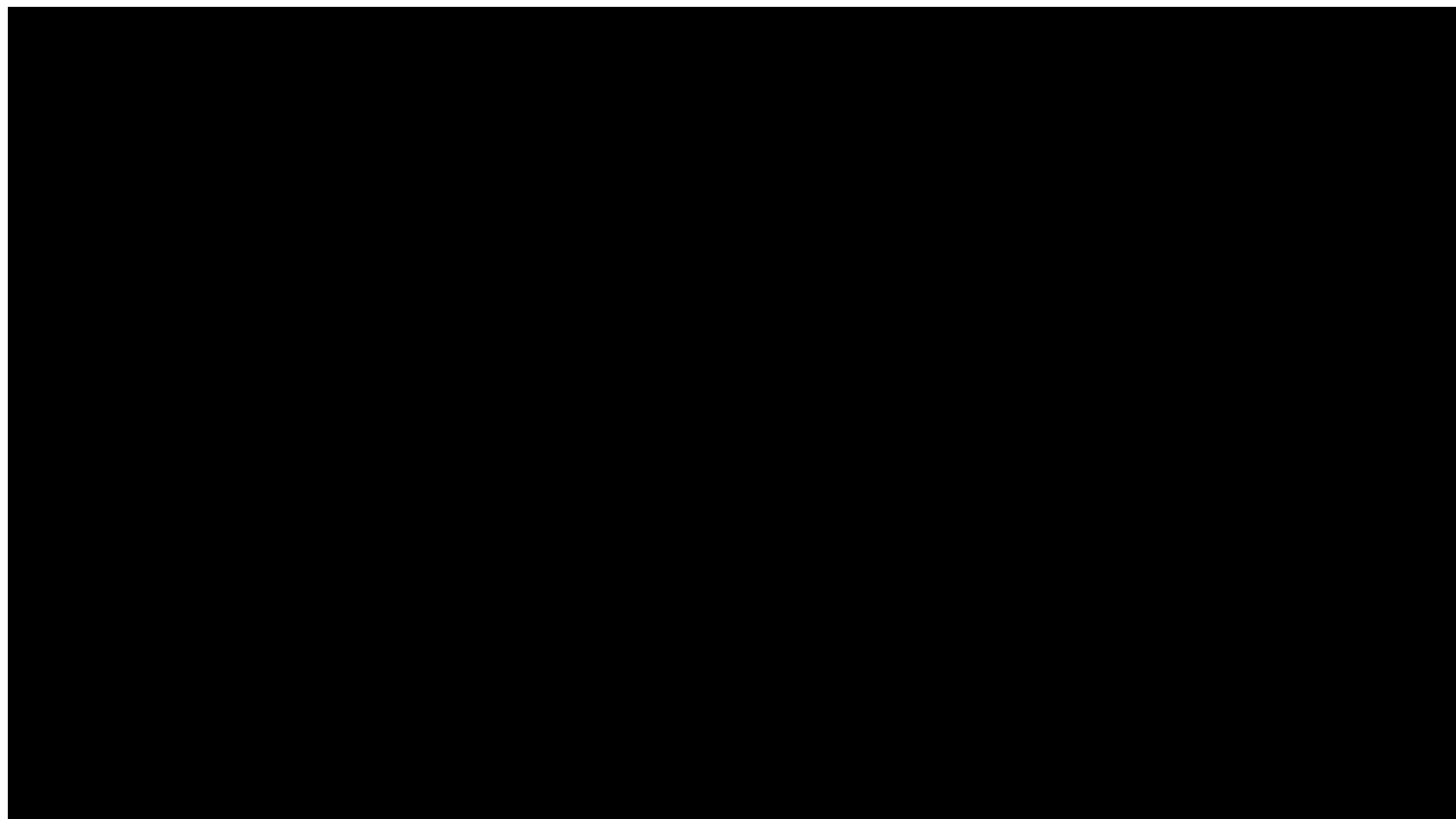


Could artificial intelligence help humanity? Two California universities think so



In the UC Berkeley laboratory of Pieter Abbeel, the PR2 robot figures out a plan for doing laundry and then executes it. (Siddharth Srivastava)

By **Amina Khan**

AUGUST 31, 2016, 1:00 PM

Call it artificial intelligence with a human touch. This week, two California universities separately announced new centers devoted to studying the ways in which AI can help humanity.

USC's Viterbi School of Engineering and its School of Social Work said Wednesday that they had joined forces to launch the Center on Artificial Intelligence for Social Solutions. A day earlier, UC Berkeley unveiled its newly minted Center for Human-Compatible Artificial Intelligence.

Even as science and technology pundits (including Stephen Hawking, Bill Gates and Elon Musk) warn of the overthrow of humanity by advanced artificial intelligence — a prospect that appears nowhere on the horizon, experts say — scientists are increasingly looking ahead to the ways in which AI might actually aid human lives.

The UC Berkeley-led center, directed by artificial intelligence researcher Stuart Russell, will seek to understand how human values can be built into AI's design, and create a mathematical framework that will help people

build AI systems that are beneficial to humanity.

[See the most-read stories in Science this hour »](#)

One of the many questions they'll be wrestling with, for example, is how to get robots to understand what humans really want (because humans are notoriously bad at communicating their objectives). Russell called it the King Midas problem. In Greek mythology, Midas asked for everything he touched to be turned to gold. As this meant his food and drink turned to metal, he died in misery and starvation. It didn't occur to Midas that he didn't really mean "everything" until it was too late.

Scientists might get around this communication problem by designing artificial intelligence that can watch humans and learn what their values are through their actions (though even that comes with some uncertainty, as humans don't always act in ways aligned with their values, Russell added).

"My objective ... is primarily to look at these long-term questions of how you design AI systems so that you are guaranteed to be happy with the outcomes," Russell said. (And if they design some useful software or devices as they do so, even better.)

The USC center, co-directed by artificial intelligence researcher Milind Tambe and social work scientist Eric Rice, seems to operate in a mindset perpendicular to the one at UC Berkeley: It seeks to harness AI's existing capabilities to solve problems in messy, complicated human contexts.

Tambe has led a workshop sponsored by the White House Office of Science and Technology Policy on using AI for "social good." He has used AI to help rangers reduce the poaching of wildlife and help LAX security officials catch more weapons, drugs and other contraband. He and Rice are working on a project that exemplifies the kind of work the center could do: using artificial intelligence to identify key people in social networks to help prevent the spread of HIV among Los Angeles' homeless youth.

AI also includes a wide range of tools, including machine learning, computer vision, natural language processing and game theory (though some may consider game theory part of another discipline, Tambe said). Some of these areas have analogs to aspects of human intelligence. Tambe said he hopes that as more researchers get involved in the center, more of these computational tool sets will be put to good use.

"An agreed upon definition of AI that is acceptable to everyone is very hard to come by," Tambe said. "But essentially all of the kinds of human reasoning that may be applied to problems, AI wants to be able to do that and more."

Rice said he saw potential for these techniques to be applied to a host of thorny problems in different human contexts, from the impact of global warming on impoverished communities to issues with the child welfare system, homelessness and healthcare access.

Though the center's founding directors have very different backgrounds, the pair's distinct skill sets complement and enhance each other, Rice explained.

“If you bring together people from social work, who have this understanding of the complexity of the real world, with people from computer science who can model incredibly complex systems, it creates a really great way of moving forward and getting traction on these complicated problems,” Rice said.

amina.khan@latimes.com

Follow [@aminawrite](#) on Twitter for more science news and "like" [Los Angeles Times Science & Health](#) on Facebook.

MORE IN SCIENCE

[3.2 million years after her death, autopsy reveals Lucy probably died after falling from a tree](#)

[Experimental drug reduces protein clumps and slows memory loss in early Alzheimer's](#)

[After Ohio passed abortion law, risk of complications nearly tripled](#)

Copyright © 2016, Los Angeles Times