# Billions for Big Brains Proposal

Daniel Qian, Nathan Vallapureddy, Anthony Wu, Jenny Wu

January 16, 2018

## 1 Abstract

Memory has been theorized to be ingrained not only in the activity–or action potentials–of the brain, but also structurally in the vast connections between neurons. Our attempts to create a map of every memory the brain has stored are so far limited by measurements of brain activity itself, but we are now approaching the capability to analyze every structural facet–every neuron and every synapse–of the mammalian brain.
Our proposal is to create a connectome of a mouse brain which will allow us to develop complex understanding of the cellular structures and connections that allow for the formation, encoding, and storage of memories. We will train a sample of mice to respond to certain stimuli in pre-programmed ways, and then select a single mouse to image. After collecting the raw image data, we will create brain graphs and, in combination with electrical activity and experimental data, will use those to build a neural network modeling the structures, connections, and activities of the original mouse brain. The network will then be converted to virtual simulations, which we can modify, conduct experiments with, and test in comparison to the original mouse.

## 2 Scientific Question

What information about memory does a comprehensive connectome of the mouse brain provide, and can we apply less precise models to generate the same results?

## 3 Hardware and Facilities

We will send out a request for proposals to teams interested in imaging an entire mouse brain, so the hardware for the project will be stored at the teams location. The majority of data collection will be conducted on a single mouse chosen from a sample of 30. This mouse will be selected based on its ability to be trained to respond to stimuli in pre-programmed ways in anticipation of reward: for example, to move towards a light, or to press a lever in response to an odor. To obtain very high resolution 3D images of the chosen mouse brain, we will use a Zeiss Crossbeam 550 Focused Ion Beam Scanning Electron Microscope (FIB-SEM). In addition, once a hyper-detailed synaptic-level model is created using electron microscopy, we will use the CLARITY light microscopy technique to analyze a wider array of mouse brain of varying performance levels and neural structures.

# 4 Data Collection Processes

The electron-beam images from the scanning electron microscope can be digitized using an analog-to-digital converter. The digitized images are then uploaded to a computer as grayscale images, from which we can analyze pixel intensity (with 0 corresponding to black and 255 corresponding to white). To create a processable image we will convert the file into a BMP (bitmap) file.

# 5 Information Extraction Processes

Because the SEM microscopes will scan the mouse brain in cuboidal sections, in order to assemble a complete model of the mouse brain we need to organize them by location. Each image, corresponding to a slice of a cube, will be indexed by which cube it belongs to and its position in the cube. Then, each image will be segmented into cell types and synapses, using a machine learning algorithm that uses not only the information from the raw image, but also previous results from adjacent images. Training data for this algorithm could be used from the Allen Brain Project.

Using these processed brain images, we will design a virtual representation of the cellular structures, connections, and activity of the original brain. This virtual network will then be tested against the original mouse: we will simulate simple reflex arcs and conditioned responses with virtual foot-shock and award-driven lever press tasks. The network will then be modified in experiments in which we turn on and off neuron clusters, pathways and brain regions, in an attempt to selectively destroy or change certain memories. With extensive study we will develop understanding of the conditions in which memory is able to function, as well as how neurological disorders affect the brains ability to store memories.

# 6 Data Upload

We will use Amazon Web Services (namely Amazon Glacier) as our intermediary for the image data taken by the FIB-SEM and the programs that convert the data into a network. We chose Amazon Glacier for this because it is the cheapest AWS for very large amounts of data. One flaw of Amazon Glacier is its long data request times (5-12 hours). However, this doesn't affect us as we can access data in large packets, so we need to request data infrequently. Uploading the data to our AWS instance is straightforward through the AWS interface. We will upload data in large packets to prevent data slowdown.

# 7 Data Storage

We will also use Amazon Glacier to store the neural network data in addition to the raw image data. We will have a copy of the neural network data on our own servers, so the slow data request time is a nonfactor here.

# 8 Cost

## 8.1 Mouse brain imaging

Average mouse brain size: $415mm^3$ (https://academic.oup.com/cercor/article/15/5/639/442213)
Cost to perform EM imaging on $1mm^3$ of brain tissue: $5M to $10M. Since we are imaging a massive amount of brain tissue, we proceed with the lower bound.
Total cost to EM image one mouse brain: $2.075 billion

## 8.2 Data Storage

We store all of our data on Amazon Glacier. This will be about 830 PB of data (2PB per $mm^3$, brain average volume of $415mm^3$), this would cost $200M ($0.004 per GB per month for 60 months) (https://aws.amazon.com/glacier/pricing/).

## 8.3 Virtual Network Representation

The Blue Brain project cost 70M ($96.5M) to process 1M neurons (http://www.artificialbrains.com/blue-brain-project). The mouse brain contains 71M neurons (http://www.pnas.org/content/103/32/12138.full), so we project a cost of 6.851B.

## 8.4 Mouse Training, Network Modification

We will hire 1700 programmers, neuroscientists, mathematicians, and biologists to make modifications to the network in a strategic way (a way that maximizes the probability of behavior modification). Each will be paid $100K per year for 5 years, a total cost of $850M.

Table 1: Cost Summary

| Item Name | Total Cost (Millions of USD) |
|---|---|
| Mouse Brain Imaging | 2075 |
| Data Storage | 200 |
| Virtual Network Representation | 6851 |
| Mouse Training, Network Modification | 850 |

Total cost: $9.976 Billion USD