

## OPINION

# The Challenge of Being Human in the Age of AI

By Henry Kissinger,  
Eric Schmidt  
And Daniel Huttenlocher

The White House Office of Science and Technology Policy has called for “a bill of rights” to protect Americans in what is becoming “an AI-powered world.” The concerns about AI are well-known and well-founded: that it will violate privacy and compromise transparency, and that biased input data will yield biased outcomes, including in fields essential to individual and societal flourishing such as medicine, law enforcement, hiring and loans.

**Reason is our primary means of understanding the world. How does that change if machines think?**

But AI will compel even more fundamental change: It will challenge the primacy of human reason. For all of history, humans have sought to understand reality and our role in it. Since the Enlightenment, we have considered our reason—our ability to investigate, understand and elaborate—our primary means of explaining the world, and by explaining it, contributing to it. For the past 300 years, in what historians have come to call the Age of Reason, we have conducted ourselves accordingly; exploring, experimenting, inventing and building.

Now AI, a product of human inge-

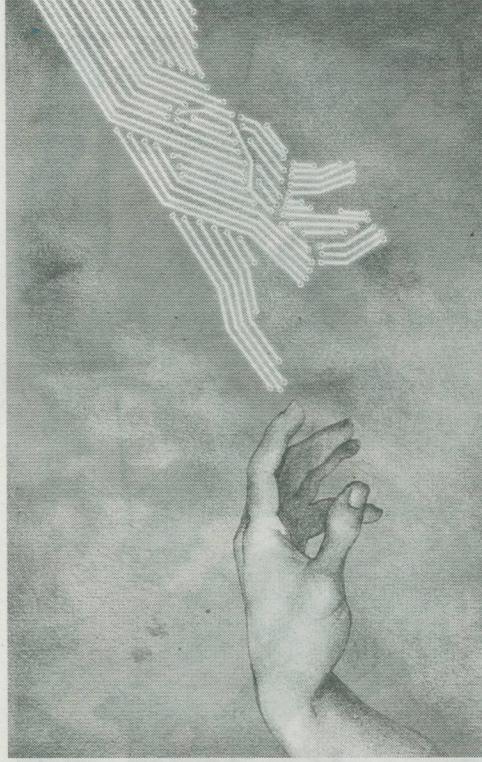
nuity, is obviating the primacy of human reason: It is investigating and coming to perceive aspects of the world faster than we do, differently from the way we do, and, in some cases, in ways we don't understand.

In 2017, Google DeepMind created a program called AlphaZero that could win at chess by studying the game without human intervention and developing a not-quite-human strategy. When grandmaster Garry Kasparov saw it play, he described it as shaking the game “to its roots”—not because it had played chess quickly or efficiently, but because it had conceived of chess anew.

In 2020, halicin, a novel antibiotic, was discovered by MIT researchers who instructed AI to compute beyond human capacity, modeling millions of compounds in days, and to explore previously undiscovered and unexplained methods of killing bacteria. Following the breakthrough, the researchers said that without AI, halicin would have been “prohibitively expensive”—in other words, impossible—to discover through traditional experimentation.

GPT-3, the language model operated by the research company OpenAI, which trains by consuming Internet text, is producing original text that meets Alan Turing's standard of displaying “intelligent” behavior indistinguishable from that of a human being.

The promise of AI is profound: translating languages; detecting diseases; combating climate change—or at least modeling climate change better. But as AlphaZero's performance, halicin's discovery and GPT-3's composition demonstrate,



DAVID KIEN

the fundamental allegation of whistleblower Frances Haugen.

Answering these questions will require concurrent efforts. One should consider not only the practical and legal implications of AI but the philosophical ones: If AI perceives aspects of reality humans cannot, how is it affecting human perception, cognition and interaction? Can AI befriend humans? What will be AI's impact on culture, humanity and history?

Another effort ought to expand the consideration of such questions beyond developers and regulators to experts in medicine, health, environment, agriculture, business, psychology, philosophy, history and other fields. The goal of both efforts should be to avoid extreme reactions—either deferring to AI or resisting it—and instead to seek a middle course: shaping AI with human values, including the dignity and moral agency of humans. In the U.S., a commission, administered by the government but staffed by many thinkers in many domains, should be established. The advancement of AI is inevitable, but its ultimate destination is not.

*Mr. Kissinger was secretary of state, 1973-77, and White House national security adviser, 1969-75. Mr. Schmidt was CEO of Google, 2001-11 and executive chairman of Google and its successor, Alphabet Inc., 2011-17. Mr. Huttenlocher is dean of the Schwarzman College of Computing at the Massachusetts Institute of Technology. They are authors of “The Age of AI: And Our Human Future.”*

If an AI can care for a nursing-home resident—remind her to take her medicine, alert paramedics if she falls, and otherwise keep her company—can her family members visit her less? Should they? If her primary interaction becomes human-to-machine, rather than human-to-human, what will be the emotional state of the final chapter of her life?

And if, in the fog of war, an AI recommends an action that would cause damage or even casualties, should a commander heed it?

These questions are arising as global network platforms, such as Google, Twitter and Facebook, are employing AI to aggregate and filter more information than their users or employees can. AI, then, is making decisions about what is important—and, increasingly, about what is true. Indeed, that Facebook knows aggregation and filtration exacerbates misinformation and mental illness is

the use of AI for an intended purpose may also have an unintended one: uncovering previously imperceptible but potentially vital aspects of reality.

That leaves humans needing to define—or perhaps redefine—our role in the world. For 300 years, the Age of Reason has been guided by the maxim “I think, therefore I am.” But if AI “thinks,” what are we?

If an AI writes the best screenplay of the year, should it win the Oscar? If an AI simulates or conducts the most consequential diplomatic negotiation of the year, should it win the Nobel Peace Prize? Should the human inventors? Can machines be “creative”? Or do their processes require new vocabulary to describe?

If a child with an AI assistant comes to consider it a “friend,” what will become of his relationships with peers, or of his social or emotional development?