

<https://slate.com/technology/2019/09/artificial-intelligence-climate-change-carbon-emissions-roy-schwartz.html>

# Artificial Intelligence Can't Think Without Polluting

But researchers are finally trying to make it greener.

By APRIL GLASER  
SEPT 20, 2019 3:17 PM

*This piece has been published as part of Slate's partnership with Covering Climate Now, a global collaboration of more than 250 news outlets to strengthen coverage of the climate story.*

Artificial intelligence is getting smarter, but it isn't getting cleaner. In order to improve at predicting the weather, sorting your social media feeds, and hailing your Uber, it needs to train on massive datasets. A few years ago, an A.I. system might have required millions of words to attempt to learn a language, but today that same system could be processing 40 billion words as it trains, according to Roy Schwartz, who researches deep learning models at the Allen Institute for Artificial Intelligence and in the University of Washington's computer science and engineering department. All that processing takes a lot of energy.

Some researchers are starting to think about how to change that. In a [paper published](#) in July, Schwartz and three other researchers—Jesse Dodge from Carnegie Mellon University, Noah Smith from the University of Washington, and Oren Etzioni from the Allen Institute—make the case for standardizing how we measure just how carbon-intensive artificial intelligence is. The team wants all future A.I. research to tack on such a measurement—and believes that doing so will lead to new innovations in making the technology less taxing on the environment. In an interview, Schwartz lamented that most A.I. systems are graded only on how much better they get at doing their job. But it can't just be about accuracy anymore, he says. It has to be about energy efficiency too. "We don't want to reach a state where A.I. will become a significant contributor to global warming," he said.

**April Glaser: Can you tell me how energy-intensive artificial intelligence systems currently are?**

**Roy Schwartz:** The industry is growing everywhere, and there's a lot of curiosity and investment being put into figuring out how far you can push A.I. systems to go. And when you try to talk about like, say, a system for answering questions or recognizing faces or whatever your favorite application is, a simple way of making a model that's already pretty good even better is to just make it bigger and consume more resources.

So think of cars, for instance. If you want to build a faster car, a simple thing you can do is just make the engine twice as large. But the problem obviously is that we're not thinking about the energetic price that comes with this increase in resources. So the field of artificial intelligence has grown very significantly in the last few years, and I started noticing that there's a troubling trend in this growth with energy consumption. There's research from the University of Massachusetts just a few months ago that took one of the biggest A.I. systems we have and conducted an experiment with it. The amount of CO<sub>2</sub> that it emitted [during this single experiment](#) came out to be about the equivalent of five years of driving your car, which was kind of mind-blowing to a lot of people in the community. And these systems are getting bigger and bigger, so this is just a starting point. We are trying to think of ways we can get the community to start thinking about, in addition to making the system more accurate, making it more efficient.

**The UMass experiment was just with one training model. Is the size and style of this model standard at universities and companies?**

No, these A.I. models are not something that any random researcher has, because these systems take up so much energy. They also cost a lot of money. And this then creates a problem of a lack of diversity within the community because most researchers are unable to afford these experiments if they don't work at one of the big companies or big research labs. So this five-car-years-worth-of-energy problem is not one that every researcher will hit, but there are many models like this running at big companies, and the problem is that these systems are continuing to grow. Soon these five cars might become 10 or 15 or 20 in the coming years, and we want to think about ways to change this. We don't want to necessarily stop these experiments because the things we learn are very valuable, but we want to encourage people to also find ways to make these experiments more efficient.

## **Can you tell me what it is about A.I. models that are so energy-intensive?**

If you think of the technology of deep learning, the technology that governs A.I., the very basics of it are that it's taking very big matrices and multiplying one by another. And that's the fundamental unit in each of these systems. And these matrices are just becoming larger and larger, and when I say larger, I mean also that these systems are becoming trained on larger sets of data. So, you can see an amazing growth in my field, which is language-understanding. We've seen that a couple of years ago, the largest system was trained on an order of magnitude of maybe several millions of words. And then it became tens of millions or hundreds of millions. And now the system is trained on 40 billion words, so it just takes more and more time. And we're growing in every possible dimension. Larger matrices that are being multiplied and take a lot of time and energy, and this process is being repeated 40 billion times. And the third criterion is that each of these full experiments is run multiple times. So if you train them twice, you might get a better result, so you have multiple dimensions in which you're growing.

**So, you're saying it's that the processing power is so high. What do you begin to think about when it comes to becoming more efficient? I ask because it's not like you can run fewer experiments. This is how A.I. learns, right?**

The first thing we want to do is bring this to the awareness of the community and make people report the amount of computing they use to generate the results in their paper, because typically this is not being done. And it's always very surprising. Think of the example with the car earlier. You can say, look, I built a faster car, but you don't say how much fuel you're consuming to get to this speed. So we're not really reporting this right now. And to be honest, there are problems in reporting because it's hard to compare between researchers at different labs and different locations. One of the things we're trying to do at the Allen Institute is to come up with a very simple way for researchers to use every time they run their code that will provide some number that will tell them how much energy they've consumed. That's something we're currently working on.

Once people start reporting it, many people will be encouraged to work to improve to find greener or more efficient solutions, whether we can make these matrices smaller using sophisticated methods, whether we can use less data or the data we have in a smarter way, or whether we can run less experiments, for instance, by recognizing early whether one expensive experiment is going to fail. So currently, we're all working on this trend of using larger and more-expensive experiments and it's kind of like, "Run everything and just pick the best number." This is because the price these systems cost [in terms of carbon] doesn't matter as much to people because of a lack of awareness.

**So to be clear, there's just no main standard right now for measuring energy consumption when it comes to building or maintaining A.I. systems?**

Well, there are several standards, some of them more widely used. The amount of time it takes to train a system is one common standard. But it's very hard to compare it. If I'm on one computer and you're on a different computer, it doesn't really say much if your program is running faster than mine. So, we're trying to develop a standard that will be easy to use and comfortable between different researchers.

**So where do you imagine will be one of the first places where efficiencies will be able to be made in A.I.?**

We are not the first one to think about this problem. For example, with your smartphone and computer vision, people have been thinking about how to make this model small so it will fit on your phone or run faster. But that's not the same as making training more efficient or identifying which version will fail faster, so we are working on projects along each of these lines and we're making progress on each of them. It does appear there are gains to be made in each of these fronts. I can't say how things will play out, but I think all of them will be quite significant.

**How far away do you think you are from having a standard that could be used by people who are building and deploying these A.I. systems?**

By the end of the year we want to have a framework for natural language processing, which is the basic building stone for many of the

language technologies out there nowadays. We're hoping to add this functionality to our NLP system so that way anyone who uses some of our tools will get this number about their environmental cost by the end of the year. And next year we're going to work on something more general that every person who develops this system will get this functionality. I'm a researcher, not an engineer, but luckily we have some of the best engineers in the world here to help us. We're hoping to get this as soon as possible.

### **What else is on your mind here?**

My goal personally is that I want to allow a better climate for researchers to think about these problems. Because currently both the academic and the industry environment is mostly rewarding scientific breakthroughs that make models more accurate. And to balance that, it would also should be rewarding efficiency. This is the biggest thing I can do at the moment for this purpose.

*[Future Tense](#) is a partnership of [Slate](#), [New America](#), and [Arizona State University](#) that examines emerging technologies, public policy, and society.*