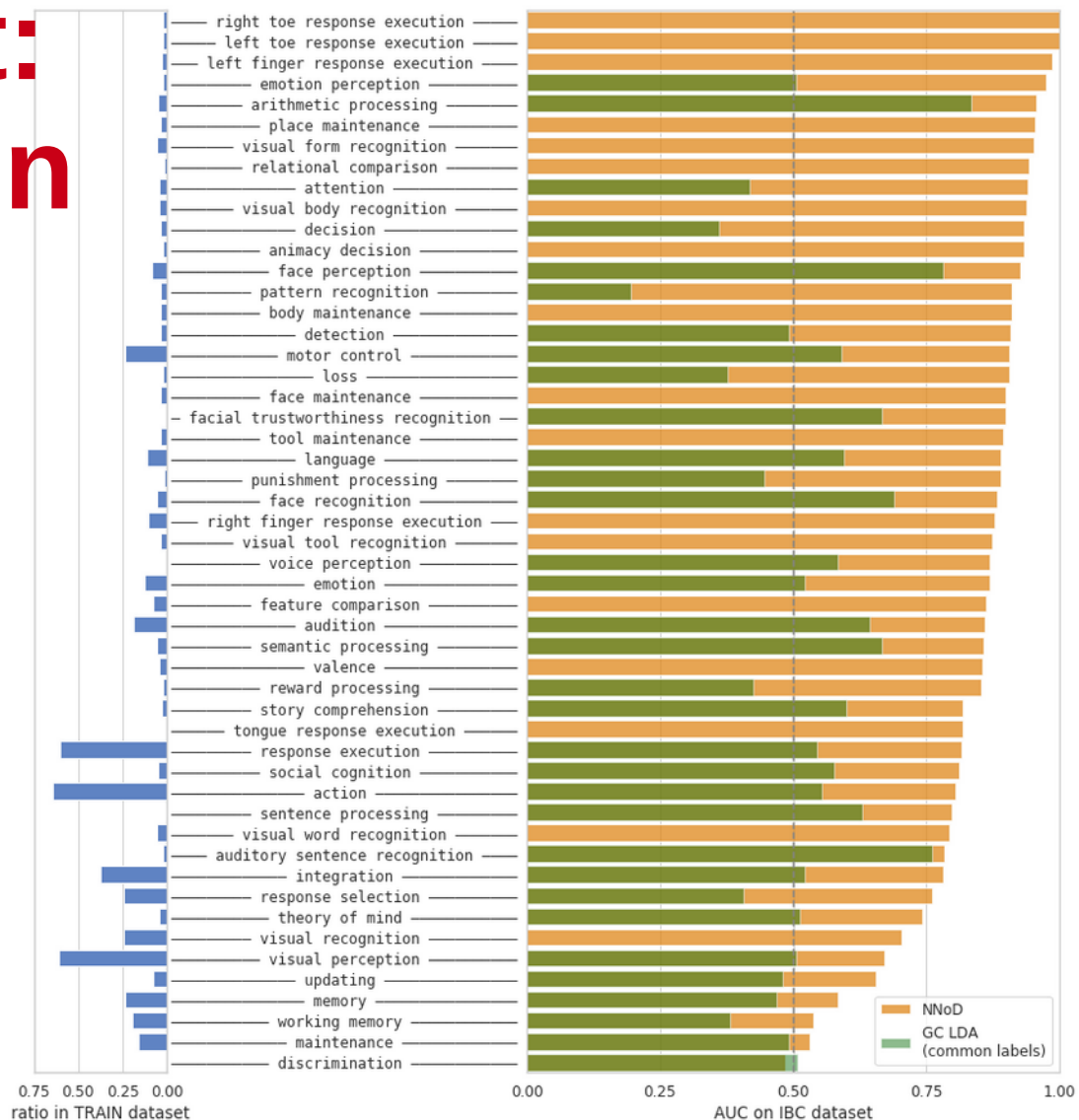


- Decoding model that generalizes across datasets
 - Relies on network models
- [Mensch et al subm]

Preliminary result: Neural networks on a Dictionary

- 51 terms decoded on NeuroVault

[Menuet et al. Subm.]



Aim 2: Assessment of coordinated (image/text) representations

- Statistical inference
 - **Aggregation approaches** (across clusterings of brain domain or concepts) [Nguyen ICML 2020]
 - **Generative approaches:**
 - Knockoffs
 - Latent cause models
 - Generalization to **non-linear** models

From linked representations to formal reasoning: integration into NeuroLang

NeuroLang → DSL for human neuroscience research
→ combine imaging data, anat. descriptions & ontologies

1. represents information in a syntax close to **natural language**
2. **Querying ontologies** w. same expressive power as current standards (SPARQL, OWL)
3. **Probabilistic** language --- graphical models allowing the implementation of many ML algorithms

Current integration of meta-analytic queries and information in NeuroLang

Aim 3: Provide neuroimaging semantic-learning tools and interfaces

NeuroQuery Brain maps by querying the neuroscience literature

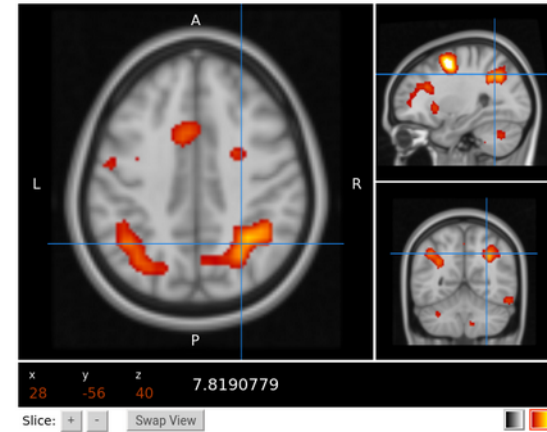
A query on neuroscience, cognition, or brain pathologies
n-back tests of autobiographical memory

Click to edit.

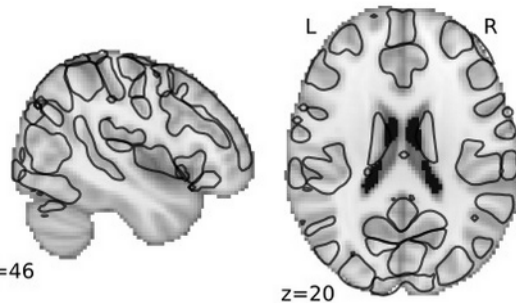
Edit query

Terms related to the query

Term	Similarity	Weight in brain map	N
ry			523
ck			3884
ng			5201
ry			8405
ry			4424
in			1670
sk			12194
tal			10017
us			5186
ral			11897
tal			10159
us			4610



DiFuMo Dictionaries of multiple dimensions



x=46

z=20

See regions for: 64 dimensions Download

DiFuMo

Dictionary of Functional Modes for brain imaging

Multi-resolution atlases of brain functional modes

Search for anatomical locations:

Google Custom Search



Code to use DiFuMo

<http://neuroquery.org>

Software development & dissemination



Nilearn:

Machine learning for Neuro-Imaging in Python

SVM Ward clustering
Searchlight ICA
Nifti IO Datasets

Google Custom Search



[Nilearn Home](#) | [User Guide](#) | [Examples](#) | [Reference](#) |

[Nipy ecosystem](#)

Nilearn is a Python module for **fast and easy statistical learning on Neuroimaging** data. It leverages the [scikit-learn](#) Python toolbox for multivariate statistics with applications such as predictive modelling, classification, decoding, or connectivity analysis.

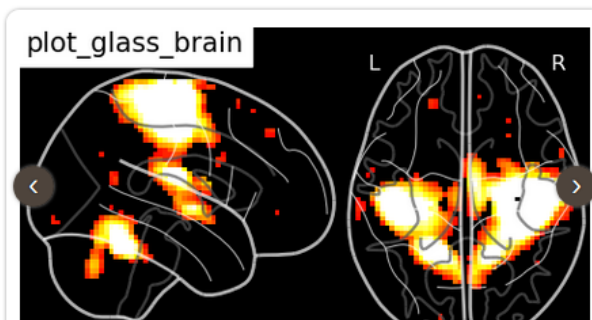
First Steps

Get started with nilearn

Examples

Visit our example gallery

User Guide



News

December 2019: Nilearn 0.6.0 released

April 17th 2019: Nilearn 0.5.2 released

April 12th 2019: Nilearn 0.5.1 released

November 2018: Nilearn 0.5.0 released

June 14th 2018: Nilearn 0.4.2 released

Ongoing: What's new.

Software

Installation

Conclusion

- 5-year project
- A new PhD student
- A starting position
- Integrate & push forward are current assets:
MODL, DiFuMo, NnoD, neuroquery,
NeuroLang, IBC

Thanks

- Gael Varoquaux
- Demian Wassermann
- Aapo Hyvärinen
- Huges Talbot
- A. Mensch, J. Dockès, R. Menuet, K. Dadi, J.A. Chevalier, Binh Nguyen, H. Richard, T. Bazeille
- All Parietal fellows