Glut of data from mice brains tests MIT's computing power

"I think there is an enormous amount of hope generated by bringing new tools into neuroscience," says Edward S. Boyden, an MIT researcher.

By Murray Carpenter | GLOBE CORRESPONDENT   JANUARY 31, 2016

Even one of the great young scientists of our time can be brought low by computer problems.

Twee   Share   2

Edward S. Boyden is an MIT neuroscientist who has won acclaim for his work into using light to control neurons, a technique known as optogenetics that is helping researchers expand the study of the brain.

But the brain, it turns out, may be the ultimate Big Data generator.

In his efforts to study the brain's electrical circuitry in granular detail, Boyden designed tiny probes that, when poked into the brain of a lab mouse, can record electronic transmissions from individual neurons. Yet even at their modest size, those probes were sucking in way more data than Boyden's computers could handle.

One day Boyden mentioned his problem to a fellow researcher, Christian Wentz. His company, Kendall Research Systems, which makes tools for optogenetics, shares office space with an unusual computer company called LeafLabs. Formed in 2009 by MIT students, LeafLabs has a simple motto: "We solve moon-shot computing challenges."

The company's chief executive, Andrew Meyer, said the sheer volume of data Boyden was collecting through thousands of channels inside the probes was daunting.

"It's really almost frightening, a terabit per second — it's huge," Meyer said. Boyden's data flow is "so big that when we try and look to other areas of technology to crib good ideas and tools, there aren't any."

LeafLabs specializes in parallel computing, or using multiple processors to crunch lots of numbers at the same time. Typically, massive computer processing projects would require a series of high-end, rack-mounted machines, each costing about $3,000. But Meyer said those types of computers were overbuilt for the job at hand. He needed storage, which is

cheap — but not all the heavy-duty processors used for graphics and video games. So his first trick in designing a computer for Boyden was to get rid of anything extra.

"We just started ripping stuff out until all we were left with was the storage, and just a tiny, tiny amount of a parallel processor that could suck up the data and put it on the storage, and that was it. It was as minimalist as we could do," Meyer said.

That resulted in a much less expensive computer, one Meyer said is more purpose-built to "push data into these hard drives as fast as they can take them."

At first, Meyer said, his computers were able to acquire about a gigabit per second. That sounds like a lot, and it is — roughly equivalent to storing the data from a DVD movie every two seconds. But that was still short of what Boyden's probes generated. Now, though, the machines are quickly getting faster — closer to 125 gigabits per second. For a sense of scale, that's about one-tenth of what Netflix generates on the entire Internet at any given moment.

Eventually, Boyden said, more powerful computers will help scientists gain a better understanding of brain function, to help people who suffer from epilepsy and Alzheimer's and Parkinson's diseases.

"The cool part of neuroengineering is that we have all these unmet needs, both scientific and clinical," Boyden said. "I think there is an enormous amount of hope generated by bringing new tools into neuroscience."

The volume of data coming out of neuroscience may be unique for now, but Meyer said many other disciplines are dealing with increasingly large troves of information that will need ever more powerful computers to process. Recent

advances in artificial intelligence, scene recognition, even self-driving cars are all predicated upon huge amounts of data.

"Independent of neuroscience, the scale at which data is useful in understanding your problem just keeps moving up and up and up," Meyer said.

LeafLabs is now producing a computing system for neurobiology called Willow that, with probes, software, hardware and a control computer, starts at $17,000. Boyden and Meyer described the system in a September paper and displayed it in October at the Society for Neuroscience meeting in Chicago.

"The idea here is to make it possible for people to do massive amounts of neural data acquisition and analysis," Boyden said. "And to maybe identify new types of patterns, like what a memory actually is, or how is an emotion represented in the brain. Those are very difficult to do if you can only look at a couple of neurons."

An unexpected benefit of his collaboration with Meyer is this new, super-fast, big-data computing system. Boyden said that's one of the fun aspects of science.

"I'm a firm believer in serendipity optimization," Boyden said. "If you can plan it, it's probably not interesting, and somebody is probably already doing it. But if you find these sort of connect-the-dots, those are much more likely to be new."

Get Today's Headlines from the Globe in your inbox: