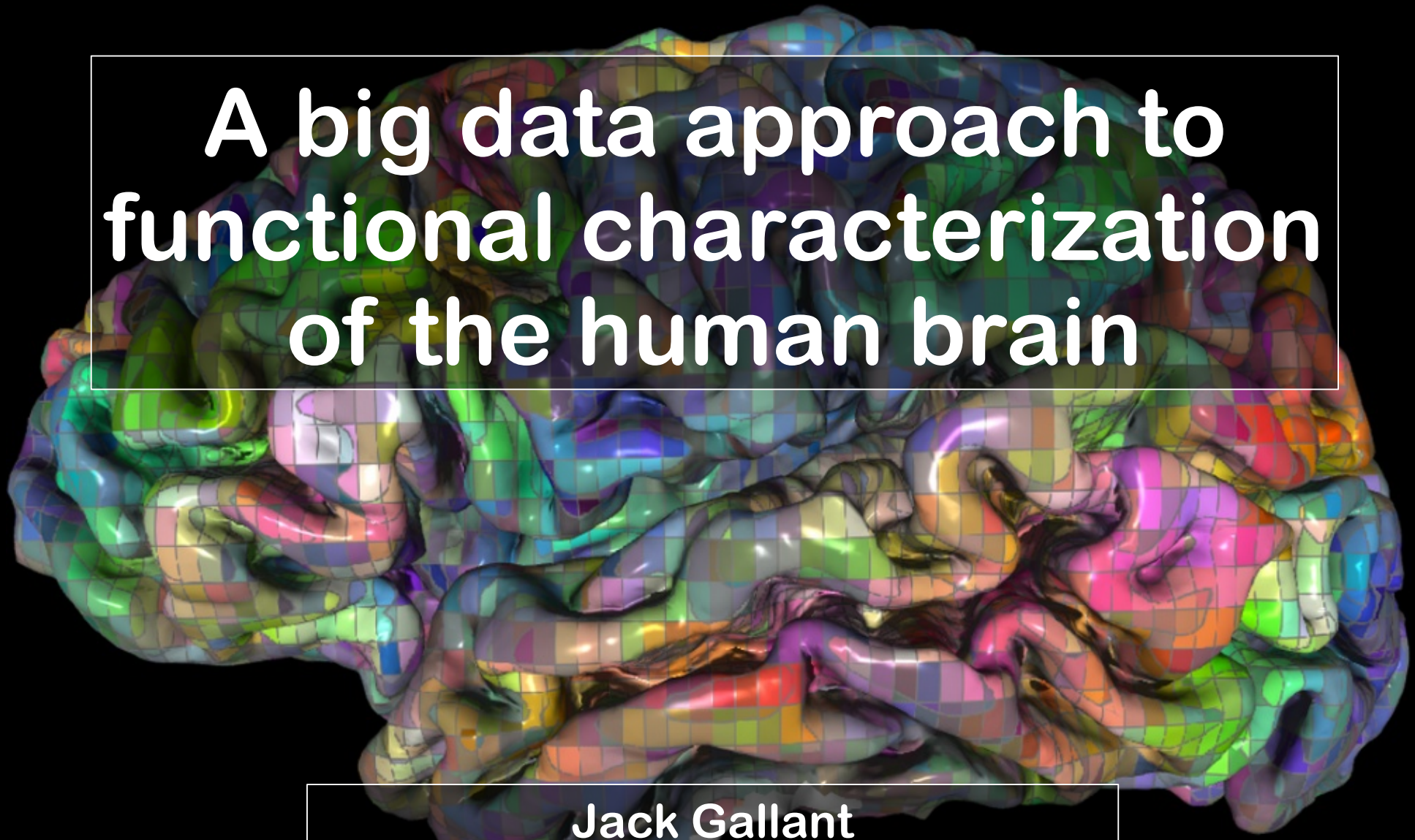
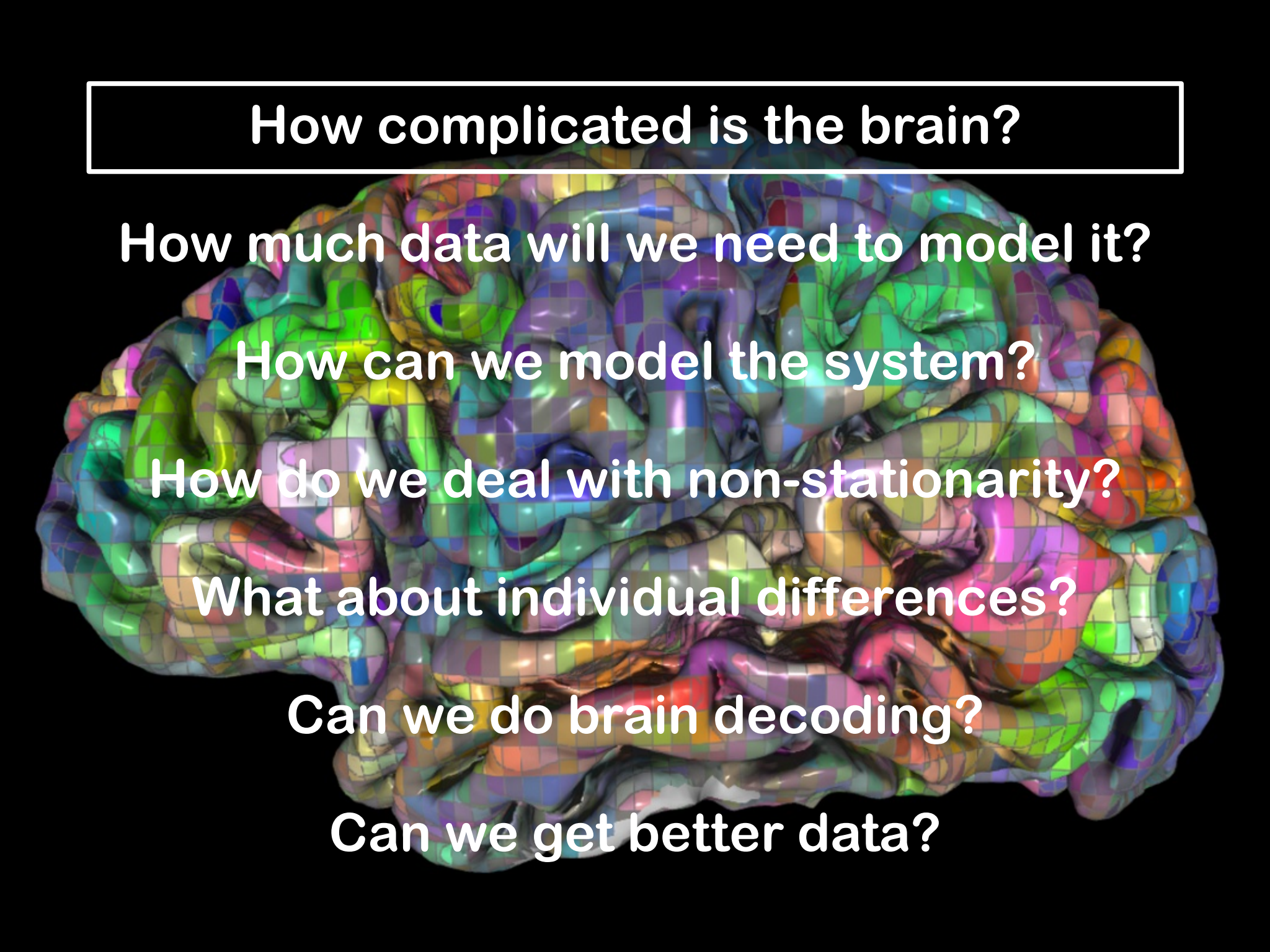
A big data approach to functional characterization of the human brain

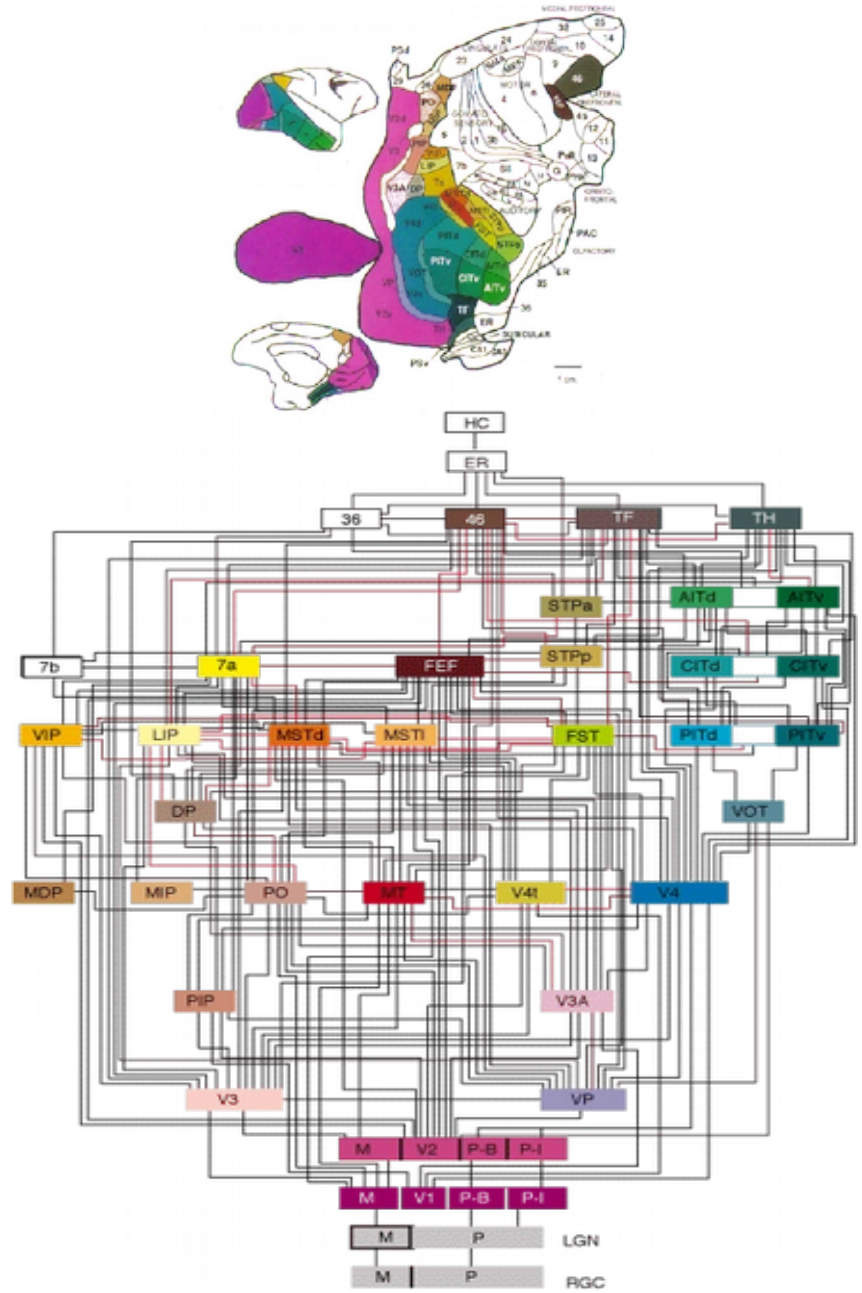


Jack Gallant
University of California at Berkeley
CCC-NSF Meeting Dec. 2014


How complicated is the brain?
How much data will we need to model it?
How can we model the system?
How do we deal with non-stationarity?
What about individual differences?
Can we do brain decoding?
Can we get better data?

Macaque vision is a typical brain subsystem

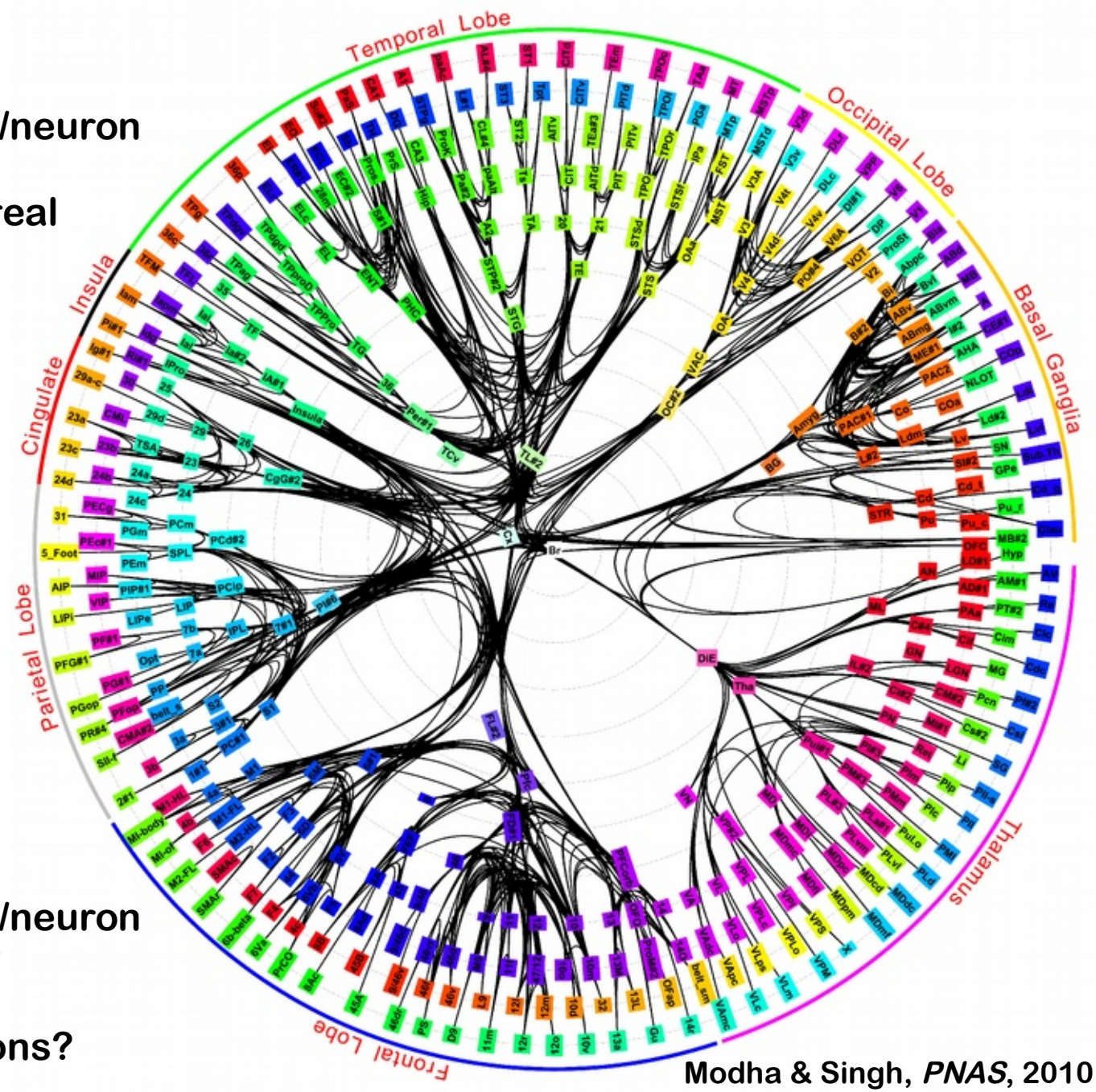
- Dozens of distinct anatomical and functional areas arranged in a hierarchical, parallel network with substantial recurrent feedback.
- Transformations between areas are nonlinear.
- Each area represents a different high-dimensional feature space projected onto the cortical surface.
- Each area is differentially affected by bottom-up and top-down information.



Felleman and Van Essen, *Cerebral Cortex,* 1992

The brain is a big place

Macaque brain:
2 billion cortical neurons
1-10 thousand connections/neuron
383 total areas and nuclei
6602 bi-directional, inter-areal connections



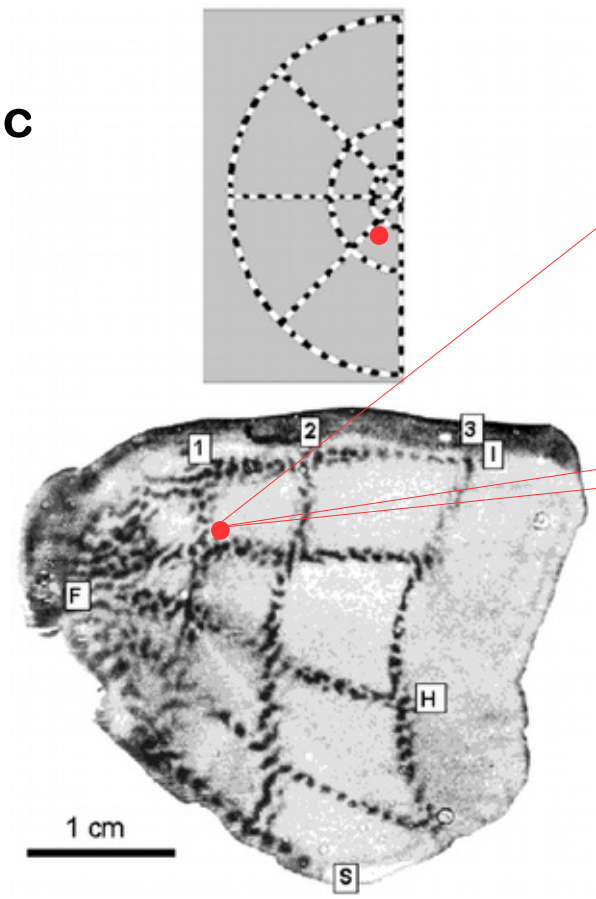
Human brain:
18 billion cortical neurons
1-10 thousand connections/neuron
5 million cortical columns ?
500 areas and nuclei?
12000 inter-areal connections?

Modha & Singh, *PNAS*, 2010

Information is organized at multiple scales

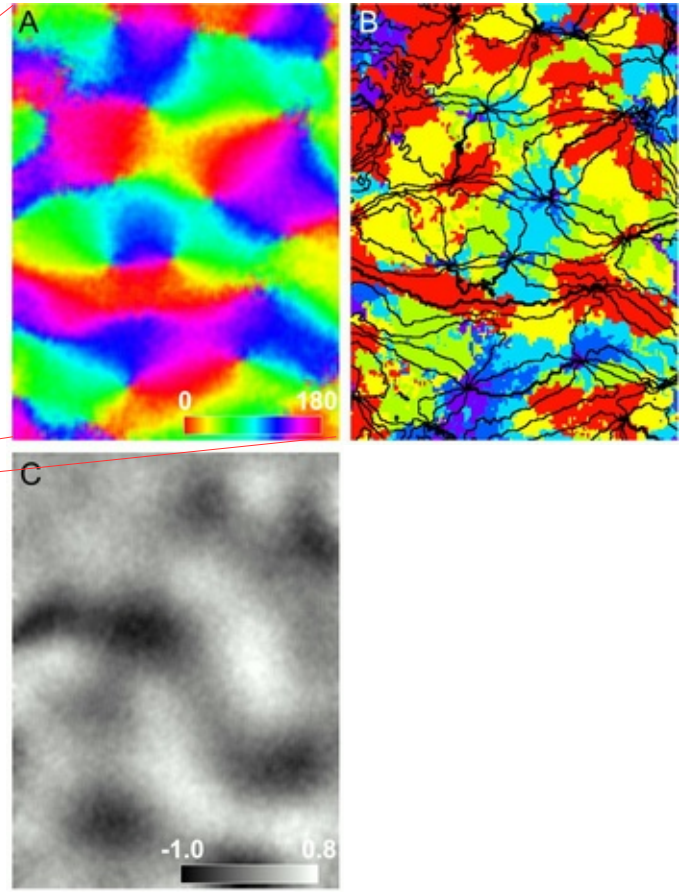
Retinotopic Mapping Stimulus

2DG map on flattened Macaque cortex



Orientation

Spatial frequency



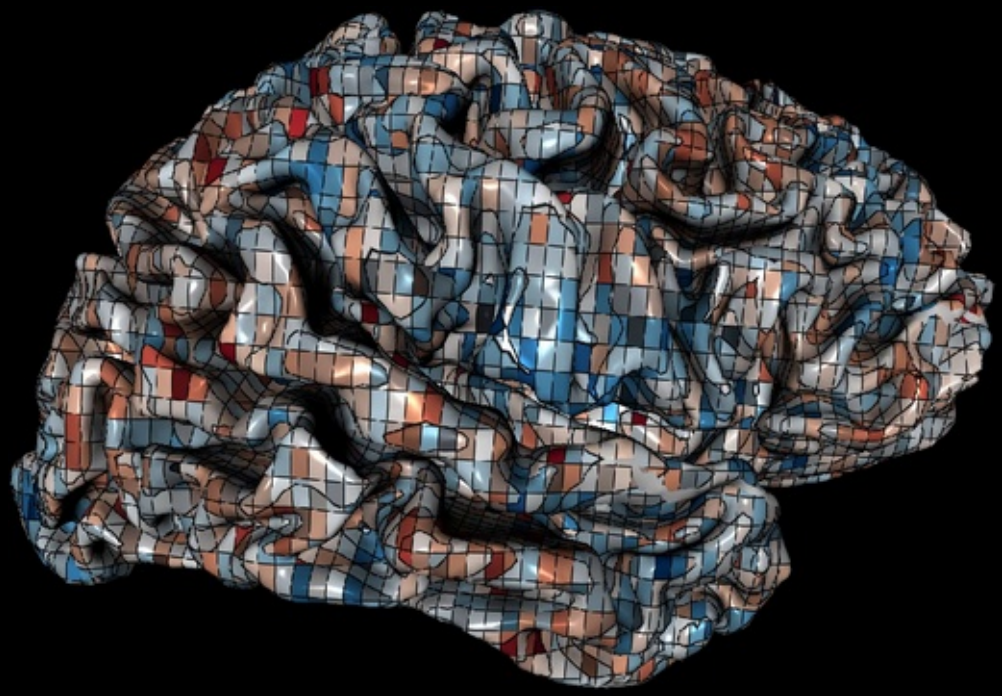
Ocular dom.

Tootel et al., *J. Neurosci* (1988)

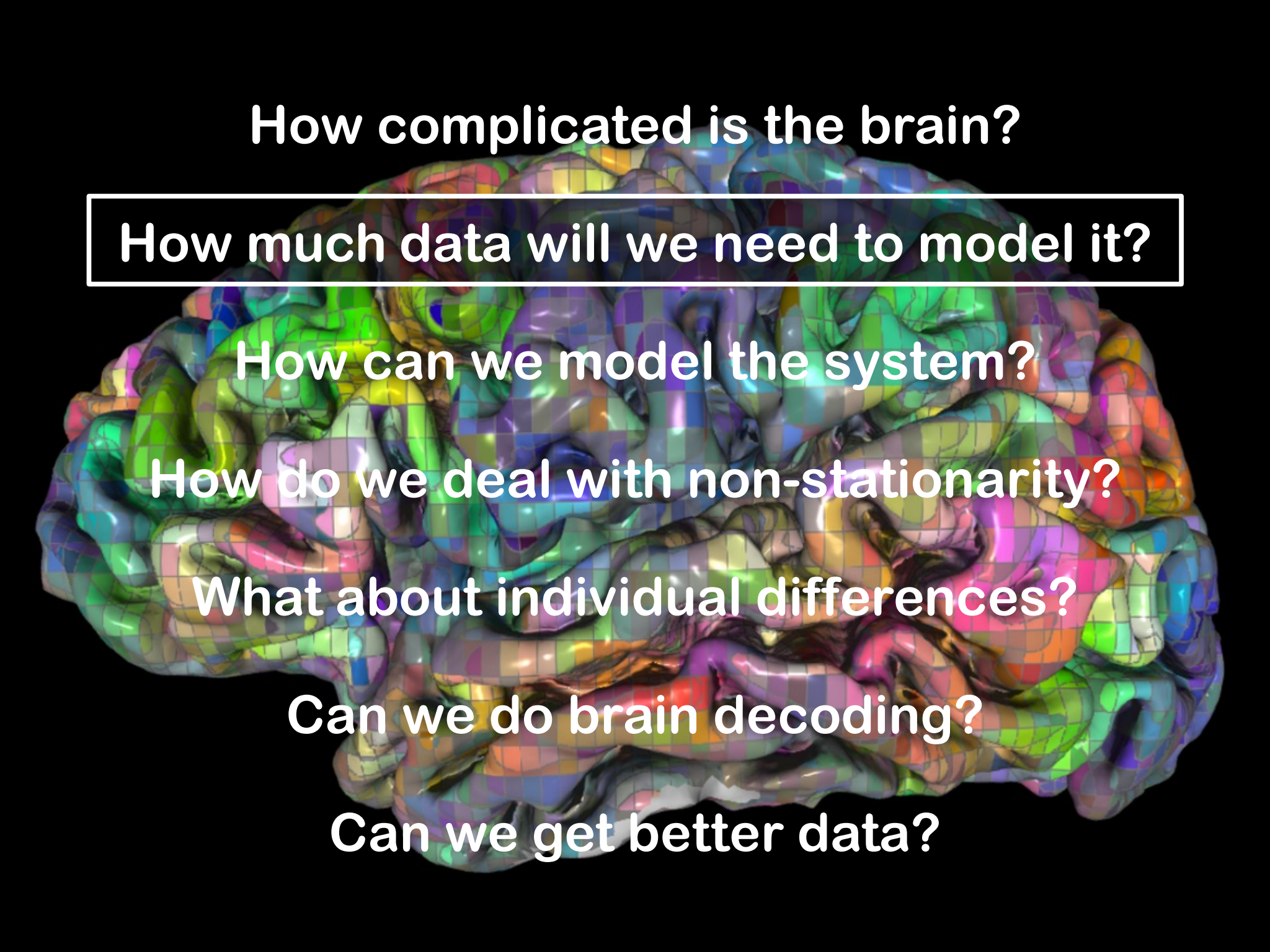
Issa, Trepel & Stryker, *J. Neurosci* (2000)

Structural and functional measurements are complimentary

Diffusion Imaging

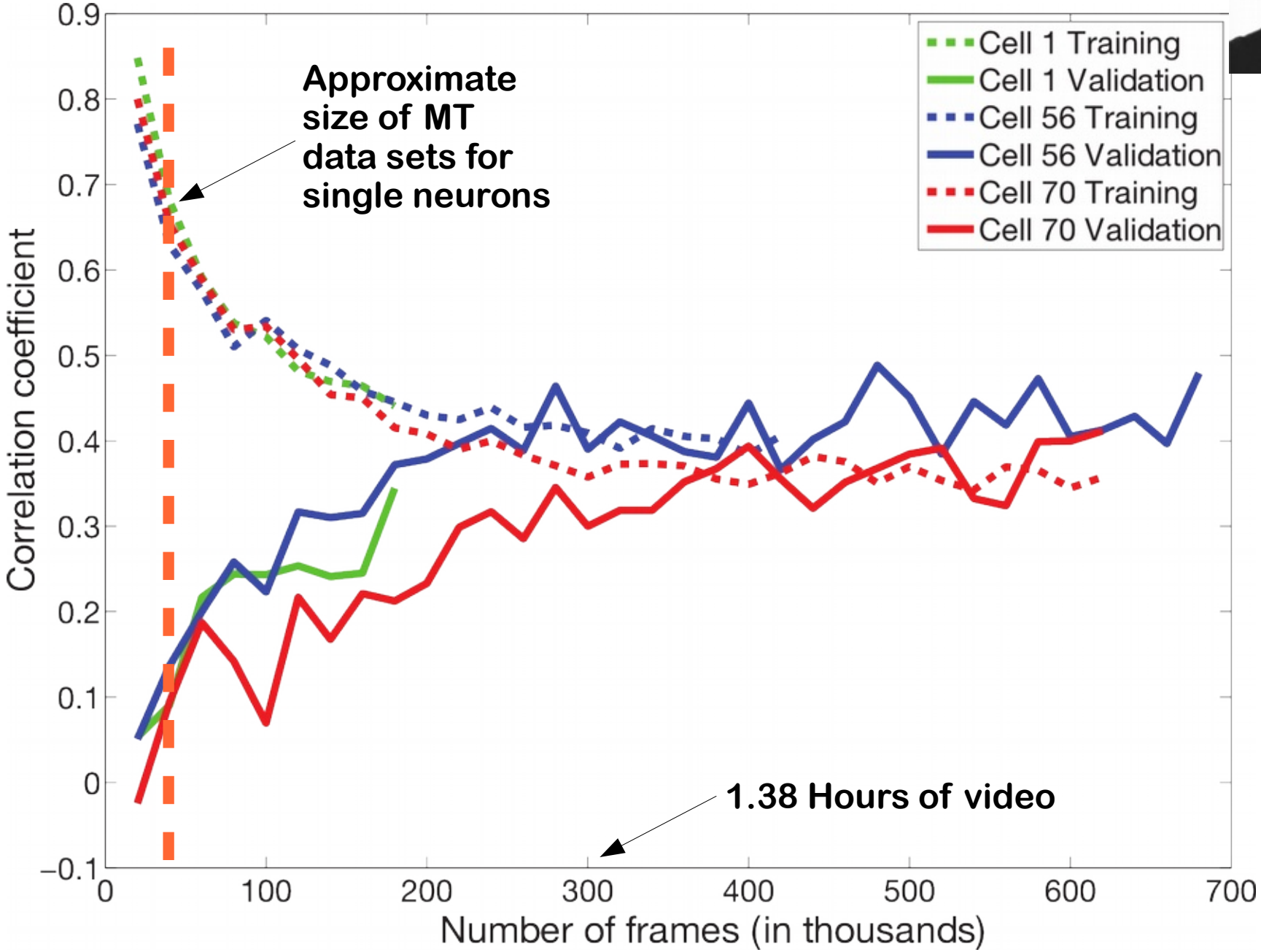


fMRI


How complicated is the brain?
How much data will we need to model it?
How can we model the system?
How do we deal with non-stationarity?
What about individual differences?
Can we do brain decoding?
Can we get better data?

Predictions are limited by data size

Mike Oliver



Oliver & Gallant, *in preparation*

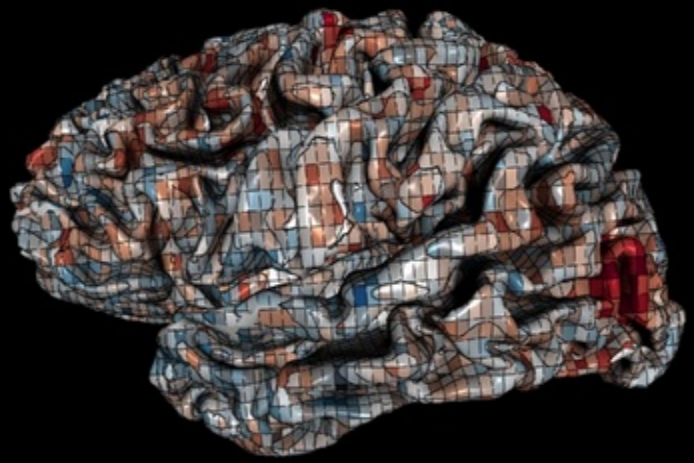
Methods are time and/or space limited

James Gao

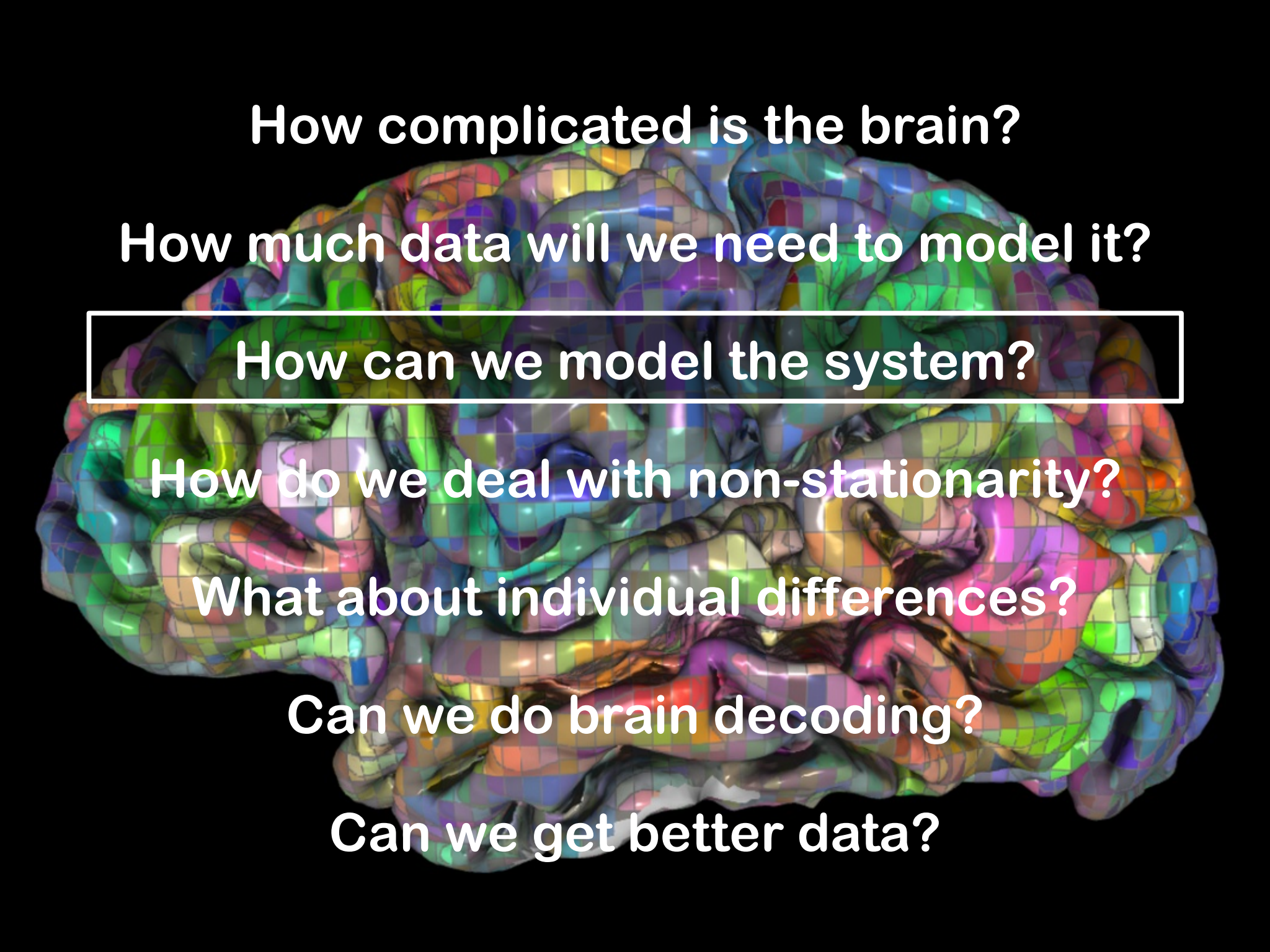
Less activity

Average activity

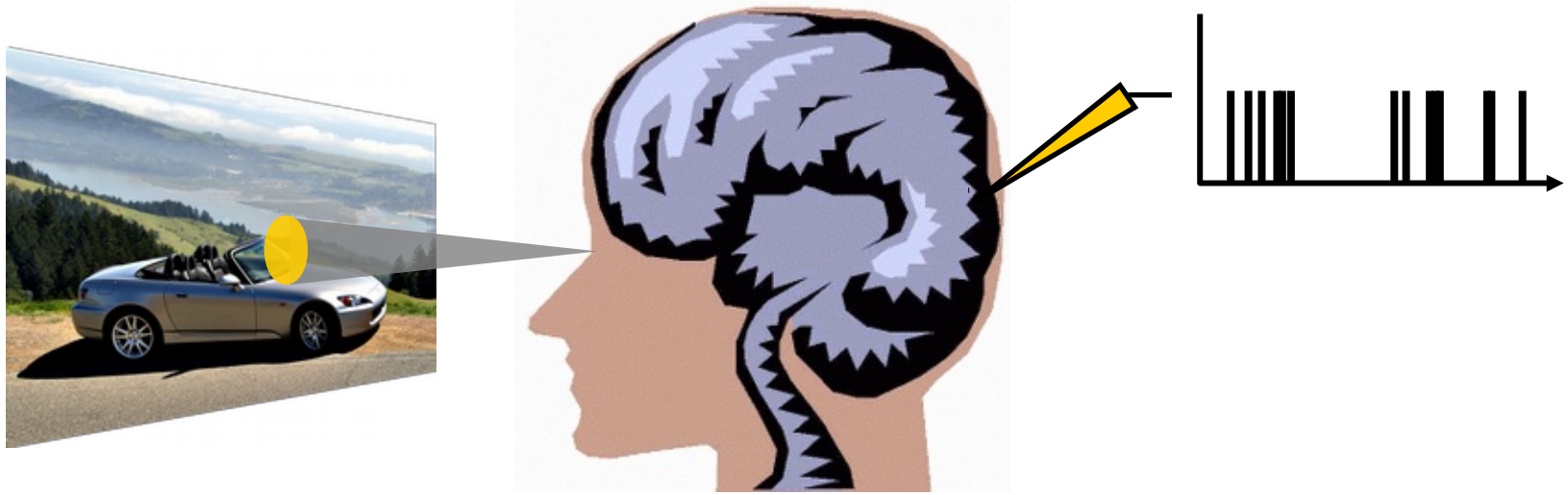
More activity



Nishimoto et al., *Current Biology*, 2011; Huth et al., *Neuron*, 2012;
Data visualization by James Gao's pyCortex (http://pycortex.org)
Online visualization available at http://gallantlab.org


How complicated is the brain?
How much data will we need to model it?
How can we model the system?
How do we deal with non-stationarity?
What about individual differences?
Can we do brain decoding?
Can we get better data?

Neuroscience as system identification

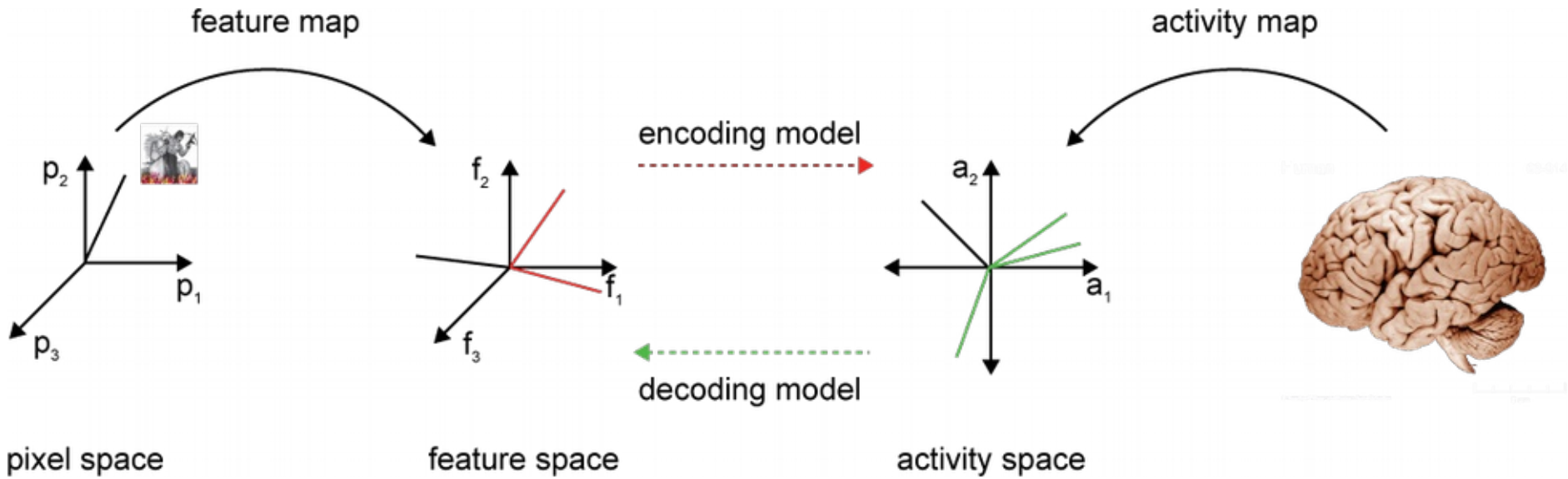


$$s(x,y,t) \longrightarrow \boxed{H(x,y,t,a)} \longrightarrow r(t)$$

The amount of data required to fit the model
Is proportional to the degree of nonlinearity.

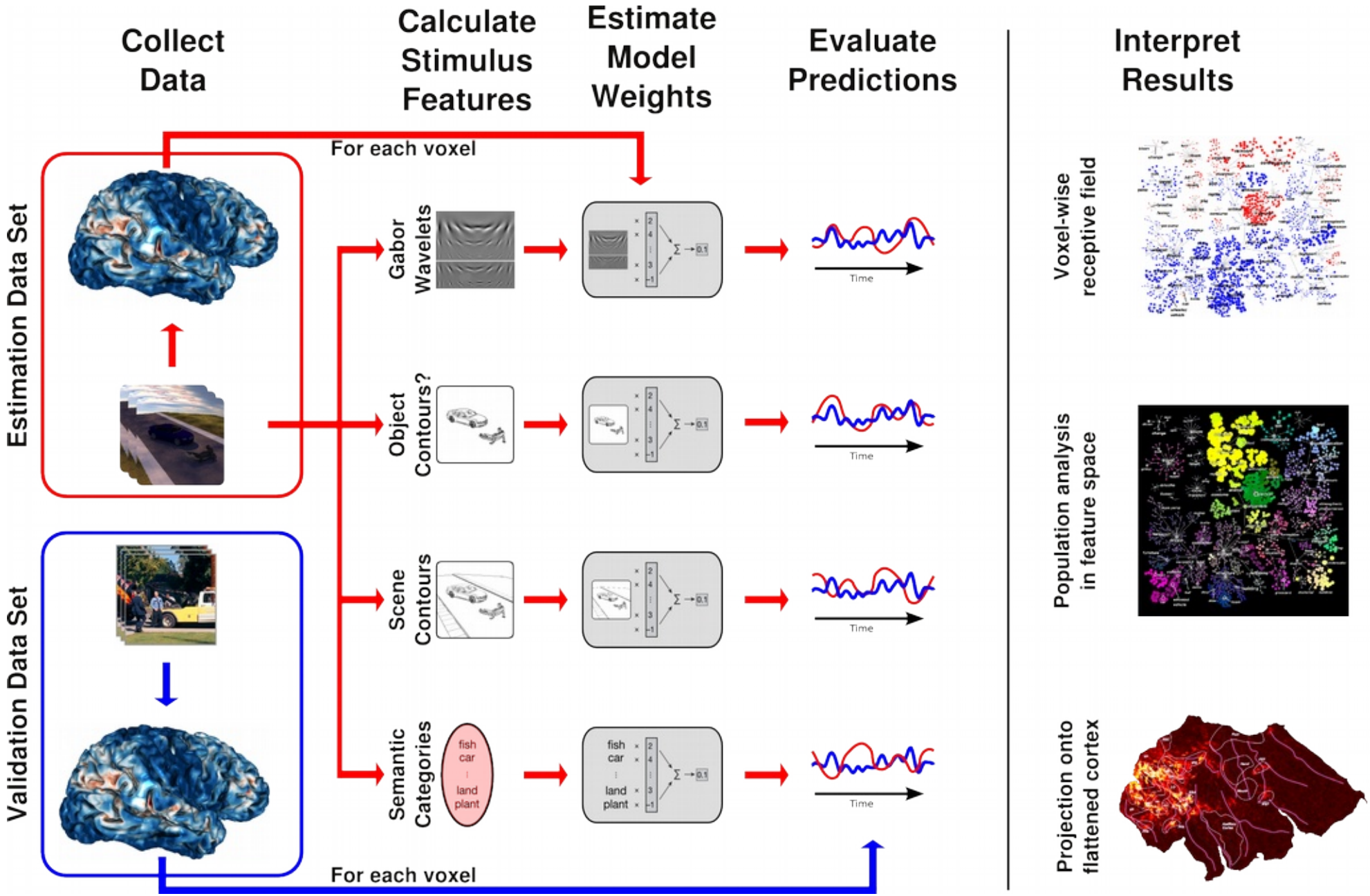
State variables (e.g. Attention) multiply
The dimensionality of the problem.

The explicit feature space approach



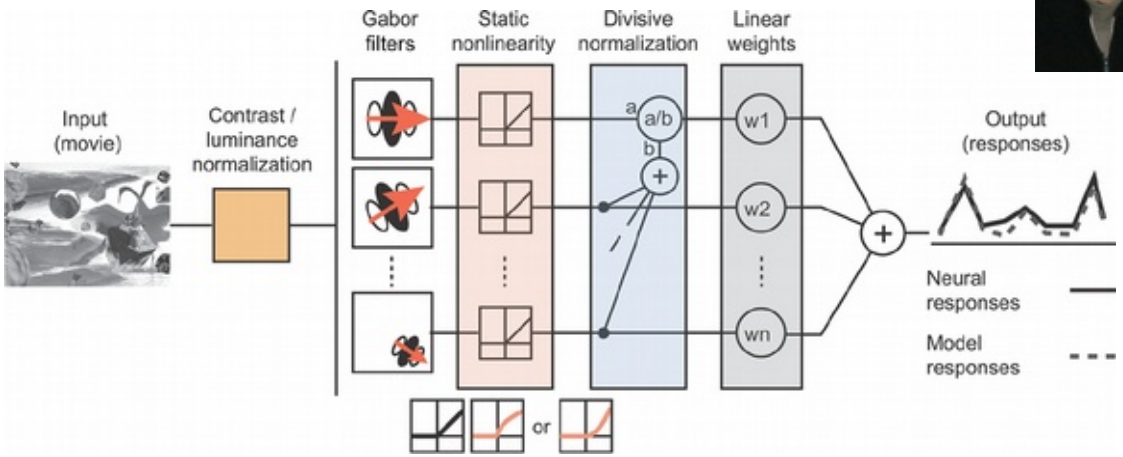
Nonlinear transformation	**Linear transformation**	**Linear or Nonlinear transformation**
Hypothesis	**Linear regression**	**Experimental measurement**

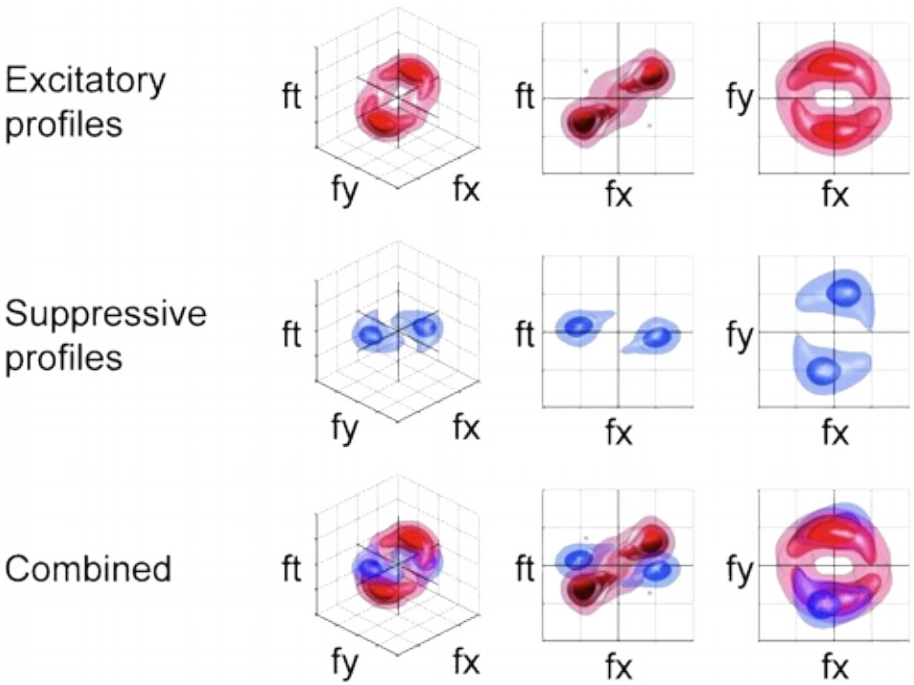
Modeling multiple, explicit feature spaces

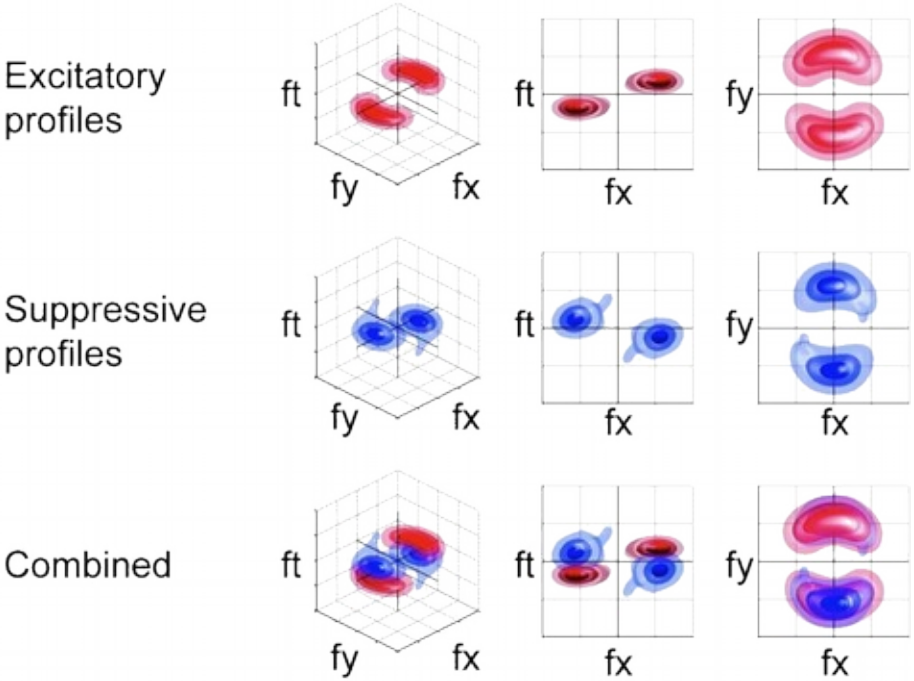


Explicit modeling of MT neurons

Shinji Nishimoto



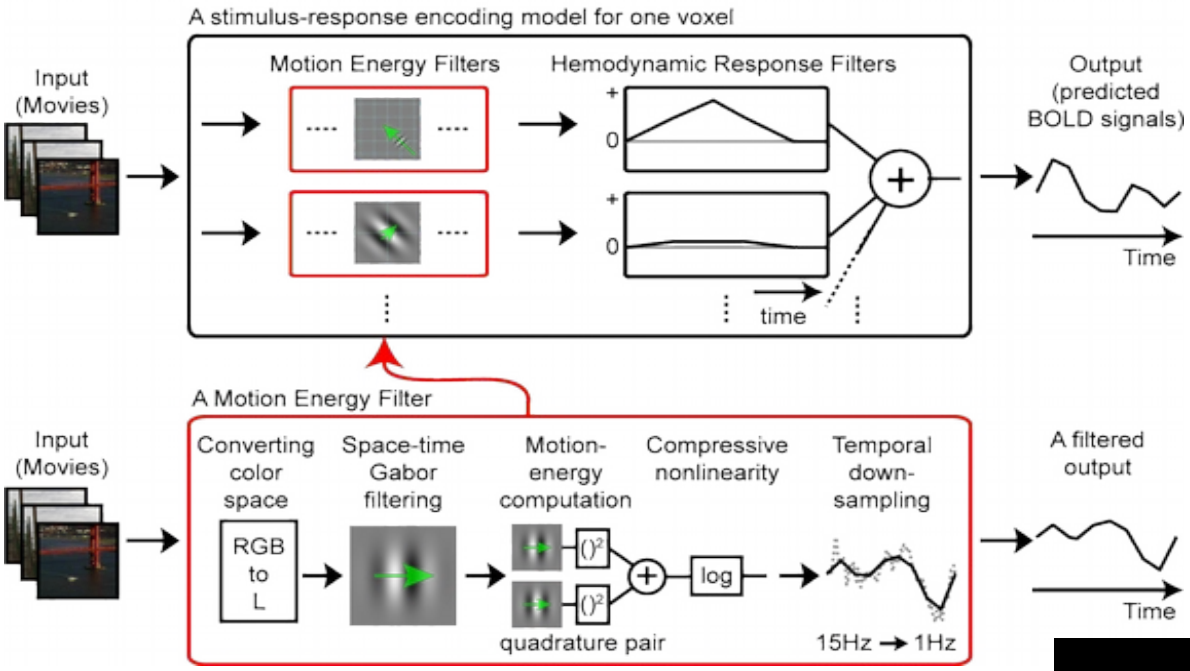


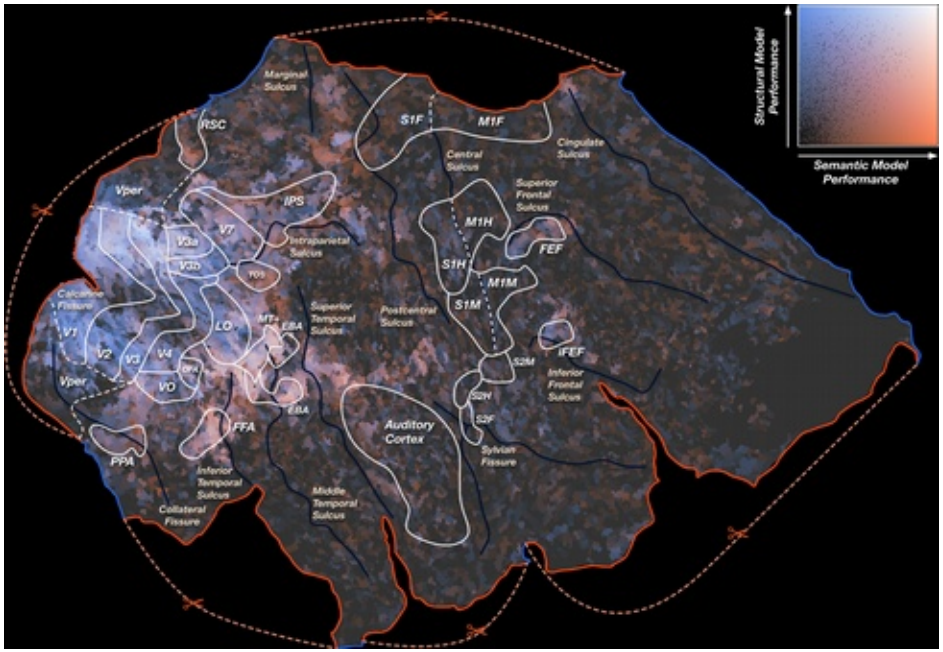


Nishimoto & Gallant, *Neuron*, 2011

A motion-energy model for low-level vision

Shinji Nishimoto





Nishimoto, Vu, Naselaris, Benjamini, Yu and Gallant, *Current Biology*, 2011

A semantic model for high-level vision

Alex Huth

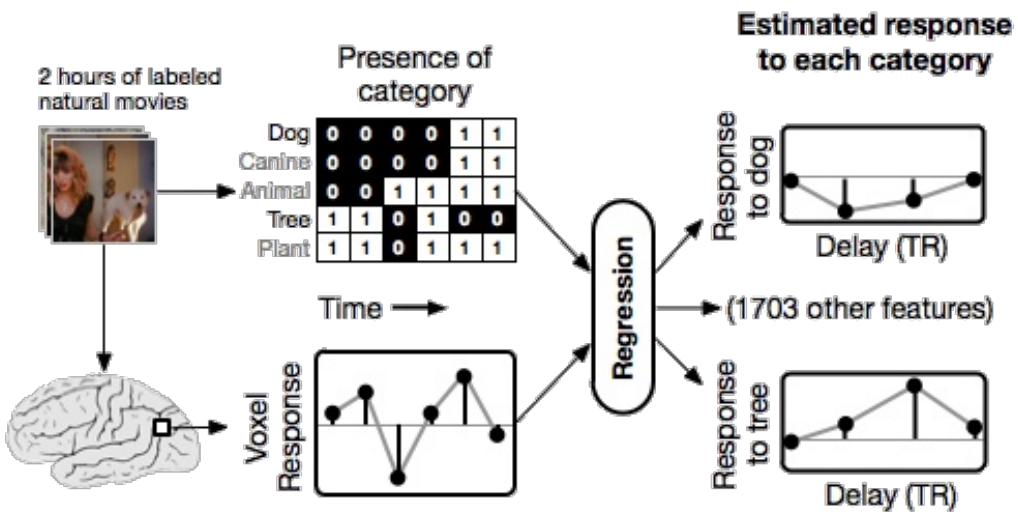
Labels:

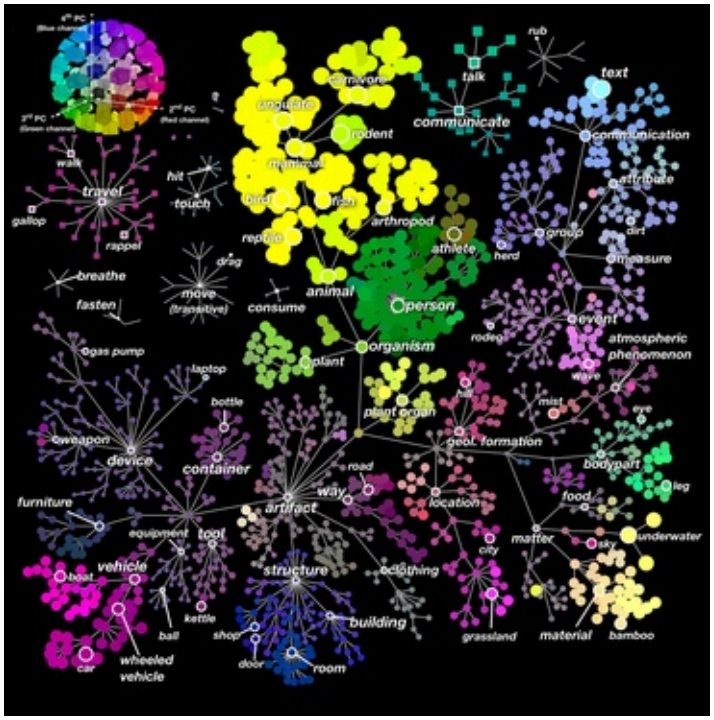
manually assigned label	inferred labels
dog	woman
canine	adult
animal	person
⋮	⋮

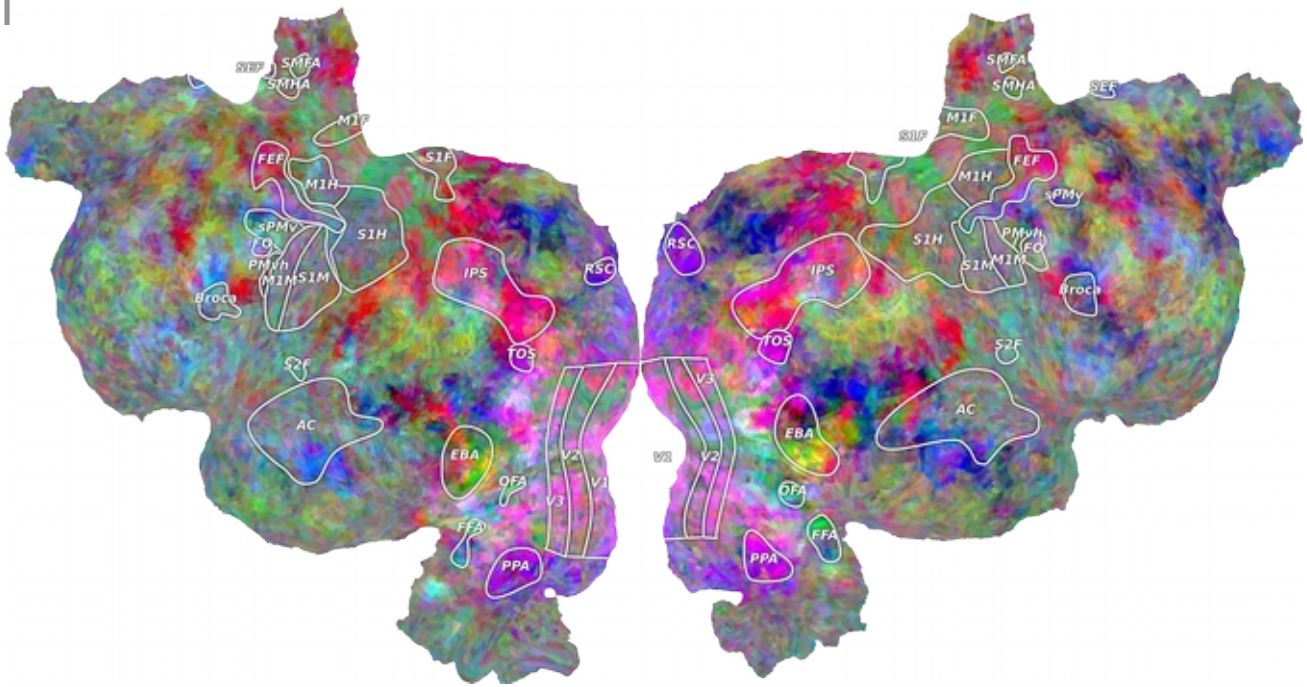
Video clips labeled every 1s

Labels:

jaguar	tree
feline	vasc. pl.
animal	plant
organism	organism
object	object
⋮	⋮

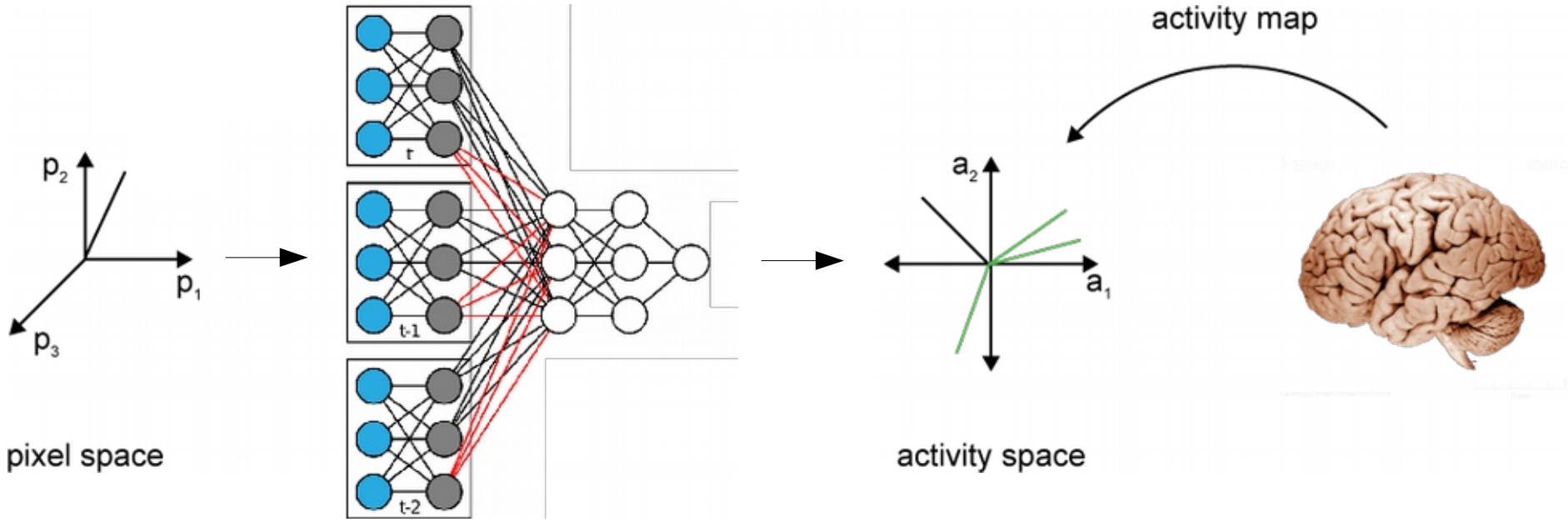






Huth, Nishimoto, Vu & Gallant, *Neuron*, 2012

The machine learning approach



Nonlinear transformation

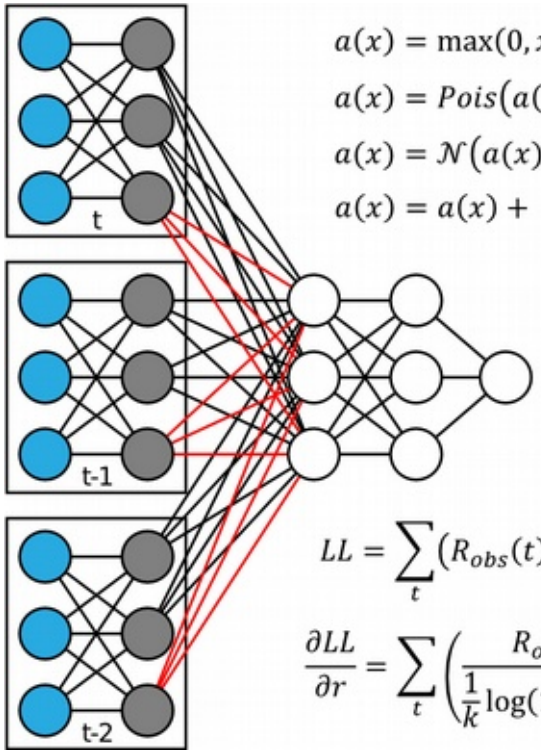
Linear or Nonlinear transformation

Learning algorithm or regression algorithm

Experimental measurement

A deep network for V2 neurons

Mike Oliver



$$a(x) = \max(0, x)$$

$$a(x) = Pois\big(a(x)\big)$$

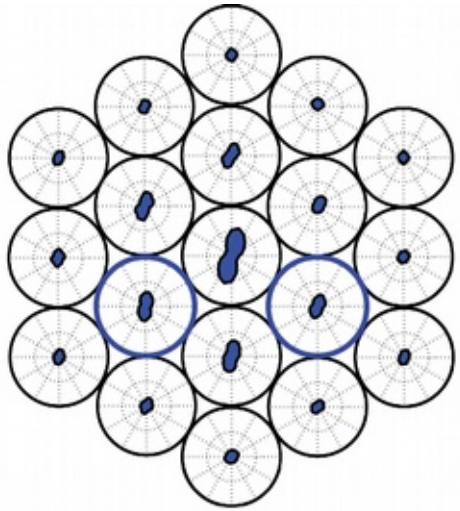
$$a(x) = \mathcal{N}\big(a(x), a(x)\big)$$

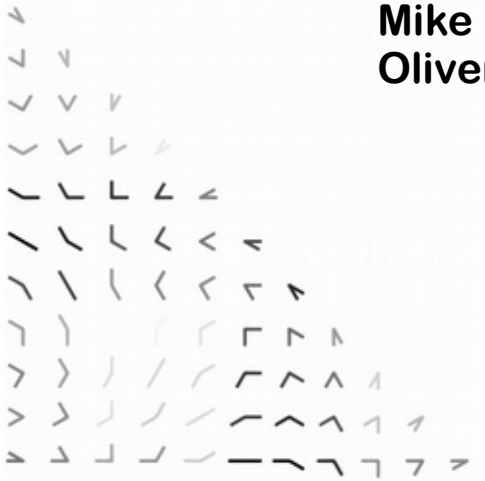
$$a(x) = a(x) + \sqrt{a(x)}\,\mathcal{N}(0,1)$$

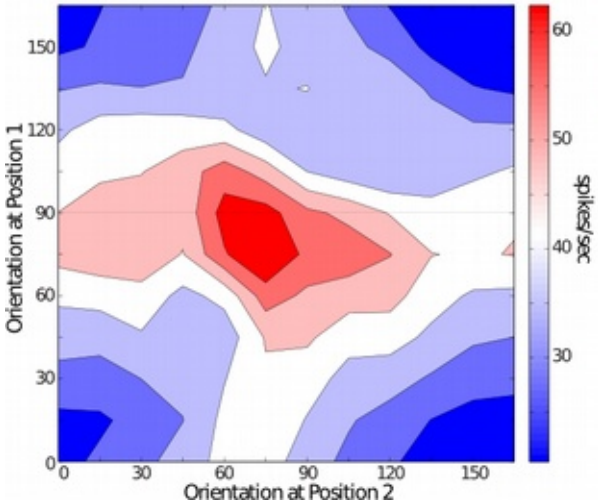
$$r(x) = \frac{1}{k}\log\big(1 + e^{k\cdot x}\big)$$

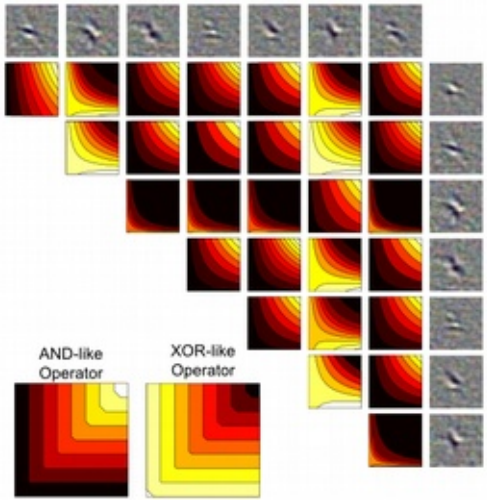
$$LL = \sum_t \big(R_{obs}(t)\log\big(r(t)\big) - r(t)\big)$$

$$\frac{\partial LL}{\partial r} = \sum_t \left(\frac{R_{obs}(t)}{\frac{1}{k}\log(1 + e^{k\cdot x})} - 1\right)\left(\frac{e^{k\cdot x}}{1 + e^{k\cdot x}}\right)$$





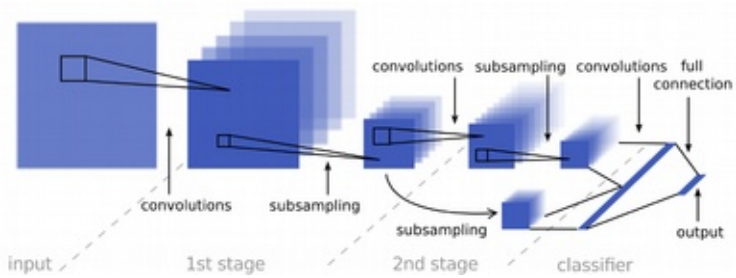




Deep time-delay network trained with DropOut and Poisson loss

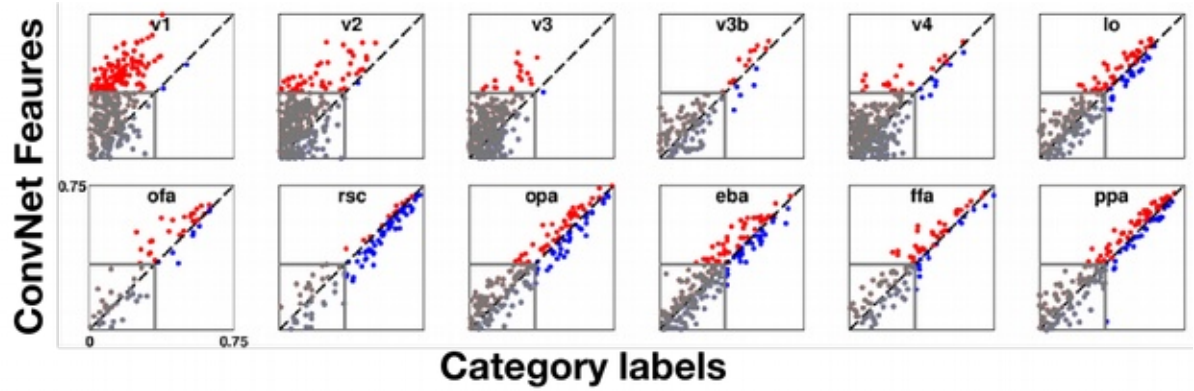
Oliver & Gallant, *in preparation*

A convolutional network for fMRI data



Note: The CNN used here contained 5 convolutional layers and 2 classifier layers.

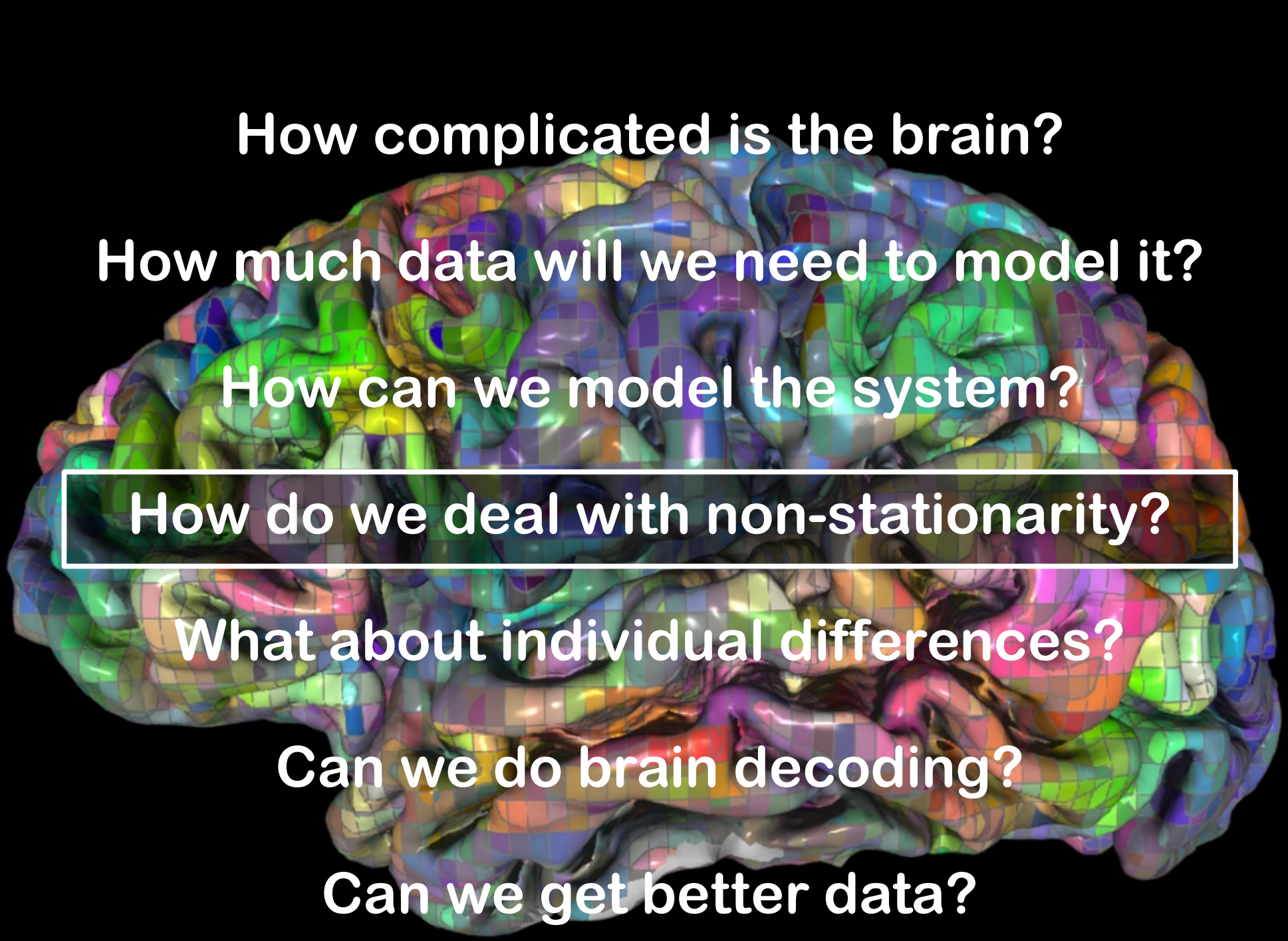
Pulkit Agrawal



Images predicted to increase BOLD

Images predicted to decrease BOLD

Agrawal, Cheung, Stansbury, Malik & Gallant, *in preparation*



How complicated is the brain?

How much data will we need to model it?

How can we model the system?

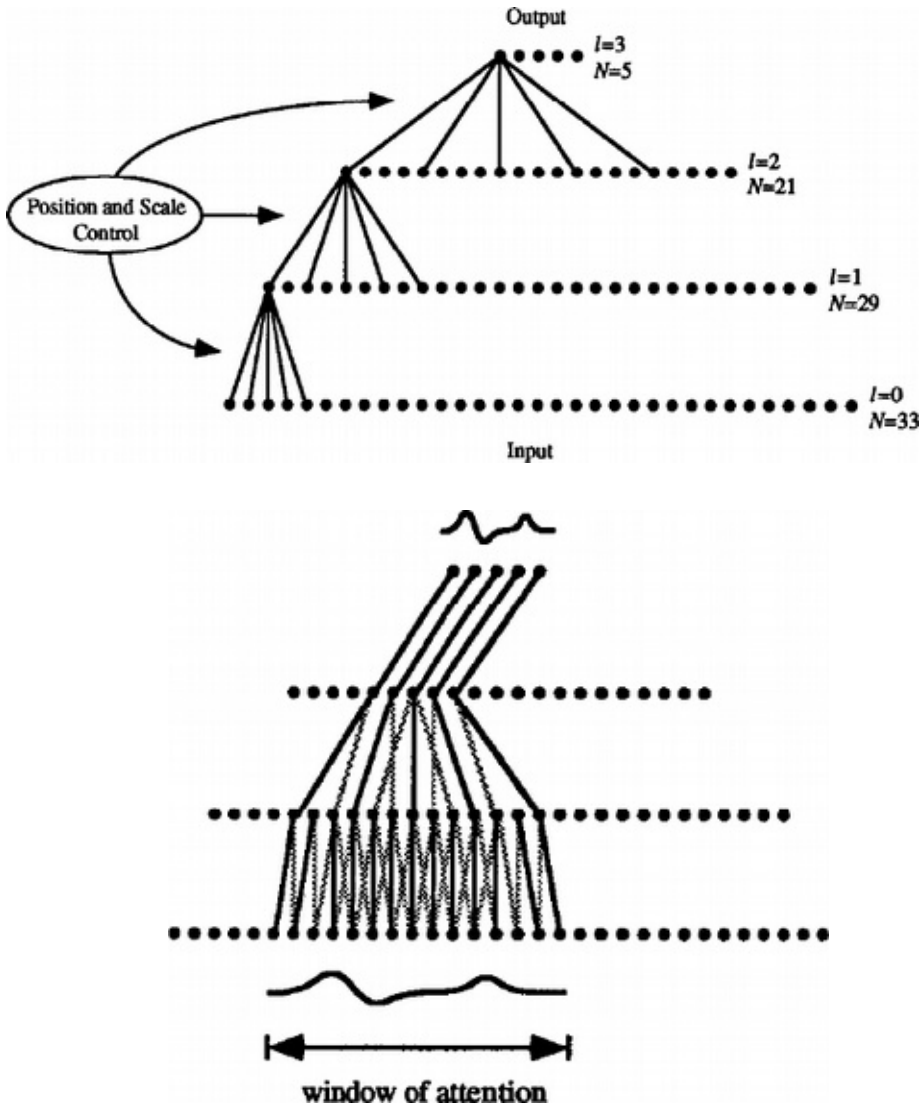
How do we deal with non-stationarity?

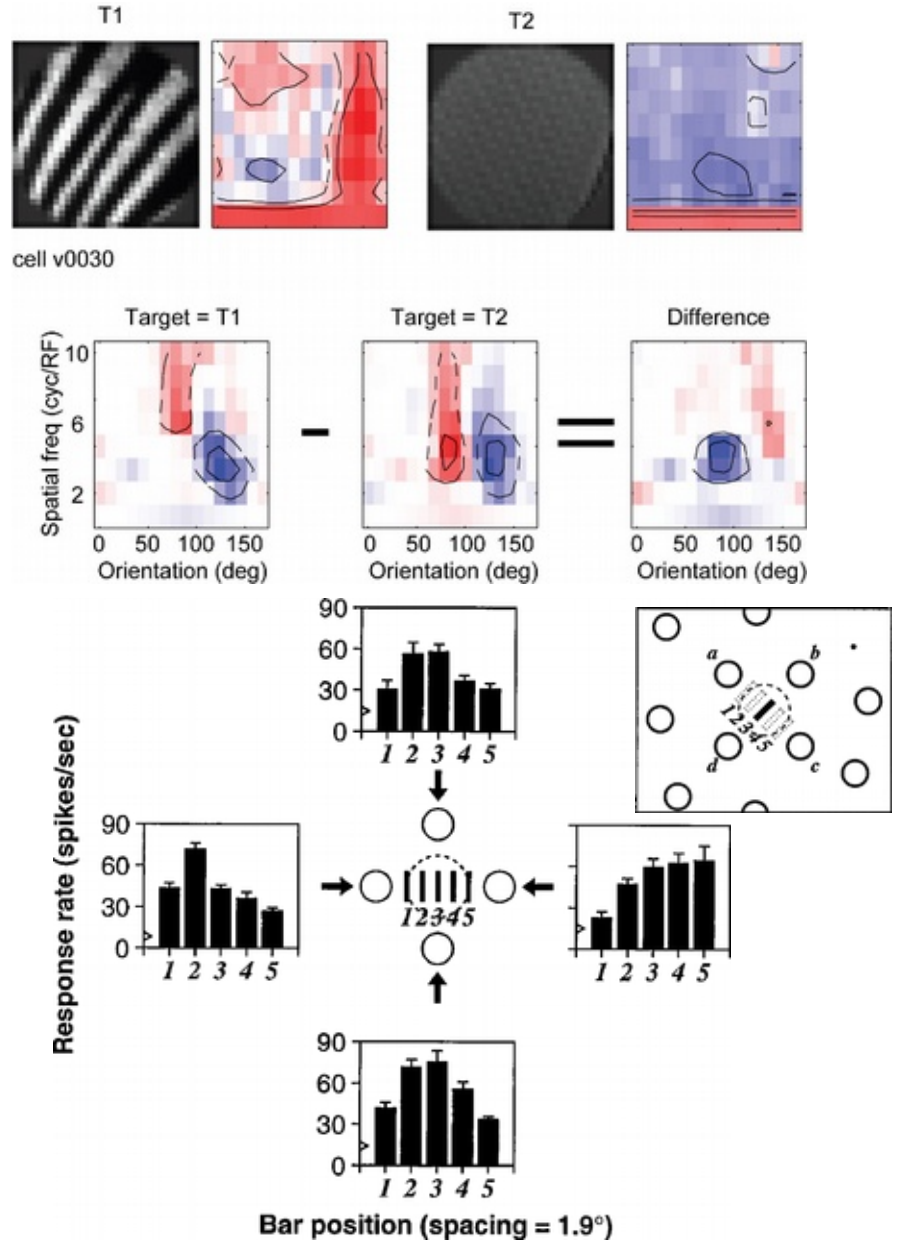
What about individual differences?

Can we do brain decoding?

Can we get better data?

Attention alters tuning in single neurons

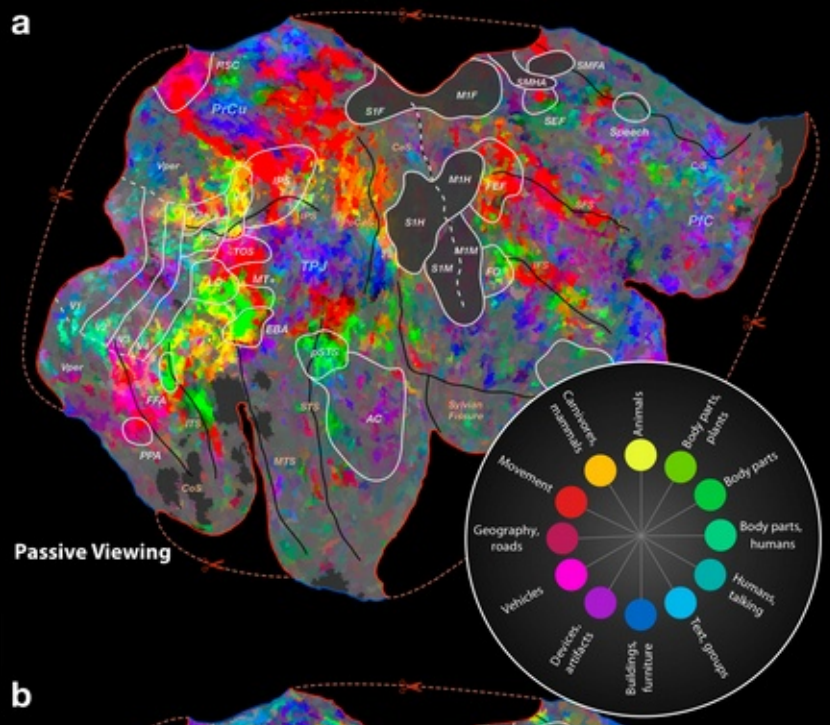


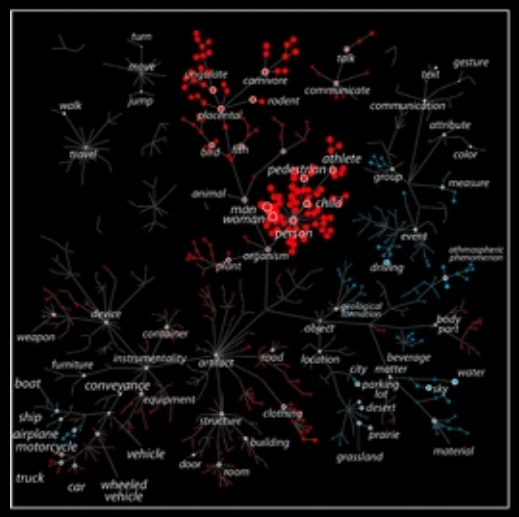


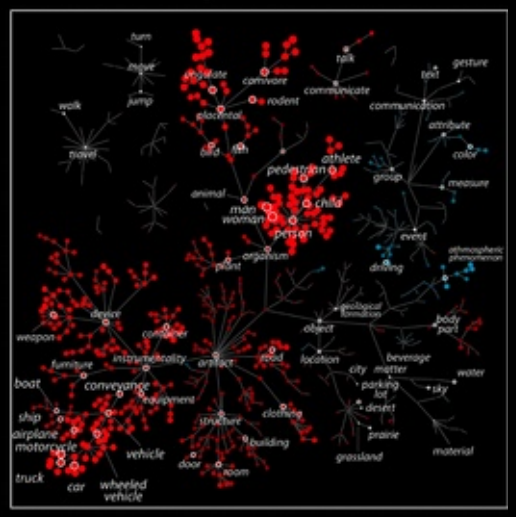
Olshausen, Anderson & Van Essen, *J. Neurosci.*, 1993

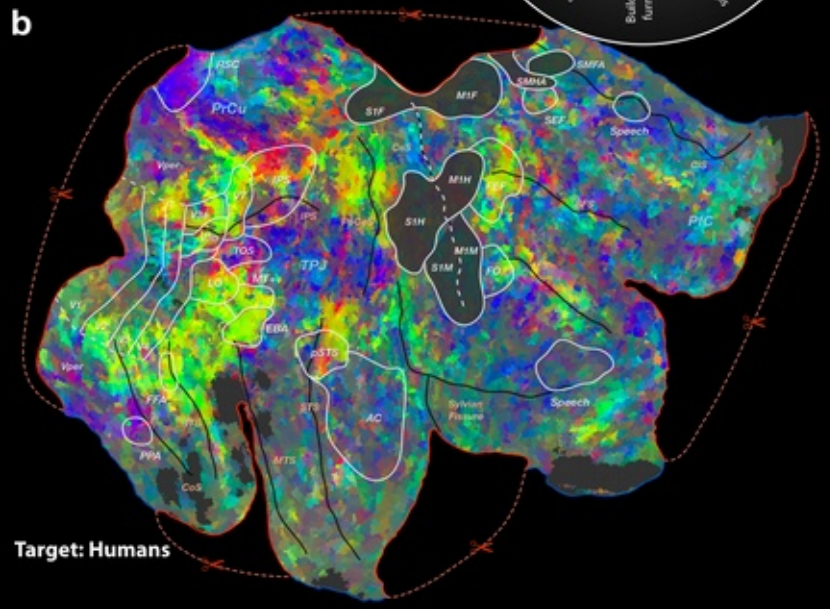
Connor, Gallant, Preddie & Van Essen, J. Neurophys., 1996
David, Mazer, Hayden & Gallant, Neuron, 2008

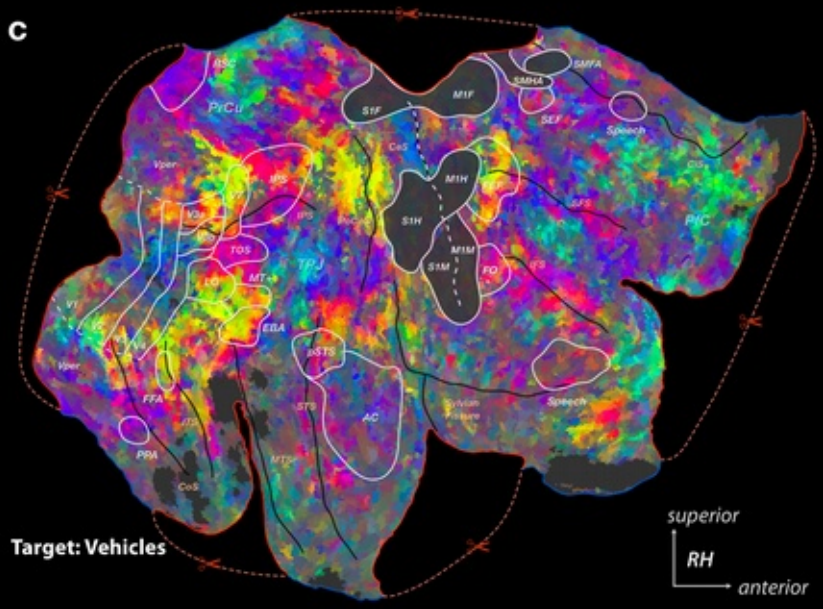
Attention changes cortical semantic maps





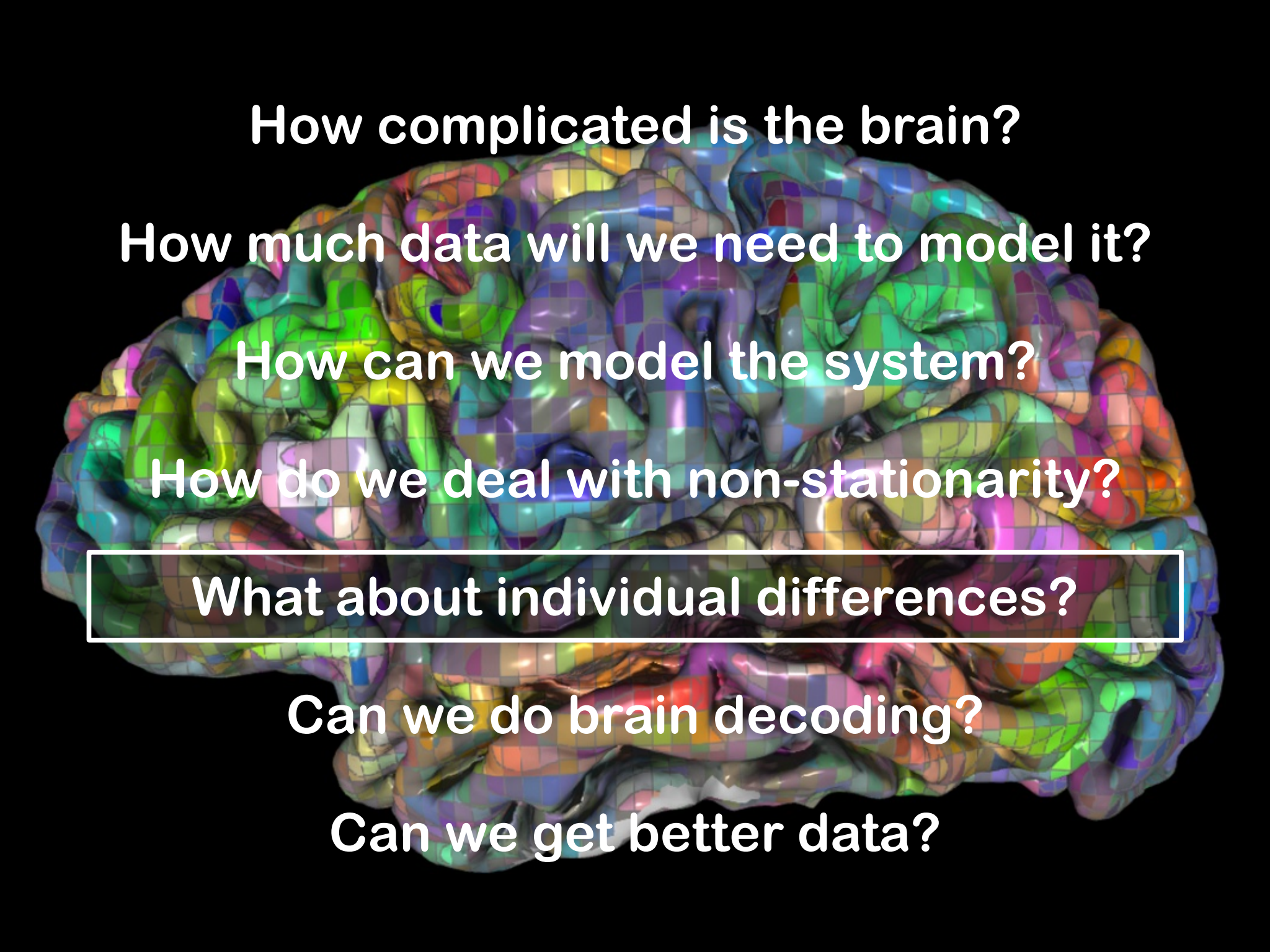




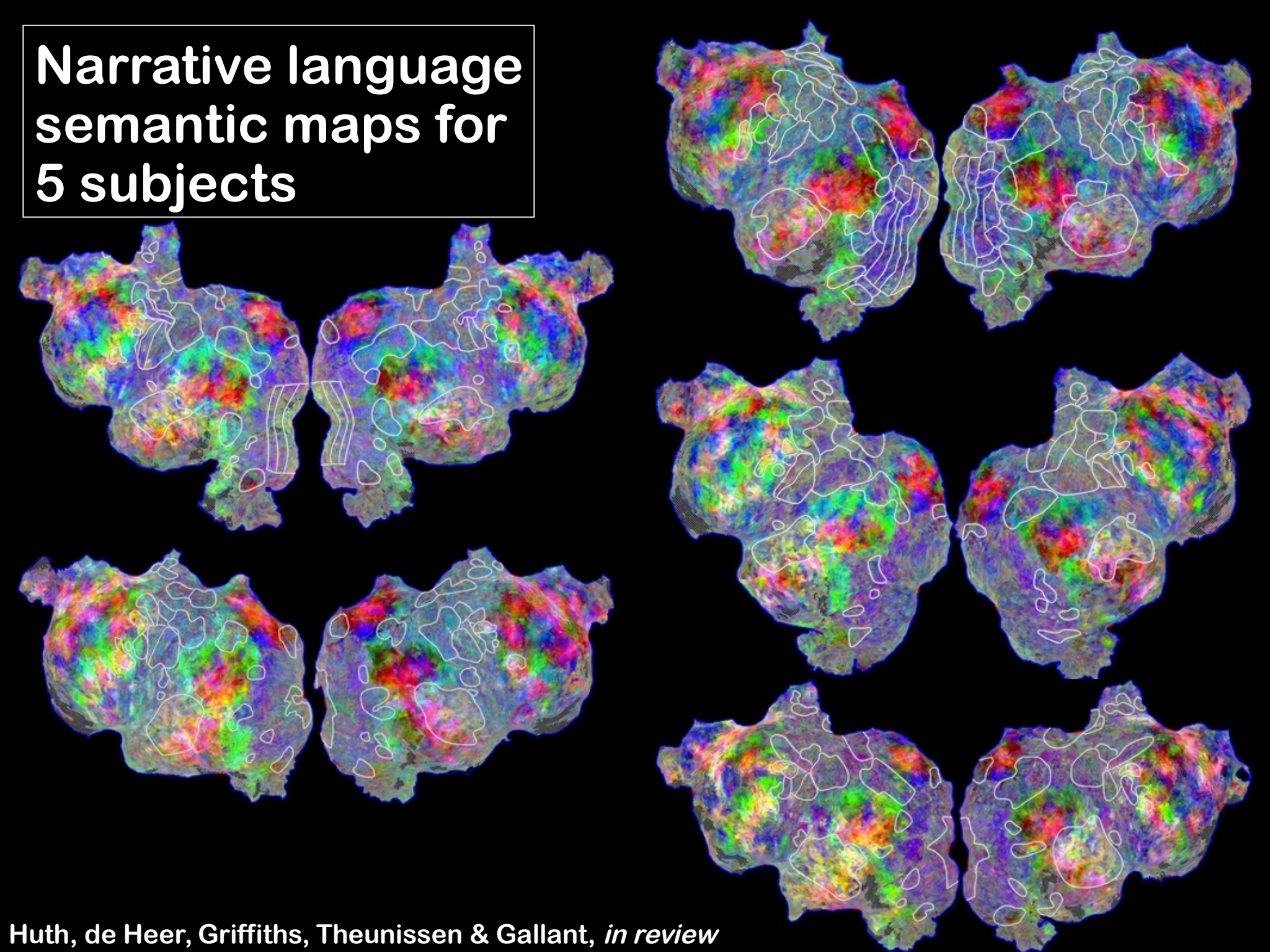


Tolga Cukur

Cukur, Nishimoto, Huth & Gallant, *Nature Neuroscience*, 2013

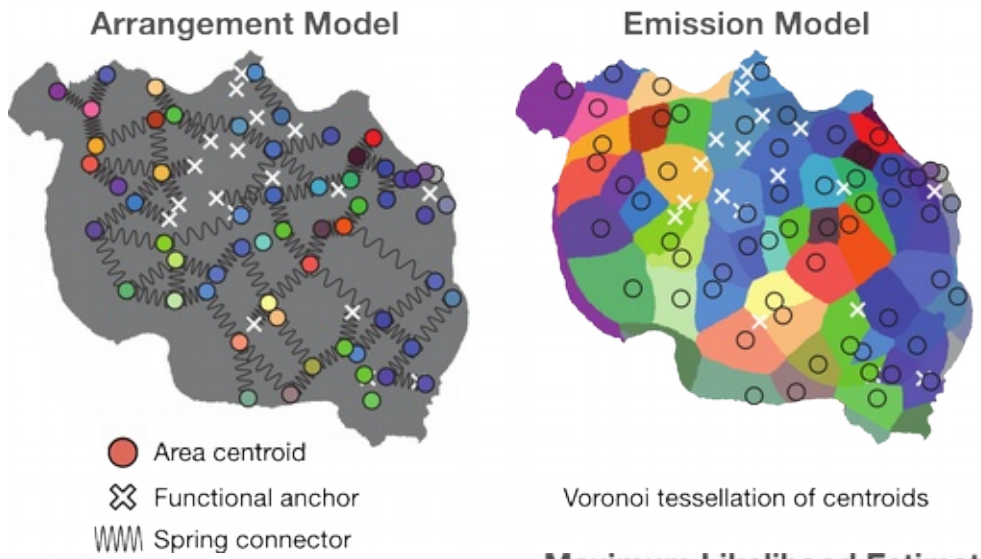

How complicated is the brain?
How much data will we need to model it?
How can we model the system?
How do we deal with non-stationarity?
What about individual differences?
Can we do brain decoding?
Can we get better data?

Narrative language semantic maps for 5 subjects



Huth, de Heer, Griffiths, Theunissen & Gallant, *in review*

PrAGMATiC: Probabilistic And Generative Model of Areas Tiling Cortex



PrAGMATiC Parameters

Shared across subjects

- ***L*** - ideal spring lengths
- ***K*** - spring constants
- ***M*** - area functional means

Unique to each subject

- ***H*** - exact centroid locations
- ***V*** - functional values on cortex

$$P(V, H; M, L, K) = P(H; L, K)\, P(V|H; M)$$

total probability = *arrangement probability* * *map probability*

$$P(H; L, K) \propto e^{-E(H;L,K)}$$

$$E(H; L, K) = \sum_{i,j,s} K_{ij}(d_{ijs} - L_{ij})^2$$

distance between centroids i and j in subject s

$$P(V|H; M) \propto e^{-E(V|H;M)}$$

$$E(V|H; M) \propto \sum_{l,s} (M_{H(s,l)} - V_{ls})^2$$

nearest centroid in subject s

Maximum Likelihood Estimation of Parameters

1. *Resample centroid locations*

$$H_j \sim P(H; L, K)$$

2. *Compute likelihood of observed functional values*

$$P(V_{obs}|H_j; M)$$

3. *Compute gradients for **K**, **L**, and **M***

$$\frac{\partial \mathscr{L}(K; V_{obs})}{\partial K} = \left\langle \frac{\partial E_H}{\partial K} \right\rangle_H - \left\langle \frac{\partial E_H}{\partial K} \right\rangle_{V_{obs}|H}$$

$$\frac{\partial \mathscr{L}(L; V_{obs})}{\partial L} = \left\langle \frac{\partial E_H}{\partial L} \right\rangle_H - \left\langle \frac{\partial E_H}{\partial L} \right\rangle_{V_{obs}|H}$$

$$\frac{\partial \mathscr{L}(M; V_{obs})}{\partial M} = -\left\langle \frac{\partial E_V}{\partial M} \right\rangle_{V_{obs}|H}$$

4. *Update parameters*

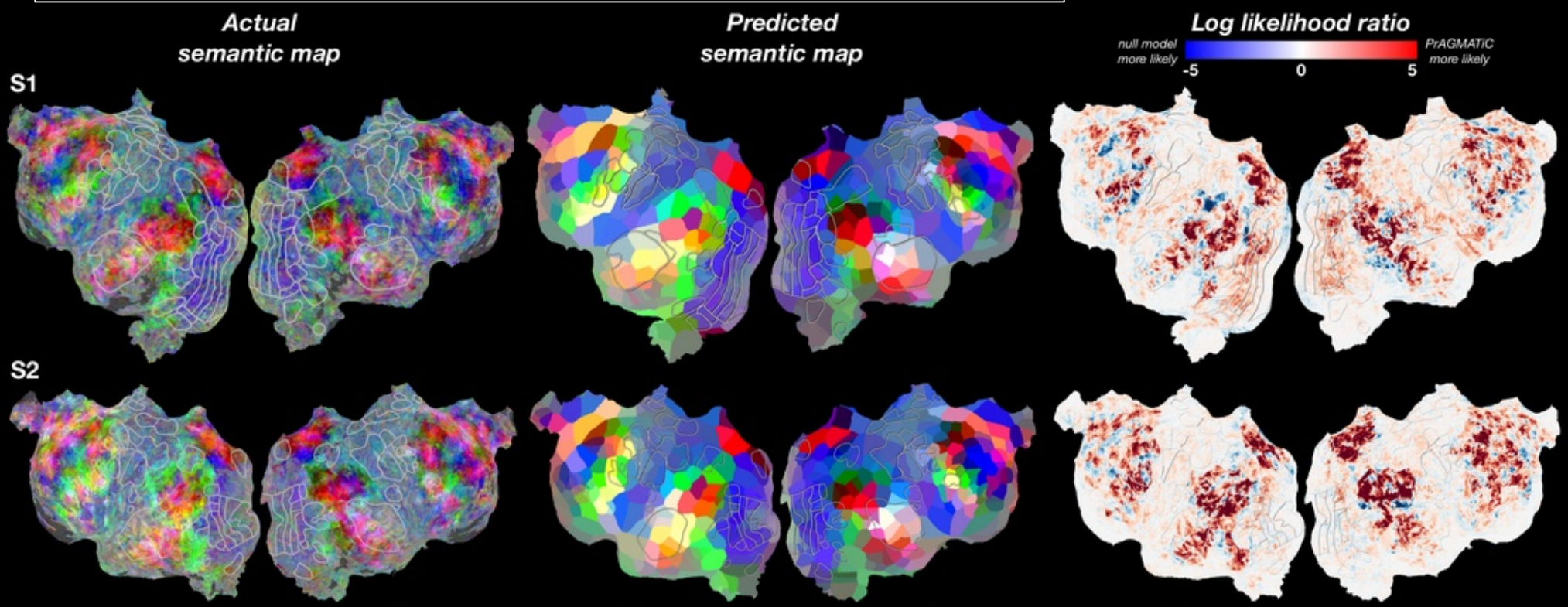
$$K^{t+1} = K^t - \epsilon \frac{\partial \mathscr{L}(K; V_{obs})}{\partial K}$$

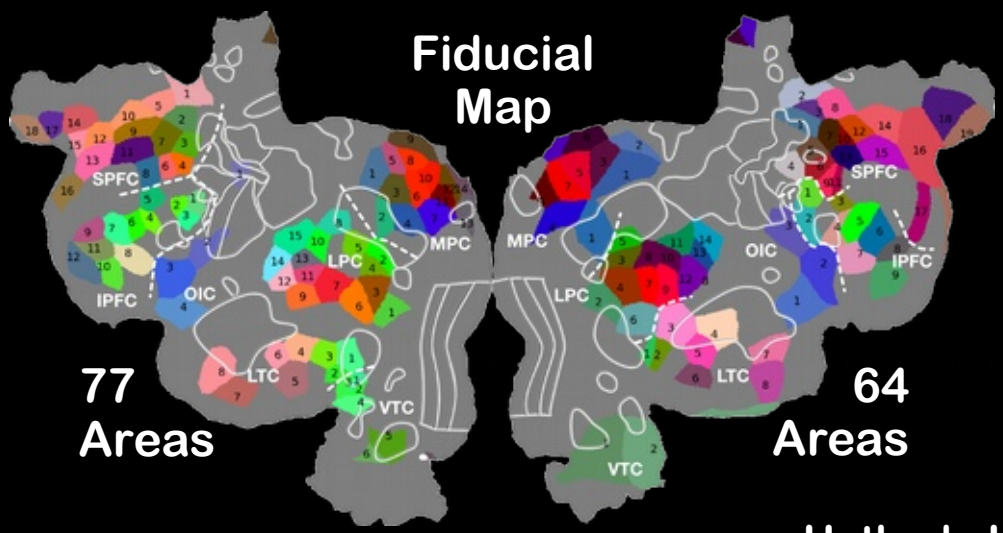
$$L^{t+1} = L^t - \epsilon \frac{\partial \mathscr{L}(L; V_{obs})}{\partial L}$$

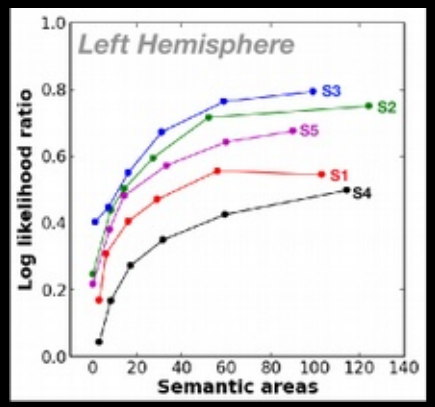
$$M^{t+1} = M^t - \epsilon \frac{\partial \mathscr{L}(M; V_{obs})}{\partial M}$$

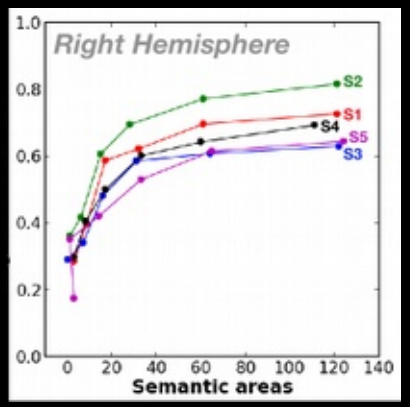
Huth, de Heer, Griffiths, Theunissen & Gallant, *in review*

Likelihood tests for PrAGMATiC

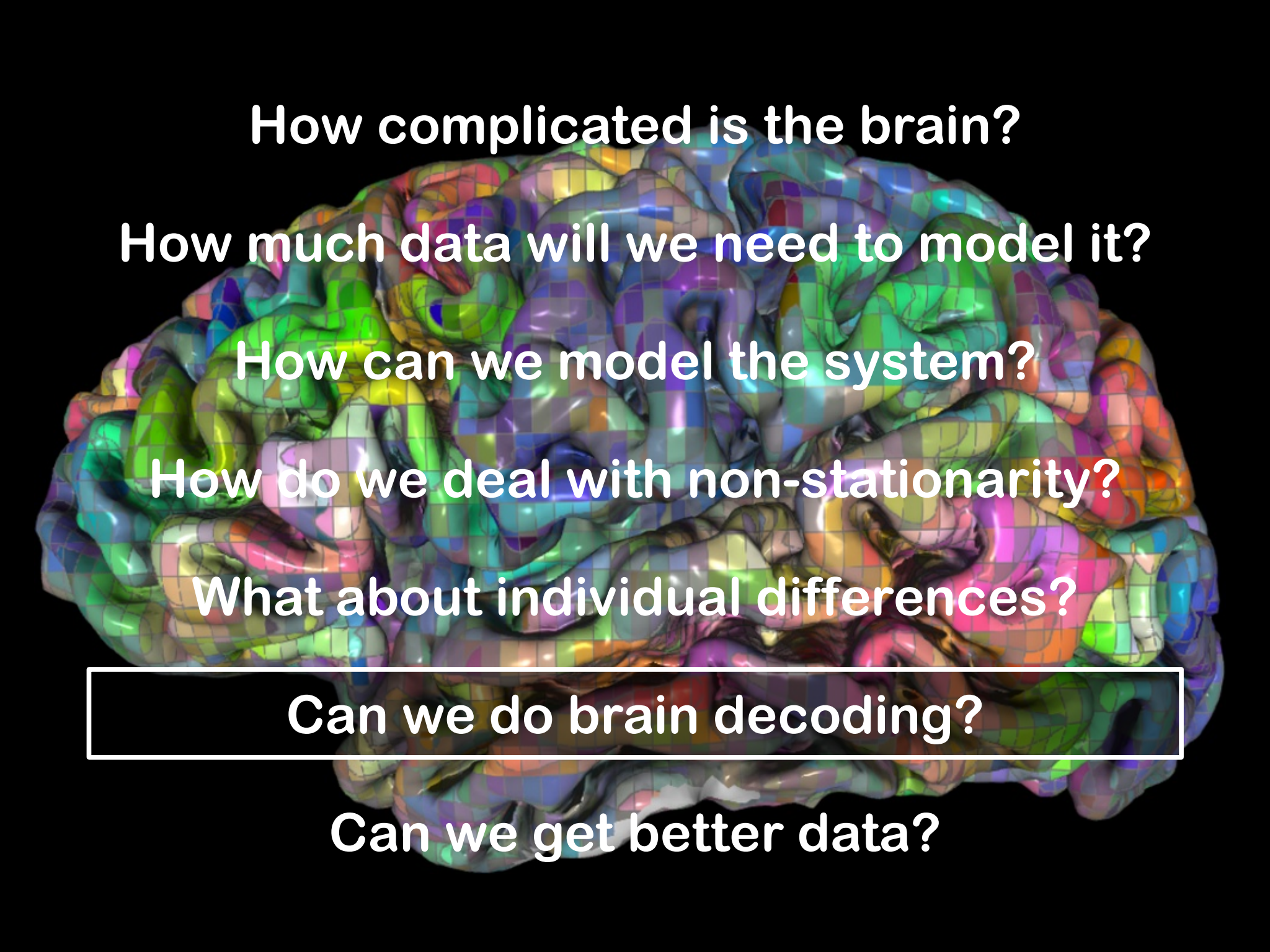




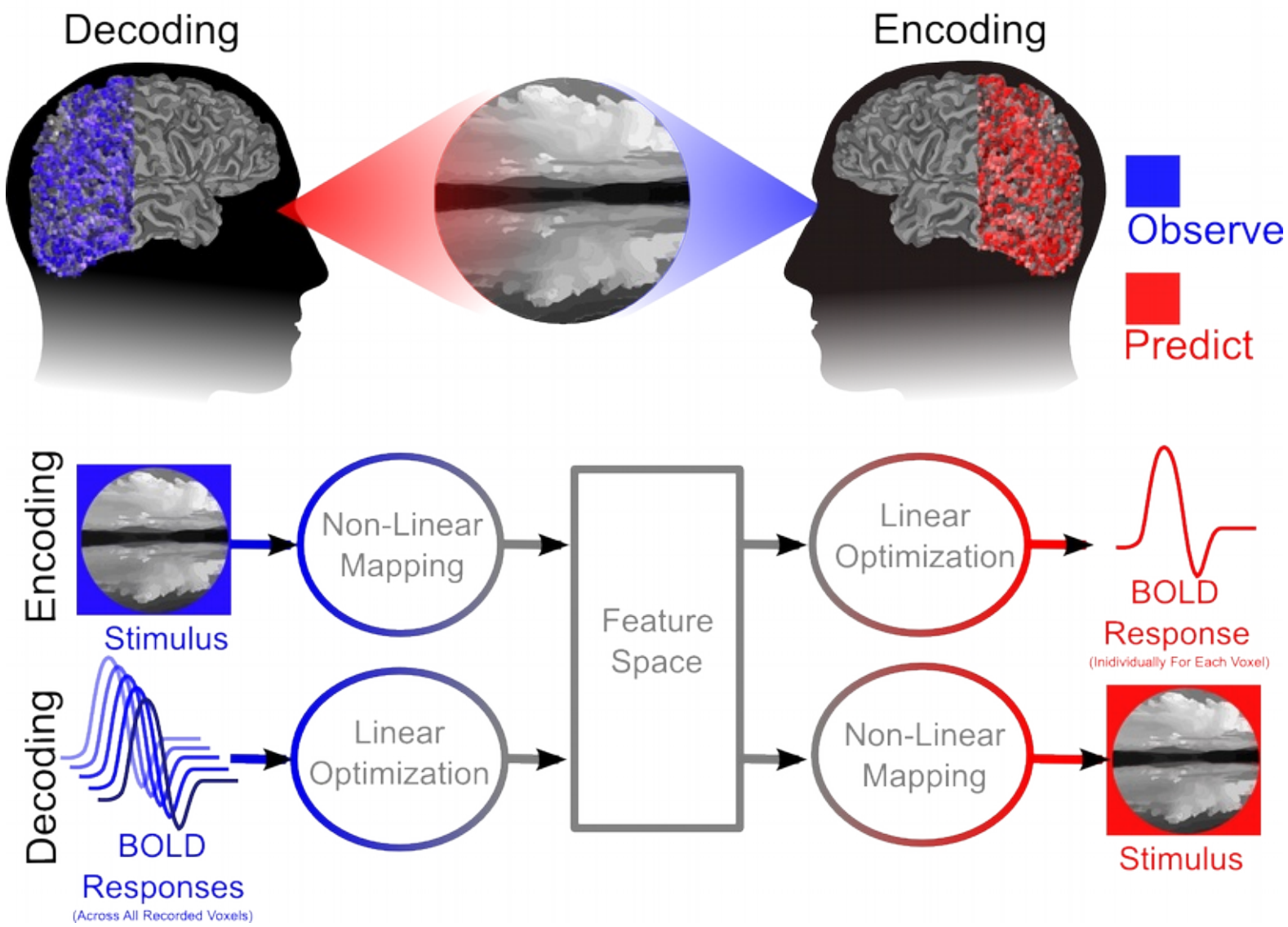




Huth, de Heer, Griffiths, Theunissen & Gallant, *in review*


How complicated is the brain?
How much data will we need to model it?
How can we model the system?
How do we deal with non-stationarity?
What about individual differences?
Can we do brain decoding?
Can we get better data?

Using encoding models to decode



$$P(f(S)|R) \propto P(R|f(S))\,P(f(S))$$

Decoding semantics from higher visual areas

Movie

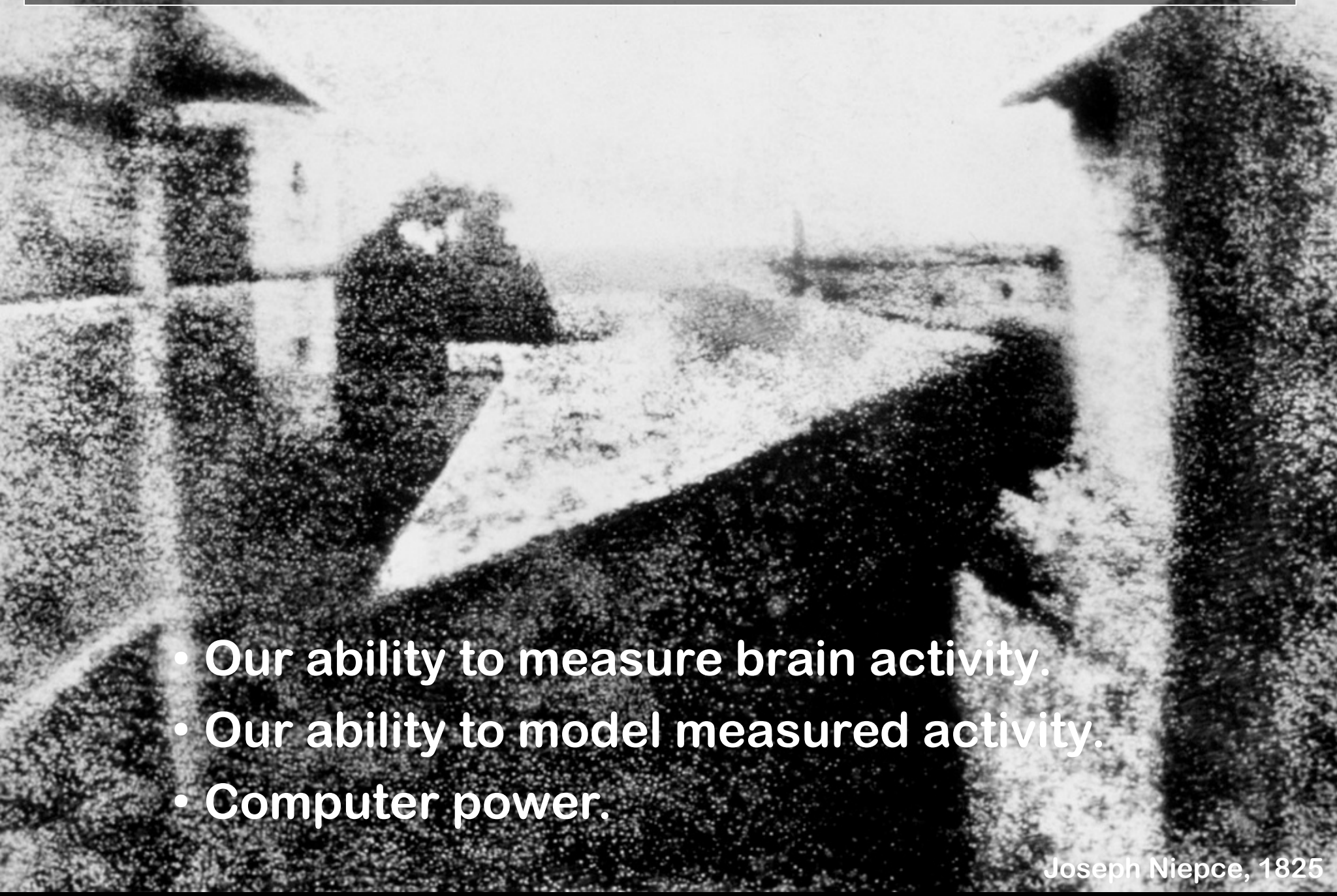
Likely
Objects and Actions

Huth, Lee, Nishimoto & Gallant, *in review*

You can decode anything that is represented in measured brain activity

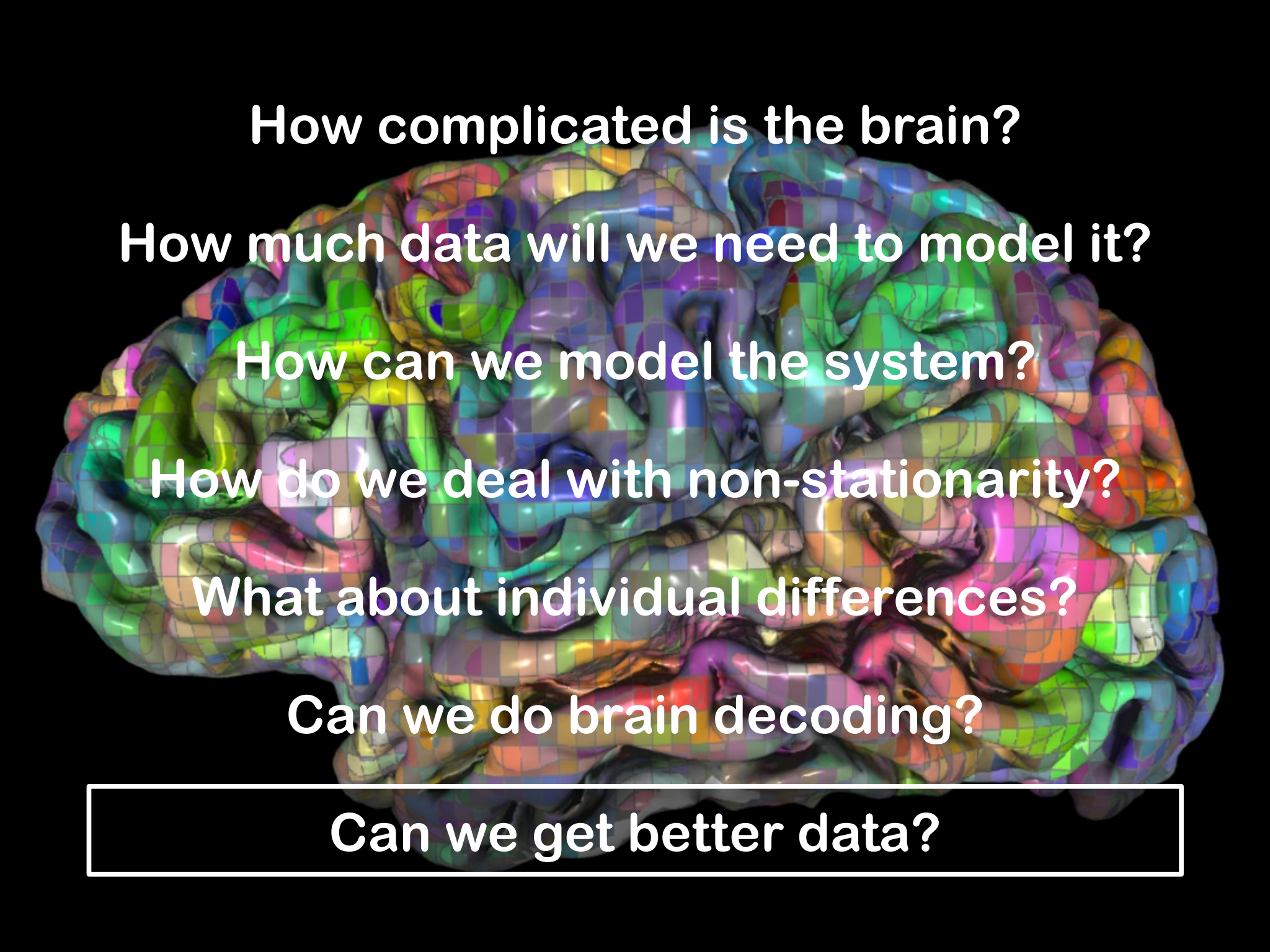
Gao & Gallant, *unpublished*

The current and future state of brain decoding



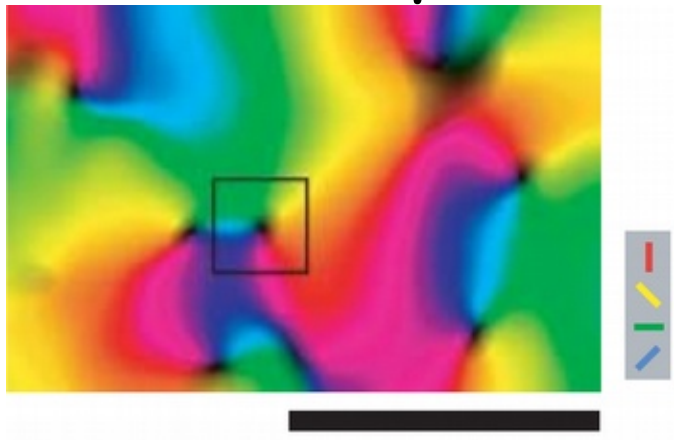
- Our ability to measure brain activity.
- Our ability to model measured activity.
- Computer power.

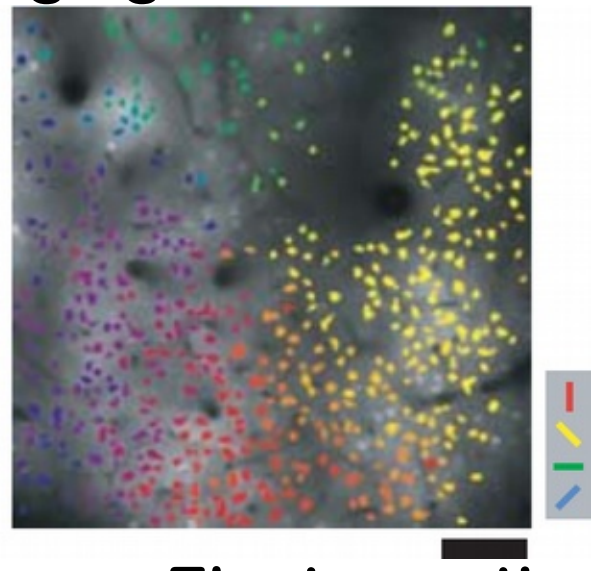
Joseph Niepce, 1825


How complicated is the brain?
How much data will we need to model it?
How can we model the system?
How do we deal with non-stationarity?
What about individual differences?
Can we do brain decoding?
Can we get better data?

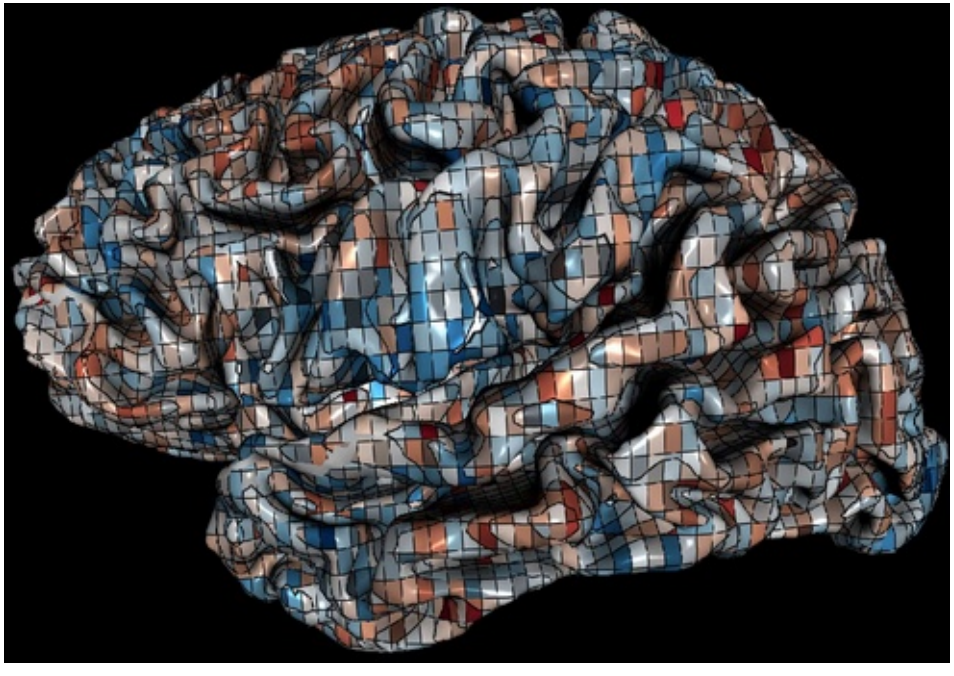
Improved methods for recording brain data

Optical imaging

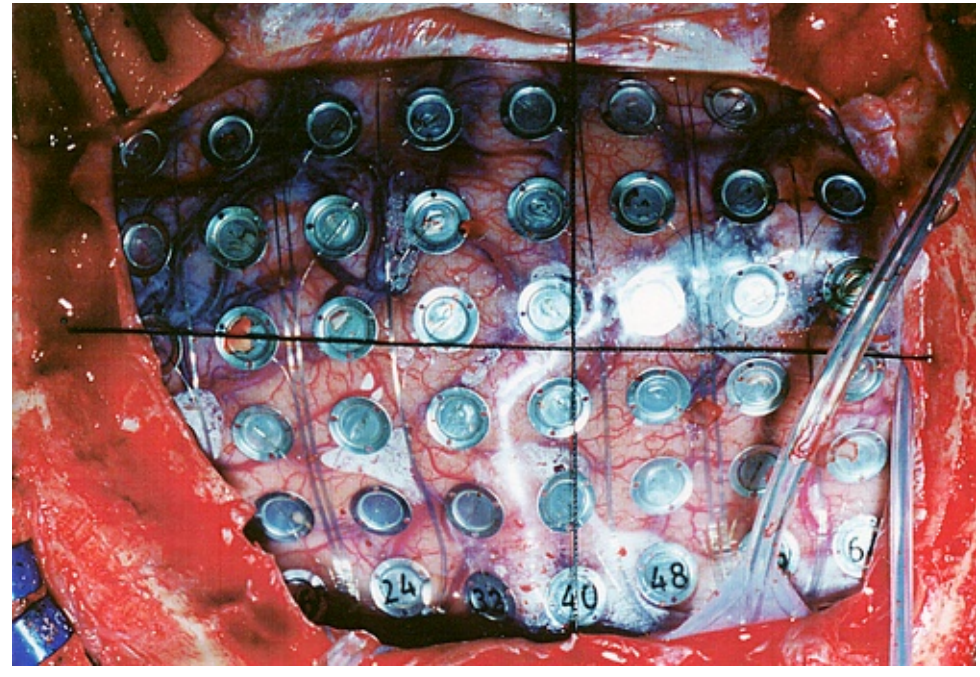




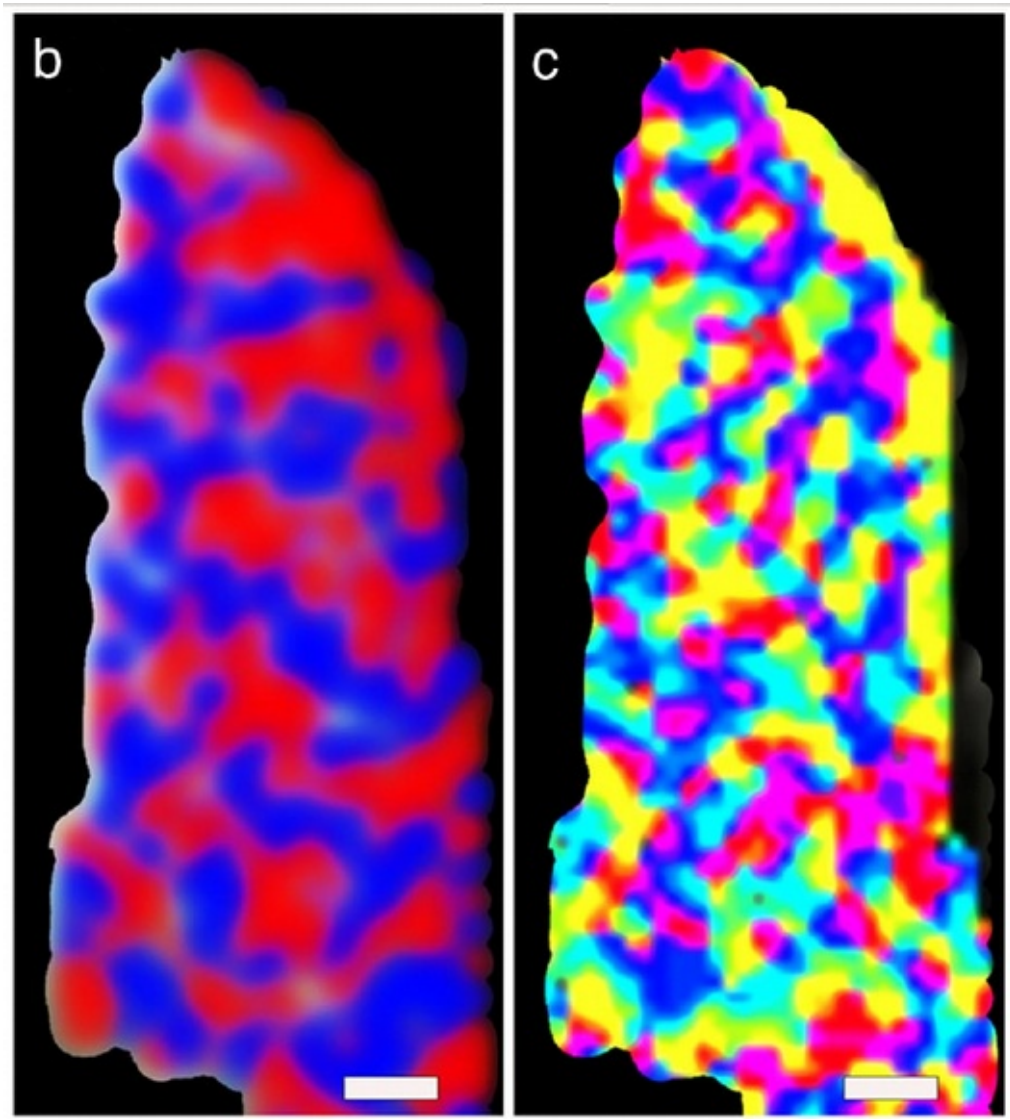
Functional MRI

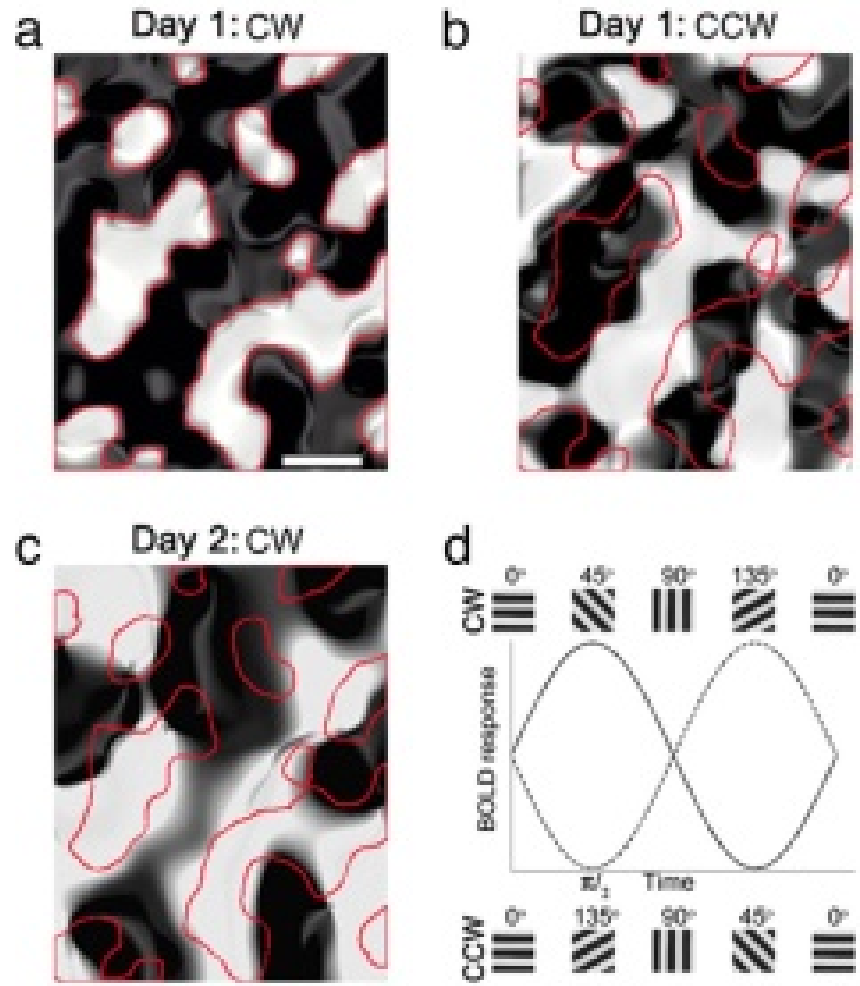


Electrocorticography



Trying to measure orientation columns





Yacoub, Harel & Ugurbil, *PNAS*, 2008

Summary and conclusions

The brain is very, very complicated structure.

Current methods for measuring the brain have limited spatial, temporal and conditional sampling.

To make progress we need to push the boundaries of spatial, temporal and conditional sampling.

If we do this then we can use a variety of computational methods to build robust computational models that predict accurately and generalize well.

When the resultant models perform poorly (which is likely), then we can focus on improving computational theory rather than mere neuroscience data collection.

Computer science versus neuroscience

We are entering the age of “Big Data” in neuroscience, and there is a pressing need for data processing algorithms that can operate on these meso-scale data.

The required algorithms fall into several domains: matrix (tensor) libraries, image processing algorithms, data mining tools, database tools and visualization tools.

The brain is a nonlinear dynamical system with feedback, but these are the least well understood sorts of systems from a theoretical point of view.

Therefore, even if we had a complete functional (or structural) map of the brain, we wouldn't know what to do with the data or how to model it!

This would be a good problem to have, and would likely drive theoretical developments in computer science and mathematics.

Current Lab Members

Natalia Bilenko
James Gao
Alex Huth
Fatma Imamoglu
Mark Lescroart
Lydia Majur
Anwar Nunez
Michael Oliver
Storm Slivcoff
Dustin Stansbury

Collaborating labs

Frederic Theunissen, Bin Yu,
Tom Griffiths, Cheryl Olman
Ralph Adolphs, Bertrand Thirion,
Essa Yacoub, Kamil Ugurbil
Peter Bandettini

Past Lab Members

Tolga Cukur (Bilkent U)
Stephen David (OHSU))
Kate Gustavsen (UTVM)
Kathleen Hansen (NIH)
Ben Hayden (URochester)
Kendrick Kay (WashU)
James Mazer (Yale)
Thomas Naselaris (MUSC)
Shinji Nishimoto (NIST Japan)
Ryan Prenger (LANL))
Ben Willmore (Cambridge)
Bill Vinje (Salk)
An Vu (UCB)
Michael Wu (Lithium Inc)

This work was supported by NEI, NIMH and NSF