Ask the AI experts: Should we be afraid of AI?

July 2017

AI researchers largely agree that fears about machines infused with artificial intelligence becoming self-acting and overpowering humans are overblown.

With advances in artificial-intelligence technology occurring more rapidly than ever, the potential for AI to assist us in nearly everything we do at work and at home has become very real. However, some fear that along with AI's tremendous upside of delivering efficiencies humans could not possibly realize on their own comes a dark side—the possibility that super-intelligent AI machines could develop complete autonomy and act against human interests. Earlier this year at the AI Frontiers conference in Santa Clara, California, we sat down with AI experts from some of the world's leading technology-first organizations to find out if fears about AI overtaking humankind have any founding. An edited version of their remarks follows.

This video is one in a five-part Ask the AI Experts series that answers top-of-mind questions about the technology:

- What's driving today's progress in AI?
- What are the applications of AI?
- What will take AI capabilities to the next level?
- What advice would you give to executives about AI?

Interview transcript

Adam Coates, *director, Baidu Research Silicon Valley AI Lab:* I do think sometimes we get carried away and start to think about sentient machines—machines that are just going to understand everything the way that we do and totally interact with us like a human. I think that stuff is pretty far away. And a lot of the scare mongering around AI taking over the world, AI doing all of these things that are negative aspects of the technology that we don't control, I think these are a little bit overwrought.

When I think about the power of AI, the thing that we're really, really good at is that we can take inputs and map them to outputs. This is a prediction problem that we're unbelievably good at. And within that framework, there are just so many positive things we can do that a lot of this other stuff about sentience feels to me like a distraction.

Li Deng, *chief AI officer, Citadel:* For the people who actually work on artificial intelligence, we worry about if there are certain kinds of behavior we want to correct, and that it takes so much thinking involved to make that happen. People talk about the danger of AI, if it is going to harm humans—I think that kind of argument is really overhyped. I think it overestimates the technology in terms of the speed of advancement.

Gary Bradski, *chief technology officer, Arraiy:* Deep nets [deep neural networks, also known as deep learning] are a data-flow architecture. You train them up, you pour something in, you get pattern recognition or large-scale pattern matching coming out the other side. There's no thought in there. There's nothing like sentience. These things are pattern recognizers. They're not something that thinks. They can innovate to an extent, where you feed it one pattern, and then it can put another pattern, or a picture, in the style of another pattern. But that's built in by its training. The networks don't wake up and say, "I'm gonna invent a new kind of game," or whatever. So there's been very little progress on what you'd call real living intelligence. It's not clear why you would want it [real living intelligence], except for space exploration and in other dangerous areas—then you do want things to be able to fend for themselves and live autonomously and repair themselves. But mostly you don't want your washing machine thinking for itself too much. ☐

Simon London is the director of McKinsey digital communications and is based in McKinsey's Silicon Valley office. **Gary Bradski** is the chief technology officer for Arraiy, **Adam Coates** is the director of the Baidu Research Silicon Valley AI Lab, and **Li Deng** is the chief AI officer for Citadel.