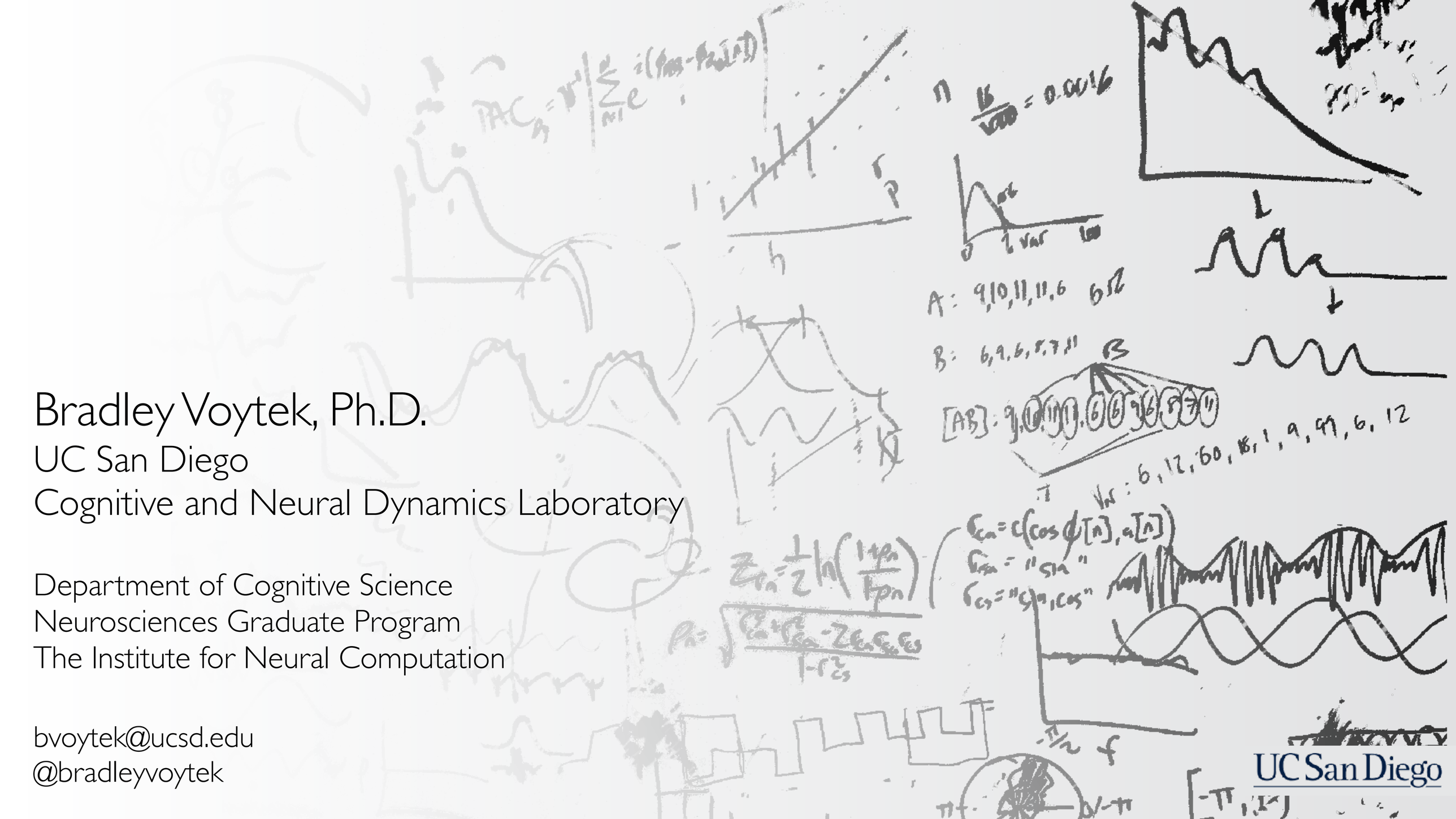


Bradley Voytek, Ph.D.

UC San Diego
Cognitive and Neural Dynamics Laboratory

Department of Cognitive Science
Neurosciences Graduate Program
The Institute for Neural Computation

bvoytek@ucsd.edu
@bradleyvoytek

UC San Diego

Sections and homeworks

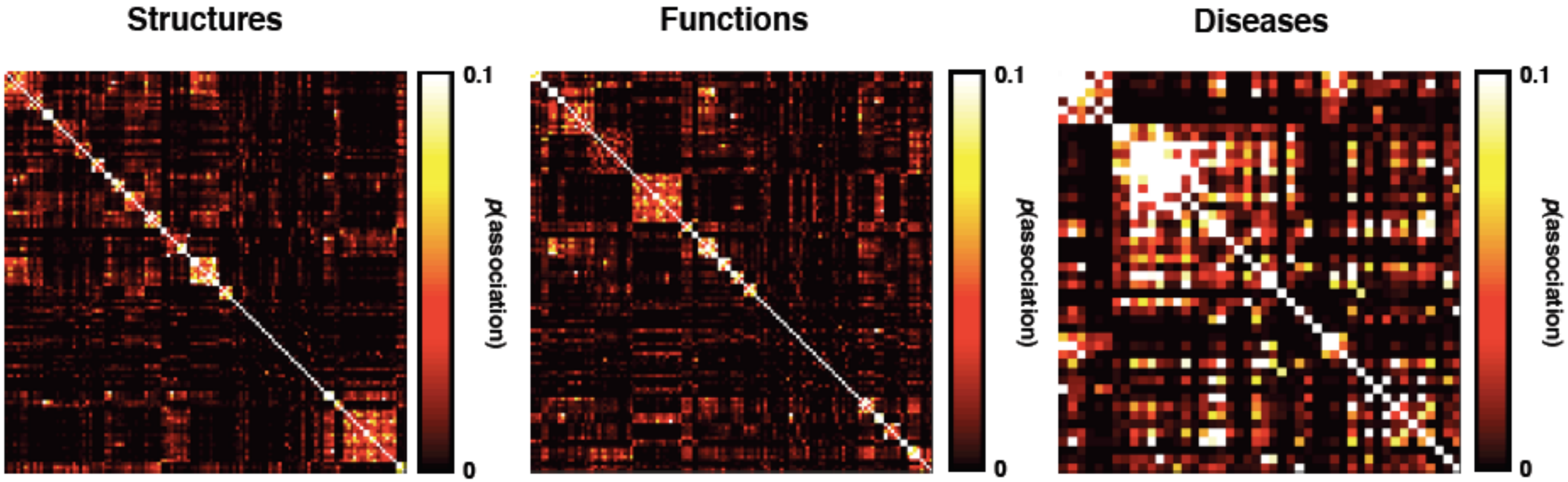
- Sections are *not* mandatory
- But bring a computer if you can
- Python, Jupyter, and git *before sections*! (We'll post instructions)
- Sections: intend to go to the section for which you're registered, but you can attend another if needed
- Sections are a great place for assignment guidance
- TA office hours will be in the computer lab, giving you more time to work on assignments
- Assignments will be released Monday mornings
- Assignments will be due Sunday nights
- Assignment solutions will be posted one week after deadline
- Late assignments graded at 50% penalty, not late assignments allowed once solutions are posted

COGS 108
Data Science in Practice

Data are awesome! (Part II)

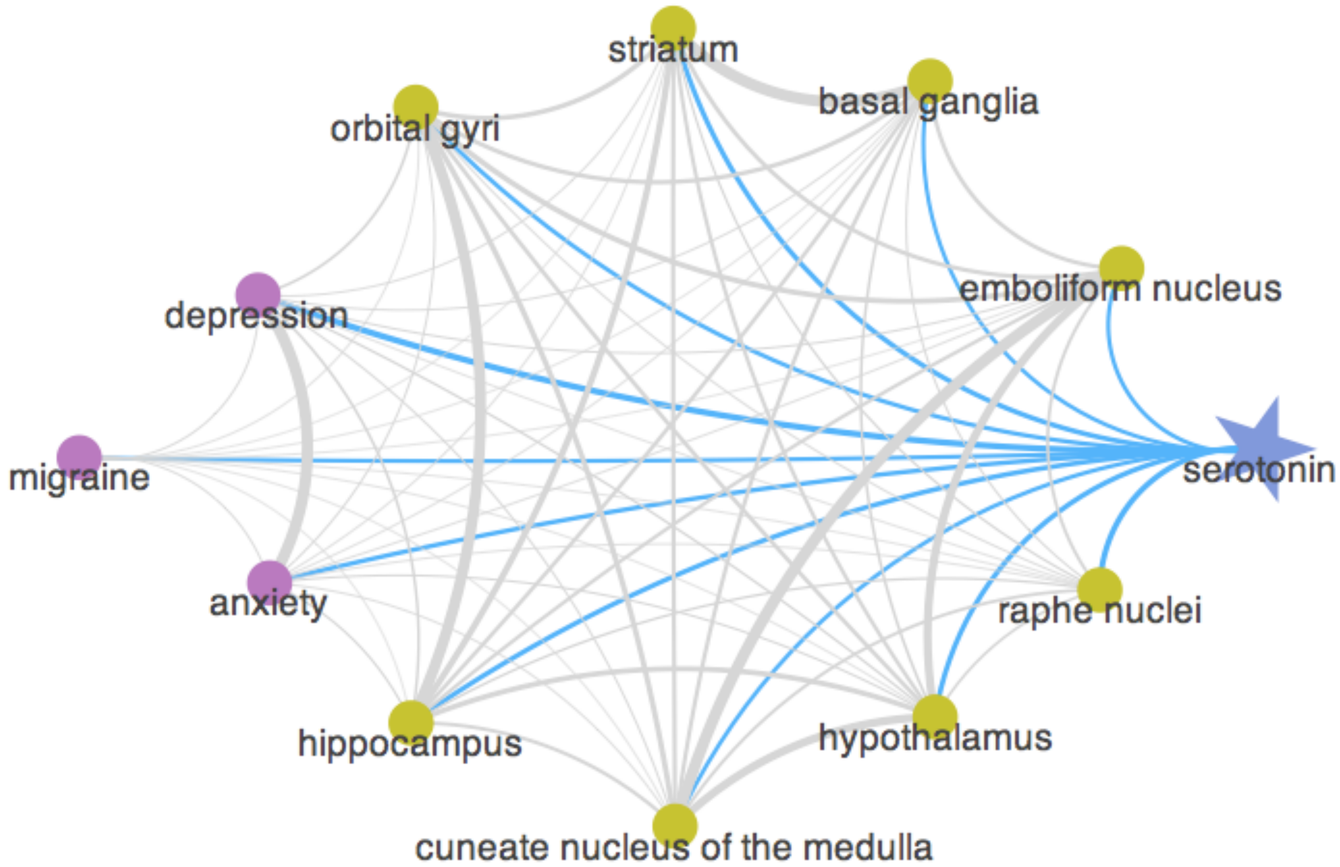
UC San Diego

Knowledge Discovery

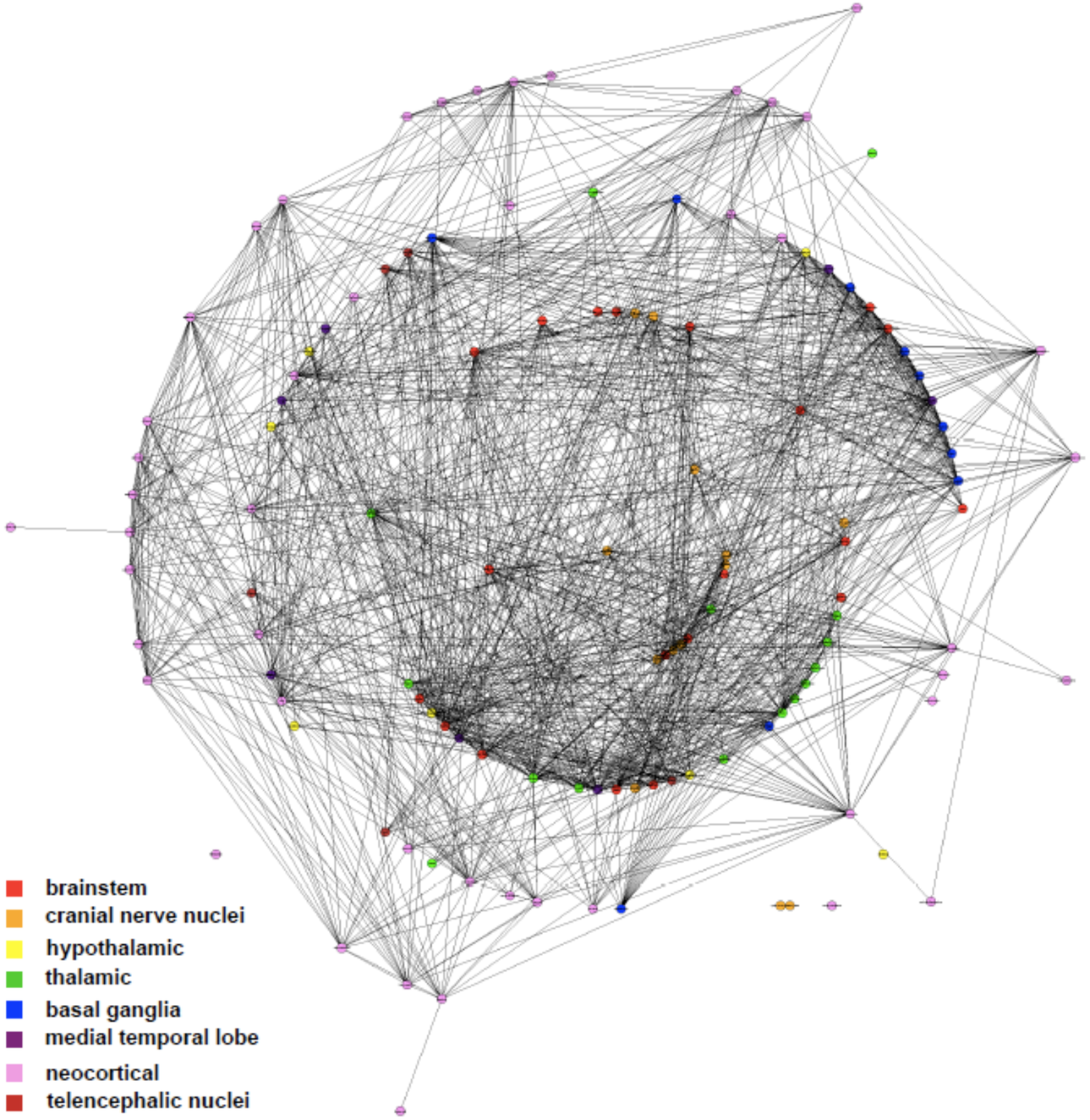


Source: Voytek & Voytek, *J Neurosci Methods* 2012

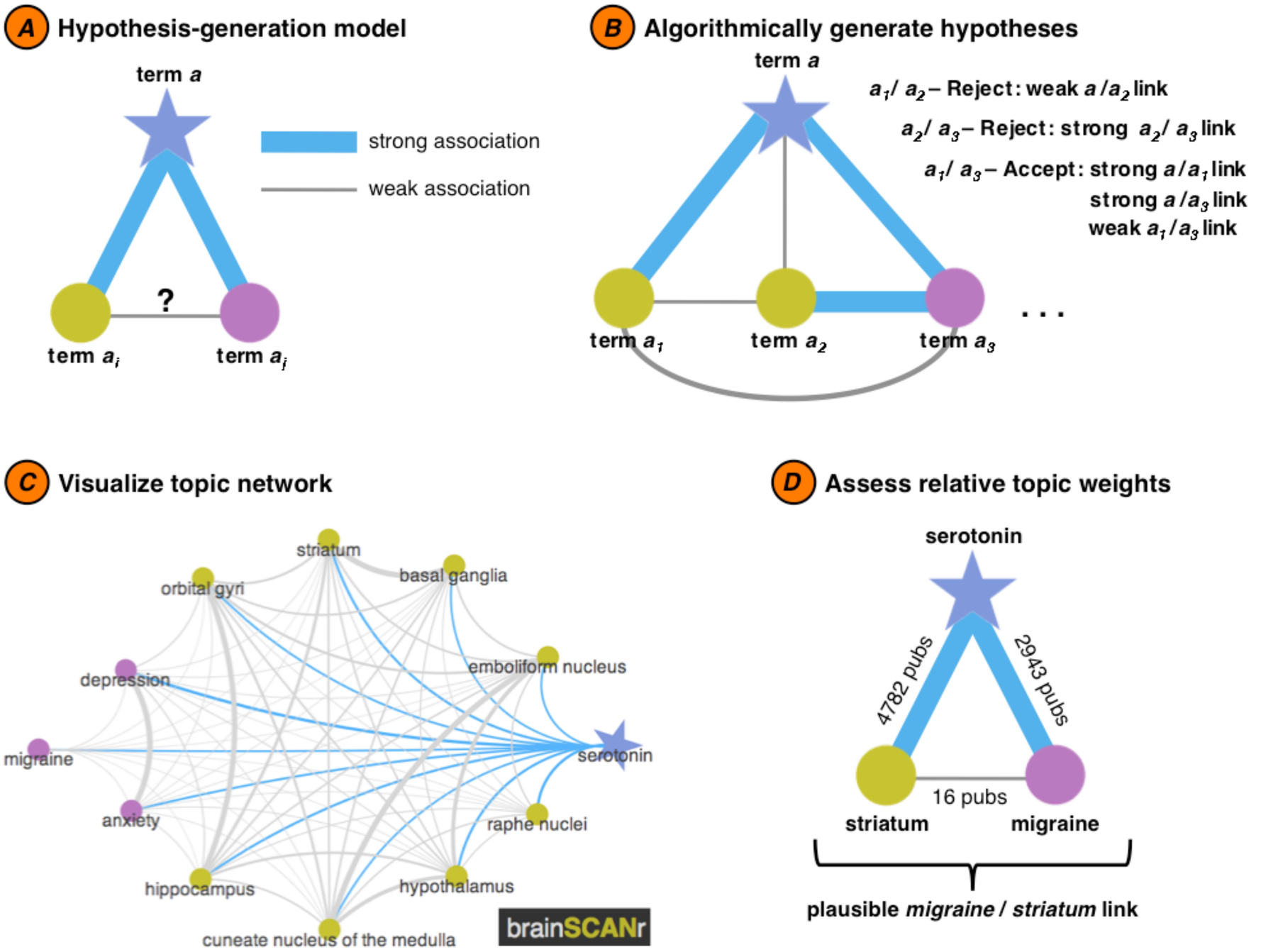
brainSCANr



Source: Voytek & Voytek, *J Neurosci Methods* 2012



Source: Voytek & Voytek, *J Neurosci Methods* 2012

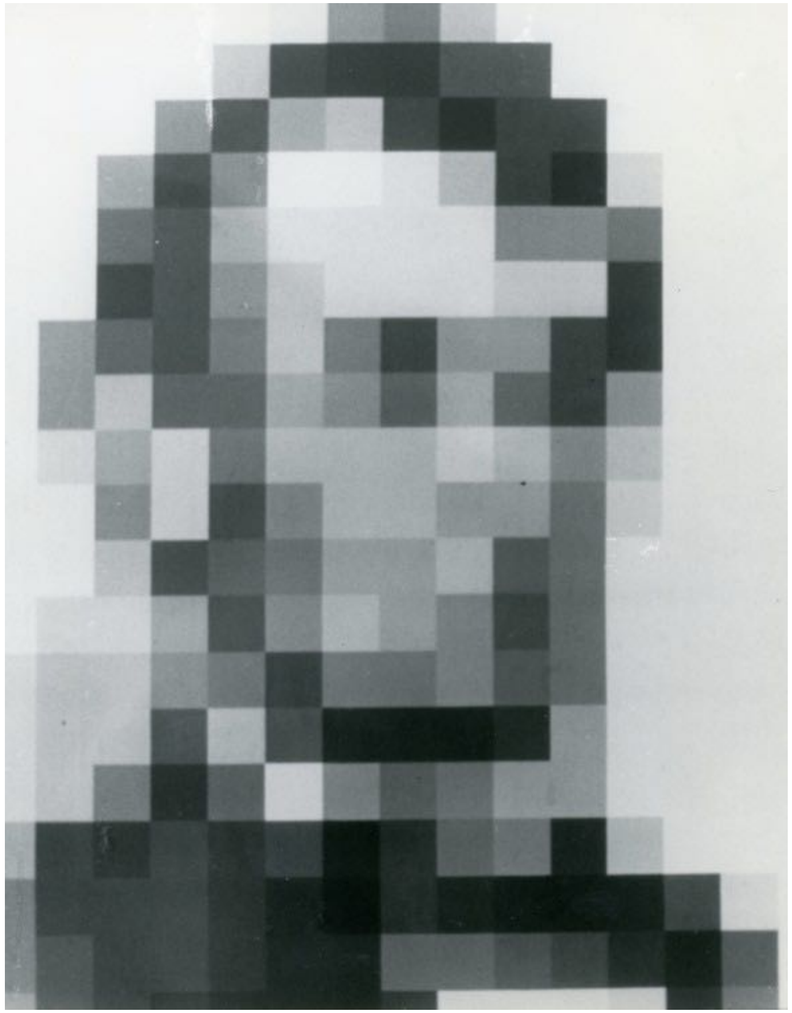


Source: Voytek & Voytek, *J Neurosci Methods* 2012

Doing the Amazing

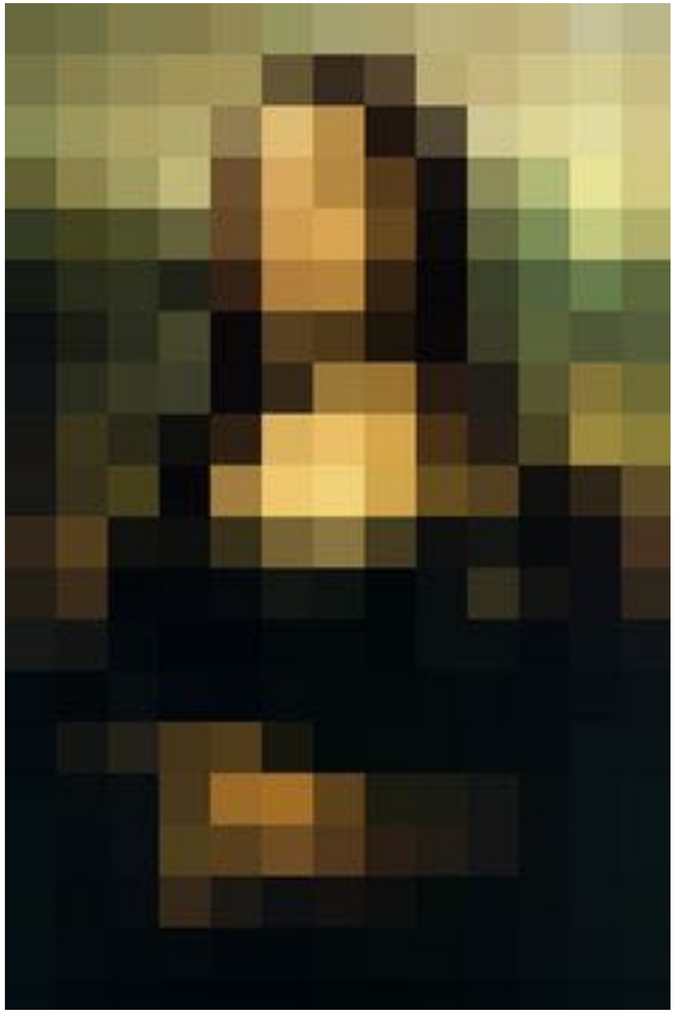
Source: Andrew Ochoa, Waverly Labs

ENHANCE!



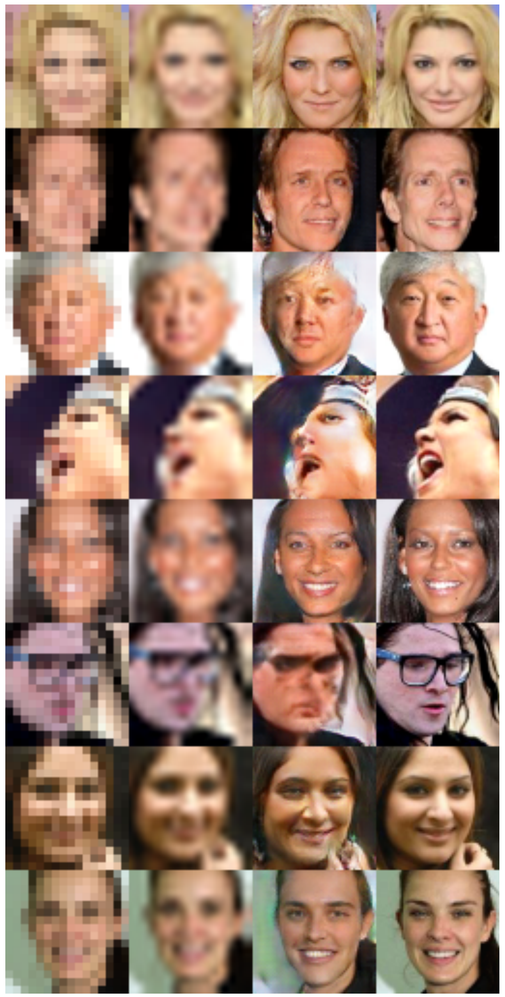
c.f. Harmon & Julesz, *Science* 1973

ENHANCE!



c.f. Harmon & Julesz, *Science* 1973

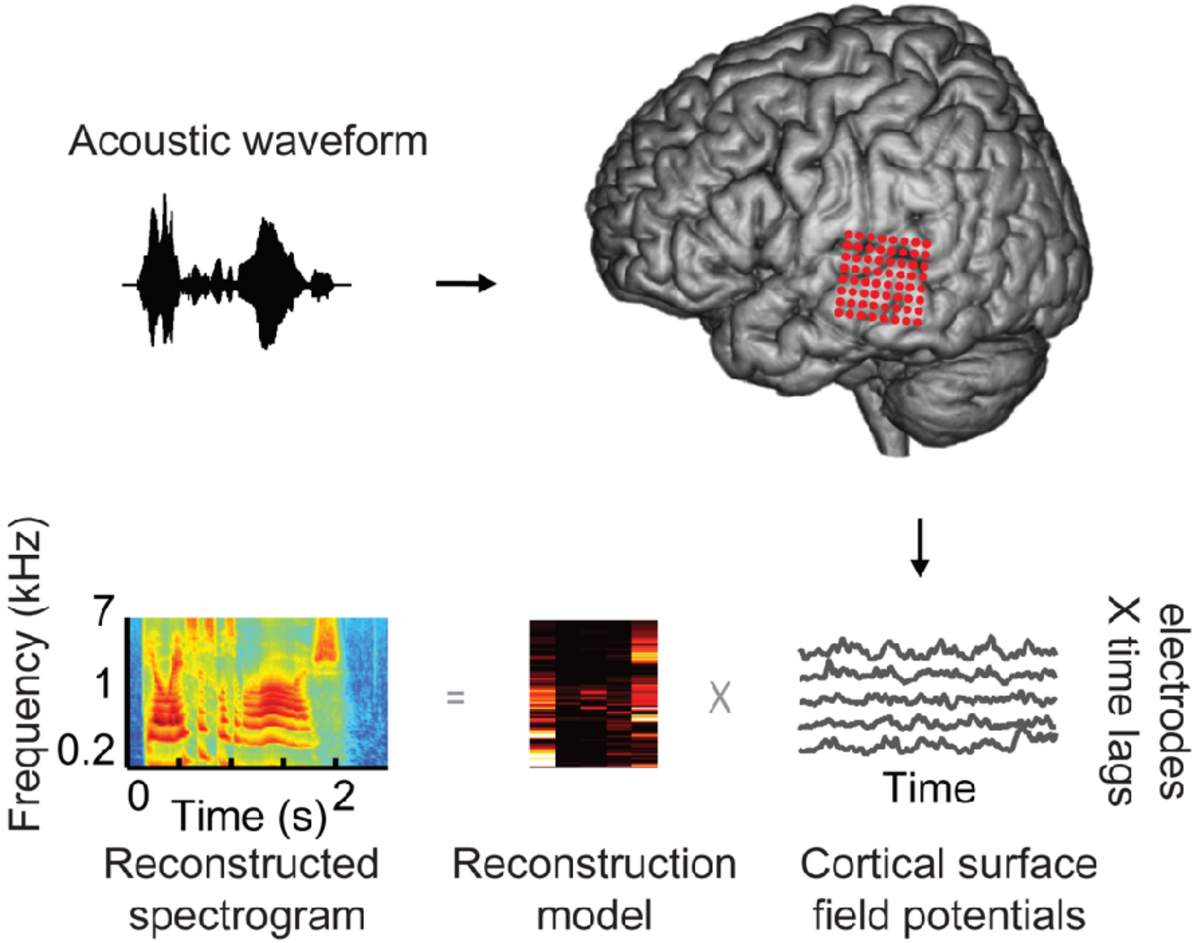
ENHANCE!



"…the first column is the 16x16 input image, the second one is what you would get from a standard bicubic interpolation, the third is the output generated by the neural net, and on the right is the ground truth."

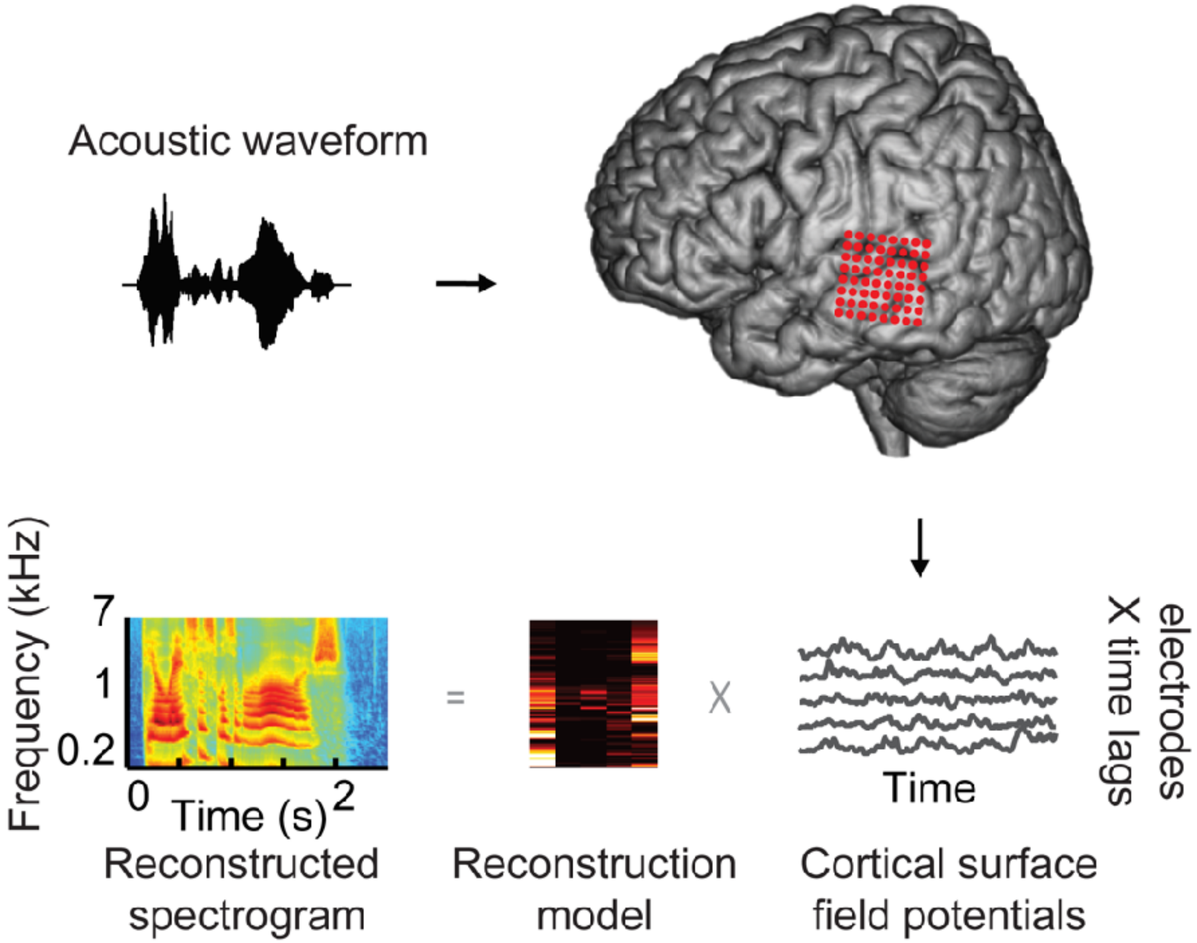
Source: David Garcia, Qualcomm

Mind reading?



Source: Holdgraf *et al.*, *Nature Comm* 2016

Mind reading?



Source: Pasley *et al.*, *PLoS Biol* 2012

COGS 108
Data Science in Practice

Python! (For great Data Science)

UC San Diego

Recommendations

- *Data Science from Scratch* - Joel Grus (Allen Institute for AI)
- *Python Data Science Handbook* - Jake VanderPlas (UW eScience Institute)

Why Python?

Python has emerged over the last couple decades as a first-class tool for scientific computing tasks, including the analysis and visualization of large datasets. This may have come as a surprise to early proponents of the Python language: the language itself was not specifically designed with data analysis or scientific computing in mind.

Source: VanderPlas, *Python Data Science Handbook*

Why Python?

The usefulness of Python for data science stems primarily from the large and active ecosystem of third-party packages: NumPy for manipulation of homogeneous array-based data, Pandas for manipulation of heterogeneous and labeled data, SciPy for common scientific computing tasks, Matplotlib for publication-quality visualizations, IPython for interactive execution and sharing of code, Scikit-Learn for machine learning, and many more tools that will be mentioned in the following pages.

Source: VanderPlas, *Python Data Science Handbook*

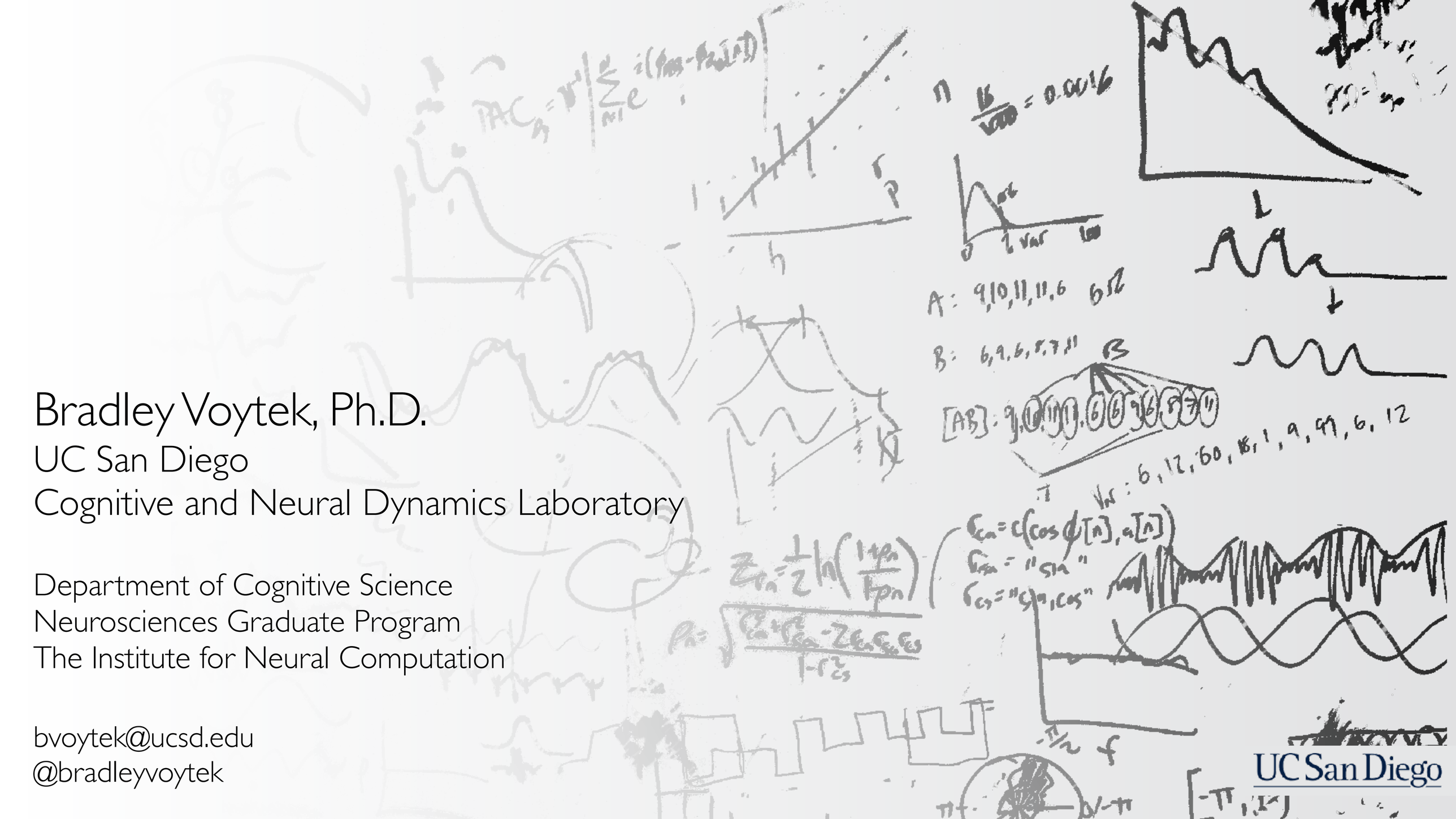
Why Python?

Python isn't the best at anything.

Why Python?

Python isn't the best at anything.

But it's the *second best* at EVERYTHING.



Bradley Voytek, Ph.D.

UC San Diego
Cognitive and Neural Dynamics Laboratory

Department of Cognitive Science
Neurosciences Graduate Program
The Institute for Neural Computation

bvoytek@ucsd.edu
@bradleyvoytek

UC San Diego