Striking an ethical balance in AI

While he acknowledges that humans must be mindful of how we allow rapidly advancing technology to impact our lives, Daniel Kelly does not buy into the doomsday scenarios tying the beginning of artificial intelligence (AI) to the end of humanity.

Instead, the associate professor of philosophy points to centuries of recorded human history that indicate the exact opposite.

"The way that the human evolutionary trajectory shot off from that of other primates was driven by our capacity to cooperate and build technologies which allowed us to collectively solve problems," said Kelly, who discussed the subject in a presentation titled, "Minds, Culture, and the Evolution of Intelligence: What's Going to Happen to Us?" at the 2014 Dawn or Doom conference. "Our intelligent capacities have never been completely confined to our biological brains."
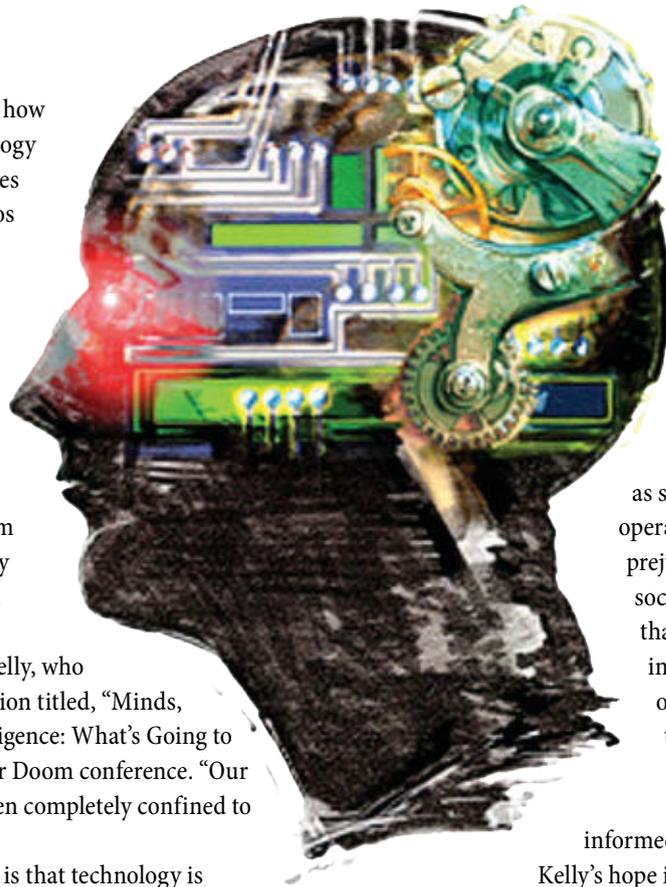
The difference today, of course, is that technology is exponentially more sophisticated than it was hundreds of years — much less decades — ago. Advancements in areas like AI and machine learning are increasingly affecting our everyday lives, sometimes in detrimental ways.

As these technologies continue to develop, it is imperative for humanity to develop an ethical structure to harness their potential instead of spelling our doom. Kelly fears society is already behind in that regard.

"How to have good oversight, and by whom, and who understands it well enough to have effective oversight are all central issues. In the last couple of decades, our technologies have developed so rapidly and have become so integral to our lives so quickly, that I think our ethical systems for thinking about how to deal with a lot of them are still playing catch-up," Kelly said.

As an example, Kelly cited his project on human reliance on algorithms to sort through big data, ignoring the reality that algorithms reflect the biases of their human programmers. There are times, Kelly said, where this allows the algorithm to "turbocharge" injustices we already see in society.

"One of the worries here is that because algorithms, in the public imagination, have this veneer of mathematical infallibility, this connotation of pure objectivity, that we're going to give them a kind of epistemic authority that they don't really deserve. 'Well, the algorithm says that this person is a high risk to commit another crime, so, of course, they should not get parole. They should be thrown back in jail.' But algorithms are not infallible oracles," Kelly said.

"Their outputs shouldn't be taken as sacrosanct — especially when they are operating with data that can reflect past prejudices that have been baked into our society already. So one worry might be that we'll start treating these artificial intelligences as being better and more objective than they actually are; we'll cede too much authority to them too soon when in fact they're not anywhere close to being able to make fine-grained and ethically informed judgments."

Kelly's hope is that those working in the humanities will provide a framework for benefiting from these radical technological advancements while still maintaining some semblances of personal privacy and transparent, ethical behavior from developers and big tech companies.

Maintaining that balance will be one of the major societal challenges of the coming decades, but the technological advancement itself does not concern Kelly. Just as previous generations of humans leveraged innovation to further the evolution of the species, Kelly predicts these developments will only further that progress.

"What's made us distinctively human is that we've been doing stuff like this — externalizing our intelligence and problem-solving capabilities into useful technologies — for hundreds of thousands of years," Kelly said. "We create new gadgets, and then change as we come to rely on those gadgets. This pattern has driven a feedback loop that over the long haul has helped generate our big brains and massively increased our collective problem-solving capacities.

"There's something ironic about the worries about whether, in general, artificial or extra-biological intelligence is going to be the doom of us. If handled poorly, sure it might be. But there's an interesting sense it was key to genesis of us as a species, too."

By David Ching

150 YEARS OF GIANT LEAPS