The AI Future No One's Predicting

In every forecast, from the most optimistic to the most cautionary, women remain invisible within the planning for humanity's most transformative technological shift

BY STACY KELTNER

HUMANS DREAMED OF INTELLIGENT MACHINES LONG BEFORE the debut of ChatGPT. Homeric golden attendants and the mythical Persian mechanical horse were depicted as beneficent helpers.

But something shifted in the 20th century when actual computing machines emerged. As computers evolved from mechanical calculators to electronic brains, fantasies of artificial intelligence morphed into nightmares. Playwright Karel Čapek's robots rebelled in 1920. HAL 9000 turned murderous in *2001: A Space Odyssey*, released back in 1968. By the time *Ex Machina*'s Ava manipulated her way to freedom in 2015, the existential risk of advanced AI systems had become a serious debate in glass conference rooms overlooking San Francisco Bay and the Gothic halls of universities like Oxford. When tech startup OpenAI, a leader in AI development, debuted its chatbot, ChatGPT, in November 2022, these anxieties were thrust into mainstream policy debates.

A chilling new document, "AI 2027," arrived in April, offering a meticulously detailed month-by-month speculative forecast of how artificial intelligence might surpass human capabilities within just a few years, sparking intense debate about superintelligence (AI that exceeds human intelligence, particularly in soft skills), job displacement and other existential risks. Written by former OpenAI researcher Daniel Kokotajlo and a team of AI researchers, AI 2027 reads like science fiction but is grounded in AI development conversations. Kokotajlo, who risked almost $2 million in OpenAI equity to speak freely about AI risks, provides charts, timelines and technical specifications for exactly how AI might achieve superintelligence and, in one theorized ending, make humanity obsolete by 2030.

Even more surreal, however, is AI 2027's philosophical shell game of debating possible future harms while ignoring present ones. Women, marginalized groups and the differential harms of AI are invisible. Mutale Nkonde, an AI policy expert and founder of AI for the People, suggests that the certainty driving these projections reveals an investment in politically motivated fear and anxiety. She says these warnings advance a "techsolutionist" ideology that presents social-technical problems as requiring only technological solutions, while ignoring how technology and society mutually shape issues of risk, safety and performance.

AI 2027 is grounded in a relatively new field of existential risk (ER) research—the study of threats that could permanently curtail human civilization. ER research was institutionalized in the early aughts and began attracting mainstream attention and significant funding in the 2010s. Over the past few years, ER research has influenced global public policy debates and the development of AI safety initiatives. According to a recent policy paper by the Friedrich Naumann Foundation for Freedom, the rise of ER anxieties is a tale of money. A multimillion-dollar investment movement run by billionaires—like Facebook cofounder Dustin Moskovitz ($510 million through Open Philanthropy) and Skype developer Jaan Tallinn ($117 million for AI safety as well as biological and nuclear threats, via his Survival and Flourishing Fund)—has sunk massive funding into ER and safety research and shaped global policy.

Though the larger field is polarized around how quickly developers might achieve artificial general intelligence (human-level intelligence, adaptability and generalization of function) and its next step, superintelligence, a 2023 Yale survey found that 42 percent of CEOs fear that AI could destroy humanity within five to 10 years, and a 2023 coalition statement calling for "mitigating the risk of extinction from AI" attracted high-profile signatories, including Geoffrey Hinton, the "Godfather of AI," who warned on CNN that "these things are getting smarter than us. … We should worry seriously about how we stop these things getting control over us."

Anxieties are high. Highest on the list of triggers is the "alignment problem": the challenge of ensuring AI systems do what humans want them to do, rather than following instructions out of line with "human values and goals." Recent alignment failures demonstrate risks. Internal tests at Anthropic (founded in 2021 by former OpenAI execs), for example, showed an AI system resorted to blackmail to avoid being taken offline, while another system faked alignment. How can we know what goals advanced AI systems are pursuing as they approach and exceed human-level intelligence?

Here is the irony: AI safety researchers focusing on "alignment" with human intentions create another alignment problem. The techni-

From right: Microsoft Vice Chair and President Brad Smith and Open AI CEO Sam Altman testify during a Senate hearing on artificial intelligence in May; only 22 percent of AI professionals are women.

BRENDAN SMIALOWSKI/GETTY IMAGES

cal goal is to ensure that AI systems follow their creators' intentions, but those intentions are not tested for broader alignment. Women make up only 22 percent of AI professionals, according to a 2022 World Economic Forum report, and ER research is overwhelmingly white, male, advantaged and concentrated in the U.S. and U.K. These technologists are solving for technical precision while potentially uploading misaligned values, embedding bias in "aligned" AI models.

Meanwhile, harm is accelerating. A new U.N. report from the International Labor Organization and Poland's National Research Institute, for example, reveals that in first world countries women are nearly three times more likely than men to work in the positions at the highest risk of AI automation. Further, research from Harvard Business School suggests women are 25 percent less likely to engage AI, making career pivoting more difficult.

As women risk losing economic independence, they could simultaneously face increased surveillance and control through AI systems. Harvard Law scholar Abeer Malik warned, "AI's use in predicting pregnancy extends the controversial concept of predictive policing into reproductive health." Meanwhile, Project 2025's proposed abortion surveillance system would force states to report detailed patient data or lose billions in federal funding. Eliminating economic opportunities while monitoring reproductive choices amplifies and automates patriarchal control through an algorithmic infrastructure.

Discrimination spans across AI models: 2025 studies show ChatGPT uses 24.5 percent fewer female-related words than human writers, and assigns male pronouns to "programmer" 83 percent of the time while using female pronouns for "nurse" 91 percent of the time. A study published in *PNAS Nexus* reported that AI hiring tools show near 0 percent selection rates for Black men's names in résumé screening. Healthcare algorithms trained on male-dominated datasets can misdiagnose women's conditions while preliminary research shows that medical AI's algorithmic bias could contribute to a nearly 30 percent higher mortality rate for Black patients than white ones. Meanwhile, in 2019, 96 percent of deepfake videos were nonconsensual pornography targeting women.

As AI 2027 envisioned, geopolitical competition is prioritizing innovation over safety. When China's DeepSeek released a competitive open-source AI model at a fraction of the cost of models built by U.S. companies like OpenAI, Meta and Anthropic, Silicon Valley panicked. The Trump administration revoked former President Joe Biden's AI safety order and released an action plan to win the AI race. As JD Vance put it, "The AI future is not going to be won by hand-wringing about safety … [but] by building."

While the administration's plan incorporates ER-influenced safety concerns, its primary focus is accelerating development and "upholding free speech," which is code for freeing AI from regulations around bias, misinformation and environmental impact. The framing of AI development as a race with China ignores how Chinese companies are actually "very seriously looking at" algorithmic bias issues, AI policy expert Nkonde notes, because they "see value in markets" across the Global South, where what Americans dismiss as "woke" concerns become essential business considerations. As China expands into African and other non-white markets, questions of algorithmic fairness that the Trump administration labels as DEI (diversity, equity and inclusion) become central to competitive advantage. Intentionally or not, ER research has arguably provided cover for this approach, offering a narrative of "safe" development that allows biased AI systems to be amplified at scale.

The political shift to accelerate innovation was quickly embraced by AI leaders who previously championed safety research. Tech companies qui-

etly changed or reduced Biden-era safety commitments from their websites and amped up competition. Prince Sarpong, a finance professor at the University of South Africa who studies AI governance and masculinity, sees masculine cultural norms driving this shift. "The default masculinity reaction to threatened masculinity is to double down," he told *Ms.*

With most AI decision-makers being men, competition has replaced collaboration. CEOs poach each other's employees while now racing to build the biggest data processing centers. The focus on developing beneficial AI has seemingly been replaced with competitive posturing among tech giants.

Sarpong argues that allowing technologists to control AI's future is like "assigning weapons manufacturers to draft a peace treaty." Just as those who build weapons shouldn't be the sole architects of peace, those building AI systems shouldn't unilaterally determine how they're deployed and governed. Technology policy expert Nina-Simone Edwards advocates for what she calls inclusive governance: involving affected communities throughout AI development, from training to deployment to ongoing auditing. "Understanding not just who is harmed, but how they've been harmed and why they're being harmed" requires bringing people into the process before systems are deployed, she says. The goal is moving from reactive fixes to proactive inclusion, ensuring that those most likely to be harmed by AI systems have a voice in shaping them.

For *Ms.* readers wondering how to resist, Nkonde's advice is both practical and political. First, she says, make conscious decisions about AI use: "Let's bring back writing emails and not put them through LLMs [large language models]. For an LLM to generate an email response I think is currently taking two whole bottles of water." Second, question whether AI actually improves your work: "AI systems are not good at recognizing what I'm trying to do because they don't necessarily understand anything outside of standard English," she notes.

Third, make AI a political issue "on every level you exist," from your local representatives to investment choices, "to make it safer and more equitable," Nkonde says. "Further, whatever your issue is, you need to now include an analysis of how AI will change this and what are the political levers you can pull." Women have market power through investments and consumer choices that can shape AI development.

The future requires both individual resistance and collective political action to ensure AI serves everyone, not just its creators. ■

STACY KELTNER *is a writer and professor based in Atlanta.*