

# Facetious blueprint

An AI researcher (Kai-Fu Lee) has written an article ("[A Blueprint for Coexistence with Artificial Intelligence](#)") that fails at its own goal, spectacularly. He assures us we need not end up in the dystopias so common in movies and literature. And I wanted to his assurances to be convincing, because I happen to feel the same way. But they aren't, because he (deliberately, I say) pulls a bait-and-switch, and so doesn't attend to the actual matter of interest.

Lee dwells on narrow AI, good for using deep learning to achieve one task (say, winning at Go). He is breezily dismissive of general AI, machine learning that can learn new tasks. It's not here (he says, correctly) and it's not close (he asserts, with little basis). I have to assume that, as a research in AI, he is aware of the history of the field, with its fits and starts and sudden, unexpected breath-taking leaps of innovation. Is it guaranteed that we will get generalized AI within a century? Of course not -- but I don't think the smart money bets against.

Restricting himself to narrow AI, Lee obviates the import of his article. No one is worried that AlphaGo is going to take over the world. And it was the [MCP's](#) chess-playing origins that made it a threat. The AI people fear -- the AI we have to worry about coexisting with -- is the AI that can think as well as and as broadly as a human ... or better.

To be fair, Lee correctly notes the actual threats that narrow AI poses for our economy. Lots of people are going to find themselves replaced by machines that perform their jobs more quickly, more surely, and more cheaply. It's not in any way obvious that those jobs will be replaced, the way that jobs lost in the Industrial Revolution were replaced. We're not swapping out muscles; we're swapping out minds. For the past two centuries, industry has paid an increasing premium on renting the supercomputer in your skull. If actual supercomputers can outperform your brain, it's not clear what economic asset of value you'll have left.

Lee's "blueprint" revolves on that. He calls for an economy that recognizes the (so far) unique capacity of the human mind for *love*. And he might be right. Surely we will need to find some other asset, and our computational capacity will not be it.

Narrow AI, tailored to narrow goals, will never threaten human dominance in that arena.

But general AI is coming, probably sooner than Lee thinks, and there is no good reason to assume that emotions will remain forever outside its ken. After all, emotions served some evolutionary purpose for *us*.