A M/EEG-fMRI fusion primer:

Resolving Human Brain Responses in Space and Time

Aude Oliva

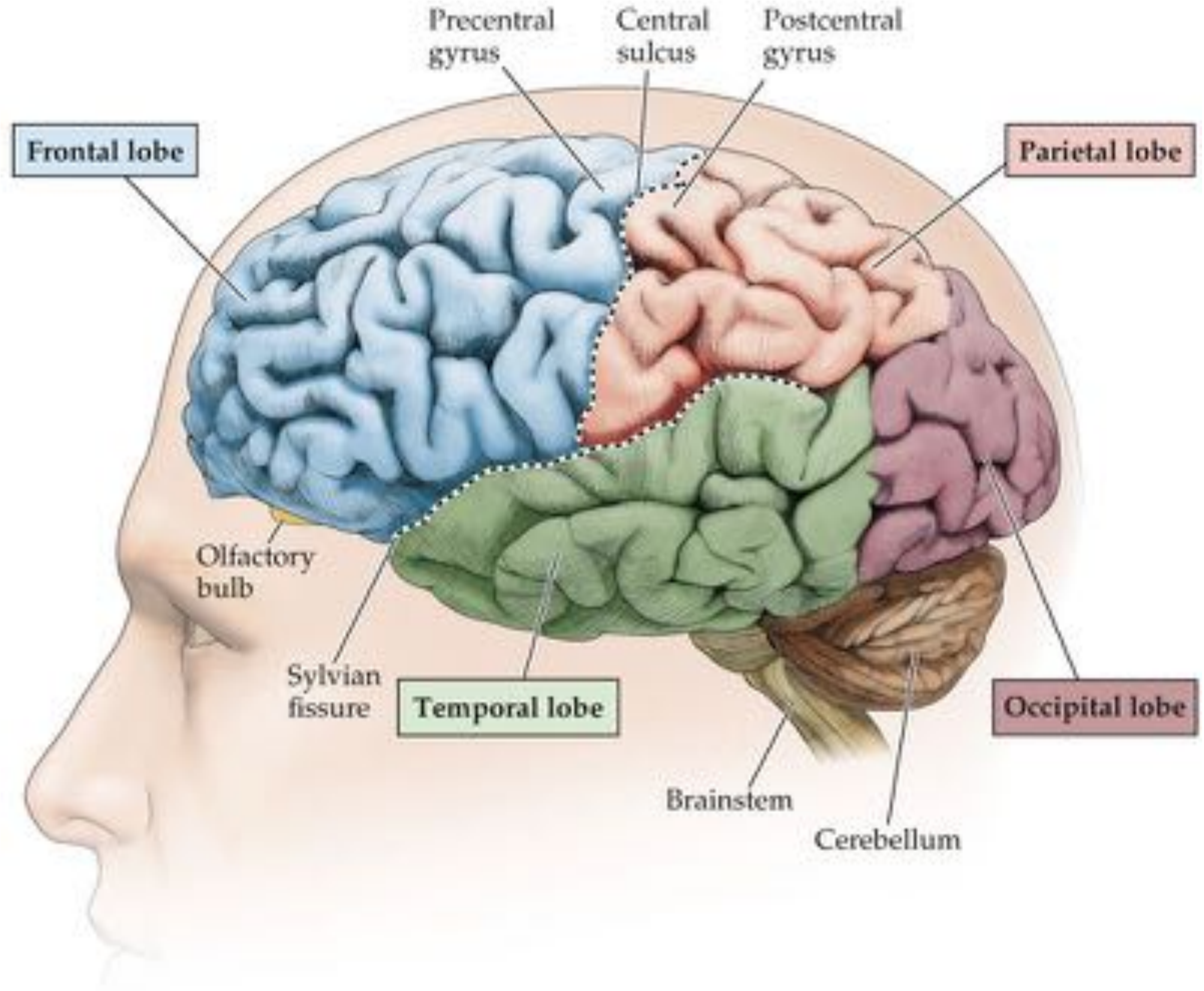
CSAIL, MIT-IBM Watson AI Lab, MIT Quest for Intelligence

MIT Schwarzman College of Computing

Massachusetts Institute of Technology

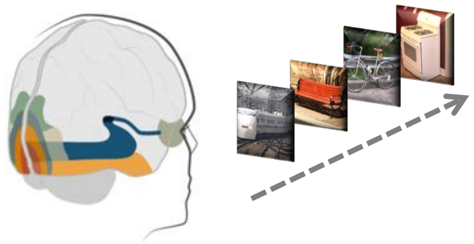
Brain: A **Dynamic** Deep Learning Network

- 100 billion neurons
  (80 billion in cerebellum)
- 100 trillion connections
  (Each neuron connects to 5000 to 200,000 other neurons)
- 10,000 different types of neurons
- 1000 > new neurons per day, all your life
- Consumes 20 watts

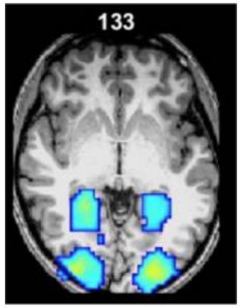


THE MIND'S MACHINE 2e, Figure 2.10

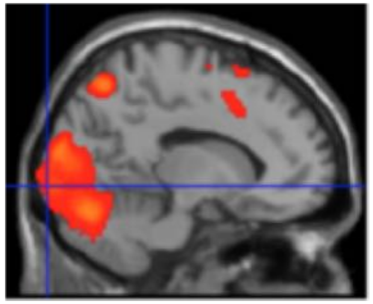
Spatio-temporal maps of human brain processes



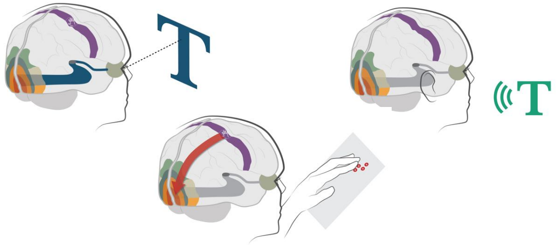
Basic sensory processing



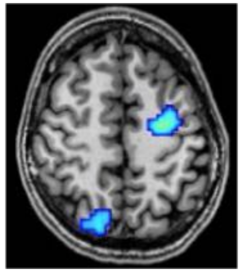
Recognition & Memory



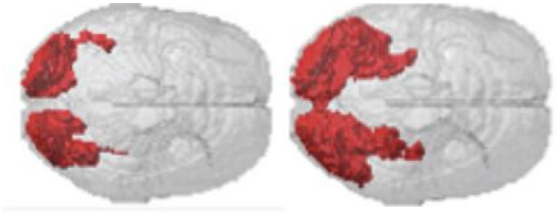
Learning & Plasticity



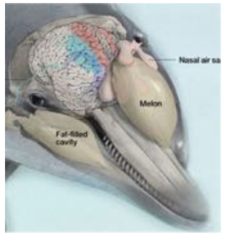
Multi-sensory integration

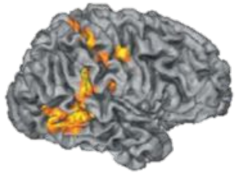


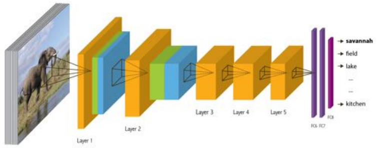
Working memory,
Attention



Clinical research



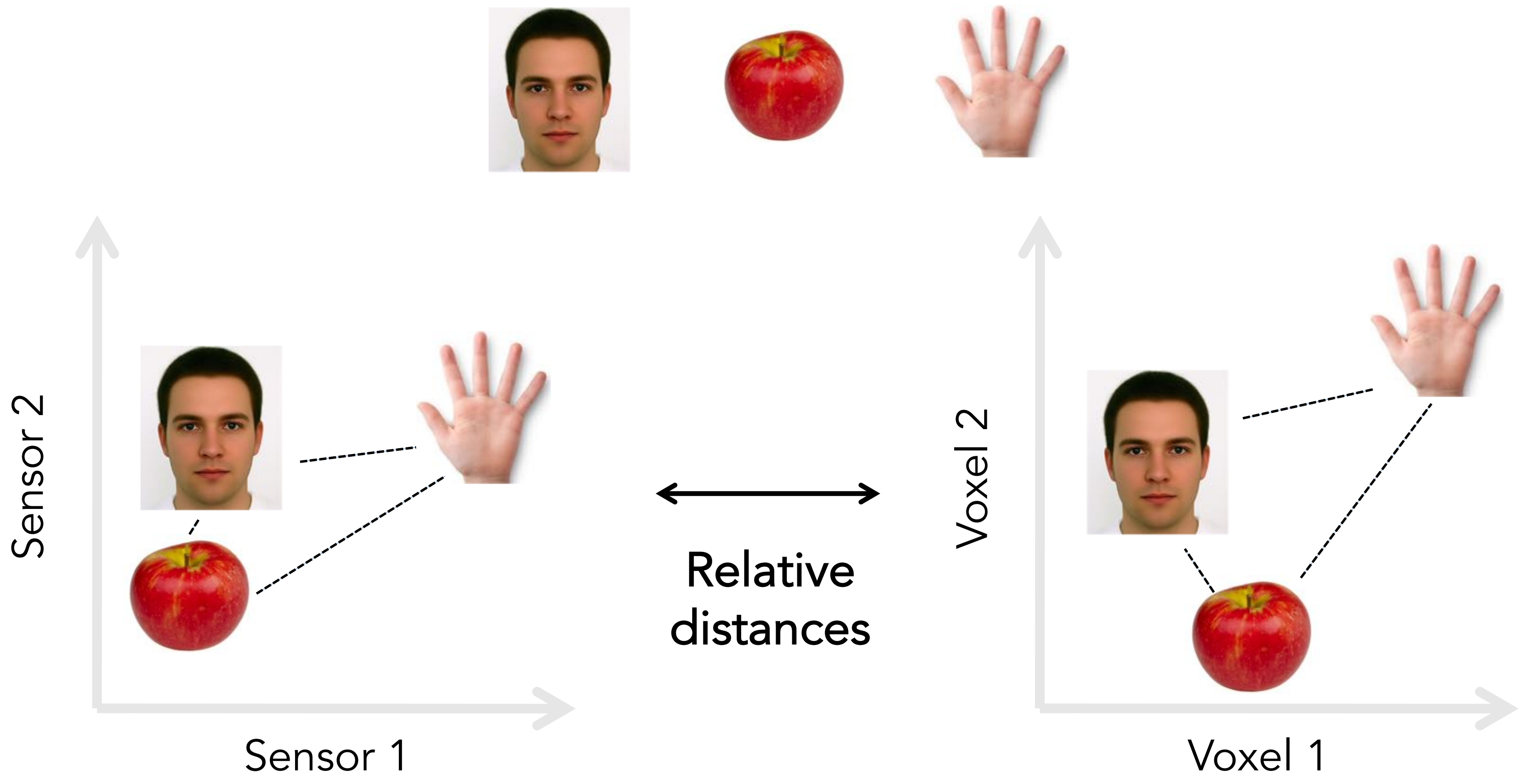




Brain-Machine Interfaces, Deep Neural Network Architectures, Comparison across Species

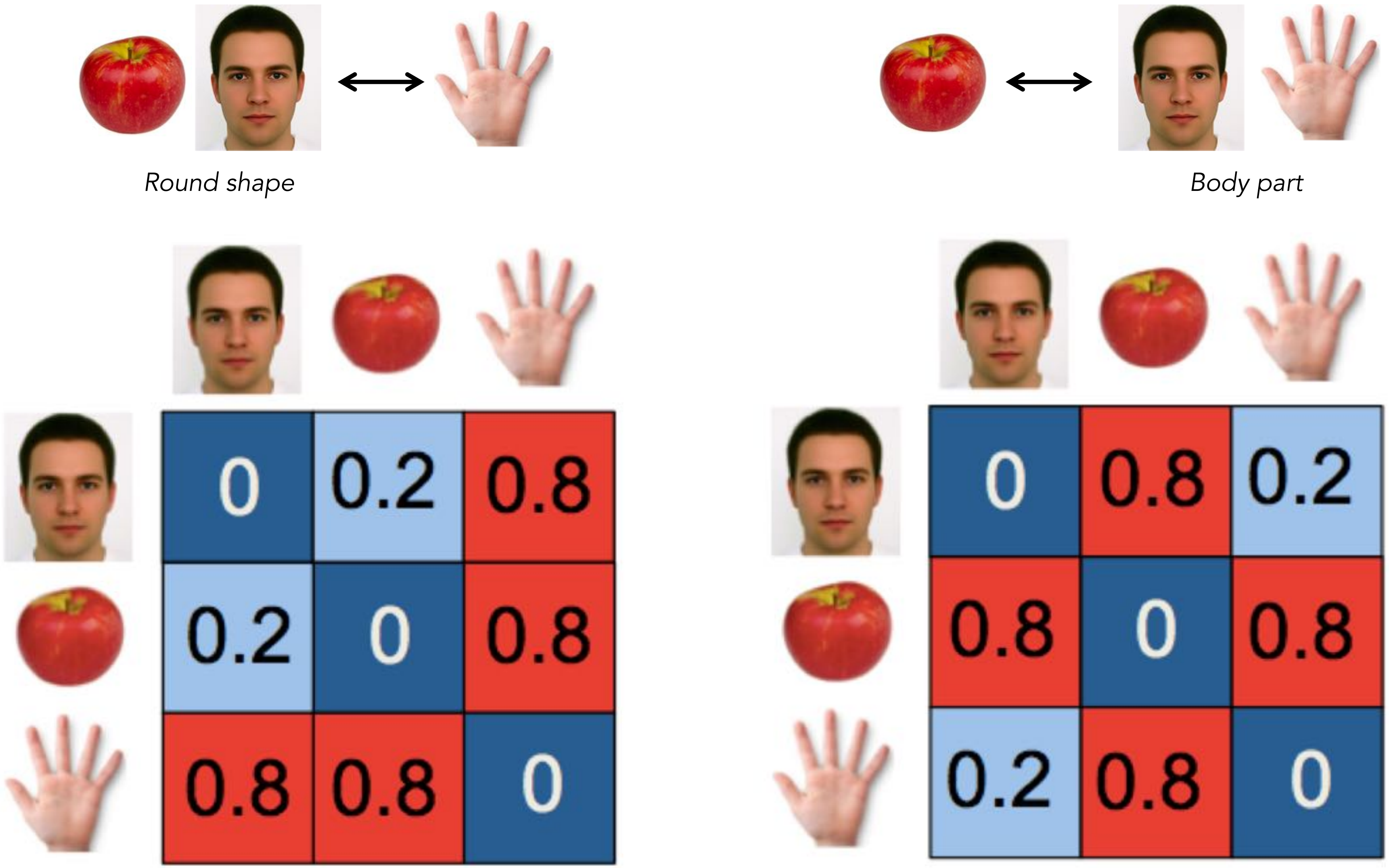

MEG: Time
Every millisecond
fMRI: Space
Each millimeter
Sensors (time)
Evoked activity
every msec
Voxels (space)
Cichy, Pantazis, Oliva (2014) Nature Neuroscience; Cichy & Oliva (2020) Neuron

Representational Geometry



Shepard et al., 1980; Kruskal and Wish., 1978; Edelman et al. 1998; Kriegeskorte et al., 2008; Mur et al., 2009; Liu et al., 2013

Representational Dissimilarity Matrices (RDMs)

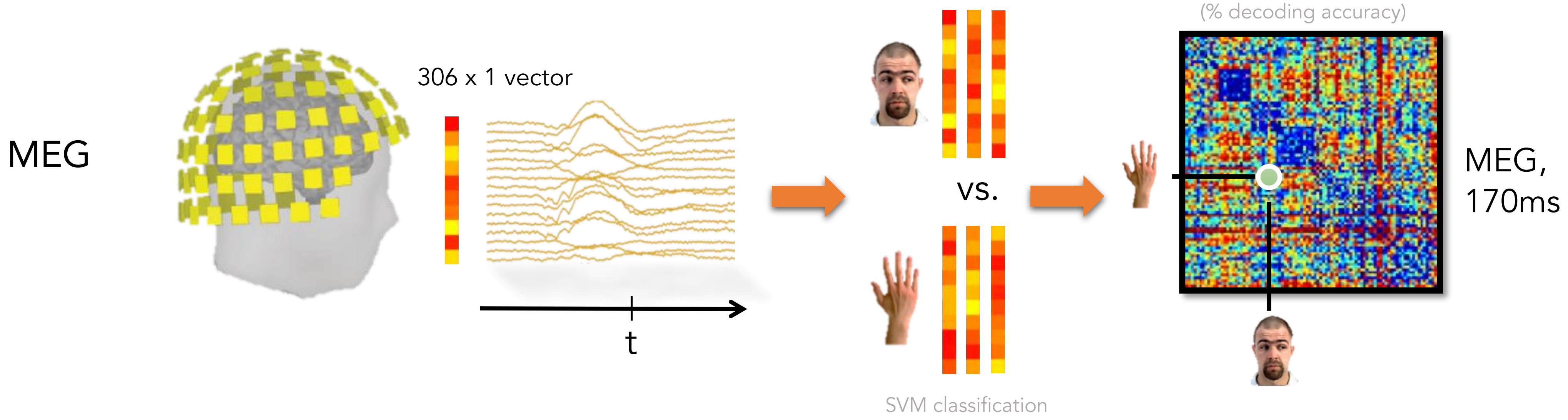


Nikolaus Kriegeskorte (2008): “RDMs as a hub to relate different representations across sensors and models”

Radoslaw
Cichy

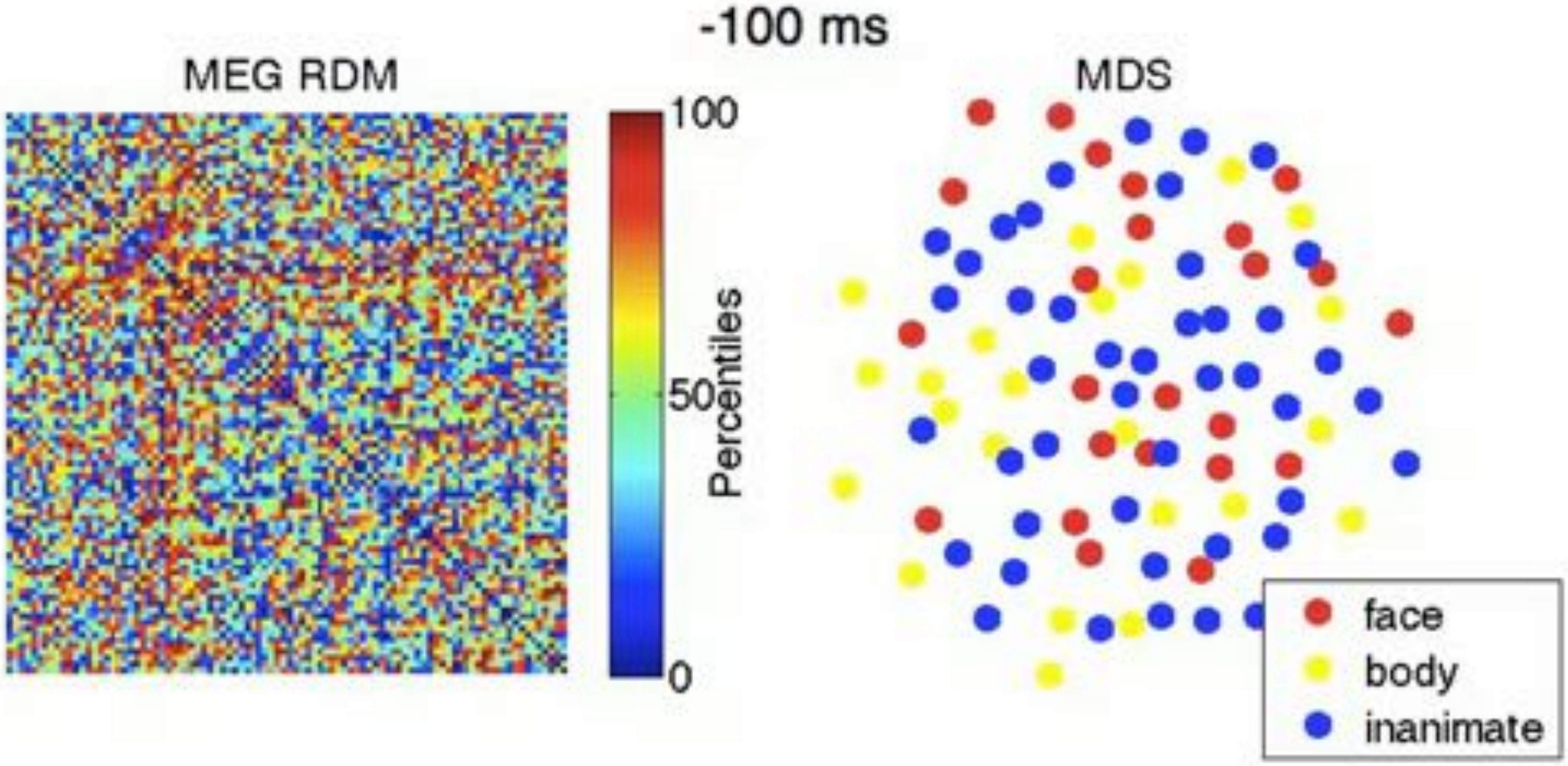
Spatio-temporal maps of brain responses

A spatially unbiased view of the relations in similarity structure
between MEG and fMRI



Cichy, Pantazis, Oliva (2014) Nature Neuroscience; (2016) Cerebral Cortex; Cichy and Oliva, (2020), Neuron.

Representational Dynamics: MEG Dissimilarity Matrices
One matrix per millisecond

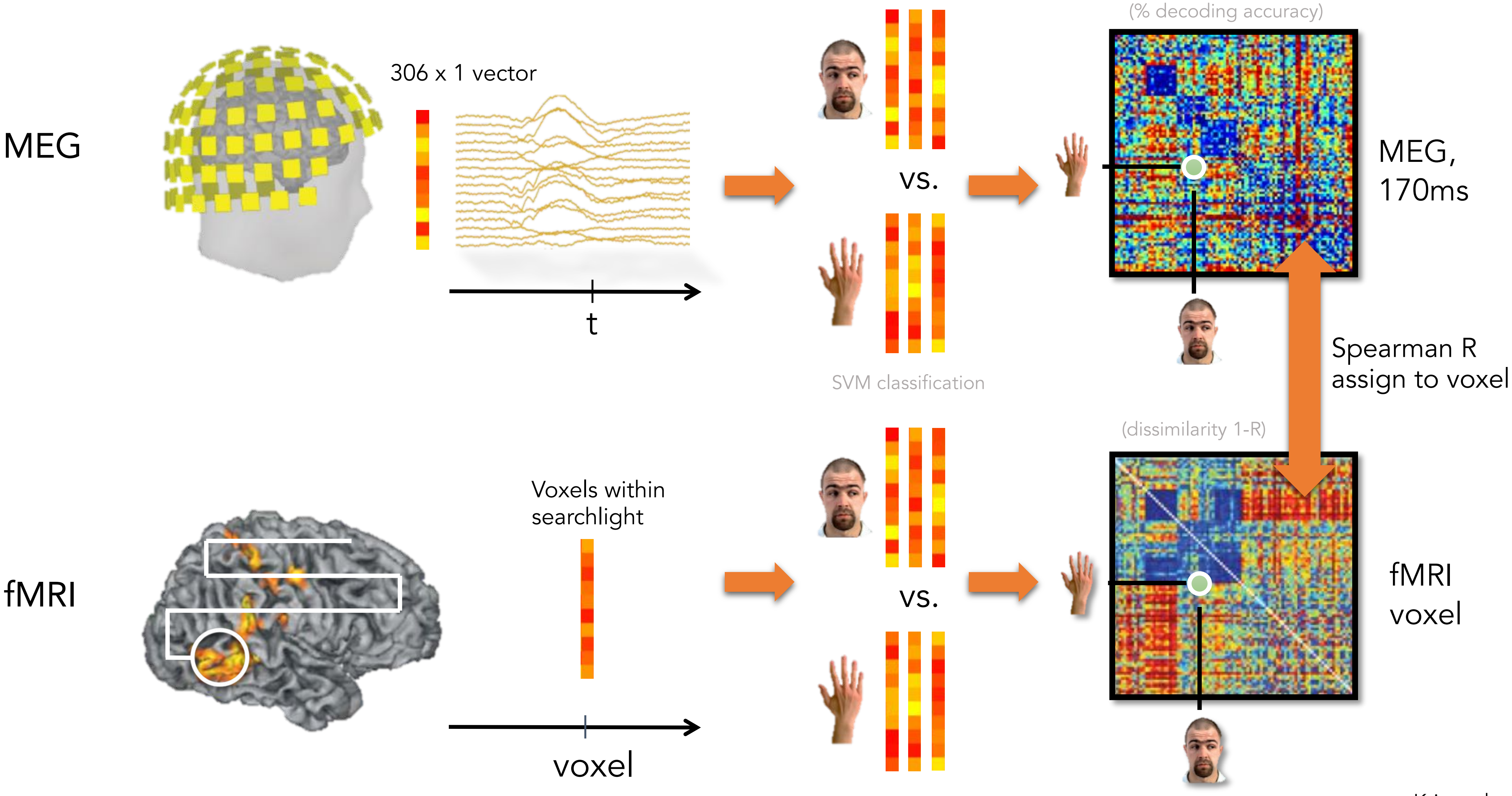


Cichy, Pantazis, Oliva (2014). Nature Neuroscience

Radoslaw Cichy

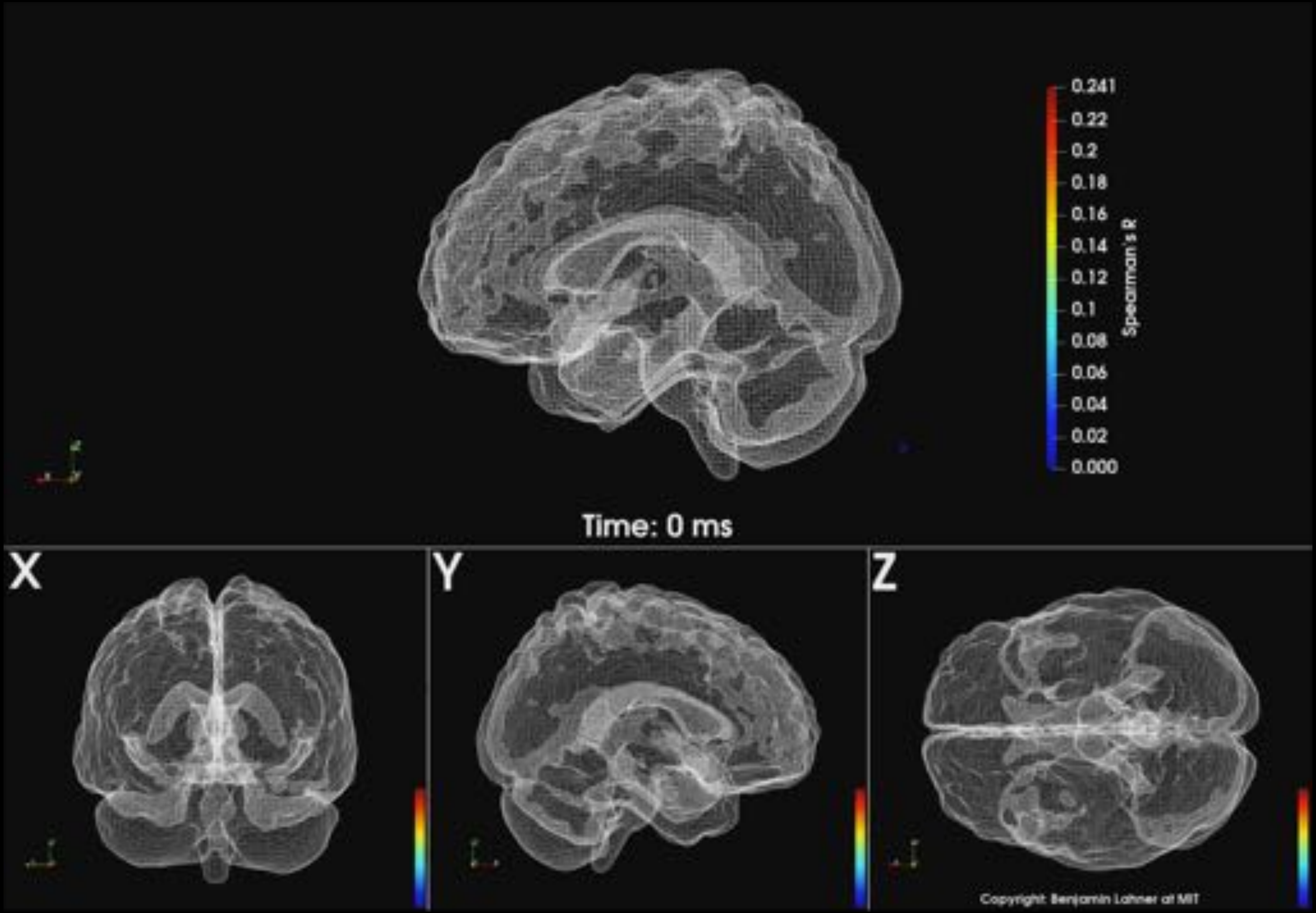
Spatio-temporal maps of brain responses

A spatially unbiased view of the relations in similarity structure between MEG and fMRI



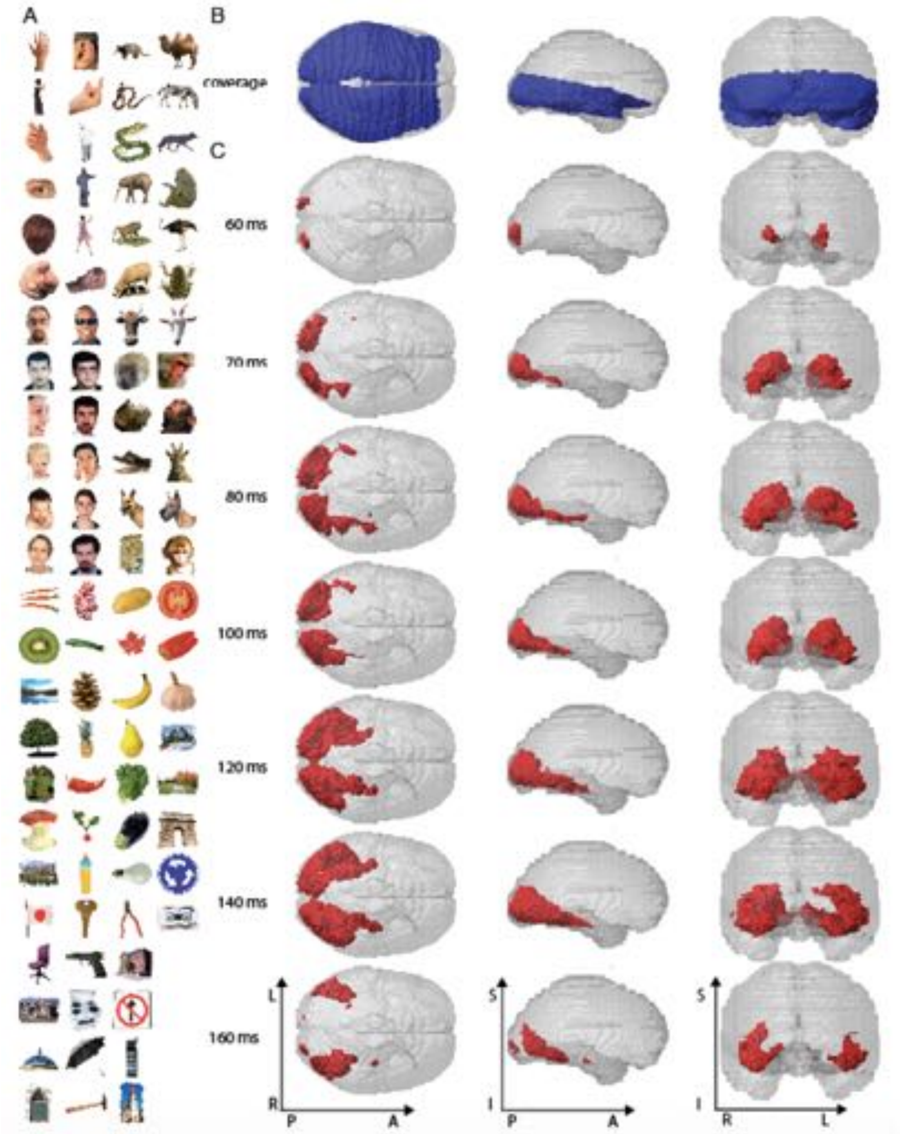
Cichy, Pantazis, Oliva (2014) Nature Neuroscience; (2016) Cerebral Cortex; Cichy and Oliva, (2020), Neuron.

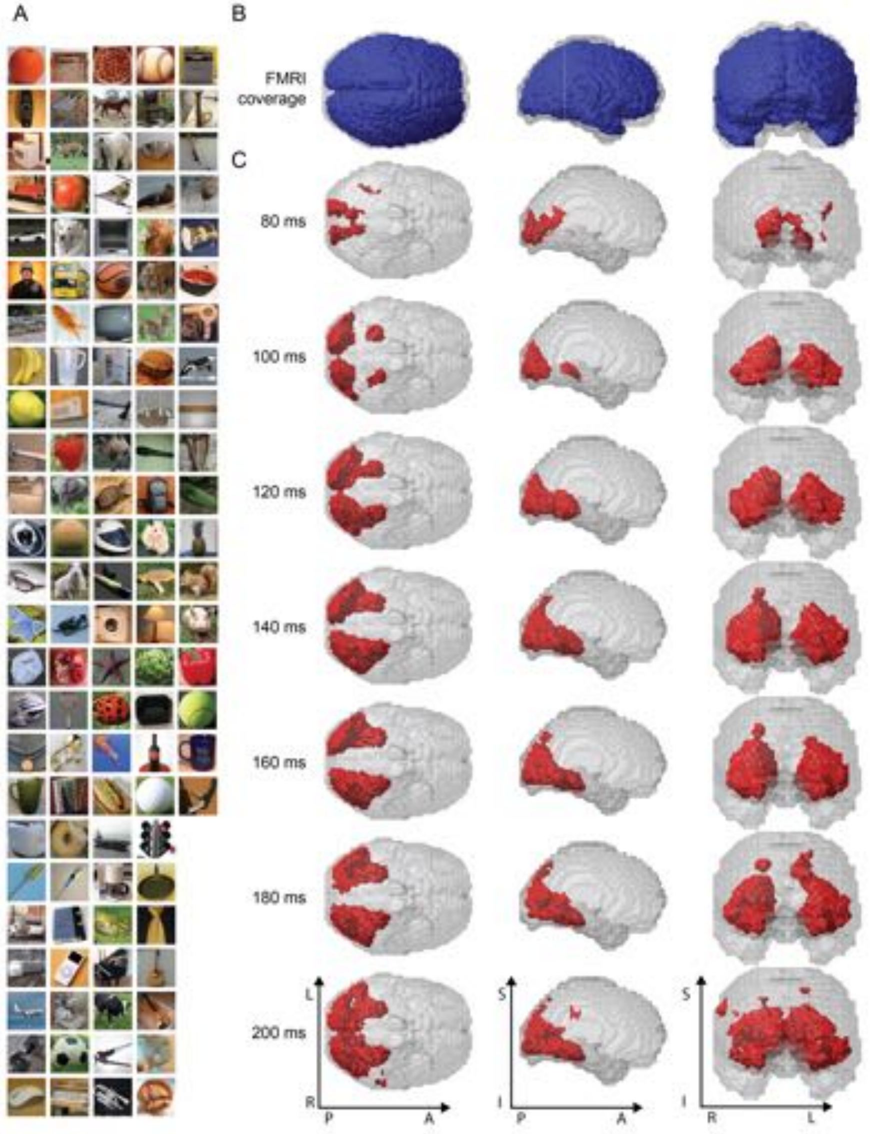
Kriegeskorte et al., 2008



Mohsenzadeh, Y., Mullin, C., Lahner, B., Cichy, R.M., & Oliva, A. (2019). Reliability and generalizability of similarity-based fusion of fMRI and MEG data in the ventral and dorsal visual streams. *Vision*, 3(1), 8

Spatio-temporal maps of brain responses





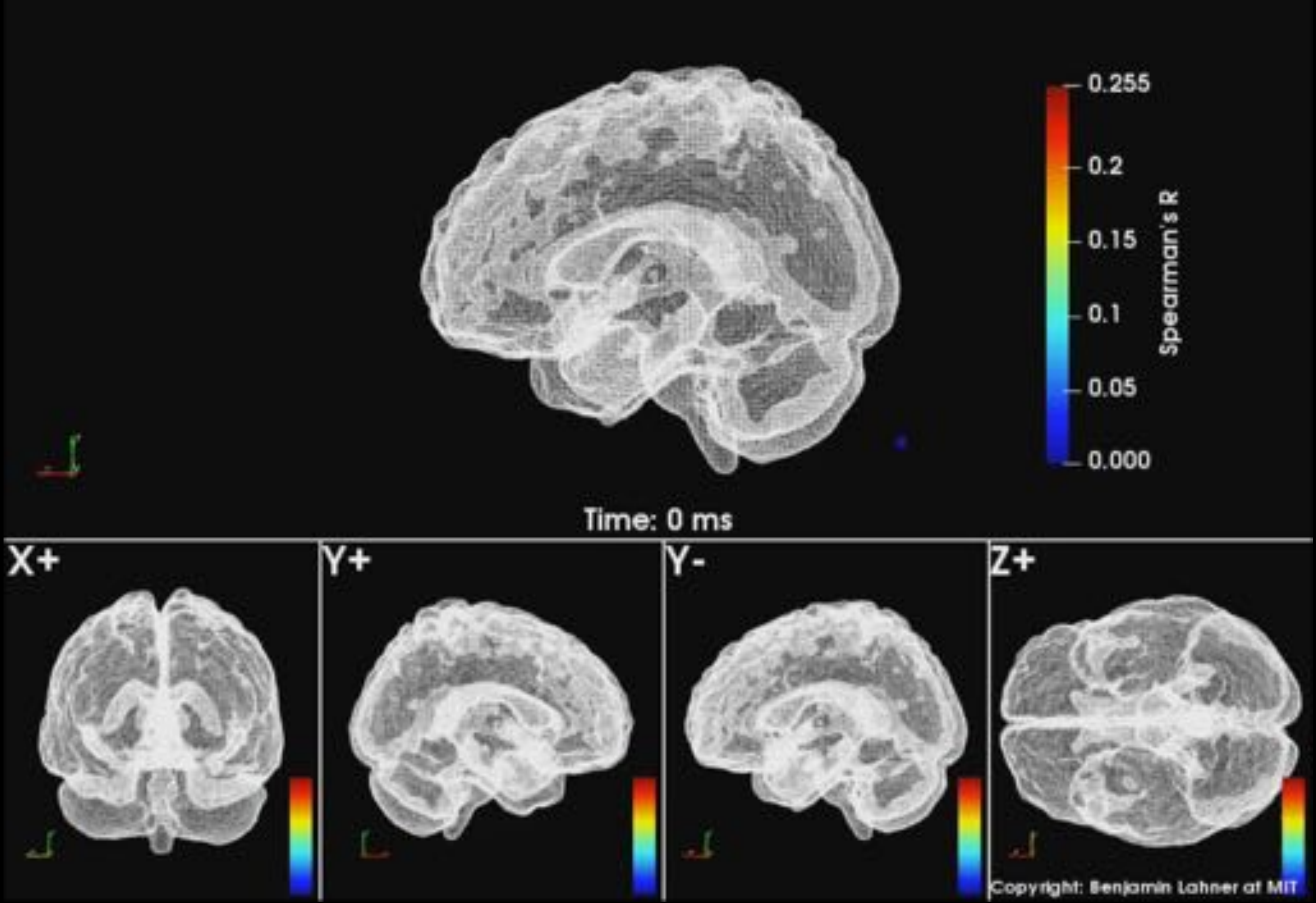
Cichy, Pantazis, Oliva (2016). Cerebral Cortex

Mohsenzadeh, Mullin, Lahner, Cichy, & Oliva, A. (2019). Vision

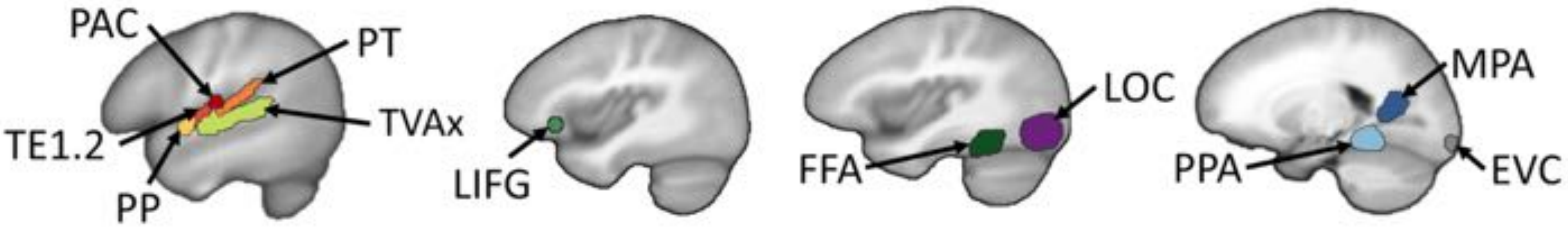
Visual Perception

- A hierarchy of responses from occipital to ventral and parietal regions during object recognition
- The spatio-temporal maps time stamps align with previous neurophysiology results
- Both ventral and dorsal responses peaks occurring around 170 msec (IPS, LO)
- **Lesson for models:** Different high-level representations of visual events after an initial common processing

Cichy, R.M., Pantazis , D., & Oliva, A. (2016). Similarity-based fusion of MEG and fMRI reveals spatio-temporal dynamics in human cortex during visual object recognition. Cerebral Cortex, 26 (8): 3563-3579

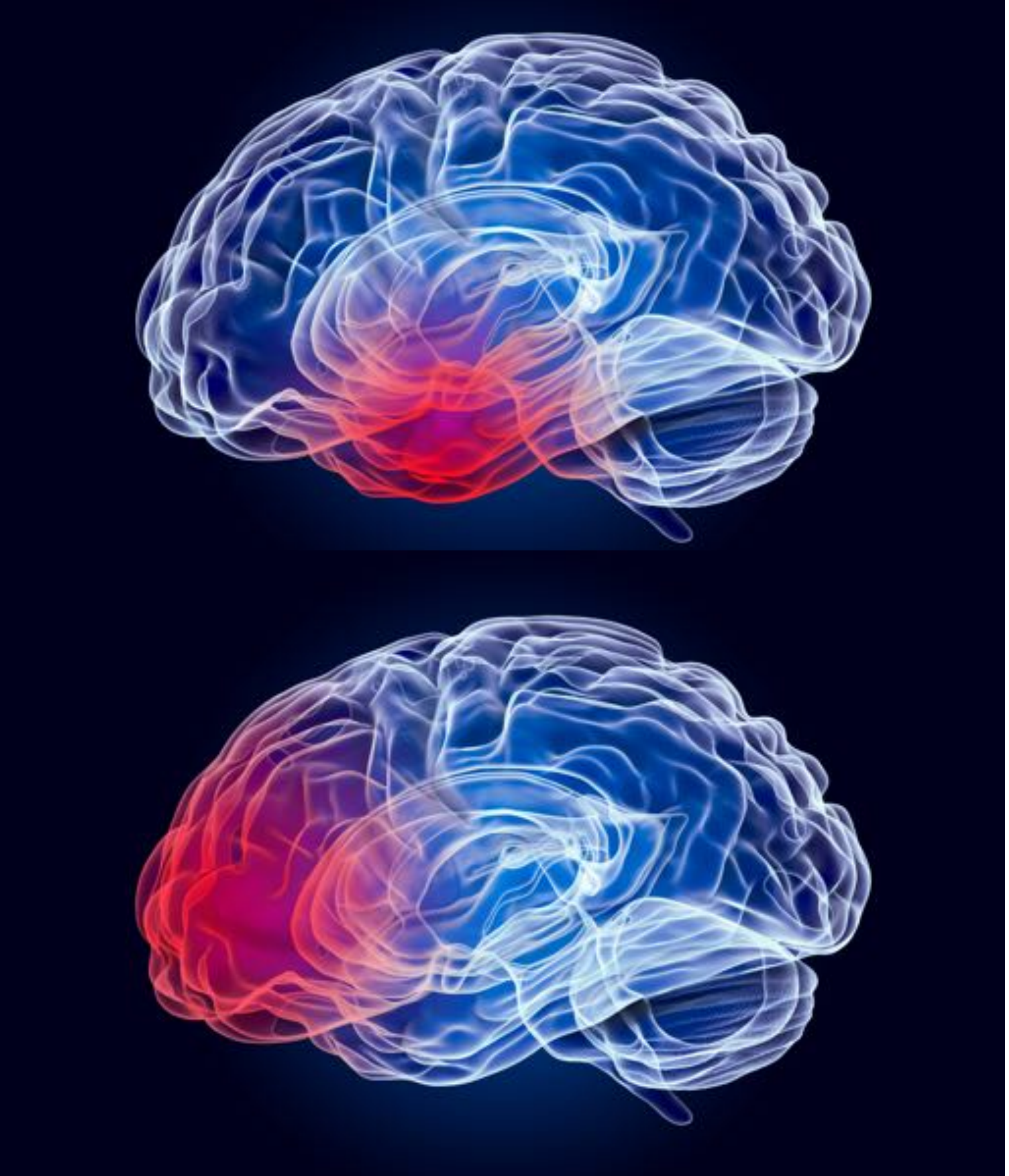


Lowe*, M., Mohsenzadeh*, Y., Lahner, B., Charest, I., Oliva, A., & Teng, S.(2020) Spatiotemporal Dynamics of Sound Representations reveal a Hierarchical Progression of Category Selectivity. bioarxiv



Region of Interest	Peak latency (ms)
PAC (TE1.0 & TE1.1)	115 (100, 191)
TE1.2	117 (111, 207)
Planum Temporale (PT)	115 (103, 197)
Planum Polare (PP)	117 (110, 207)
Temporal Voice Area (TVAx)	199 (188, 251)
Left Inferior Frontal Gyrus (LIFG)	196 (188, 380)
Fusiform Face Area (FFA)	298 (200, 368)
Parahippocampal Place Area (PPA)	293 (219, 380)
Medial Place Area (MPA)	199 (196, 400)
Lateral Occipital Complex (LOC)	377 (169, 410)

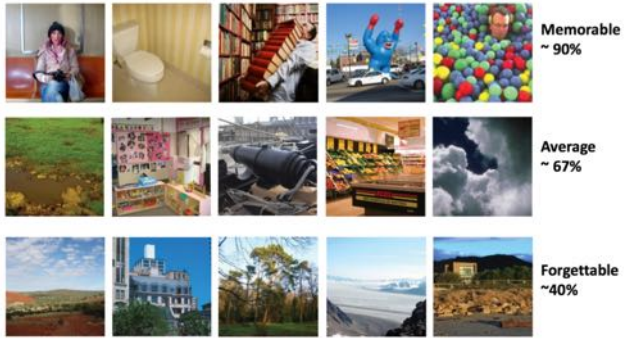
Lowe*, M., Mohsenzadeh*, Y., Lahner, B., Charest, I., Oliva, A., & Teng, S.(2020) Spatiotemporal Dynamics of Sound Representations reveal a Hierarchical Progression of Category Selectivity. bioarxiv



Auditory Perception

- Timing of auditory cortex, pre-frontal and higher-level regions responses in human
- Specificity of human voices in auditory, pre-frontal and fusiform gyrus with time stamps
- A late response (~ 300 msec) for auditory stimuli in regions classically associated with vision
- **Lesson for models:** Human voices have specialized modules, being separated early on; audition is mixed with visual features at a later stage in the hierarchy of processing

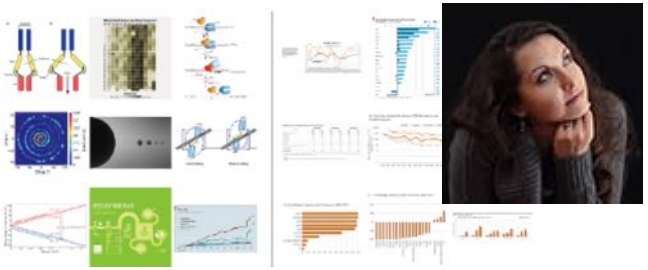
Memorability



Phillip Isola (CVPR 2011, NIPS 2011, PAMI 2014)

Aditya Khosla (ICCV, 2013; ICCV 2015)

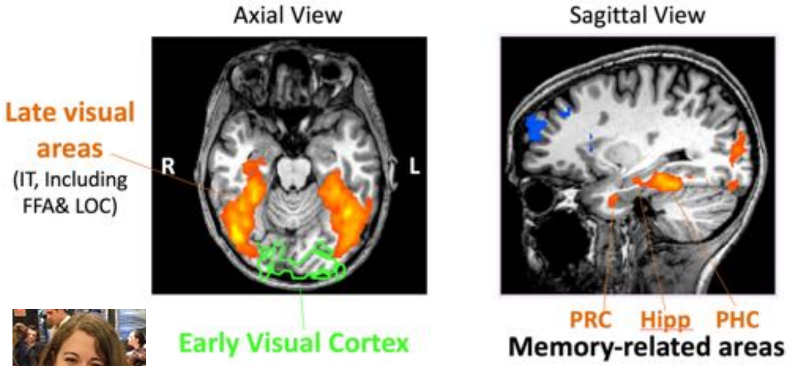
Memorable Words (with Fedorenko and Gibson)



Memorable Infographics
(Zoya Bylinskii, 2013, 2015)

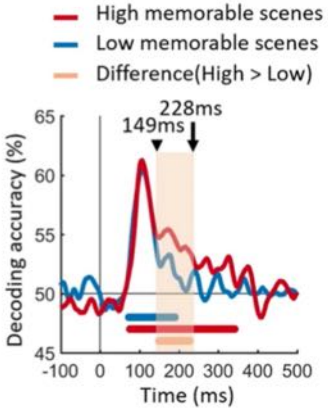
ICCV 2019

Lore Goetschalckx, Alex Andonian, Phillip Isola

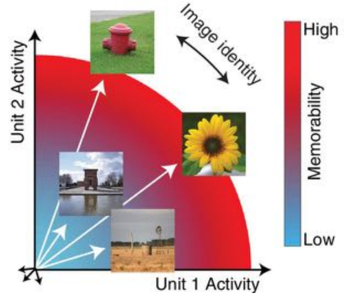


Wilma Bainbridge (2016)

Yalda Mohsenzadeh (Scientific Reports, 2019)



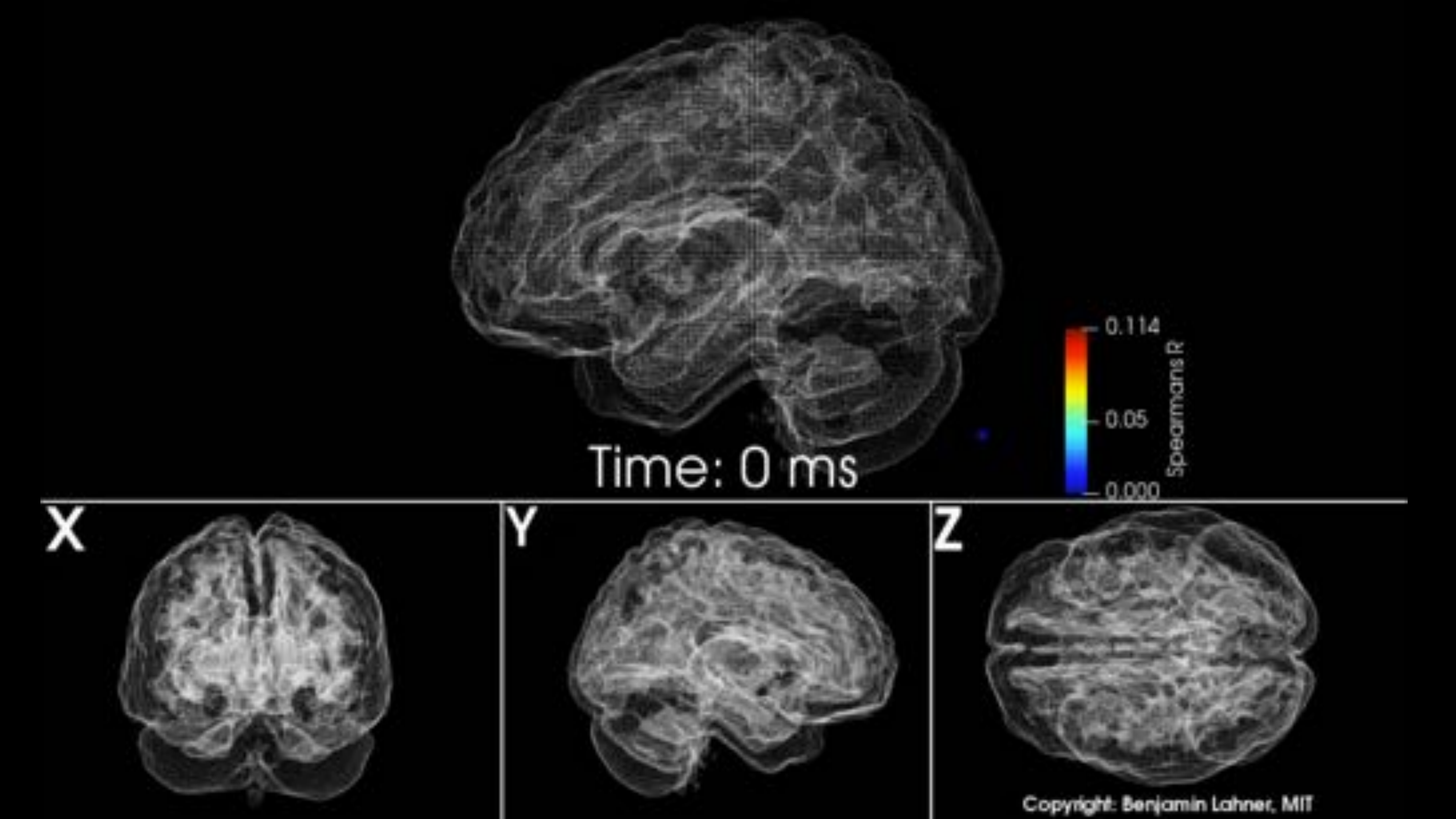
Nicole Rust (eLife, 2019)



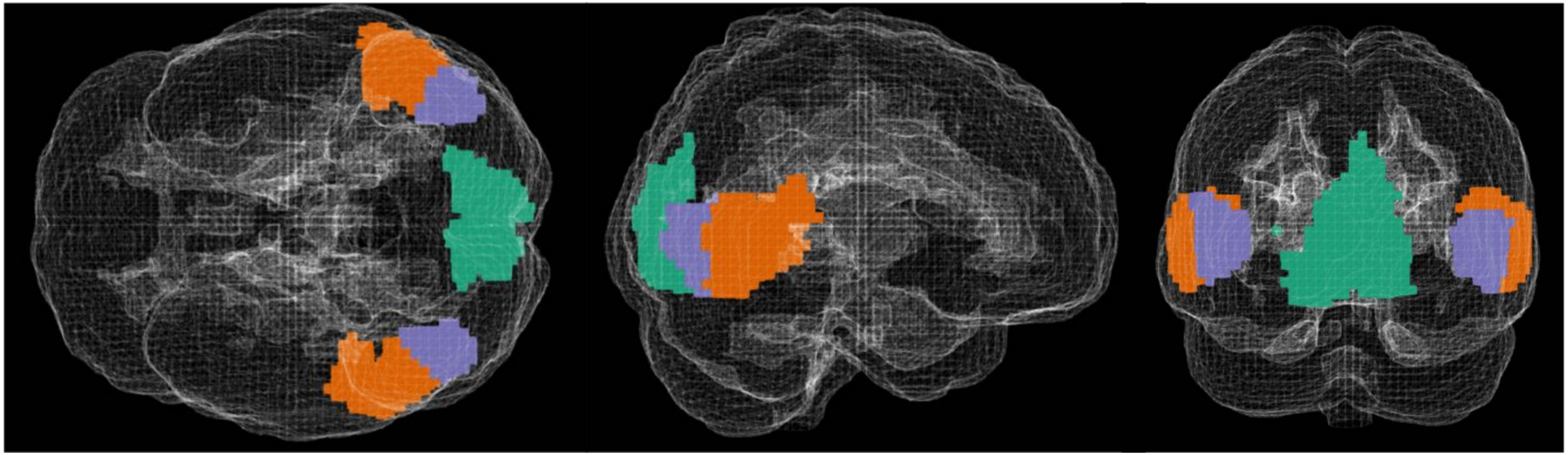
Visual memorability is a consequence of the optimizations required for visual processing

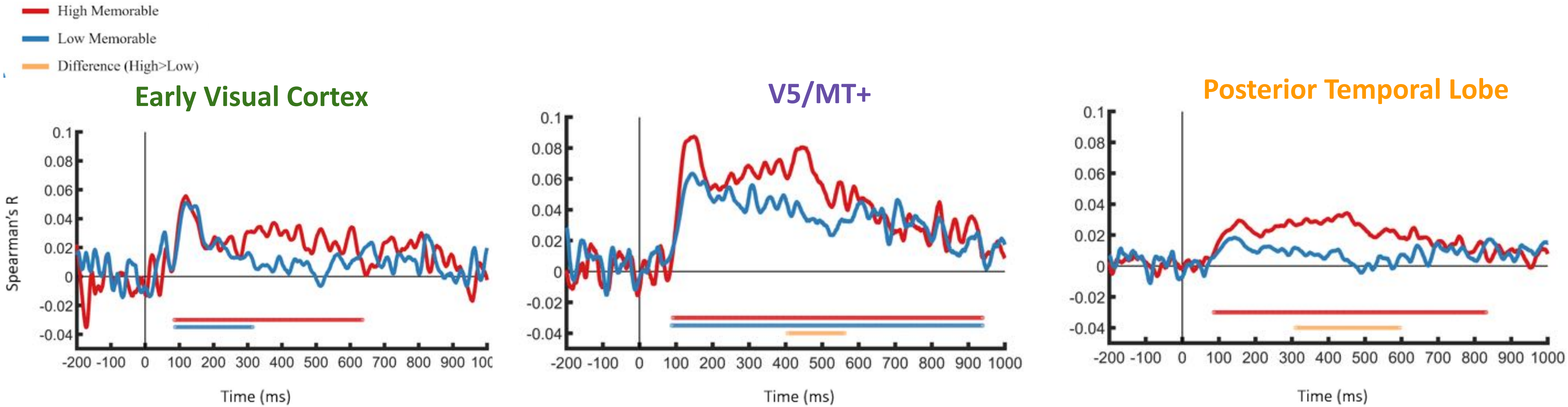
More memorable

Less memorable


0.114
0.05
0.000
Spearmans R
Time: 0 ms
X
Y
Z
Copyright: Benjamin Lahner, MIT

Early Visual Cortex, V5/MT+, Posterior Temporal Lobe





With Y. Mohsenzadeh, B. Lahner , C. Mullin

Memorability

- Late responses for memorable images in higher cortical regions associated with object recognition
- Memorable images have a stronger signals during visual processing
- Visual memorability is a consequence of the optimizations required for visual processing
- **Lesson for models:** Memorability as a measure of the *utility* of information

Human & Artificial Cognition

- Characterizing the **bandwidth of human perception and cognition** is critical
- **A new field of investigation:** Cognitive / Clinical / Social / Perceptual Computational Experimentalist / Synthetic Neuroscientist
- **Studying the implementation** that works best for performing specific tasks
- **Exploring the alternatives** that have not been taken by biological systems

Computational Perception & Cognition

http://olivalab.mit.edu/

HOME TEAM PROJECTS PUBLICATIONS DATASETS & MODELS COURSES

Datasets and Models

Memento10k: Video Memorability Dataset (to be released soon)
Memento10k is the largest in-the-wild video memorability dataset to date, with more than 10,000 videos and close to 1 million human annotations. Videos represent varied everyday events, captured in a homemade fashion. Each video was annotated through our Memento Memory Game and possesses 90 human annotations on average. We also release action labels, as well as 5 detailed captions for each video.

Moments in Time and Multi-Moments in Time
Moments in Time is a research project aiming to build a very large-scale dataset to help AI systems recognize events in videos. The first release includes one million 3 second videos each with one activity label (covering >300 action categories). The second release *Multi-Moments in Time* (M-MiT) includes over 2 million labels. The third version *Spoken Moments* will be released by ECCV 2020.

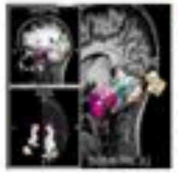
GANalyze: Generate Memorable and Forgettable Images
A framework that uses Generative Adversarial Networks (GANs) to study cognitive properties like memorability, aesthetics, and emotional valence. GANs allow us to generate a manifold of natural-looking images with fine-grained differences in their visual attributes. By navigating this manifold in directions that increase memorability, we can visualize what it looks like for a particular generated image to become more or less memorable.

The Algonauts Project: 2019 Edition
The **Algonauts Project** brings biological and artificial intelligence researchers together on a common platform to exchange ideas and advance both fields. Our first challenge *Explaining the Human Visual Brain*, focused on building computer vision models that simulate how the brain recognizes objects. The released dataset includes multiple image sets with fMRI and MEG human brain data. The second Algonauts challenge will be announced during Summer 2020.

MEG and fMRI Data of Images
We recorded magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI) data while 15 participants viewed a set of 156 natural images. These images can be subdivided into five categories (faces, bodies, animals, objects, scenes) or two twinsets of 78 images each.

Places Datasets
Places2 contains more than 10 million images comprising 400+ unique scene categories. The dataset features 5000 to 30,000 training images per class, consistent with real-world frequencies of occurrence. The Places dataset allows learning of deep scene features for many scene recognition tasks, and comparing models' performances on large-scale image benchmarks. First version of Places dataset can be found **here**.

MiniPlaces Dataset
A subset of the **Places2 dataset** used for the image classification benchmark MiniPlaces Challenge. It contains 100 scene categories and consists of 100,000 images for training, 10,000 images for validation, and 10,000 images for testing, with all images resized to 128x128 pixels.



A Human Brain Atlas: Medial Temporal Lobe & Ventral Visual Stream ROI Collection
Forty sets of human medial temporal lobe anatomical and 70 sets of ventral visual stream functional region of interest (ROI) definitions, collected over the course of fMRI image recognition experiments. They include bilateral perirhinal cortex, entorhinal cortex, amygdala, hippocampus head, body, and tail, and parahippocampal cortex among others.

Network Dissection: Models
Network Dissection is our method for quantifying interpretability of individual units in a deep CNN. It works by measuring the alignment between unit response and a set of concepts drawn from a broad and dense segmentation dataset called Broden.

BubbleView: An Interface for Crowdsourcing Image Importance Maps
BubbleView is an alternative methodology for eye tracking using discrete mouse clicks to measure which information people consciously choose to examine. Here you can find all the data and analysis codes used in our BubbleView experiments.

Visually29K: A Large-Scale Curated Infographics Dataset
We curated a subset of 28,973 infographics to cover a fixed set of 391 tags (filtered down from free-form text). There are at least 50 infographic instances for each of 391 different tags. Infographics have an average of 2 tags each (fine-grained topics) and are additionally annotated with 1 of 26 categories (coarse topics). We have split the files into training and test sets for prediction tasks.

MEG and fMRI Data of Two Object Datasets
We recorded MEG and fMRI data while participants viewed images of objects from two different image sets (N=92 and N=118). For each object image set, you can find here the visual stimuli, MEG RDMs, MEG epoched raw data, and fMRI beta- and t-value maps.