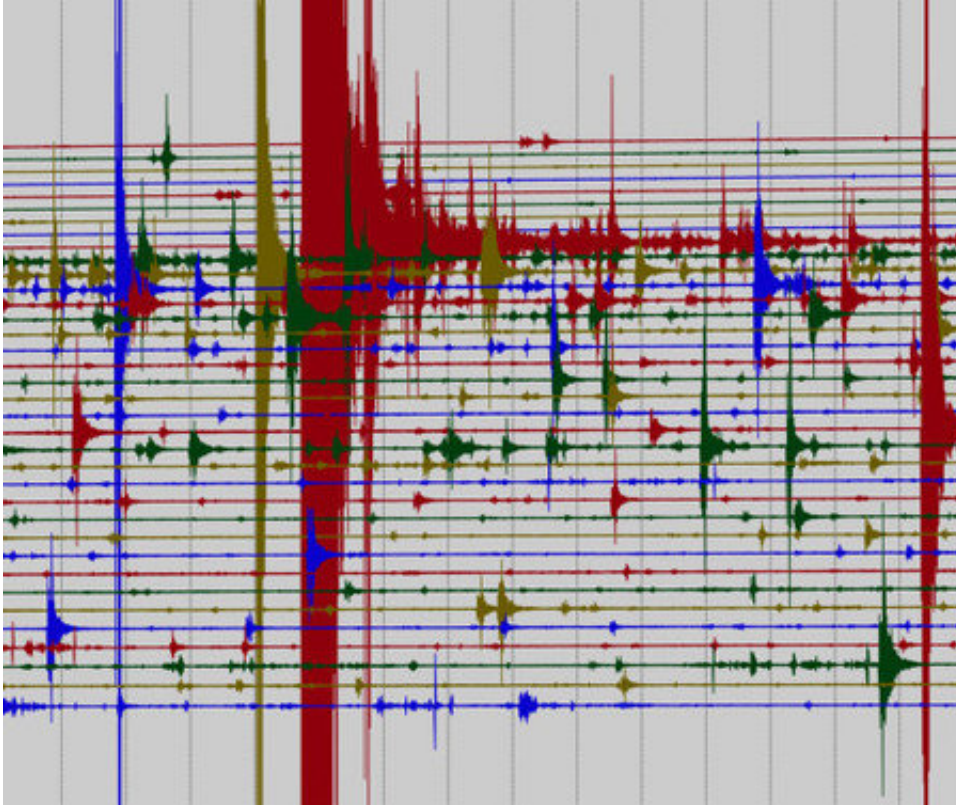# Creating a 'virtual seismologist'

January 4 2019, by Elise Cutts

A snapshot of seismic data taken at a single station during the peak of an aftershock sequence. Credit: Zachary Ross/Caltech

Understanding earthquakes is a challenging problem—not only because they are potentially dangerous but also because they are complicated phenomena that are difficult to study. Interpreting the massive, often convoluted data sets that are recorded by earthquake monitoring networks is a herculean task for seismologists, but the effort involved in

producing accurate analyses could significantly improve the development of reliable earthquake early-warning systems.

A promising new collaboration between Caltech seismologists and computer scientists using artificial intelligence (AI)—computer systems capable of learning and performing tasks that previously required humans—aims to improve the automated processes that identify earthquake waves and assess the strength, speed, and direction of shaking in real time. The collaboration includes researchers from the divisions of Geological and Planetary Sciences and Engineering and Applied Science, and is part of Caltech's AI4Science Initiative to apply AI to the big-data problems faced by scientists throughout the Institute. Powered by advanced hardware and machine-learning algorithms, modern AI has the potential to revolutionize seismological data tools and make all of us a little safer from earthquakes.

Recently, Caltech's Yisong Yue, an assistant professor of computing and mathematical sciences, sat down with his collaborators, Research Professor of Geophysics Egill Hauksson, Postdoctoral Scholar in Geophysics Zachary Ross, and Associate Staff Seismologist Men-Andrin Meier, to discuss the new project and future of AI and earthquake science.

What seismological problem inspired you to include AI in your research?

Meier: One of the things that I work on is earthquake early warning. Early warning requires us to try to detect earthquakes very rapidly and predict the shaking that they will produce later so that you can get a few seconds to maybe tens of seconds of warning before the shaking starts.

Hauksson: It has to be done very quickly—that's the game. The

earthquake waves will hit the closest monitoring station first, and if we can recognize them immediately, then we can send out an alert before the waves travel farther.

Meier: You only have a few seconds of seismogram to decide whether it is an earthquake, which would mean sending out an alert, or if it is instead a nuisance signal—a truck driving by one of our seismometers or something like that. We have too many false classifications, too many false alerts, and people don't like that. This is a classic machine-learning problem: you have some data and you need to make a realistic and accurate classification. So, we reached out to Caltech's computing and mathematical science (CMS) department and started working on it with them.

Why is AI a good tool for improving earthquake monitoring systems?

Yue: The reasons why AI can be a good tool have to do with scale and complexity coupled with an abundant amount of data. Earthquake monitoring systems generate massive data sets that need to be processed in order to provide useful information to scientists. AI can do that faster and more accurately than humans can, and even find patterns that would otherwise escape the human eye. Furthermore, the patterns we hope to extract are hard for rule-based systems to adequately capture, and so the advanced pattern-matching abilities of modern deep learning can offer superior performance than existing automated earthquake monitoring algorithms.

Ross: In a big aftershock sequence, for example, you could have events that are spaced every 10 seconds, rapid fire, all day long. We use maybe 400 stations in Southern California to monitor earthquakes, and the waves caused by each different earthquake will hit them all at different

times.

Yue: When you have multiple earthquakes, and the sensors are all firing at different locations, you want to be able to unscramble which data belong to which earthquake. Cleaning up and analyzing the data takes time. But once you train a machine-learning algorithm—a computer program that learns by studying examples as opposed to through explicit programing—to do this, it could make an assessment really quickly. That's the value.

How else will AI help seismologists?

Yue: We are not just interested in the occasional very big earthquake that happens every few years or so. We are interested in the earthquakes of all sizes that happen every day. AI has the potential to identify small earthquakes that are currently indistinguishable from background noise.

Ross: On average we see about 50 or so earthquakes each day in Southern California, and we have a mandate from the U.S. Geological Survey to monitor each one. There are many more, but they're just too small for us to detect with existing technology. And the smaller they are, the more often they occur. What we are trying to do is monitor, locate, detect, and characterize each and every one of those events to build "earthquake catalogs." All of this analysis is starting to reveal the very intricate details of the physical processes that drive earthquakes. Those details were not really visible before.

Why hasn't anyone applied AI to seismology before?

Ross: Only in the last year or two has seismology started to seriously consider AI technology. Part of it has to do with the dramatic increase in computer processing power that we have seen just within the past

decade.

What is the long-term goal of this collaboration?

Meier: Ultimately, we want to build an algorithm that mimics what human experts do. A human seismologist can feel an earthquake or see a seismogram and immediately tell a lot of things about that [earthquake](#) just from experience. It was really difficult to teach that to a computer. With [artificial intelligence](#), we can get much closer to how a human expert would treat the problem. We are getting much closer to creating a "virtual seismologist."

Why do we need a "virtual seismologist?"

Yue: Fundamentally both in seismology and beyond, the reason that you want to do this kind of thing is scale and complexity. If you can train an AI that learns, then you can take a specialized skill set and make it available to anyone. The other issue is complexity. You could have a human look at detailed seismic data for a long time and uncover small earthquakes. Or you could just have an algorithm learn to pick out the patterns that matter much faster.

Meier: The detailed information that we're gathering helps us figure out the physics of earthquakes—why they fizzle out along certain faults and trigger big quakes along others, and how often they occur.

Will creating a "virtual seismologist" mean the end of human seismologists?

Ross: Having talked to a range of students, I can say with fairly high confidence that most of them don't want to do cataloguing work. [Laughs.] They would rather be doing more exciting work.

Yue: Imagine that you're a musician and before you can become a musician, first you have to build your own piano. So you spend five years building your piano, and then you become a musician. Now we have an automated way of building pianos—are we going to destroy musicians' jobs? No, we are actually empowering a new generation of musicians. We have other problems that they could be working on.

Provided by California Institute of Technology