

## THE TURBULENT PAST AND UNCERTAIN FUTURE OF ARTIFICIAL INTELLIGENCE

Is there a way out of AI's boom-and-bust cycle?

ELIZA STRICKLAND

30 SEP 2021

<https://spectrum.ieee.org/history-of-ai>

IN THE SUMMER OF 1956, a group of mathematicians and computer scientists took over the top floor of the building that housed the math department of Dartmouth College. For about eight weeks, they imagined the possibilities of a new field of research. John McCarthy, then a young professor at Dartmouth, had coined the term "artificial intelligence" when he wrote his proposal for the workshop, which he said would explore the hypothesis that "every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it."

The researchers at that legendary meeting sketched out, in broad strokes, AI as we know it today. It gave rise to the first camp of investigators: the "symbolists," whose expert systems reached a zenith in the 1980s. The years after the meeting also saw the emergence of the "connectionists," who toiled for decades on the artificial neural networks that took off only recently. These two approaches were long seen as mutually exclusive, and competition for funding among researchers created animosity. Each side thought it was on the path to artificial general intelligence.

A look back at the decades since that meeting shows how often AI researchers' hopes have been crushed—and how little those setbacks have deterred them. Today, even as AI is revolutionizing industries and threatening to upend the global labor market, many experts are wondering if today's AI is reaching its limits. As Charles Choi delineates in "Seven Revealing Ways AIs Fail," the weaknesses of today's deep-learning systems are becoming more and more apparent. Yet there's little sense of doom among researchers. Yes, it's possible that we're in for yet another AI winter in the not-so-distant future. But this might just be the time when inspired engineers finally usher us into an eternal summer of the machine mind.

Researchers developing symbolic AI set out to explicitly teach computers about the world. Their founding tenet held that knowledge can be represented by a set of rules, and computer programs can use logic to manipulate that knowledge. Leading symbolists Allen Newell and Herbert Simon argued that if a symbolic system had enough structured facts and premises, the aggregation would eventually produce broad intelligence.

The connectionists, on the other hand, inspired by biology, worked on "artificial neural networks" that would take in information and make sense of it themselves. The pioneering example was the perceptron, an experimental machine built by the Cornell psychologist Frank Rosenblatt with funding from the U.S. Navy. It had 400 light sensors that together acted as a retina, feeding information to about 1,000 "neurons" that did the processing and produced a

single output. In 1958, a New York Times article quoted Rosenblatt as saying that "the machine would be the first device to think as the human brain."

Unbridled optimism encouraged government agencies in the United States and United Kingdom to pour money into speculative research. In 1967, MIT professor Marvin Minsky wrote: "Within a generation...the problem of creating 'artificial intelligence' will be substantially solved." Yet soon thereafter, government funding started drying up, driven by a sense that AI research wasn't living up to its own hype. The 1970s saw the first AI winter.

True believers soldiered on, however. And by the early 1980s renewed enthusiasm brought a heyday for researchers in symbolic AI, who received acclaim and funding for "expert systems" that encoded the knowledge of a particular discipline, such as law or medicine. Investors hoped these systems would quickly find commercial applications. The most famous symbolic AI venture began in 1984, when the researcher Douglas Lenat began work on a project he named Cyc that aimed to encode common sense in a machine. To this very day, Lenat and his team continue to add terms (facts and concepts) to Cyc's ontology and explain the relationships between them via rules. By 2017, the team had 1.5 million terms and 24.5 million rules. Yet Cyc is still nowhere near achieving general intelligence.

In the late 1980s, the cold winds of commerce brought on the second AI winter. The market for expert systems crashed because they required specialized hardware and couldn't compete with the cheaper desktop computers that were

becoming common. By the 1990s, it was no longer academically fashionable to be working on either symbolic AI or neural networks, because both strategies seemed to have flopped.

But the cheap computers that supplanted expert systems turned out to be a boon for the connectionists, who suddenly had access to enough computer power to run neural networks with many layers of artificial neurons. Such systems became known as deep neural networks, and the approach they enabled was called deep learning. Geoffrey Hinton, at the University of Toronto, applied a principle called back-propagation to make neural nets learn from their mistakes (see "How Deep Learning Works").

One of Hinton's postdocs, Yann LeCun, went on to AT&T Bell Laboratories in 1988, where he and a postdoc named Yoshua Bengio used neural nets for optical character recognition; U.S. banks soon adopted the technique for processing checks. Hinton, LeCun, and Bengio eventually won the 2019 Turing Award and are sometimes called the godfathers of deep learning.

But the neural-net advocates still had one big problem: They had a theoretical framework and growing computer power, but there wasn't enough digital data in the world to train their systems, at least not for most applications. Spring had not yet arrived.

Over the last two decades, everything has changed. In particular, the World Wide Web blossomed, and suddenly, there was data everywhere. Digital cameras and

then smartphones filled the Internet with images, websites such as Wikipedia and Reddit were full of freely accessible digital text, and YouTube had plenty of videos. Finally, there was enough data to train neural networks for a wide range of applications.

The other big development came courtesy of the gaming industry. Companies such as Nvidia had developed chips called graphics processing units (GPUs) for the heavy processing required to render images in video games. Game developers used GPUs to do sophisticated kinds of shading and geometric transformations. Computer scientists in need of serious compute power realized that they could essentially trick a GPU into doing other tasks—such as training neural networks. Nvidia noticed the trend and created CUDA, a platform that enabled researchers to use GPUs for general-purpose processing. Among these researchers was a Ph.D. student in Hinton's lab named Alex Krizhevsky, who used CUDA to write the code for a neural network that blew everyone away in 2012.

He wrote it for the ImageNet competition, which challenged AI researchers to build computer-vision systems that could sort more than 1 million images into 1,000 categories of objects. While Krizhevsky's AlexNet wasn't the first neural net to be used for image recognition, its performance in the 2012 contest caught the world's attention. AlexNet's error rate was 15 percent, compared with the 26 percent error rate of the second-best entry. The neural net owed its runaway victory to GPU power and a "deep" structure of multiple layers containing

650,000 neurons in all. In the next year's ImageNet competition, almost everyone used neural networks. By 2017, many of the contenders' error rates had fallen to 5 percent, and the organizers ended the contest.

Deep learning took off. With the compute power of GPUs and plenty of digital data to train deep-learning systems, self-driving cars could navigate roads, voice assistants could recognize users' speech, and Web browsers could translate between dozens of languages. AIs also trounced human champions at several games that were previously thought to be unwinnable by machines, including the ancient board game Go and the video game StarCraft II. The current boom in AI has touched every industry, offering new ways to recognize patterns and make complex decisions.

A look back across the decades shows how often AI researchers' hopes have been crushed—and how little those setbacks have deterred them.

But the widening array of triumphs in deep learning have relied on increasing the number of layers in neural nets and increasing the GPU time dedicated to training them. One analysis from the AI research company OpenAI showed that the amount of computational power required to train the biggest AI systems doubled every two years until 2012—and after that it doubled every 3.4 months. As Neil C. Thompson and his colleagues write in "Deep Learning's Diminishing Returns," many researchers worry that AI's computational needs are on an unsustainable trajectory. To avoid busting the planet's energy budget, researchers need to bust out of the established ways of constructing these systems.

While it might seem as though the neural-net camp has definitively tromped the symbolists, in truth the battle's outcome is not that simple. Take, for example, the robotic hand from OpenAI that made headlines for manipulating and solving a Rubik's cube. The robot used neural nets and symbolic AI. It's one of many new neuro-symbolic systems that use neural nets for perception and symbolic AI for reasoning, a hybrid approach that may offer gains in both efficiency and explainability.

Although deep-learning systems tend to be black boxes that make inferences in opaque and mystifying ways, neuro-symbolic systems enable users to look under the hood and understand how the AI reached its conclusions. The U.S. Army is particularly wary of relying on black-box systems, as Evan Ackerman describes in "How the U.S. Army Is Turning Robots Into Team Players," so Army researchers are investigating a variety of hybrid approaches to drive their robots and autonomous vehicles.

Imagine if you could take one of the U.S. Army's road-clearing robots and ask it to make you a cup of coffee. That's a laughable proposition today, because deep-learning systems are built for narrow purposes and can't generalize their abilities from one task to another. What's more, learning a new task usually requires an AI to erase everything it knows about how to solve its prior task, a conundrum called catastrophic forgetting. At DeepMind, Google's London-based AI lab, the renowned roboticist Raia Hadsell is tackling this problem with a variety of sophisticated techniques. In "How DeepMind Is Reinventing the Robot," Tom

Chivers explains why this issue is so important for robots acting in the unpredictable real world. Other researchers are investigating new types of meta-learning in hopes of creating AI systems that learn how to learn and then apply that skill to any domain or task.

All these strategies may aid researchers' attempts to meet their loftiest goal: building AI with the kind of fluid intelligence that we watch our children develop. Toddlers don't need a massive amount of data to draw conclusions. They simply observe the world, create a mental model of how it works, take action, and use the results of their action to adjust that mental model. They iterate until they understand. This process is tremendously efficient and effective, and it's well beyond the capabilities of even the most advanced AI today.

Although the current level of enthusiasm has earned AI its own Gartner hype cycle, and although the funding for AI has reached an all-time high, there's scant evidence that there's a fizzle in our future. Companies around the world are adopting AI systems because they see immediate improvements to their bottom lines, and they'll never go back. It just remains to be seen whether researchers will find ways to adapt deep learning to make it more flexible and robust, or devise new approaches that haven't yet been dreamed of in the 65-year-old quest to make machines more like us.