

Intelligent Machines

Microsoft Thinks AI Will Fill Your Blind Spots, Not Take Over Your Job

The company is looking to improve the way AI and humans get along, but it says we should think differently about how we ask machines to explain themselves.

by Jamie Condliffe July 12, 2017



Eric Horvitz, managing director of Microsoft Research, speaking in London.

Hot on the heels of Google, Microsoft has launched an initiative that it hopes will enable humans and artificial intelligence to complement each other more effectively.

At an event in London on Wednesday, Microsoft announced that it's bringing together a new team of 100 engineers and researchers under the umbrella of Microsoft Research AI at its headquarters in Redmond, Washington. The company says that it's an effort to break down barriers between people who have until now been working across separate areas of AI. Speaking at the event, Eric Horvitz, the managing director of Microsoft Research, said that he thinks the initiative will put Microsoft on "the path to understanding the mysteries of human intellect."

A big part of the initiative is to help improve human-AI collaboration. "When computers can speak human [and] balance the smarts of IQ with the empathy of EQ ... then every human will be able collaborate with computers," explained Harry Shum, executive vice president of Microsoft's AI and Research Group. The move echoes a similar announcement earlier this week from Google ([50 Smartest Companies 2017](#)), which [launched its new People + AI Research \(PAIR\) program](#) to try to improve the way humans and machines work with each other.

Ensuring that humans and AI can neatly coexist will be hugely important for business. Right now, you might think of algorithms as simplistic aides, while in the future, it's often said, they'll steal away jobs. But there's a gulf between those realities, and the truth is that humans and machines [will labor together for decades to come](#).

For its part, Microsoft ([50 Smartest Companies 2017](#)) says it wants to focus on how AI can help fill the gaps in human intelligence, rather than simply re-creating it in silico. So its new team aims to lean on cognitive psychology in order to identify holes in human intellect—such as our propensity to forget things or be easily distracted—and use those to build AIs that complement the blind spots. As an example, the team pointed to a project it's working on that uses machine learning to digest historical medical cases and alert doctors to potential problems that they may have missed when making a diagnosis or discharging a

patient. The implication is that AI shouldn't necessarily take over from humans but, rather, help them do a better a job.

The company will also try to develop new ways to test its machine-learning tools so that they don't go haywire in the real world even when the worked in the lab (see, for instance, [Microsoft's accidentally neo-Nazi sexbot, Tay](#)) and iron out biases that [creep into AIs via the data sets they're trained on](#).

Finally, it hopes to explore the thorny issue of getting AIs to explain themselves. Currently, it's [incredibly difficult to understand how a deep-learning system has reached a decision](#), and that's a huge concern when artificial intelligence is increasingly used to make decisions that affect people—from loan approvals at banks to law enforcement in the courts.

"I think there's a lot that can be done that's not taken to be what we [usually] mean by explanation," Horvitz said in response to a question from *MIT Technology Review*. "It may be more like the answer to a question: What if? In a medical diagnosis, what if I hadn't had hepatitis? What if I was a woman versus a man? These are called sensitivity analyses, and to visualize how robust or how unstable a recommendation is to different inputs [in this way] is another kind of explanation. Our teams are looking at many different dimensions of explanation."

Ultimately, though, he reckons that our current nervousness about understanding each and every decision made by AI might be transient, and may fade once we're familiar with the technology. "I think someday we'll discover that most people are happy to know that an expert certified as best practices the data, the inference, the reasoning, the machinery [that form AI]," he said. "Just the same way that you trust a carburetor in your car: you don't need an explanation every morning of how it's going to work today."

(Read more: ["Your Best Teammate Might Someday Be an Algorithm," "The Dark Secret at the Heart of AI," "Skype's Gone Multilingual"](#))

Hear more about artificial intelligence at EmTech MIT 2017.
[Register now](#)