# Relating Brain Structures To Open-Ended Descriptions Of Cognition

Jérôme Dockès, Olivier Grisel, Joan Massich, Fabian Suchanek, Bertrand Thirion, Gaël Varoquaux

## HAL Id: hal-01699045
## https://imt.hal.science/hal-01699045

Submitted on 1 Mar 2018

# Relating Brain Structures To Open-Ended Descriptions Of Cognition

**Jérôme Dockès (jerome.dockes@inria.fr)**
INRIA, 1 Rue Honoré d'Estienne d'Orves, 91120 Palaiseau, France

**Olivier Grisel (olivier.grisel@inria.fr)**
INRIA, 1 Rue Honoré d'Estienne d'Orves, 91120 Palaiseau, France

**Joan Massich (mailsik@gmail.com)**
INRIA, 1 Rue Honoré d'Estienne d'Orves, 91120 Palaiseau, France

**Fabian Suchanek (fabian@suchanek.name)**
Télécom ParisTech, 46 Rue Barrault, 75013 Paris, France

**Bertrand Thirion (bertrand.thirion@inria.fr)**
INRIA, 1 Rue Honoré d'Estienne d'Orves, 91120 Palaiseau, France

**Gaël Varoquaux (gael.varoquaux@inria.fr)**
INRIA, 1 Rue Honoré d'Estienne d'Orves, 91120 Palaiseau, France

## Abstract

Finding correspondences between mental processes and brain structures is a central goal in cognitive neuroscience. Neuroimaging provides brain-activity maps associated with cognitive task, however each study explores only a handful of mental processes. Here we use literature mining to bridge results across studies, establishing a bidirectional mapping between brain and mind. Our goal is to work on an open-ended set of terms describing cognitive processes. Moreover we introduce a validation framework using information retrieval metrics to ensure the accuracy of such correspondences, with a clear focus on relative frequencies, ignored in previous studies, to capture the relative importance of cognitive concepts. We show that this approach enables open-ended encoding and decoding.

**Keywords:** brain mapping; text mining; cognition;

## Introduction: The quest of brain-cognition mapping

An essential question for cognitive neuroscience is mapping cognitive functions to the brain territories that implement them. The literature describing associations between mind and brain activity is so extensive that looking for this mapping calls for large-scale, automated meta-analyses. The advent of the NeuroSynth database (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011) has been key in that respect: it automates extraction of cognitive terms and locations from neuroscience articles, opening the possibility of system-level studies of brain functional organization. Neurosynth can be used in *encoding* settings, mapping the probability of activations in the brain associated with a given cognitive term (Naselaris, Kay, Nishimoto, & Gallant, 2011), or in *decoding* settings, mapping brain locations in which activity implies a given behavior (POLDRACK, 2006). Recently, (Rubin et al., 2016) improved open-ended mappings with a framework that allows seeding such models with arbitrary priors.

Here, we build an open-ended engine for decoding and encoding from text mining on a very large body of literature. The main contribution of this work is (i) to capture rich descriptions of publications that *weight* their various cognitive concepts. (ii) formalize the evaluation of open-ended decoding using concepts from information retrieval. We perform a quantitative assessment of the power of encoding and decoding models in open-ended settings. On thousands of publications, we demonstrate for the first time that both encoding and decoding accuracies, from and to textual description, are far beyond chance. Finally, we show that these brain-mind associations also capture a meaningful structure across concepts of cognitive science.

## Experiments

Figure 1 outlines the idea behind the experiments reported in this work. Each article contains activation coordinates and free text. The activation coordinates can be transformed into a brain volume, and therefore can be represented by a point in the vector space of brain voxels. The free text is represented in terms of the Cognitive Atlas (http://www.cognitiveatlas.org) ontology using Term Frequency - Inverse Document Frequency (TFIDF) features, which represents the text as a point in the ontology span. Our goal is to assess the link between these two feature spaces. To do so, we take one representation as features and try to predict the other using Ridge Regression. We call predicting the activations from the text *encoding*, and predicting the TFIDF from the activations *decoding*. For both experiments, we compute scores across 10-fold cross-validation, and we perform permutation tests to compare our results to chance.

**Metrics** In the encoding setting, we want to be able to predict which part of the brain will be most active, rather than the shape of the distribution of activations. Ranking metrics are therefore adapted and we report Spearman correlations.
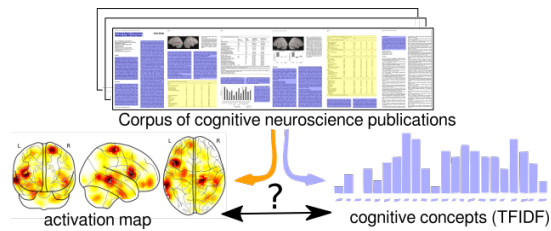
Figure 1: Quantitative analysis workflow for the discovery of brain/mind statistical associations.

| k | 5 | 10 | 15 | 20 |
|---|---|---|---|---|
| weighted recall @ k | 0.283 | 0.381 | 0.443 | 0.493 |

Table 1: Weighted Recall at k obtained in decoding cross-validation.

The decoding task can be seen as a tagging problem: given a query image, which terms are most relevant to describe it? The labels that we have are not a discrete set of relevant tags, but TFIDF, and we use the standard ranking metrics for this setting. Normalized Discounted Cumulative Gain (NDCG) is used to compare the ranking induced by a prediction to the frequencies contained in a ground truth, and Weighted Recall at k (WRec@k) measures what proportion of the mass of the ground truth is covered by the first $k$ terms of the prediction (Suchanek, Vojnovic, & Gunawardena, 2008).

## Results

**Encoding** The explained variance is $0.065$, and the mean Spearman correlation is $0.395$. This is significantly above chance: for 1000 permutations, the mean Spearman correlation is 0.35, and the variance across permutations is $1.4e{-7}$. An example of encoding result is given in Figure 2, and shows for instance that the language system is easily captured in an article related to reading.

The coefficients learnt by the Ridge Regression provide embedding of the articles' text in brain space (see Figure 3).

**Decoding** The mean NDCG is $0.503$. The values of WRec@k for $k = 5, 10, 15, 20$ are reported on Table 1. These scores are significantly above chance: for 700 permutations, the mean NDCG is $0.46$, and the variance is $1.7e{-7}$.

The weight maps of the Ridge Regression for each term provide reverse-inference maps (see Figure 3). They reliably map known brain structures associated with the corresponding term, such as the hypocampus for memory.

For the sake of comparison, we also performed a 10-fold cross-validation of the Neurosynth model on our data: the mean NDCG is 0.19. Neurosynth's predictions are based only on similarities of the query image with meta-analytic maps. This approach does not take any prior belief into account. The rankings predicted by the Neurosynth decoder are well below chance level, because it ignores the mean trends in the literature. This simply means that this model was designed to
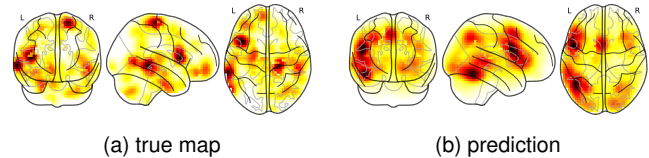


(a) true map          (b) prediction

Figure 2: True map and prediction for article "The neural basis of hyperlexic reading: an FMRI case study" . the lateralized language network is well recovered, while the appearent motor response was not anticipated by the semantic model.
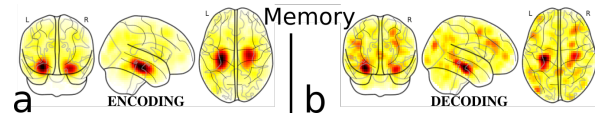


Figure 3: **Encoding and decoding results** for "memory". **a. encoding map**: map predicted by the model for "memory". **b. decoding weight map**: hyperplane of the ridge associated with "memory"

provide a way to quantitatively compare a query with a set of canonical maps, and not for tagging or open-ended decoding.

## Conclusion

Building upon the seminal work of Neurosynth (Yarkoni et al., 2011), we have used text mining to model the statistical link between brain activations and free-text descriptions of cognitive processes. We have proposed a quantitative validation framework for such open-ended mappings, which was missing from previous publications; with this we have shown that our encoding and decoding models perform significantly better than chance.

## References

Naselaris, T., Kay, K. N., Nishimoto, S., & Gallant, J. L. (2011). Encoding and decoding in fmri. *Neuroimage*, *56*(2), 400–410.

POLDRACK, R. (2006, Feb). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*(2), 59–63. doi: 10.1016/j.tics.2005.12.004

Rubin, T. N., Koyejo, O., Gorgolewski, K. J., Jones, M. N., Poldrack, R. A., & Yarkoni, T. (2016). Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition. *bioRxiv*, 059618.

Suchanek, F. M., Vojnovic, M., & Gunawardena, D. (2008). Social tags: meaning and suggestions. In *Proceedings of the 17th acm conference on information and knowledge management* (pp. 223–232).

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.