# Can this guy help Intel catch the AI wave?

September 28 2017, by Russ Mitchell, Los Angeles Times

To get a sense of computer scientist Naveen Rao, just take a look at his hands.

The 42-year-old has busted all 10 of his fingers over a lifetime of skiing, skateboarding, bicycling, rollerblading, race-car driving, wrestling and hoops.

He's not a clod; he's a risk taker who pushes physical and mental boundaries. On the mental side, he's trying to quicken the computer industry's move into a new age of artificial intelligence by creating chips and software inspired by the structure of the human brain.

What sets Rao apart from others attempting the same thing is the fact that Intel last year bought his San Diego company, Nervana, for $400 million.

That's some stamp of approval. Intel is a computer-chip industry giant, with sales topping $60 billion a year. But it's an aging giant. Intel turns 50 next year, and everyone agrees it requires some revitalization, having infamously missed the industry's massive shift to mobile computing.

Now Intel is trying to catch the industry's next rising wave: artificial intelligence, or, more precisely, a subset of AI known as deep learning.

Cars that drive themselves, personal devices that converse with their owners and carry out tasks, social media features that can identify your friends in group photographs - all are made possible by deep learning. So

are computers that can diagnose CAT-scan images on their own or pick stocks for [hedge fund managers](#).

Snap a picture of a road sign in Spain and your phone translates it into English. That app is based on deep learning. Recommendations just for you on Netflix and Spotify? Deep learning plays a role in those features, too.

"We can now solve problems that we couldn't solve before," said Peter Norvig, director of research at Google and a key figure in the history of AI.

Deep learning is a new, more marketing-friendly term for a concept that's been around for decades, known as neural networks. Rather than run calculations in serial fashion, like traditional computers, neural networks mimic the behavior of nerve clusters in the human brain - firing signals and stifling them, arranging data into patterns roughly analogous to human memory.

Why is deep learning suddenly on the rise? The enormous amount of data being created around the world from sensors mounted in everything from smartphones to drones to surveillance cameras and more, matched with ever more powerful chips and other computer hardware, make it possible.

As a result, the computer industry is changing "faster than Intel or anyone else expected it to," said Mario Morales, semiconductor analyst at IDC. Experts say artificial intelligence is now a major growth business that will rake in billions of dollars for companies that provide the best hardware and software.

That's where Rao comes in. His job is to help Intel move beyond the central processing unit, or CPU - the product that's ruled the

semiconductor industry for several decades. The CPU is at the heart of every desktop and laptop computer built since the start of the personal computer revolution in the 1970s. Two companies dominated that field: Microsoft, with its operating system software, and Intel, with its "Intel Inside" central processors.

Deep learning software runs on CPUs, but inefficiently. CPUs are inherently general purpose devices. Maximum computing power required for deep learning is possible only with new kinds of special purpose chips.

The holy grail is a new kind of chip tailor-made for deep learning. That's what Rao and Nervana are attempting. If they succeed, Intel too will bear the fruit.

Rao sports an athlete's build in his T-shirt and blue jeans. A child of Indian immigrants, he grew up in a tiny eastern Kentucky town called Whitesburg. In the late 1970s, his father, a physicist turned physician, drove Rao and his brother to a Radio Shack in Hazard, 40 minutes away, where they wrote programs in Basic on floor-model computers. "We didn't have a computer at home yet," he says.

After a rural childhood that mixed outdoor sports with Dungeons & Dragons and novels by Asimov, Heinlein and Tolkien, Rao attended Duke University. There, he was attracted to neural networks after learning how the human eye detected edges on visual objects. Edge detection at the time was a cutting-edge problem in computer vision, with a number of solutions being offered up, including neural networks.

He cut his teeth on computer chips at Sun Microsystems in Silicon Valley. Still drawn by all things neural, he earned a Ph.D. at Brown University under neuroscience pioneer John Donoghue. He headed back West to chipmaker Qualcomm in San Diego, on a team conducting

neural net research.

Qualcomm is doing interesting work, Rao said. But he had specific ideas of his own and craved the freedom of a startup. In 2014, he founded Nervana. It wasn't long before larger companies took an interest. When Intel found out another big tech company was sniffing around, Rao says, it moved quickly. Which was fine with Rao. More than anything, he sought a company that excelled at chip manufacturing. "The best at that is Intel," he said.

Analysts concur. "They have the best manufacturing facilities in the world," said Linley Gwennap, founder of The Linley Group and longtime semiconductor industry analyst. "With Intel, you also have the advantage of walking into a market with Intel's name, Intel's resources."

Nervana is crafting a neural net computer chip that Intel will release by the end of the year, known as Lake Crest. Just as important, Intel and Nervana are coming up with sets of software tools developers can use to write deep learning programs.

A chip company called Nvidia, a rising star in Silicon Valley, has an early lead on Intel and all other neural net newcomers, including Advanced Micro Devices.

Gamers know Nvidia well. The company started out making graphics accelerators that worked with Intel chips to keep computer games from lagging. Nvidia continued to improve the chips as gamers demanded action in real time so they could compete with players online.

The company's graphical processing units, or GPUs, were designed to cram thousands of "cores" into one chip that processed information in parallel. It turned out GPUs lend themselves to neural networks, too.

Nvidia discovered this around the time of what's become known as the Google Cat Project, considered a milestone in deep learning research. Show a neural network thousands of pictures of cats, and it learns how to recognize photos that include images of cats. The Google cat project in 2012 was carried out on 16,000 Intel central processors inside Google's vast "farms" of computer servers. It took a month. Google could afford to do that, but not many others could.

Around the same time, forward-thinking hedge fund managers were trying to adapt deep learning to stock picking. They began using chips from Nvidia, whose graphic chips lent themselves to neural net processing better than anything made by Intel.

When Nvidia found out hedge funds and others were using its chips for deep learning, it made a quick strategic move: tailor its chips and develop software tools to support neural networks. Almost every major automaker is using Nvidia chips to develop driverless-car technology. Google, Amazon and others have been adding Nvidia chips to their data centers at a furious clip.

A few numbers illustrate the challenge Intel faces. Nvidia stock is up more than 165 percent over the past 52 weeks, closing at $171.96 on Tuesday. Over the same period, Intel stock is up 2.2 percent, closing on Tuesday at $37.47.

Nvidia revenue grew to $6.9 billion in 2016, up 38 percent. Intel was up 7.3 percent, to $59.3 billion. Nvidia does have a much smaller revenue base. But consider profits: Nvidia up 171 percent to $1.6 billion, Intel down 9.7 percent to $10.3 billion.

Rao said the deep learning "kind of fell into Nvidia's lap." He added: "To their credit, they took full advantage of it."

In a statement, Nvidia affirmed its commitment to artificial intelligence.

"Artificial intelligence is driving the greatest technology advances known to humankind," the company said. "From diagnosing skin cancer using a photo to making our roads safer with self-driving cars, AI will automate intelligence and spur a wave of social progress unmatched since the industrial revolution."

Rao sees a way to surpass Nvidia with chips designed not for computer games, but specifically for [neural networks](#).

He'll have to integrate them into the rest of Intel's business. Artificial intelligence chips won't work on their own. For a time, they'll be tied into Intel's CPUs at cloud data centers around the world, where Intel CPUs still dominate - often in concert with Nvidia chips.

Intel is on an acquisition spree. In 2016, it acquired Movidius, a Silicon Valley company that specializes in making smart vision chips for consumer devices, including drones. Earlier this year, it paid a whopping $15.3 billion for Mobileye, an Israeli maker of camera, chip and software systems for driverless cars.

It's also partnered with a company that specializes in custom chips for specific applications, and bought another that makes chips whose "firmware" can be reprogrammed depending on the job at hand.

It needs to pull all those pieces together. When Intel bought Nervana, it deemed the small company the "foundation" of its foray into [artificial intelligence](#). It put Rao in charge of all of Intel's AI efforts, reporting directly to Intel Chief Executive Brian Krzanich. (Krzanich was unavailable for an interview.)

But pulling off the merger will be tricky, in terms of culture and

technical execution. Intel suffers a poor record with startup acquisitions, analyst Gwennap said. Intel had a wireless-device chip unit before Apple introduced the iPhone, he noted, but sold it to Marvell Technology Group in 2006.

To try to correct its error, Intel bought the wireless unit of Infineon Technologies in 2010, but culture clashes and disagreements about technology knocked Intel out of the mobile game. Nimbler companies such as ARM Holdings (now part of SoftBank) ran away with mobile chips business.

Steve Jurvetson of DFJ Venture Capital, an early investor in Nervana, said Rao's got the chops to help pull off a cultural change at Intel, if Intel will let him. "He's a polymath, and he's brilliant in his ability to integrate ideas across a lot of disciplines," said Steve Jurvetson of DFJ Venture Capital. "He has a warm professorial kind of manner. He enjoys teaching others and debating the viability of their ideas. But he won't hold back if he thinks something's a bad idea."

Rao said he's encountered resistance from some of Intel's old guard, but believes the company has left its Intel-always-knows-best past behind it.

Buying Nervana "is not something the Intel of five years ago would have done," he said. "It's an open culture. I can say, 'Guys, you aren't getting this.'" Sometimes, he's found himself pounding on the table in meetings. But he said he's confident that Krzanich has his back.

"Intel has an opportunity to match or surpass the kind of performance Nvidia is talking about," Gwennap said. "But just because you start from a blank sheet of paper doesn't mean you're going to come up with a masterpiece. Nothing is guaranteed."