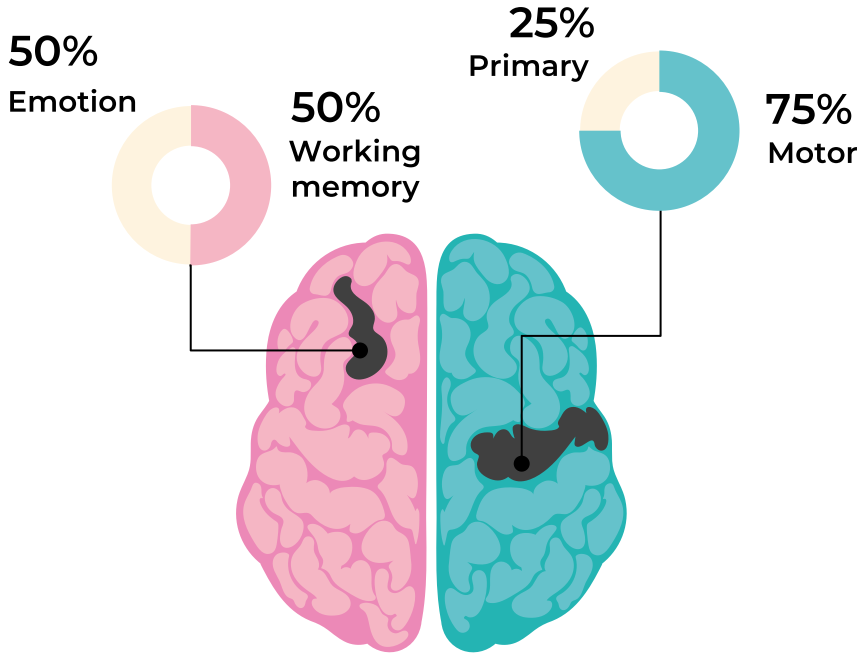
Meta-analytic

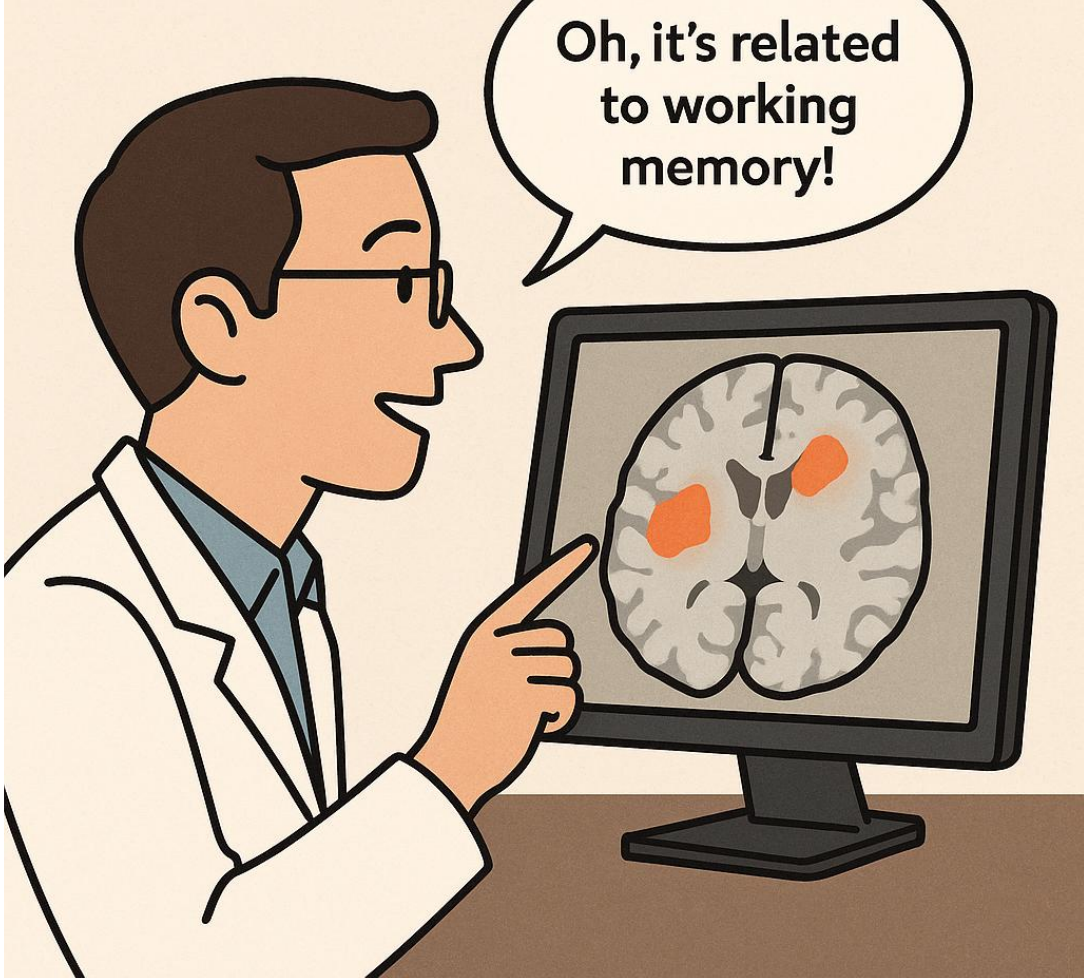
# Functional Decoding

Youngjo Song

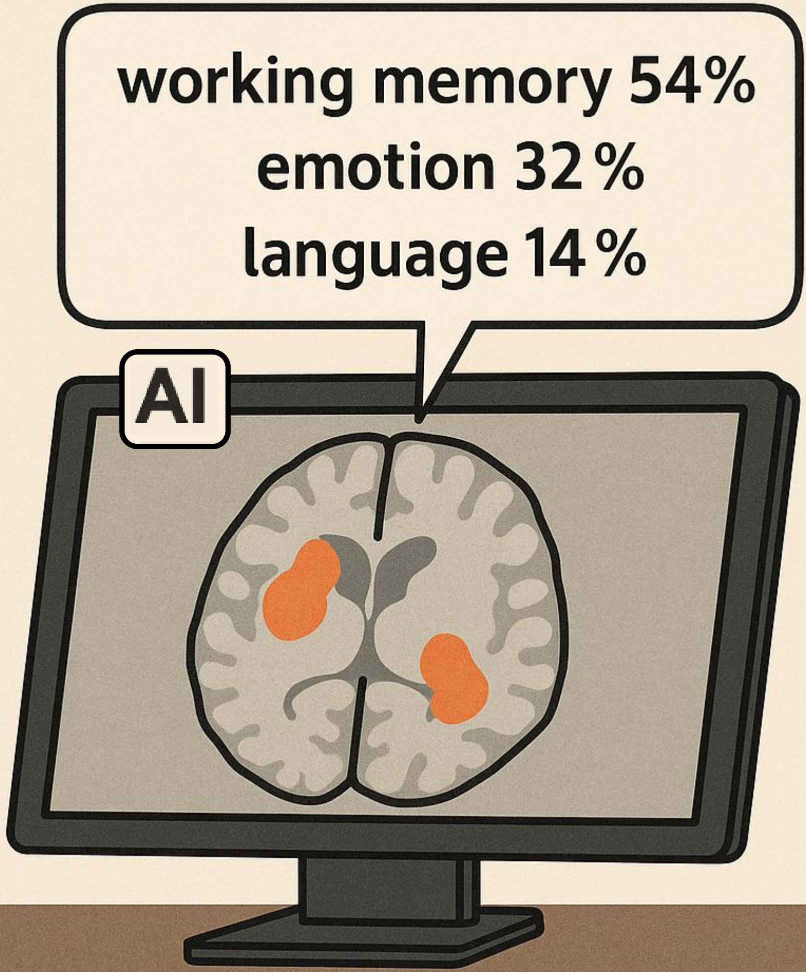



For example codes: https://github.com/YJ-0000/NiMARE_Functional_Decoding_Examples

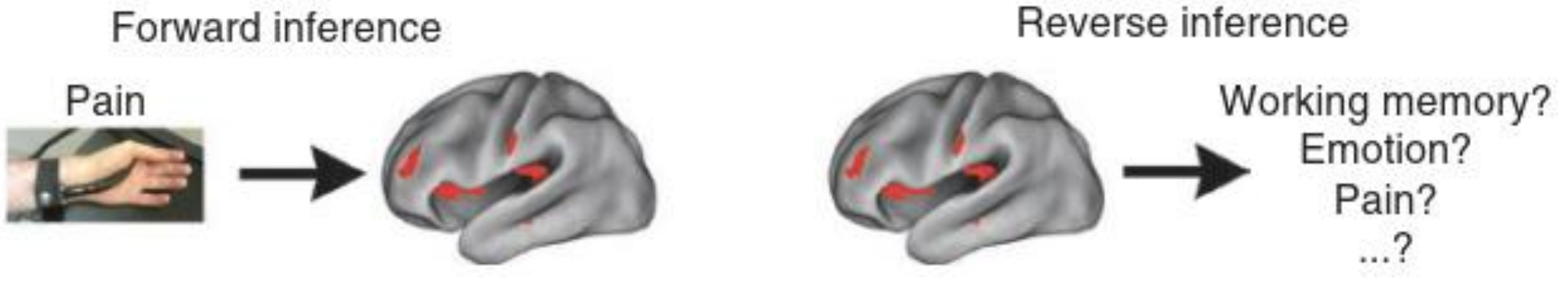
# In Old Days,…



But Now,



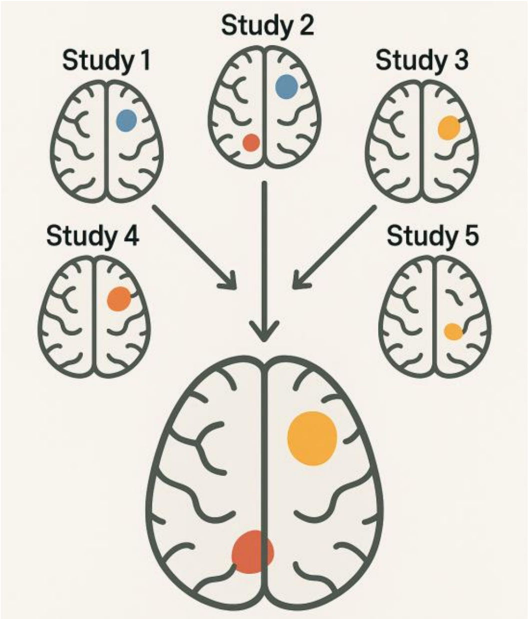
# Forward/Reverse Inference




Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nature methods, 8(8), 665-670.

# Why We Need Meta-Analysis in fMRI Research

- Rapid growth of fMRI studies has produced **massive amounts of data.**
- Individual studies are often underpowered and prone to false positives.
- **Reliable conclusions require evidence across multiple studies.**
- Meta-analysis enables **large-scale data aggregation and synthesis.**

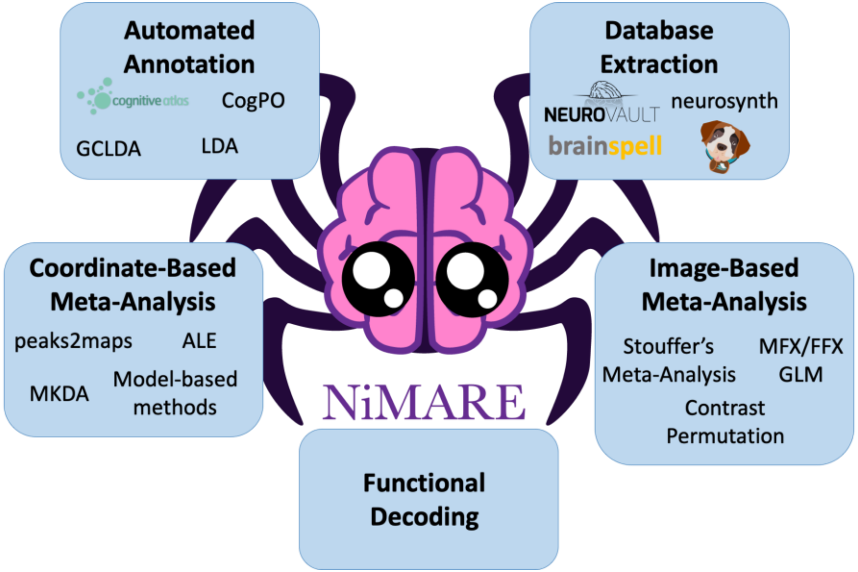



Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nature methods, 8(8), 665-670.

# NiMARE: Neuroimaging Meta-Analysis Research Environment

- **NiMARE** is a **Python package** for conducting meta-analyses of neuroimaging studies.

- It provides **a standardized, flexible** interface for multiple algorithms and databases

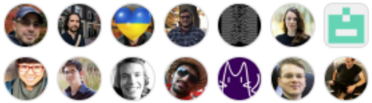



https://nimare.readthedocs.io/

github.com/neurosynth/neurosynth

| | | |
|---|---|---|
| Changelog.md | version bump | 7 years ago |
| Dockerfile | Fix issues with PR. | 8 years ago |
| LICENSE | initial commit | 13 years ago |
| MANIFEST | docstring edits for PEP8 compliance | 11 years ago |
| MANIFEST.in | RF: Move resources/ under neurosynth + do not rely on os-s... | 12 years ago |
| README.md | Update README.md | last year |
| requirements.txt | Update requirements.txt | 4 years ago |
| setup.cfg | setup/testing tweaks | 10 years ago |
| setup.py | Add biopython to dependencies | 8 years ago |

Contributors 16



+ 2 contributors

Languages

Python 99.1% Dockerfile 0.9%

README MIT license

**Note: this package is no longer actively maintained; most of its functionality has been integrated into the much more expansive NiMARE package, which we recommend using instead.**

# What is Neurosynth?

Neurosynth is a Python package for large-scale synthesis of functional neuroimaging data.

https://github.com/neurosynth/neurosynth

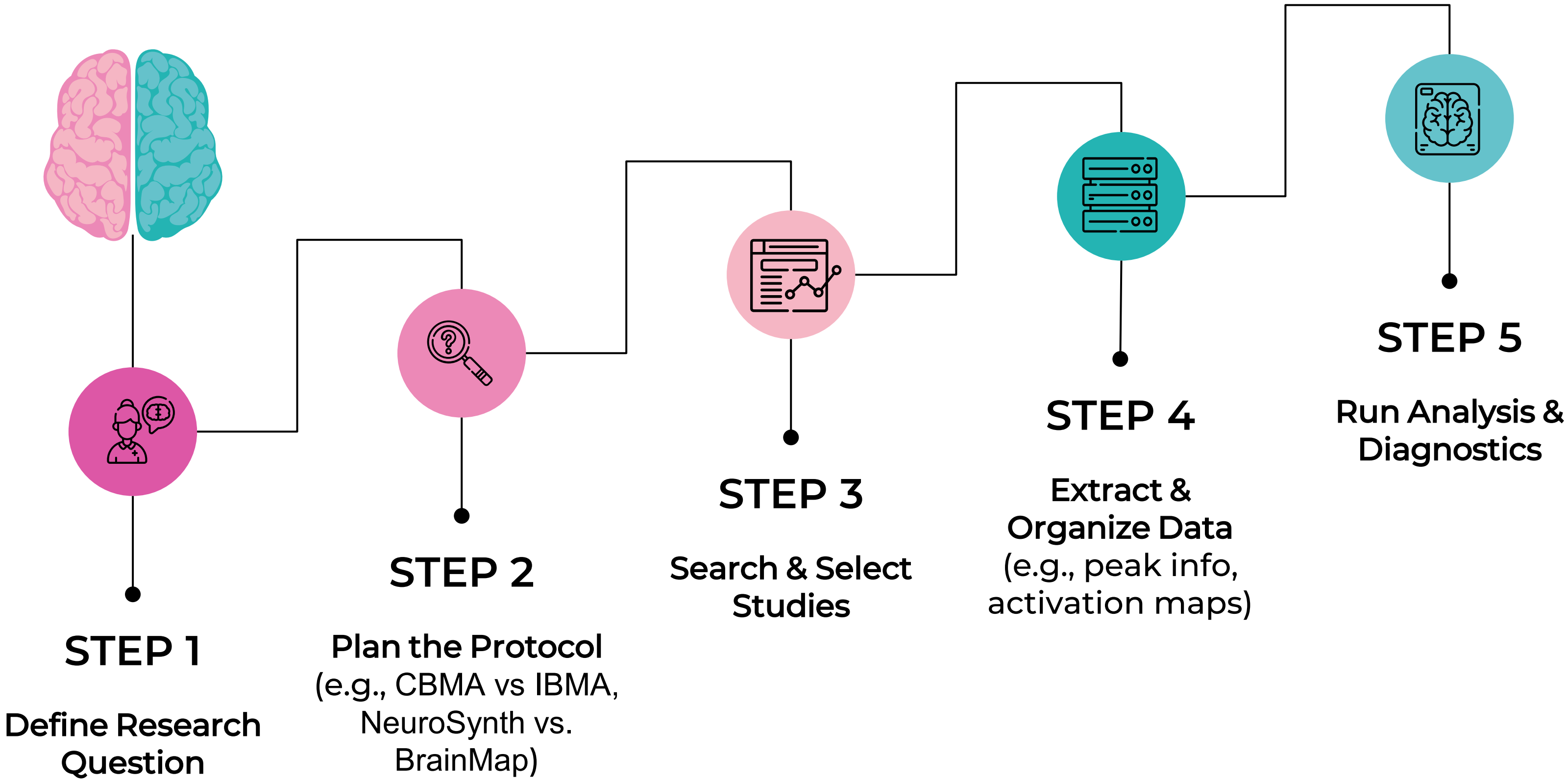
# Contents

## fMRI Meta-analysis via NiMARE

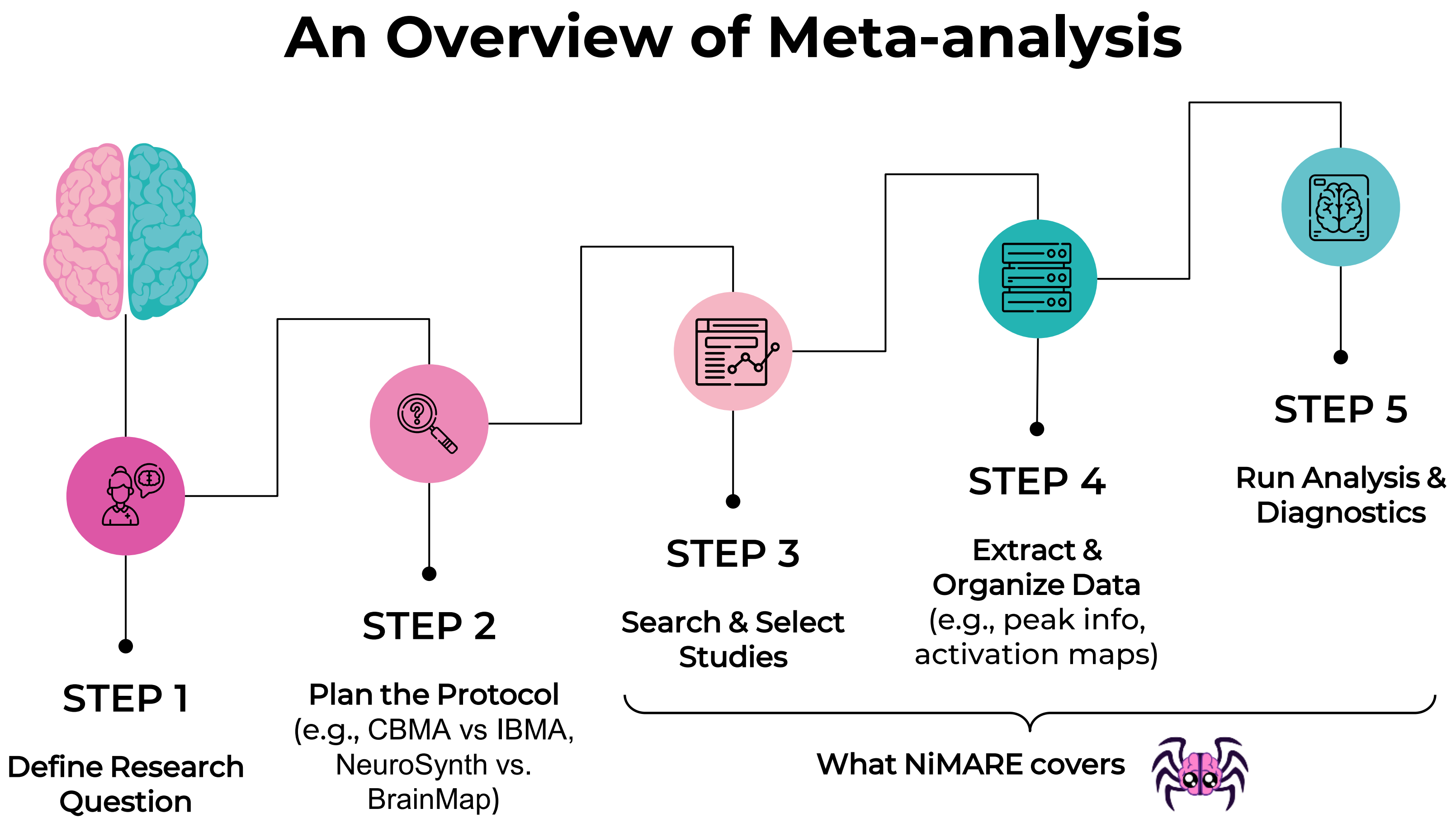
# 01

# Overview of fMRI Meta-Analysis Workflow

# An Overview of Meta-analysis

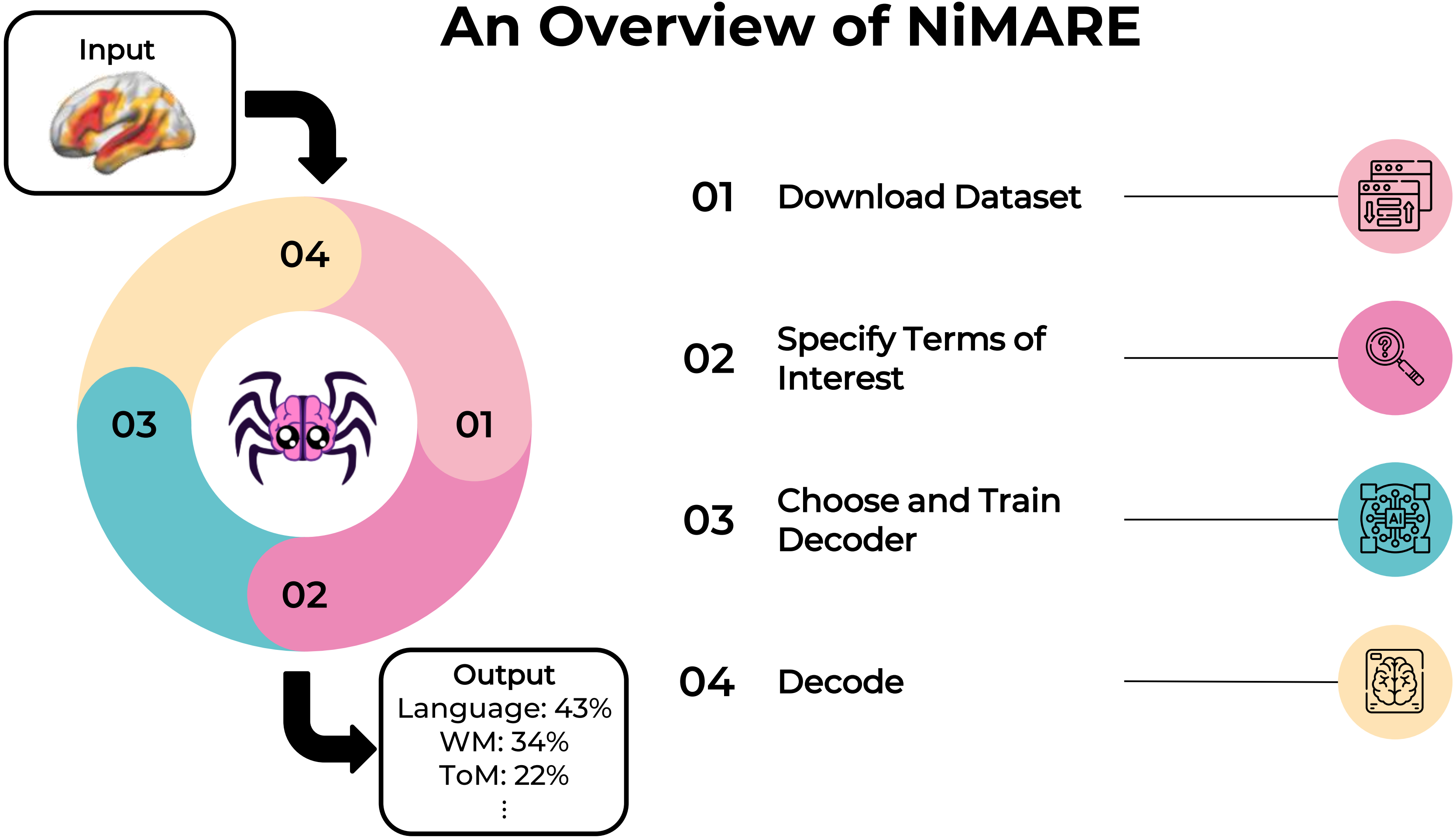



Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., ... & Eickhoff, S. B. (2018). Ten simple rules for neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews*, *84*, 151-161.



An Overview of Meta-analysis
STEP 1
Define Research Question
STEP 2
Plan the Protocol (e.g., CBMA vs IBMA, NeuroSynth vs. BrainMap)
STEP 3
Search & Select Studies
STEP 4
Extract & Organize Data (e.g., peak info, activation maps)
STEP 5
Run Analysis & Diagnostics
What NiMARE covers

# An Overview of NiMARE



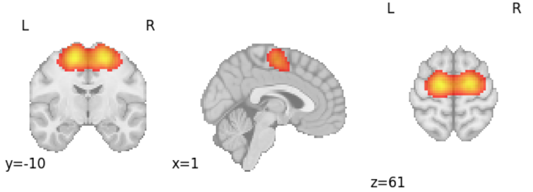
# 02

# Using NiMARE: Practical Guide and Example

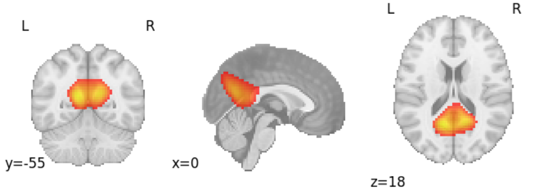
# Goal: Functionally Characterizing a Given Spatial Map

- We aim to interpret three spatial maps.
    - These maps are three different components from NeuroMark 2.2 template.

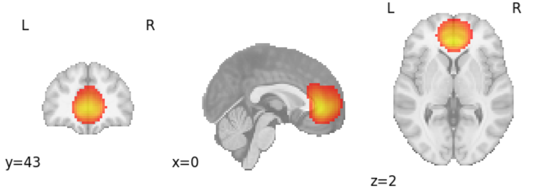
Sensorimotor



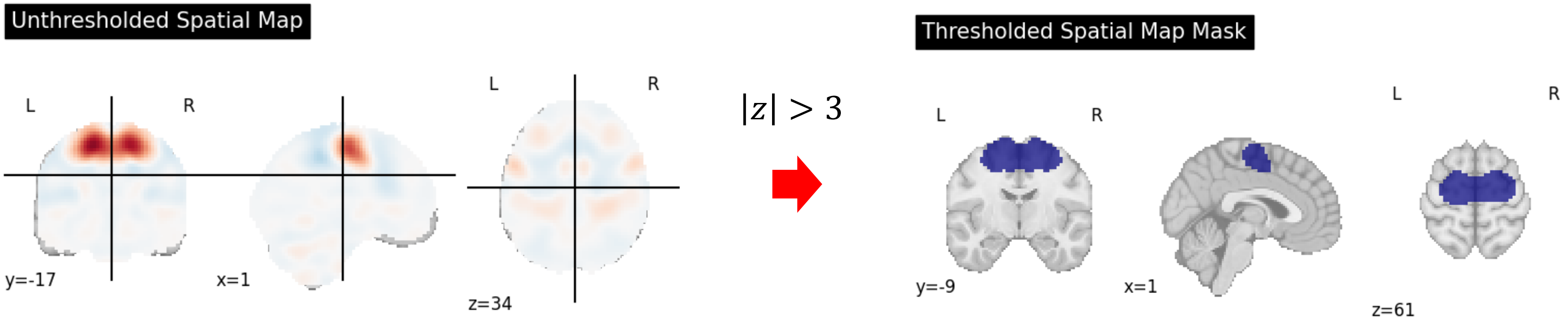

DMN (Precu)




DMN (MPFC)



# You need to threshold the spatial maps to retain only the most significant regions.

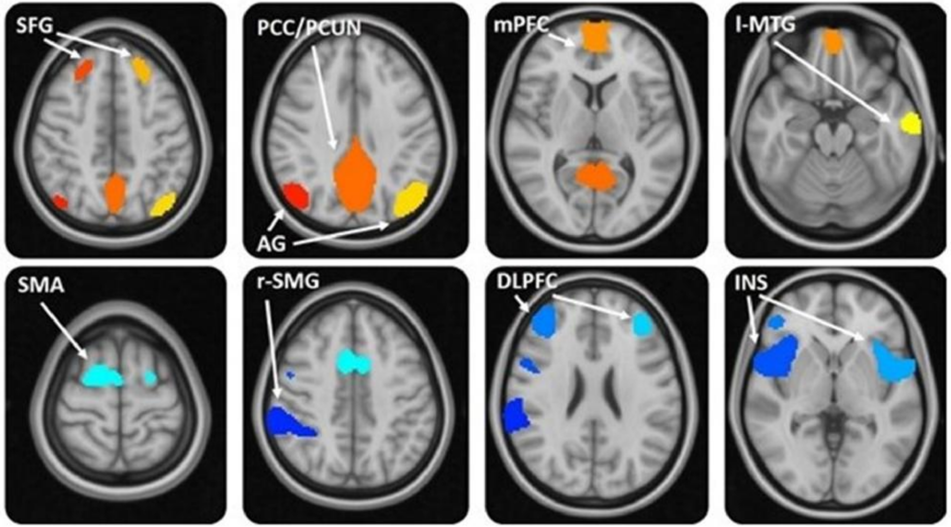



This step is required in order to use the NeuroSynth Decoder.

# How can we collect data from the studies for meta-analysis?

# Activation patterns are reported based on the local maxima (i.e., peaks).

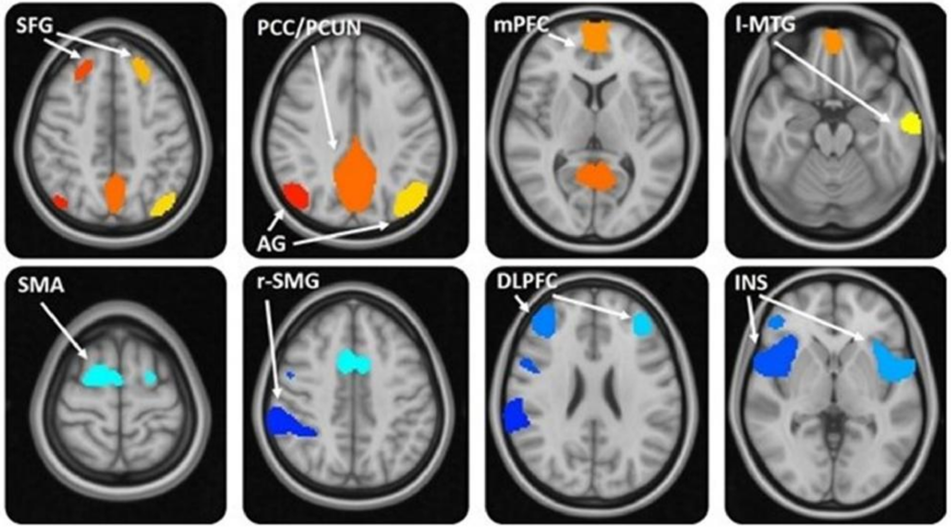
| Region of interest | Peak coordinates | | | $N_{vox}$ |
|---|---|---|---|---|
| | $MNI_x$ | $MNI_y$ | $MNI_z$ | |
| Default Mode Network | | | | |
| PCC/PCUN | -1 | -57 | 26 | 5561 |
| mPFC | -1 | 59 | 2 | 2537 |
| right-AG | 48 | -66 | 31 | 1024 |
| left-AG | -45 | -69 | 31 | 1712 |
| right-SFG | 24 | 33 | 45 | 170 |
| left-SFG | -22 | 34 | 45 | 267 |
| left-MTG | -61 | -12 | -21 | 287 |
| Anti-Correlated Network | | | | |
| SMA | 7 | 4 | 53 | 1630 |
| right-SMG | 56 | -38 | 38 | 2142 |
| right-INS | 48 | 10 | 3 | 2093 |
| left-INS | -47 | 6 | -0 | 1406 |
| right-DLPFC | 40 | 42 | 19 | 1074 |
| left-DLPFC | -37 | 41 | 22 | 347 |




Groot, J. M., Boayue, N. M., Csifcsák, G., Boekel, W., Huster, R., Forstmann, B. U., & Mittner, M. (2021). Probing the neural signature of mind wandering with simultaneous fMRI-EEG and pupillometry. *NeuroImage*, *224*, 117412.

# Activation patterns are reported based on the local maxima (i.e., peaks).

| Region of interest | Peak coordinates | | | $N_{vox}$ |
|---|---|---|---|---|
| | $MNI_x$ | $MNI_y$ | $MNI_z$ | |
| Default Mode Network | | | | |
| PCC/PCUN | -1 | -57 | 26 | 5561 |
| mPFC | -1 | 59 | 2 | 2537 |
| right-AG | 48 | -66 | 31 | 1024 |
| left-AG | -45 | -69 | 31 | 1712 |
| right-SFG | 24 | 33 | 45 | 170 |
| left-SFG | -22 | 34 | 45 | 267 |
| left-MTG | -61 | -12 | -21 | 287 |
| Anti-Correlated Network | | | | |
| SMA | 7 | 4 | 53 | 1630 |
| right-SMG | 56 | -38 | 38 | 2142 |
| right-INS | 48 | 10 | 3 | 2093 |
| left-INS | -47 | 6 | -0 | 1406 |
| right-DLPFC | 40 | 42 | 19 | 1074 |
| left-DLPFC | -37 | 41 | 22 | 347 |

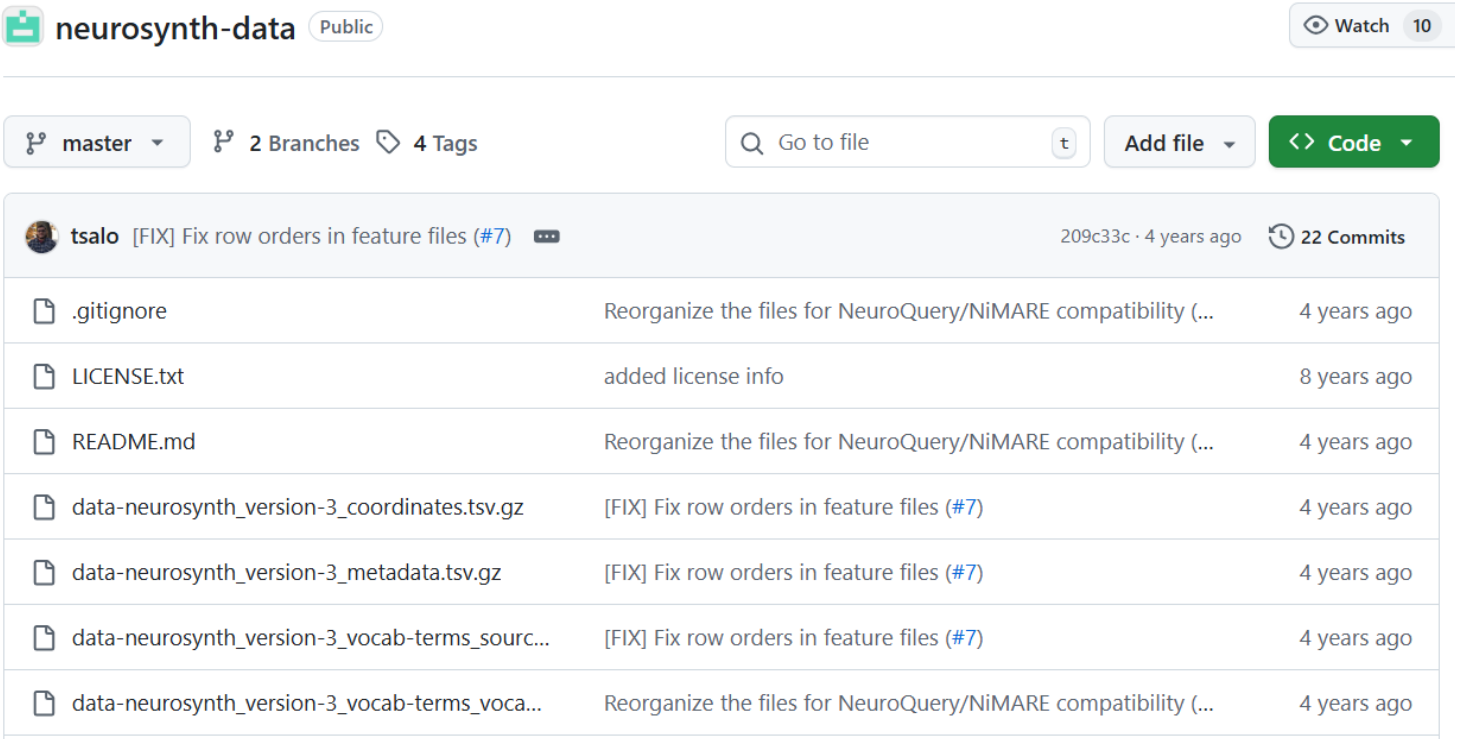



We need to extract these.

Groot, J. M., Boayue, N. M., Csifcsák, G., Boekel, W., Huster, R., Forstmann, B. U., & Mittner, M. (2021). Probing the neural signature of mind wandering with simultaneous fMRI-EEG and pupillometry. *NeuroImage*, *224*, 117412.

# Good news is ...

# NeuroSynth provides a database that stores the extracted peaks

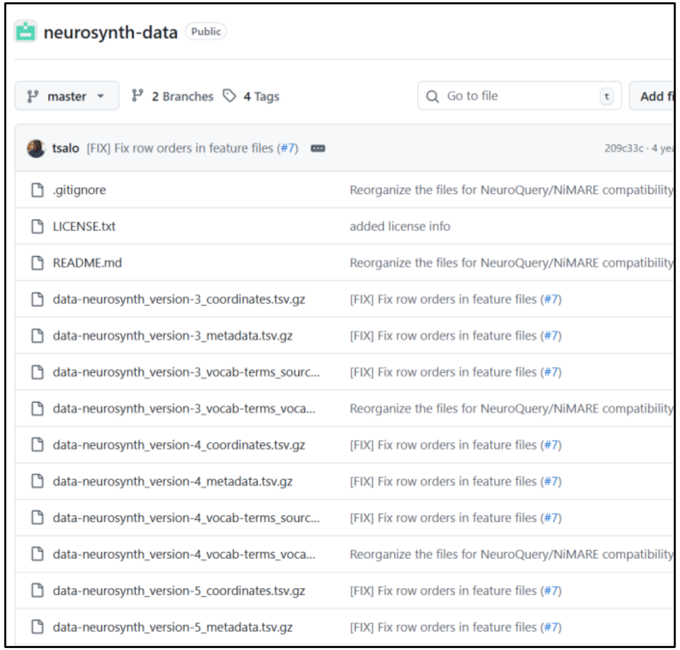



https://github.com/neurosynth/neurosynth-data

All we need is just downloading the data.

# STEP 1. Download datasets

- First, we need to download the dataset that will be used to train the NeuroSynth decoder.
  - The activation peak coordinates extracted and study-annotated data are provided by NeuroSynth via GitHub.




https://github.com/neurosynth/neurosynth-data

Check “Code01_Download_Datasets.ipynb”

```python
###############################################################################
# Download Neurosynth
# ----------------------------------------------------------------------------
# Neurosynth's data files are stored at https://github.com/neurosynth/neurosynth-data.

from nimare import extract
import os
from nimare import io

# The overall procedure may take a while (~10 minutes)
# Download the desired version of Neurosynth from GitHub.
files = extract.fetch_neurosynth(
    data_dir='./neurosynth_data',
    version="7",
    source="abstract",
    vocab="terms",
    overwrite=True,
)
print(files)
neurosynth_db = files[0]

# Convert the files to a Dataset.
neurosynth_dset = io.convert_neurosynth_to_dataset(
    coordinates_file=neurosynth_db["coordinates"],
    metadata_file=neurosynth_db["metadata"],
    annotations_files=neurosynth_db["features"],
)

# Save the Dataset for later use.
neurosynth_dset.save(os.path.join('./neurosynth_data', "neurosynth_dataset.pkl.gz"))
```

```
INFO:nimare.extract.utils:Dataset created in ./neurosynth_data/neurosynth
INFO:nimare.extract.extract:Searching for any feature files matching the following criteria: [('source-abstract', 'vocab-terms', 'data-neurosynth', 'version-7')]
Downloading data-neurosynth_version-7_coordinates.tsv.gz
Downloading data-neurosynth_version-7_metadata.tsv.gz
Downloading data-neurosynth_version-7_vocab-terms_source-abstract_type-tfidf_features.npz
Downloading data-neurosynth_version-7_vocab-terms_vocabulary.txt
[{'coordinates': '/home/users/ysong30/Documents/Projects/Functional_Decoding_Tutorial/neurosynth_data/neurosynth/data-neurosynth_version-7_coordinates.tsv.gz', 'metadata': '/h
ome/users/ysong30/Documents/Projects/Functional_Decoding_Tutorial/neurosynth_data/neurosynth/data-neurosynth_version-7_metadata.tsv.gz', 'features': [{'features': '/home/user
s/ysong30/Documents/Projects/Functional_Decoding_Tutorial/neurosynth_data/neurosynth/data-neurosynth_version-7_vocab-terms_source-abstract_type-tfidf_features.npz', 'vocabular
y': '/home/users/ysong30/Documents/Projects/Functional_Decoding_Tutorial/neurosynth_data/neurosynth/data-neurosynth_version-7_vocab-terms_vocabulary.txt'}]}]
WARNING:nimare.utils:Not applying transforms to coordinates in unrecognized space 'UNKNOWN'
```

```python
# Checks how many studies are included in this dataset.
# All coordinates are represented in MNI152 space.
print(neurosynth_dset)
```

```
Dataset(14371 experiments, space='mni152_2mm')
```

```python
# Check which studies are included.
neurosynth_dset.metadata.head()
```

| | id | study_id | contrast_id | authors | journal | year | title |
|---|---|---|---|---|---|---|---|
| **0** | 10022492-1 | 10022492 | 1 | Callicott JH, Mattay VS, Bertolino A, Finn K, ... | Cerebral cortex (New York, N.Y. : 1991) | 1999 | Physiological characteristics of capacity cons... |
| **1** | 10022494-1 | 10022494 | 1 | Toni I, Schluter ND, Josephs O, Friston K, Pas... | Cerebral cortex (New York, N.Y. : 1991) | 1999 | Signal-, set- and movement-related activity in... |
| **2** | 10022496-1 | 10022496 | 1 | Lockwood AH, Salvi RJ, Coad ML, Arnold SA, Wac... | Cerebral cortex (New York, N.Y. : 1991) | 1999 | The functional anatomy of the normal human aud... |
| **3** | 10051677-1 | 10051677 | 1 | Denton D, Shade R, Zamarippa F, Egan G, Blair-... | Proceedings of the National Academy of Science... | 1999 | Correlation of regional cerebral blood flow an... |
| **4** | 10191322-1 | 10191322 | 1 | Chee MW, Tan EW, Thiel T | The Journal of neuroscience : the official jou... | 1999 | Mandarin and English single word processing st... |

⋮

```python
# The coordinates of activation peaks are extracted for each study.
neurosynth_dset.coordinates.head()
```

| | id | study_id | contrast_id | x | y | z | space |
|---|---|---|---|---|---|---|---|
| **1483** | 10022492-1 | 10022492 | 1 | 36.0 | -58.0 | 52.0 | mni152_2mm |
| **1499** | 10022492-1 | 10022492 | 1 | 48.0 | 24.0 | 20.0 | mni152_2mm |
| **1498** | 10022492-1 | 10022492 | 1 | -42.0 | 26.0 | 20.0 | mni152_2mm |
| **1497** | 10022492-1 | 10022492 | 1 | -36.0 | 30.0 | 16.0 | mni152_2mm |
| **1496** | 10022492-1 | 10022492 | 1 | -30.0 | 32.0 | 0.0 | mni152_2mm |

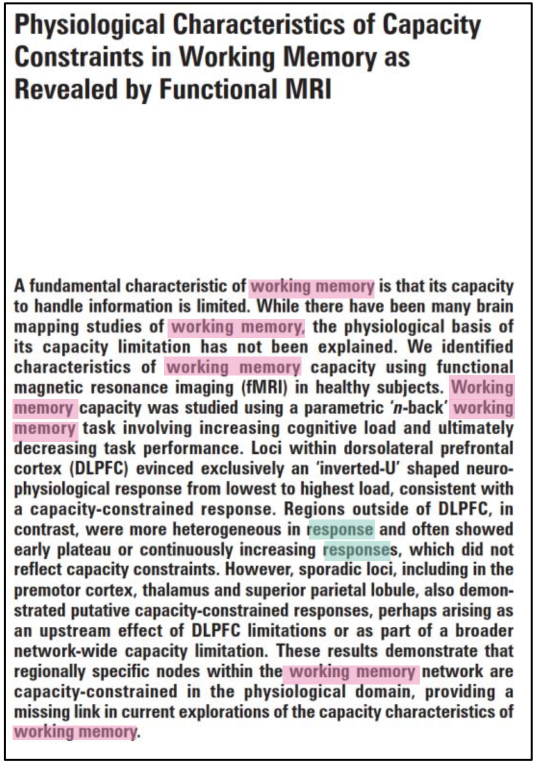
⋮

How can we know each study or each peak is related to which cognitive aspect?

# Article annotation

- The content of each article is represented as a **TF-IDF** vector, where each element corresponds to a cognitive term (or label).

- The raw text used to annotate the NeuroSynth database is the article **abstract**.



**Physiological Characteristics of Capacity Constraints in Working Memory as Revealed by Functional MRI**

A fundamental characteristic of working memory is that its capacity to handle information is limited. While there have been many brain mapping studies of working memory, the physiological basis of its capacity limitation has not been explained. We identified characteristics of working memory capacity using functional magnetic resonance imaging (fMRI) in healthy subjects. Working memory capacity was studied using a parametric '*n*-back' working memory task involving increasing cognitive load and ultimately decreasing task performance. Loci within dorsolateral prefrontal cortex (DLPFC) evinced exclusively an 'inverted-U' shaped neuro-physiological response from lowest to highest load, consistent with a capacity-constrained response. Regions outside of DLPFC, in contrast, were more heterogeneous in response and often showed early plateau or continuously increasing responses, which did not reflect capacity constraints. However, sporadic loci, including in the premotor cortex, thalamus and superior parietal lobule, also demonstrated putative capacity-constrained responses, perhaps arising as an upstream effect of DLPFC limitations or as part of a broader network-wide capacity limitation. These results demonstrate that regionally specific nodes within the working memory network are capacity-constrained in the physiological domain, providing a missing link in current explorations of the capacity characteristics of working memory.

TF-IDF vector

$$\begin{array}{cl} \left[\begin{array}{c} \vdots \\ 0 \\ 0 \\ 0.32 \\ 0 \\ 0.1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{array}\right] & \begin{array}{l} \\ \text{pain} \\ \text{visual} \\ \text{working memory} \\ \text{motor} \\ \text{response} \\ \text{social} \\ \text{ToM} \\ \text{reward} \\ \\ \end{array} \end{array}$$

**TF-IDF:** term frequency–inverse document frequency

✓ A measure of importance of a word to a document in a collection or corpus, adjusted for the fact that some words appear more frequently in general.

# Cognitive Terms

- You can check the all the terms are used to annotate the studies.

- A frequency vector of each term represents a single study.

```
# Cognitive Labels to describe studies
neurosynth_dset.annotations.columns.tolist()[100:110]

['terms_abstract_tfidf__actively',
 'terms_abstract_tfidf__activities',
 'terms_abstract_tfidf__acts',
 'terms_abstract_tfidf__actual',
 'terms_abstract_tfidf__actually',
 'terms_abstract_tfidf__acute',
 'terms_abstract_tfidf__ad',
 'terms_abstract_tfidf__adaptation',
 'terms_abstract_tfidf__adapted',
 'terms_abstract_tfidf__adaptive']
```

https://github.com/neurosynth/neurosynth-data

```python
# This dataset includes annotations for each study using TF-IDF (Term Frequency-Inverse Document Frequency).
neurosynth_dset.annotations.head()
```

| | id | study_id | contrast_id | terms_abstract_tfidf__001 | terms_abstract_tfidf__01 | terms_abstract_tfidf__05 | terms_abstract_tfidf__10 | terms_abstract_tfidf__100 | terms_abstract_tfidf__11 | terr |
|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 10022492-1 | 10022492 | 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| **1** | 10022494-1 | 10022494 | 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| **2** | 10022496-1 | 10022496 | 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| **3** | 10051677-1 | 10051677 | 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| **4** | 10191322-1 | 10191322 | 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |

5 rows × 3231 columns

3228 terms are used to annotate each study.

# TF-IDF

| terms_abstract_tfidf__words | terms_abstract_tfidf__work | terms_abstract_tfidf__working | terms_abstract_tfidf__working memory |
|---|---|---|---|
| 0.0 | 0.0 | 0.317592 | 0.320417 |
| 0.0 | 0.0 | 0.000000 | 0.000000 |
| 0.0 | 0.0 | 0.000000 | 0.000000 |
| 0.0 | 0.0 | 0.000000 | 0.000000 |
| 0.0 | 0.0 | 0.000000 | 0.000000 |

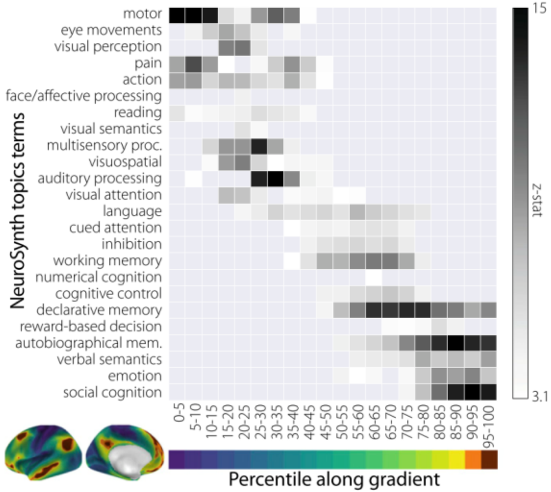
TF-IDF value for working memory

Check “Code02_Decoding_with_neurosynth_data.ipynb”

# STEP 2. Specify Terms of Interest

- More than 3,000 terms are used for annotation in the NeuroSynth dataset.
  - However, these include non-cognitive terms (e.g., brain regions or demographic information) and semantically duplicated terms (e.g., *short-term memory* vs. *working memory*).
- Refining these terms can improve both decoding efficiency and interpretability

**Only 24 topic terms were used in Margulias et al. (2016)'s work.**




Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., ... & Smallwood, J. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences*, *113*(44), 12574-12579.

```python
#### Specify the scope of cognitive terms ####
neuroquery_dset_copy = neuroquery_dset.copy()
neuroquery_dset_copy.annotations=neuroquery_dset_copy.annotations.iloc[:, [0,1,2,
                                                        71-1, 452-1, 478-1, 511-1, 1084-1, 1806-1, 1956-1, 1967-1, 2721-1, 2932-1, 3347-1,
                                                        3509-1, 3557-1, 3777-1, 3964-1, 4674-1, 4850-1, 5259-1, 6143-1,
                                                        6150-1, 6168-1, 6184-1, 6280-1]]

# display the selected cognitive terms
column_names = neuroquery_dset_copy.annotations.columns.tolist()
new_column_names = column_names[:3]
for i in range(3,len(column_names)):
    cleaned_term = column_names[i].split("__")[1]
    print(cleaned_term)
    new_column_names.append(cleaned_term)

# Replace column names
neuroquery_dset_copy.annotations.columns = new_column_names
```

```
action
attention
auditory
autobiographical
cognitive control
emotional
eye movement
face
inhibition
language
memory
motor
multisensory
numerical
pain
reading
reward
social cognition
visual
visual attention
visual perception
visuospatial
working memory
```

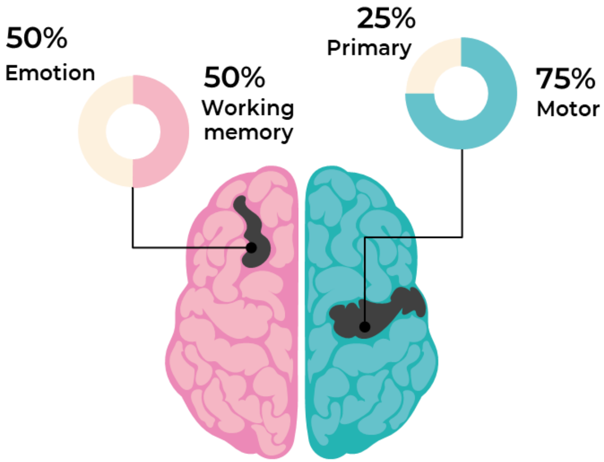
**You should select cognitive terms for decoding.**

**Here, for demonstration purposes, I used terms similar to those in Margulias et al. (2016).**

# STEP 3. Choose and Train a Decoder

- Several decoding algorithms are available in NiMARE:
    - NeuroSynth decoder
    - BrainMap decoder
    - ROI Association Decoder
    - Correlation Decoder

- These algorithms are based on different assumptions and inference approaches.



Since ROI decoders essentially perform Bayesian inference for functional decoding, there isn’t much to fit (i.e., train) although NiMARE uses the basic fit–transform.

```python
#### Training NeuroSynth Decoder ####

from nimare.decode import discrete

# Get studies with voxels in the mask
ids = neurosynth_dset.get_studies_by_mask(ICN_SM_thres_mni152)

# Train the decoder
decoder = discrete.NeurosynthDecoder(correction=None)
decoder.fit(neurosynth_dset_copy)
```

# STEP 4. Decode

- Perform Bayesian inference to quantify the degree of association between term use and voxel activation.
    - If probReverse is high, a given region activates more consistently in studies that mention the current term than in those that do not.

Ref: https://neurosynth.org/faq/#q18

```
#### Decode ####

decoded_df = decoder.transform(ids=ids)
decoded_df.sort_values(by="probReverse", ascending=False)
```

Major outputs for reports

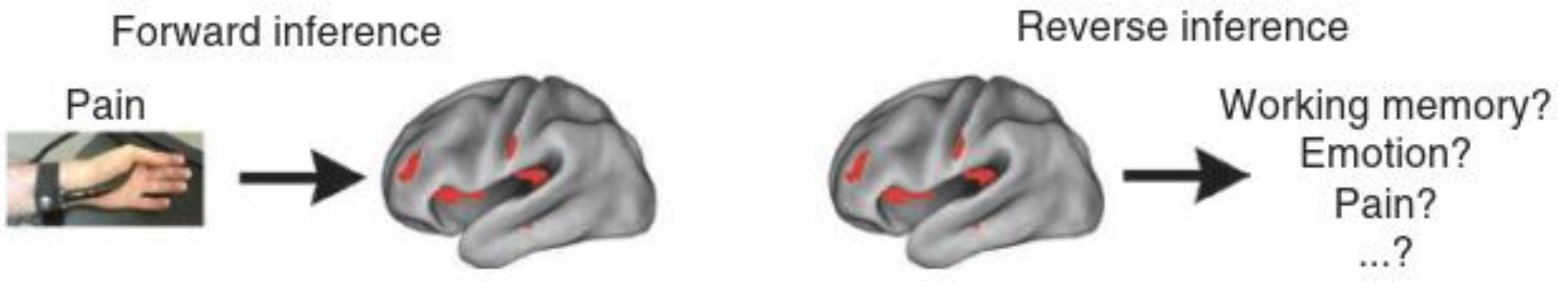
| Term | pForward | zForward | probForward | pReverse | zReverse | probReverse |
|---|---|---|---|---|---|---|
| motor | 0.000000e+00 | inf | 0.416853 | 5.714893e-180 | 28.605651 | 0.674786 |
| eye movements | NaN | NaN | 0.436507 | 1.255302e-07 | 5.285254 | 0.631976 |
| action | 5.912026e-14 | 7.510016 | 0.400493 | 1.467755e-27 | 10.877983 | 0.610139 |
| visual attention | NaN | NaN | 0.382826 | 3.987689e-03 | 2.879134 | 0.579309 |
| visuospatial | NaN | NaN | 0.372239 | 4.167178e-04 | 3.529264 | 0.568481 |
| attention | 3.370160e-116 | 22.913990 | 0.348409 | 6.875312e-09 | 5.793942 | 0.548643 |
| pain | 9.181554e-30 | -11.331315 | 0.355107 | 1.024070e-03 | 3.283829 | 0.548476 |
| working memory | 4.255896e-09 | 5.873927 | 0.348996 | 3.509021e-05 | 4.137636 | 0.543657 |
| inhibition | 1.708363e-15 | -7.960868 | 0.350032 | 2.608555e-03 | 3.010456 | 0.541904 |
| visual perception | NaN | NaN | 0.348176 | 3.253452e-01 | 0.983533 | 0.536358 |
| cognitive control | 2.971511e-19 | -8.969764 | 0.341351 | 4.210679e-02 | 2.032463 | 0.529080 |
| visual | 0.000000e+00 | inf | 0.330133 | 9.353038e-03 | 2.598873 | 0.518645 |
| numerical | NaN | NaN | 0.323850 | 9.742068e-01 | 0.032333 | 0.501098 |
| memory | 3.035747e-231 | 32.467821 | 0.321518 | 5.946423e-01 | -0.532121 | 0.495893 |
| multisensory | NaN | NaN | 0.319151 | 8.337580e-01 | -0.209884 | 0.493604 |
| social cognition | NaN | NaN | 0.313993 | 5.520877e-01 | -0.594635 | 0.484956 |
| auditory | 3.268939e-04 | 3.592999 | 0.314631 | 1.357842e-01 | -1.491676 | 0.483603 |
| verbal | 2.444843e-15 | -7.916401 | 0.312389 | 1.778519e-01 | -1.347399 | 0.480830 |
| emotion | 1.525824e-02 | -2.426191 | 0.302649 | 1.481385e-03 | -3.178306 | 0.460416 |
| reading | 2.043242e-60 | -16.395978 | 0.300651 | 2.009105e-02 | -2.324643 | 0.459656 |

What's the difference?

```
#### Decode ####

decoded_df = decoder.transform(ids=ids)
decoded_df.sort_values(by="probReverse", ascending=False)
```

| Term | pForward | zForward | probForward | pReverse | zReverse | probReverse |
|---|---|---|---|---|---|---|
| **motor** | 0.000000e+00 | inf | 0.416853 | 5.714893e-180 | 28.605651 | 0.674786 |
| **eye movements** | NaN | NaN | 0.436507 | 1.255302e-07 | 5.285254 | 0.631976 |
| **action** | 5.912026e-14 | 7.510016 | 0.400493 | 1.467755e-27 | 10.877983 | 0.610139 |
| **visual attention** | NaN | NaN | 0.382826 | 3.987689e-03 | 2.879134 | 0.579309 |

# Forward/Reverse Inference




**a.k.a. Uniformity test**

The degree to which each voxel is consistently activated in studies that use a given term.

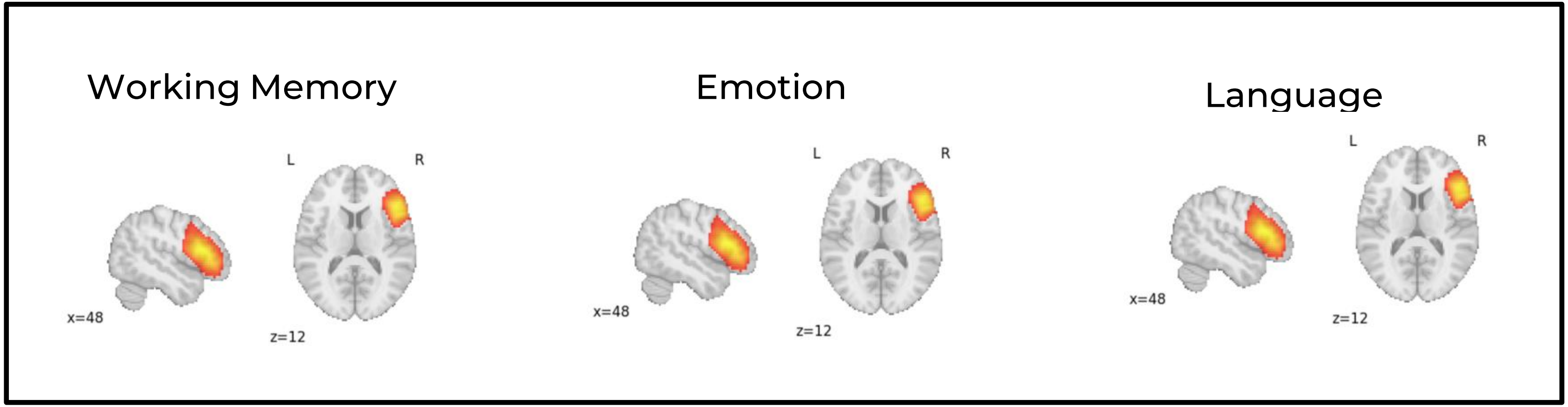
**a.k.a. Association test**

The degree to which a given region activates more consistently in studies that mention the current term than in those that do not

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nature methods, 8(8), 665-670.

# Forward/Reverse Inference for Emotion

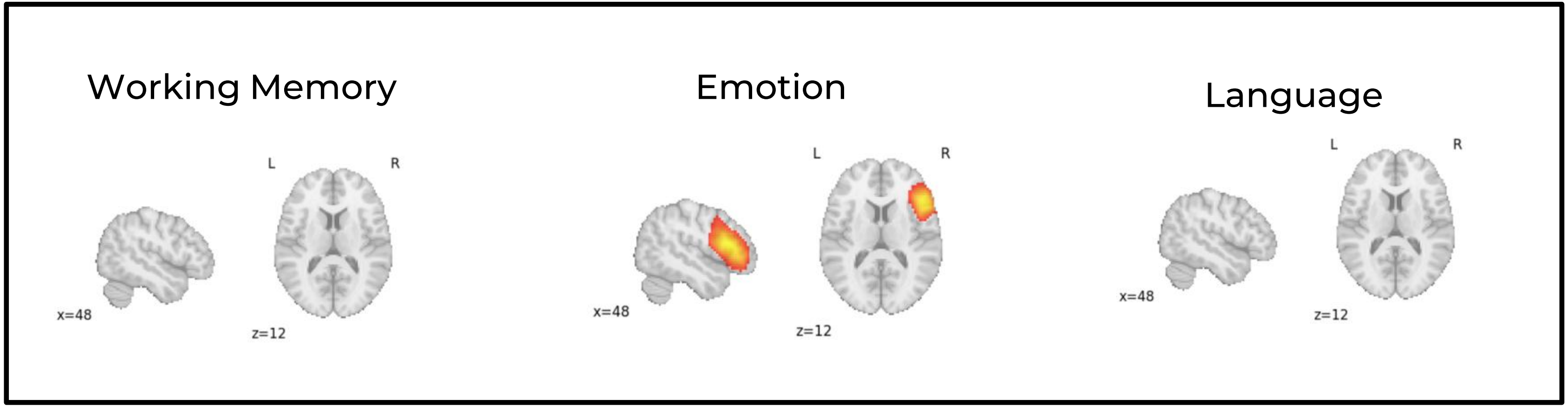
If a given brain area is all activated during working memory-, emotion-, and language-mentioned studies, ...




Forward probability ($p(\text{activaiton}|\text{term})$) of this area is high.
Reverse probability ($p(\text{term}|\text{activation})$) of this area is low.

# Forward/Reverse Inference for Emotion

If a given brain area is specifically activated during emotion-mentioned task, ...vv

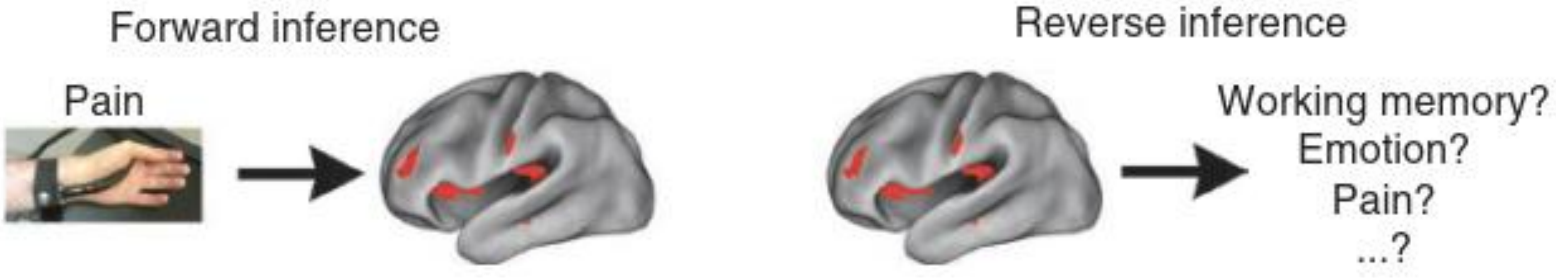



Forward probability ($p(\text{activaiton}|\text{term})$) of this area is high.
Reverse probability ($p(\text{term}|\text{activation})$) of this area is high.

Caution: The default prior for this inference is set primarily for practical purposes. Therefore, it should not be interpreted as a true probability but rather as a degree of association.

# Forward/Reverse Inference

I highly recommend you to understand the formula to get the nuiances in these two probability.




**a.k.a. Uniformity test**

**a.k.a. Association test**

$$p(A_j = 1 | T_k = 1) = \left[\sum_i A_{ij} T_{ik} + 1\right] / \left[\sum_i T_{ik} + 2\right]$$

$$p(T_k = 1 | A_j = 1) = p(A_j = 1 | T_k = 1) p(T_k = 1) / p(A_j = 1)$$

$$p(T_k = 1) = p(T_k = 0) = 0.5$$

$T_{ik}$: Whether term $k$ is present in study $i$ *(0 or 1).*
$A_{ij}$*:* Whether voxel $j$ is activated in study $i$ *(0 or 1)*
$A_j$*:* Whether voxel $j$ is activated *(0 or 1).*
$T_k$: Whether term $k$ is present *(0 or 1).*

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nature methods, 8(8), 665-670.

# STEP 4. Decode

- NeuroSynth decoder accesses statistical significance by performing a **chi-square test** to determine if presence of the label and selection are independent.

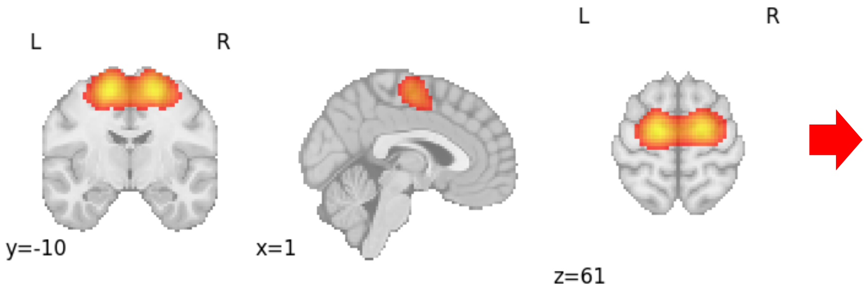
Ref:
https://nimare.readthedocs.io/en/0.0.11/decoding.html#the-neurosynth-roi-association-approach

Statistical significance

```
#### Decode ####

decoded_df = decoder.transform(ids=ids)
decoded_df.sort_values(by="probReverse", ascending=False)
```

| Term | pForward | zForward | probForward | pReverse | zReverse | probReverse |
|---|---|---|---|---|---|---|
| **motor** | 0.000000e+00 | inf | 0.416853 | 5.714893e-180 | 28.605651 | 0.674786 |
| **eye movements** | NaN | NaN | 0.436507 | 1.255302e-07 | 5.285254 | 0.631976 |
| **action** | 5.912026e-14 | 7.510016 | 0.400493 | 1.467755e-27 | 10.877983 | 0.610139 |
| **visual attention** | NaN | NaN | 0.382826 | 3.987689e-03 | 2.879134 | 0.579309 |
| **visuospatial** | NaN | NaN | 0.372239 | 4.167178e-04 | 3.529264 | 0.568481 |
| **attention** | 3.370160e-116 | 22.913990 | 0.348409 | 6.875312e-09 | 5.793942 | 0.548643 |
| **pain** | 9.181554e-30 | -11.331315 | 0.355107 | 1.024070e-03 | 3.283829 | 0.548476 |
| **working memory** | 4.255896e-09 | 5.873927 | 0.348996 | 3.509021e-05 | 4.137636 | 0.543657 |
| **inhibition** | 1.708363e-15 | -7.960868 | 0.350032 | 2.608555e-03 | 3.010456 | 0.541904 |
| **visual perception** | NaN | NaN | 0.348176 | 3.253452e-01 | 0.983533 | 0.536358 |
| **cognitive control** | 2.971511e-19 | -8.969764 | 0.341351 | 4.210679e-02 | 2.032463 | 0.529080 |
| **visual** | 0.000000e+00 | inf | 0.330133 | 9.353038e-03 | 2.598873 | 0.518645 |
| **numerical** | NaN | NaN | 0.323850 | 9.742068e-01 | 0.032333 | 0.501098 |
| **memory** | 3.035747e-231 | 32.467821 | 0.321518 | 5.946423e-01 | -0.532121 | 0.495893 |
| **multisensory** | NaN | NaN | 0.319151 | 8.337580e-01 | -0.209884 | 0.493604 |
| **social cognition** | NaN | NaN | 0.313993 | 5.520877e-01 | -0.594635 | 0.484956 |
| **auditory** | 3.268939e-04 | 3.592999 | 0.314631 | 1.357842e-01 | -1.491676 | 0.483603 |
| **verbal** | 2.444843e-15 | -7.916401 | 0.312389 | 1.778519e-01 | -1.347399 | 0.480830 |
| **emotion** | 1.525824e-02 | -2.426191 | 0.302649 | 1.481385e-03 | -3.178306 | 0.460416 |
| **reading** | 2.043242e-60 | -16.395978 | 0.300651 | 2.009105e-02 | -2.324643 | 0.459656 |

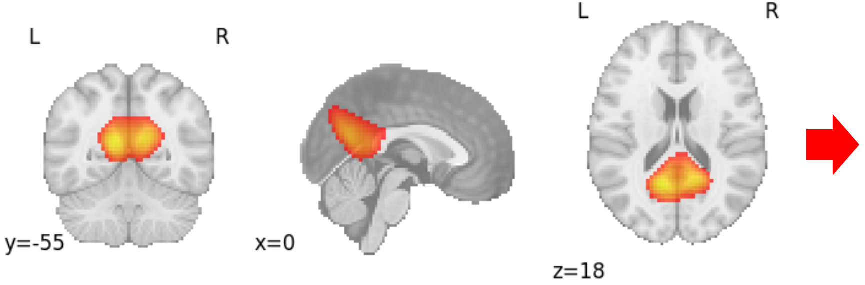
# Decoding Results – sensorimotor




| Term | pForward | zForward | probForward | pReverse | zReverse | probReverse |
|---|---|---|---|---|---|---|
| **motor** | 0.000000e+00 | inf | 0.416853 | 5.714893e-180 | 28.605651 | 0.674786 |
| **eye movements** | NaN | NaN | 0.436507 | 1.255302e-07 | 5.285254 | 0.631976 |
| **action** | 5.912026e-14 | 7.510016 | 0.400493 | 1.467755e-27 | 10.877983 | 0.610139 |
| **visual attention** | NaN | NaN | 0.382826 | 3.987689e-03 | 2.879134 | 0.579309 |
| **visuospatial** | NaN | NaN | 0.372239 | 4.167178e-04 | 3.529264 | 0.568481 |
| **attention** | 3.370160e-116 | 22.913990 | 0.348409 | 6.875312e-09 | 5.793942 | 0.548643 |
| **pain** | 9.181554e-30 | -11.331315 | 0.355107 | 1.024070e-03 | 3.283829 | 0.548476 |
| **working memory** | 4.255896e-09 | 5.873927 | 0.348996 | 3.509021e-05 | 4.137636 | 0.543657 |
| **inhibition** | 1.708363e-15 | -7.960868 | 0.350032 | 2.608555e-03 | 3.010456 | 0.541904 |
| **visual perception** | NaN | NaN | 0.348176 | 3.253452e-01 | 0.983533 | 0.536358 |
| **cognitive control** | 2.971511e-19 | -8.969764 | 0.341351 | 4.210679e-02 | 2.032463 | 0.529080 |
| **visual** | 0.000000e+00 | inf | 0.330133 | 9.353038e-03 | 2.598873 | 0.518645 |
| **numerical** | NaN | NaN | 0.323850 | 9.742068e-01 | 0.032333 | 0.501098 |

Only displaying terms with probReverse > 0.5

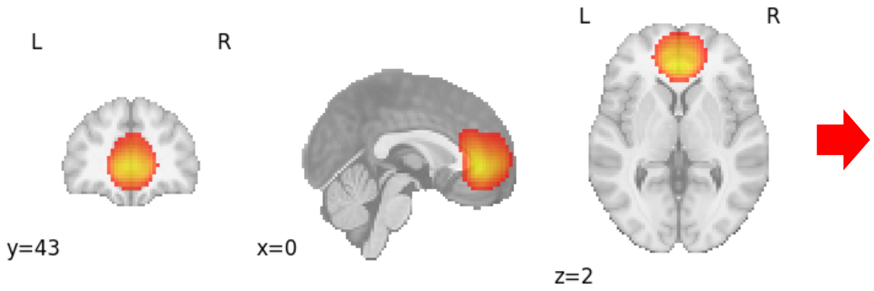
# Decoding Results – DMN (Precuneus)




| Term | pForward | zForward | probForward | pReverse | zReverse | probReverse |
|---|---|---|---|---|---|---|
| **autobiographical memory** | NaN | NaN | 0.542011 | 6.442063e-20 | 9.136644 | 0.713832 |
| **social cognition** | NaN | NaN | 0.378158 | 1.934541e-05 | 4.272315 | 0.588979 |
| **memory** | 0.000000e+00 | inf | 0.346241 | 6.281016e-28 | 10.955098 | 0.577842 |
| **cognitive control** | 1.091679e-11 | -6.793866 | 0.321618 | 3.273422e-01 | 0.979481 | 0.514747 |
| **emotion** | 2.689682e-05 | 4.198265 | 0.320259 | 2.513005e-01 | 1.147196 | 0.513387 |
| **attention** | 8.750443e-93 | 20.431645 | 0.318295 | 2.138231e-01 | 1.243121 | 0.511326 |
| **working memory** | 8.573893e-07 | 4.921829 | 0.317382 | 4.714466e-01 | 0.720127 | 0.508284 |
| **reward** | 4.277601e-01 | 0.793030 | 0.314683 | 7.979939e-01 | 0.255944 | 0.503210 |

Only displaying terms with probReverse > 0.5

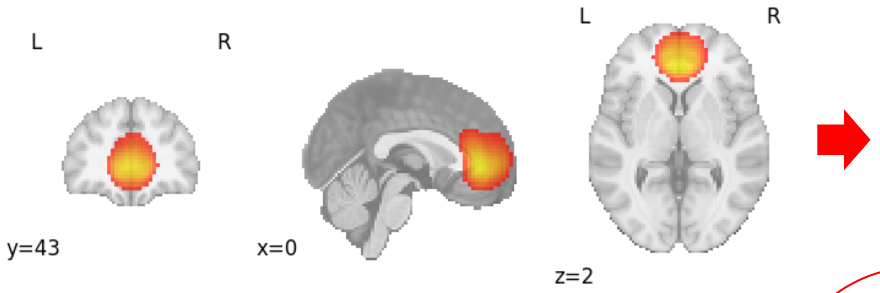
# Decoding Results – DMN (MPFC)




| Term | pForward | zForward | probForward | pReverse | zReverse | probReverse |
|---|---|---|---|---|---|---|
| **autobiographical memory** | NaN | NaN | 0.535312 | 9.027011e-12 | 6.821221 | 0.667168 |
| **social cognition** | NaN | NaN | 0.472863 | 4.280846e-13 | 7.246372 | 0.624821 |
| **reward** | 8.437706e-42 | 13.545379 | 0.443557 | 5.480242e-33 | 11.964079 | 0.610085 |
| **emotion** | 2.670773e-41 | 13.460507 | 0.416818 | 1.069867e-18 | 8.827557 | 0.581853 |
| **pain** | 5.044729e-10 | -6.217707 | 0.406055 | 1.815443e-06 | 4.772953 | 0.563180 |
| **inhibition** | 9.747957e-04 | -3.297702 | 0.397881 | 1.641752e-05 | 4.308756 | 0.554101 |
| **cognitive control** | 1.803367e-05 | -4.287941 | 0.389350 | 7.586440e-04 | 3.367459 | 0.543312 |
| **memory** | 0.000000e+00 | inf | 0.364064 | 7.328978e-02 | 1.791022 | 0.512517 |
| **attention** | 1.196019e-107 | 22.039824 | 0.358597 | 9.702710e-01 | 0.037268 | 0.500312 |

Only displaying terms with probReverse > 0.5

# Decoding Results – DMN (MPFC)




| Term | pForward | zForward | probForward | pReverse | zReverse | probReverse |
|---|---|---|---|---|---|---|
| **autobiographical memory** | NaN | NaN | 0.535312 | 9.027011e-12 | 6.821221 | 0.667168 |
| **social cognition** | NaN | NaN | 0.472863 | 4.280846e-13 | 7.246372 | 0.624821 |
| **reward** | 8.437706e-42 | 13.545379 | 0.443557 | 5.480242e-33 | 11.964079 | 0.610085 |
| **emotion** | 2.670773e-41 | 13.460507 | 0.416818 | 1.069867e-18 | 8.827557 | 0.581853 |
| **pain** | 5.044729e-10 | -6.217707 | 0.406055 | 1.815443e-06 | 4.772953 | 0.563180 |
| **inhibition** | 9.747957e-04 | -3.297702 | 0.397881 | 1.641752e-05 | 4.308756 | 0.554101 |
| **cognitive control** | 1.803367e-05 | -4.287941 | 0.389350 | 7.586440e-04 | 3.367459 | 0.543312 |
| **memory** | 0.000000e+00 | inf | 0.364064 | 7.328978e-02 | 1.791022 | 0.512517 |
| **attention** | 1.196019e-107 | 22.039824 | 0.358597 | 9.702710e-01 | 0.037268 | 0.500312 |

Compared to the Precuneus, Reward/Emotion shows a higher rank and statistical significance in the MPFC.

Only displaying terms with probReverse > 0.5

# 03
# Alternative Options and Key Details

# Alternative Option – NeuroQuery Dataset

- **NeuroQuery** is a tool for meta-analysis of neuroimaging studies.

- Unlike **NeuroSynth**, **NeuroQuery** is focused on producing a brain map that ***predicts*** where in the brain a study on the topic of interest is likely report observations.

https://neuroquery.org/

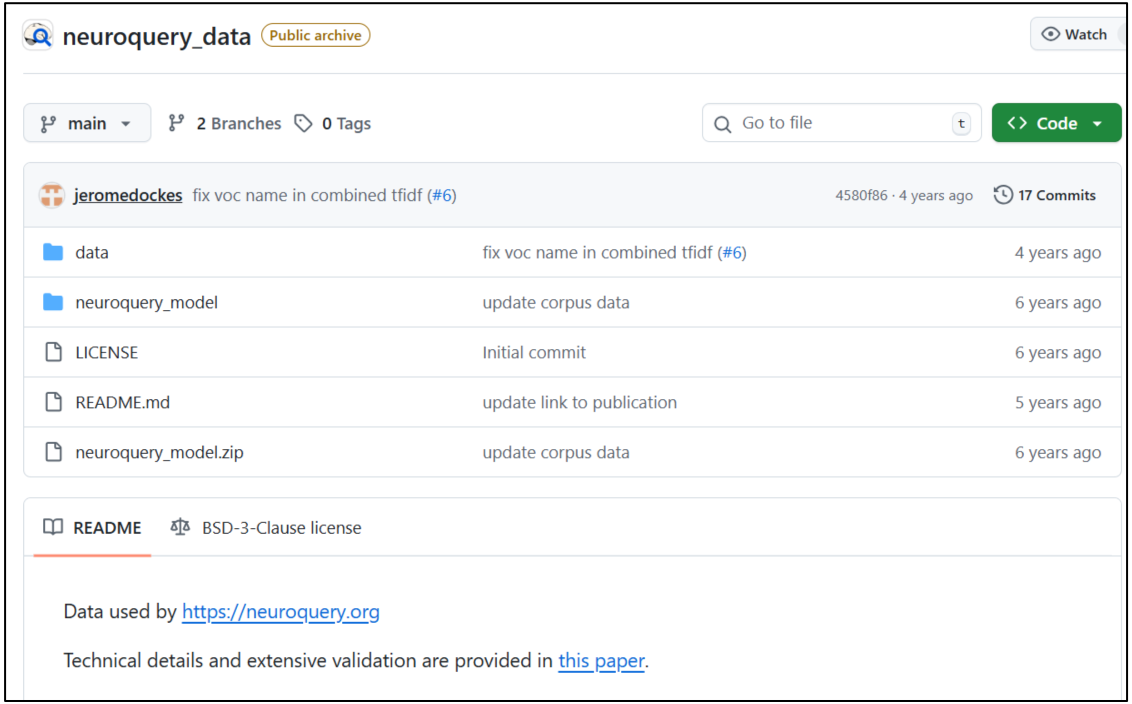
# Alternative Option – NeuroQuery Dataset

- The dataset used in NeuroQuery includes more diverse cognitive terms for annotation and covers a wider range of journals.

| | NeuroSynth | NeuroQuery |
|---|---|---|
| Dataset size | | |
| articles | 14 371 | 13 459 |
| terms | 3 228 (1 335 online) | 7 547 |
| journals | 60 | 458 |
| raw text length (words) | ≈4 M | ≈75 M |
| unique term occurrences | 1 063 670 | 5 855 483 |
| unique term occurrences in voc intersection | 677 345 | 3 089 040 |
| coordinates | 448 255 | 418 772 |
| Coordinate extraction errors on conflicting articles | | |
| articles with false positives / 40 | 20 | 3 |
| articles with false negatives / 40 | 28 | 8 |

Dockès, J., Poldrack, R. A., Primet, R., Gözükan, H., Yarkoni, T., Suchanek, F., ... & Varoquaux, G. (2020). NeuroQuery, comprehensive meta-analysis of human brain mapping. elife, 9, e53385.

# They provide the dataset through GitHub.



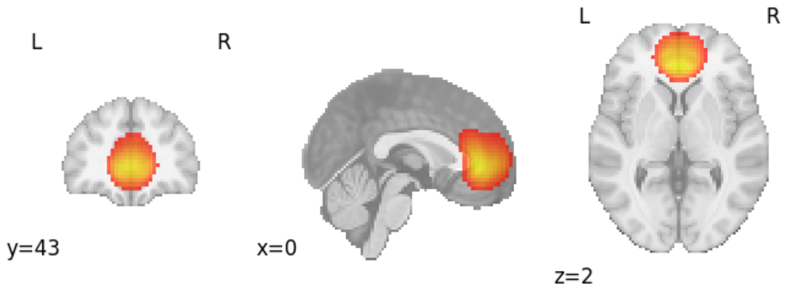

**We can use this for decoding!**

https://github.com/neuroquery/neuroquery_data/

Check “Code03_Decoding_with_neuroquery_data.ipynb”

# Decoding Results – DMN (MPFC)

**NeuroSynth Decoder trained with the NeuroQuery Dataset**




| Term | pForward | zForward | probForward | pReverse | zReverse | probReverse |
|---|---|---|---|---|---|---|
| **autobiographical** | NaN | NaN | 0.499960 | 4.829034e-47 | 14.404773 | 0.625802 |
| **emotional** | 0.000000e+00 | inf | 0.400049 | 1.192973e-92 | 20.416506 | 0.608044 |
| **reward** | 9.894976e-01 | -0.013163 | 0.435347 | 2.677134e-51 | 15.066814 | 0.589664 |
| **social cognition** | NaN | NaN | 0.447569 | 3.299004e-21 | 9.452795 | 0.578420 |
| **pain** | 0.000000e+00 | -inf | 0.421356 | 6.375812e-11 | 6.534651 | 0.549733 |
| **cognitive control** | 7.322371e-12 | -6.851218 | 0.412854 | 3.366900e-14 | 7.583371 | 0.547715 |
| **inhibition** | 2.706134e-11 | 6.661735 | 0.398485 | 6.474637e-06 | 4.510266 | 0.526218 |
| **attention** | 0.000000e+00 | inf | 0.384278 | 5.011440e-05 | 4.055093 | 0.523276 |
| **memory** | 0.000000e+00 | inf | 0.387311 | 5.985336e-03 | 2.748584 | 0.515232 |
| **face** | 2.016254e-16 | 8.221112 | 0.391470 | 2.858343e-01 | 1.067305 | 0.506159 |
| **action** | 2.413422e-14 | 7.626431 | 0.389724 | 9.894885e-01 | 0.013175 | 0.500076 |

**Results should be similar.**

Only displaying terms with probReverse > 0.5

# Alternative Option - BrainMap Decoder

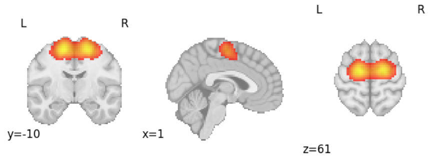
- BrainMap decoder accounts for the number of activation foci per study.


https://www.brainmap.org/
https://nimare.readthedocs.io/en/0.0.11/decoding.html#the-brainmap-approach

```python
#### BrainMap Decoder ####

from nimare.decode import discrete

# Get studies with voxels in the mask
ids = neurosynth_dset.get_studies_by_mask(ICN_SM_thres_mni152)

# Decode
decoder = discrete.BrainMapDecoder(correction="fdr_bh")
decoder.fit(neurosynth_dset_copy)
decoded_df = decoder.transform(ids=ids)
decoded_df.sort_values(by="probReverse", ascending=False)
```

```
/home/users/ysong30/anaconda3/envs/fmri/lib/python3.11/site-packages/numpy/_core/fromnumeric.py:84: Futur
in a future version this will reduce over both axes and return a scalar. To retain the old behavior, pass
  return reduction(axis=axis, out=out, **passkwargs)
```

| Term | pForward | zForward | likelihoodForward | pReverse | zReverse | probReverse |
|---|---|---|---|---|---|---|
| **motor** | 0.000000e+00 | inf | 1.441339 | 5.714893e-180 | 28.605651 | 0.163733 |
| **visual** | 4.085129e-01 | 8.265135e-01 | 1.006461 | 9.353038e-03 | 2.598873 | 0.138624 |
| **memory** | 9.691869e-01 | 3.862815e-02 | 0.938963 | 5.946423e-01 | -0.532121 | 0.114108 |
| **attention** | 1.838037e-05 | 4.283708e+00 | 1.171541 | 6.875312e-09 | 5.793942 | 0.095001 |
| **working memory** | 5.194925e-03 | 2.794692e+00 | 1.134622 | 3.509021e-05 | 4.137636 | 0.054822 |
| **action** | 2.170422e-10 | 6.348771e+00 | 1.366431 | 1.467755e-27 | 10.877983 | 0.053617 |
| **auditory** | 9.854723e-01 | 1.820873e-02 | 0.894990 | 1.357842e-01 | -1.491676 | 0.049625 |
| **emotion** | 9.798428e-01 | -2.526605e-02 | 0.887025 | 1.481385e-03 | -3.178306 | 0.040738 |
| **language** | 9.999966e-01 | -4.204964e-06 | 0.780283 | 6.593925e-04 | -3.405929 | 0.038047 |
| **reward** | 9.327199e-01 | -8.442331e-02 | 0.909637 | 1.069239e-03 | -3.271645 | 0.037143 |
| **inhibition** | 3.175586e-04 | -3.600539e+00 | 1.259354 | 2.608555e-03 | 3.010456 | 0.033520 |
| **cognitive control** | 3.352167e-03 | -2.933451e+00 | 1.204387 | 4.210679e-02 | 2.032463 | 0.031897 |




Check “Code02_Decoding_with_neurosynth_data.ipynb”

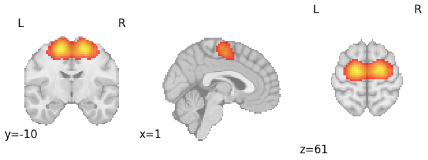
# Alternative Option – ROI Association Decoder

- The ROI association decoding method computes the correlation between averaged modeled activation within a target ROI and term weights for all labels, yielding one correlation coefficient per label

https://nimare.readthedocs.io/en/0.0.11/decoding.html#the-neurosynth-roi-association-approach

```python
#### ROI Assoicaiton Decoder (used in Margulias et al. 2016) ####

from nimare.decode import discrete

decoder = discrete.ROIAssociationDecoder(ICN_SM_thres_mni152)
decoder.fit(neurosynth_dset_copy)

# The `transform` method doesn't take any parameters.
decoded_df = decoder.transform()
decoded_df.sort_values(by="r", ascending=False)
```

| feature | r |
|---|---|
| motor | 0.300255 |
| action | 0.089180 |
| eye movements | 0.072698 |
| attention | 0.034519 |
| working memory | 0.029189 |
| visuospatial | 0.026658 |
| inhibition | 0.022357 |
| pain | 0.011743 |
| visual attention | 0.006437 |
| visual | 0.003127 |
| cognitive control | -0.001027 |
| visual perception | -0.002997 |
| numerical | -0.003374 |
| reading | -0.007402 |
| social cognition | -0.009577 |
| multisensory | -0.010492 |
| verbal | -0.013694 |




Check "Code02_Decoding_with_neurosynth_data.ipynb"

# Useful references...


- NeuroSynth paper: Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nature methods, 8(8), 665-670.

- NiMARE paper: https://preprint.neurolibre.org/10.55458/neurolibre.00007/00_abstract.html

- NiMARE docs: https://nimare.readthedocs.io/en/stable/index.html

# Enjoy exploring it!