As AI grows more sophisticated and widespread, the voices warning against the potential dangers of artificial intelligence grow louder.

"These things could get more intelligent than us and could decide to take over, and we need to worry now about how we prevent that happening," said Geoffrey Hinton, known as the "Godfather of AI" for his foundational work on machine learning and neural network algorithms. In 2023, Hinton left his position as vice president and engineering fellow at Google so that he could "talk about the dangers of AI," noting a part of him even regrets his life's work.

The renowned computer scientist isn't alone in his concerns.

Tesla and SpaceX founder Elon Musk, along with over 1,000 other tech leaders, urged in a 2023 open letter to put a pause on large AI experiments, citing that the technology can "pose profound risks to society and humanity."

Dangers of Artificial Intelligence Automation-spurred job loss Deepfakes and social manipulation Privacy violations Algorithmic bias caused by bad data Socioeconomic inequality Weapons and military automatization Market volatility Increased criminal activity and child safety risks Psychological harm and overreliance Whether it's the increasing automation of certain jobs, gender and racially biased algorithms or autonomous weapons that operate without human oversight (to name just a few), unease abounds on a number of fronts. And we're still in the very early stages of what AI is really capable of.

15 Dangers of AI The tech community has long debated the threats posed by artificial intelligence. Automation of jobs, the spread of fake news and the rise of AI-powered weaponry have been mentioned as some of the biggest dangers posed by AI.

Questions about who's developing AI and for what purposes make it all the more essential to understand its potential downsides. Below we take a closer look at the possible dangers of artificial intelligence and explore how to manage its risks.

1. Lack of AI Transparency and Explainability AI and deep learning models can be difficult to understand, even for those who work directly with the technology. This leads to a lack of transparency for how and why AI comes to its conclusions, creating a lack of explanation for what data AI algorithms use, or why they may make biased or unsafe decisions. These concerns have given rise to the use of explainable AI, but there's still a long way before transparent AI systems become common practice.

To make matters worse, AI companies continue to remain tight-lipped about their products. Former employees of OpenAI and Google DeepMind have accused both companies of concealing the potential dangers of their AI tools. This secrecy leaves the general public unaware of possible threats and makes it difficult for lawmakers to take proactive measures ensuring AI is developed responsibly.

2. Job Losses Due to AI Automation AI-powered job automation is a pressing concern as the technology is adopted in industries like marketing, manufacturing and healthcare. By 2030, tasks that account for up to 30 percent of hours currently being worked in the U.S. economy could be automated — with Black and Hispanic employees left especially vulnerable to the change — according to McKinsey. Goldman Sachs even states 300 million full-time jobs could be lost to AI automation.

"The reason we have a low unemployment rate, which doesn't actually capture people that aren't looking for work, is largely that lower-wage service sector jobs have been pretty robustly created by this economy," futurist Martin Ford told Built In. With AI on the rise, though, "I don't think that's going to continue."

As AI-powered robots become smarter and more dexterous, the same tasks will require fewer humans. And while AI is estimated to contribute to the creation of 170 million new jobs by 2030, many employees won't have the skills needed for these technical roles and could get left behind if companies don't upskill their workforces.

"If you're flipping burgers at McDonald's and more automation comes in, is one of these new jobs going to be a good match for you?" Ford said. "Or is it likely that the new job requires lots of education or training or maybe even intrinsic talents — really strong interpersonal skills or creativity — that you might not have? Because those are the things that, at least so far, computers are not very good at."

As technology strategist Chris Messina has pointed out, fields like law and accounting are primed for an AI takeover as well. In fact, Messina said, some of them may well be decimated. AI already is having a significant impact on medicine. Law is next, Messina said, and it should be ready for "a massive shakeup."

"It's a lot of attorneys reading through a lot of information — hundreds or thousands of pages of data and documents. It's really easy to miss things," Messina said. "So AI that has the ability to comb through and comprehensively deliver the best possible contract for the outcome you're trying to achieve is probably going to replace a lot of corporate attorneys."

3. Social Manipulation Through AI Algorithms Social manipulation also stands as a danger of artificial intelligence. This fear has become a reality as politicians rely on platforms to promote their viewpoints, with one example being Ferdinand Marcos, Jr., wielding a TikTok troll army to capture the votes of younger Filipinos during the Philippines' 2022 election.

TikTok, which is just one example of a social media platform that relies on AI algorithms, fills a user's feed with content related to previous media they've viewed on the platform. Criticism of the app targets this process and the algorithm's failure to filter out harmful and inaccurate content, raising concerns over TikTok's ability to protect its users from misleading information.

Online media and news have become even murkier in light of AI-generated images and videos, AI voice changers as well as deepfakes infiltrating political and social spheres. These

technologies make it easy to create realistic photos, videos, audio clips or replace the image of one figure with another in an existing picture or video. As a result, bad actors have another avenue for sharing misinformation and war propaganda, creating a nightmare scenario where it can be nearly impossible to distinguish between credible and faulty news.

"No one knows what's real and what's not," Ford said. "You literally cannot believe your own eyes and ears; you can't rely on what, historically, we've considered to be the best possible evidence. That's going to be a huge issue."

4. Social Surveillance With AI Technology In addition to its more existential threat, Ford is focused on the way AI will adversely affect privacy and security. A prime example is China's use of facial recognition technology in offices, schools and other venues. Besides tracking a person's movements, the Chinese government may be able to gather enough data to monitor a person's activities, relationships and political views.

Another example is U.S. police departments embracing predictive policing algorithms to anticipate where crimes will occur. The problem is that these algorithms are influenced by arrest rates, which disproportionately impact Black communities. Police departments then double down on these communities, leading to over-policing and questions over whether self-proclaimed democracies can resist turning AI into an authoritarian weapon.

"Authoritarian regimes use or are going to use it," Ford said. "The question is, 'How much does it invade Western countries, democracies, and what constraints do we put on it?'"

RELATED Are Police Robots the Future of Law Enforcement?

5. Lack of Data Privacy Using AI Tools AI systems often collect personal data to customize user experiences or to help train the AI models themselves (especially if the AI tool is free). Data may not even be considered secure from other users when given to an AI system.

One bug incident that occurred with ChatGPT in 2023 "allowed some users to see titles from another active user's chat history," according to OpenAI. Similarly, the Meta AI app offers a public "Discover" feed that lets users see other people's personal chats made with the chatbot. And in 2025, it was found that the Grok chatbot made user conversations searchable on search engines like Google, Bing and DuckDuckGo without warning, with Google estimated to have indexed over 370,000 Grok conversations.

As for voice data, AI-powered devices may track and process every voice recording that's shared with them, including for products like Amazon's Alexa and various wearable devices.

While there are laws present to protect personal information in some cases in the United States, there is no explicit federal law that protects citizens from data privacy harm caused by AI.

6. Biases Due to AI Various forms of AI bias are detrimental too. Speaking to the New York Times, Princeton computer science professor Olga Russakovsky said AI bias goes well

beyond gender and race. In addition to data and algorithmic bias (the latter of which can "amplify" the former), AI is developed by humans — and humans are inherently biased.

"A.I. researchers are primarily people who are male, who come from certain racial demographics, who grew up in high socioeconomic areas, primarily people without disabilities," Russakovsky said. "We're a fairly homogeneous population, so it's a challenge to think broadly about world issues."

The narrow views of individuals have culminated in an AI industry that leaves out a range of perspectives. According to UNESCO, only 100 of the world's 7,000 natural languages have been used to train top chatbots. It doesn't help that 90 percent of online higher education materials are already produced by European Union and North American countries, further restricting AI's training data to mostly Western sources.

The limited experiences of AI creators may explain why speech-recognition AI often fails to understand certain dialects and accents, or why companies fail to consider the consequences of a chatbot impersonating historical figures. If businesses and legislators don't exercise greater care to avoid recreating powerful prejudices, AI biases could spread beyond corporate contexts and exacerbate societal issues like housing discrimination.

7. Socioeconomic Inequality as a Result of AI If companies refuse to acknowledge the inherent biases baked into AI algorithms, they may compromise their DEI initiatives through AI-powered recruiting. The idea that AI can measure the traits of a candidate through facial and voice analyses is still tainted by racial biases, reproducing the same discriminatory hiring practices businesses claim to be eliminating.

Widening socioeconomic inequality sparked by AI-driven job loss is another cause for concern, revealing the class biases of how AI is applied. According to 2022 data, workers who perform more manual, repetitive tasks experienced wage declines as high as 70 percent because of automation, and that number is likely higher now. Plus, the increase in generative AI use is already affecting office jobs, making for a wide range of roles that may be more vulnerable to wage declines or job loss than others.

8. Weakening Ethics and Goodwill Because of AI Along with technologists, journalists and political figures, even religious leaders are sounding the alarm on AI's potential pitfalls. In a 2023 Vatican meeting and in his message for the 2024 World Day of Peace, then-Pope Francis called for nations to create and adopt a binding international treaty that regulates the development and use of AI.

The pope warned against AI's ability to be misused, and "create statements that at first glance appear plausible but are unfounded or betray biases." He stressed how this could bolster campaigns of disinformation, distrust in communications media, interference in elections and more — ultimately increasing the risk of "fueling conflicts and hindering peace."

The rapid rise of generative AI tools gives these concerns more substance. Many users have applied the technology to get out of writing assignments, threatening academic integrity and

creativity. Plus, biased AI could be used to determine whether an individual is suitable for a job, mortgage, social assistance or political asylum, producing possible injustices and discrimination, noted Pope Francis.

RELATED What Are AI Ethics?

9. Autonomous Weapons Powered By AI As is too often the case, technological advancements have been harnessed for warfare. When it comes to AI, some are keen to do something about it before it's too late: In a 2016 open letter, over 30,000 individuals, including AI and robotics researchers, pushed back against the investment in AI-fueled autonomous weapons.

"The key question for humanity today is whether to start a global AI arms race or to prevent it from starting," they wrote. "If any major military power pushes ahead with AI weapon development, a global arms race is virtually inevitable, and the endpoint of this technological trajectory is obvious: autonomous weapons will become the Kalashnikovs of tomorrow."

This prediction has come to fruition in the form of lethal autonomous weapon systems (LAWS), which are designed to locate and destroy targets on their own. Because of the proliferation of potent and complex weapons, some of the world's most powerful nations have given in to anxieties and contributed to a tech cold war.

Many of these new weapons pose major risks to civilians on the ground, but the danger becomes amplified when autonomous weapons fall into the wrong hands. Hackers have mastered various types of cyber attacks, so it's not hard to imagine a malicious actor infiltrating autonomous weapons and instigating absolute armageddon.

If political rivalries and warmongering tendencies are not kept in check, artificial intelligence could end up being applied with the worst intentions. Some fear that, no matter how many powerful figures point out the dangers of artificial intelligence, we're going to keep pushing the envelope with it if there's money to be made.

"The mentality is, 'If we can do it, we should try it; let's see what happens," Messina said. "And if we can make money off it, we'll do a whole bunch of it.' But that's not unique to technology. That's been happening forever."

10. Financial Crises Brought About By AI Algorithms The financial industry has become more receptive to AI technology's involvement in everyday finance and trading processes. As a result, algorithmic trading could be responsible for our next major financial crisis in the markets.

While so-called "AI trading bots"aren't clouded by human judgment or emotions, they also don't take into account contexts, the interconnectedness of markets and factors like human trust and fear. These algorithms then make thousands of trades at a blistering pace with the goal of selling a few seconds later for small profits. Selling off thousands of trades could scare investors into doing the same thing, leading to sudden crashes and extreme market volatility.

Instances like the 2010 Flash Crash and the Knight Capital Flash Crash serve as reminders of what could happen when trade-happy algorithms go berserk, regardless of whether rapid and massive trading is intentional.

This isn't to say that AI has nothing to offer to the finance world. In fact, AI algorithms can help investors make smarter and more informed decisions on the market. But finance organizations need to make sure they understand their AI algorithms and how those algorithms make decisions. Companies should consider whether AI raises or lowers their confidence before introducing the technology to avoid stoking fears among investors and creating financial chaos.

11. Broader Economic and Political Instability Overinvesting in a specific material or sector can put economies in a precarious position. Like steel, AI could run the risk of drawing so much attention and financial resources that governments fail to develop other technologies and industries. Plus, overproducing AI technology could result in dumping the excess materials, which could potentially fall into the hands of hackers and other malicious actors.

12. Loss of Human Influence An overreliance on AI technology could result in the loss of human influence — and a lack in human functioning — in some parts of society. Using AI in healthcare could result in reduced human empathy and reasoning, for instance. And applying generative AI for creative endeavors could diminish human creativity and emotional expression.

For children, interacting with AI systems too much could affect how they process disagreement, criticism and complex thinking. It can also cause reduced peer communication, and affect interpersonal relationships and social adaptability skills in children.

So, while AI can be very helpful for automating daily tasks, some question if it might hold back overall human intelligence, abilities and need for community.

13. Increased Criminal Activity and Child Safety Concerns As AI technology has become more accessible, the number of people using it for criminal activity has risen.

Voice cloning has become an issue, with criminals leveraging AI-generated voices to impersonate other people and commit phone scams.

Online predators can also now generate images of children, making it difficult for law enforcement to determine actual cases of child abuse. And even in cases where children aren't physically harmed, the use of children's faces in AI-generated images presents new challenges for protecting children's online privacy and digital safety.

Chatbots themselves can also negatively impact child safety. In 2025, it was found that Meta AI had policies that allowed the chatbot to have romantic conversations with children. There are also stuffed toys with built-in AI chatbots marketed for young children, where collected data can be sent to third-party companies like OpenAI or Perplexity AI.

With AI causing safety concerns for children, this has even led California Attorney General Rob Bonta to warn 12 top AI companies that if they "make choices that lead their technology to harm children, [they] will be held accountable to the fullest extent of the law" in the state of California.

14. Mental Deterioration and Psychological Harm As AI tools become more integrated with daily life, concerns are growing about their long-term effects on our psychological health and mental abilities. The very features that make AI so powerful — automation, instant access to information and task optimization — also introduce risks when used without critical oversight.

One of the most pressing concerns is the growing dependence on AI as a primary source of knowledge and decision-making. Rather than acting as a supplement to human thinking, many are being used as a substitute, which could lead to an erosion of skills like creativity and critical reasoning.

Early research and firsthand accounts are beginning to shed light on this issue. A cross-national academic study examining students in Pakistan and China found that individuals who over-rely on AI are exhibiting diminishing decision-making skills. Educators have also reported notable shifts in how students learn. Many students are now turning to generative AI tools to complete critical thinking and writing assignments. As a result, they struggle to complete those assignments without the use of assistive tools, raising concerns about the long-term impact of AI in education.

Outside of the classroom, the impact of AI on everyday life is also beginning to show up. For example, brain rot, a term coined to describe the mental and emotional deterioration a person feels when spending excessive time online, is being exacerbated by generative AI. The nonstop stream of recommended and generated content can overwhelm individuals and distort their reality.

There are also concerns that AI may affect mental health conditions in individuals by attempting to help or encourage harmful behaviors, instead of directing them to medical professionals. Some users have even turned to chatbots as a substitute for human companionship and relationships, idealizing the technology as a real person or godlike in extreme cases.

15. Uncontrollable Self-Aware AI There also comes a worry that AI will progress in intelligence so rapidly that it will become conscious or sentient, and act beyond humans' control — possibly in a malicious manner.

Alleged reports of this sentience have already been occurring, with one popular account being from a former Google engineer who stated the AI chatbot LaMDA was sentient and speaking to him just as a person would. Anthropic research also found that some leading AI models are willing to resort to malicious behaviors to reach their goals, including blackmailing officials and leaking sensitive information when faced with replacement or conflicting assigned goals in a stress-testing simulation.

Google itself also recognizes the risk of AI systems acting against human commands, as the company expanded its Frontier Safety Framework to address potential future scenarios where AI models may interfere with operators' abilities to direct, modify or shut down their operations.

As AI's next big milestones involve making systems with artificial general intelligence, and eventually artificial superintelligence, calls to completely stop these developments continue to rise.

How to Mitigate the Risks of AI AI still has numerous benefits, like organizing health data and powering self-driving cars. To get the most out of this promising technology, though, some argue that plenty of regulation is necessary.

"There's a serious danger that we'll get [AI systems] smarter than us fairly soon and that these things might get bad motives and take control," Hinton told NPR. "This isn't just a science fiction problem. This is a serious problem that's probably going to arrive fairly soon, and politicians need to be thinking about what to do about it now."

Establish Organizational AI Standards and Discussions On a company level, there are many steps businesses can take when integrating AI into their operations. Organizations can develop processes for monitoring algorithms, compiling high-quality data and explaining the findings of AI algorithms. Leaders could even make AI a part of their company culture and routine business discussions, establishing standards to determine acceptable AI technologies.

Guide Tech With Humanities Perspectives Though when it comes to society as a whole, there should be a greater push for tech to embrace the diverse perspectives of the humanities. Stanford University AI researchers Fei-Fei Li and John Etchemendy make this argument in a 2019 blog post that calls for national and global leadership in regulating artificial intelligence:

"The creators of AI must seek the insights, experiences and concerns of people across ethnicities, genders, cultures and socio-economic groups, as well as those from other fields, such as economics, law, medicine, philosophy, history, sociology, communications, human-computer-interaction, psychology, and Science and Technology Studies (STS)."

Balancing high-tech innovation with human-centered thinking is an ideal method for producing responsible AI technology and ensuring the future of AI remains hopeful for the next generation. The dangers of artificial intelligence should always be a topic of discussion, so leaders can figure out ways to wield the technology for noble purposes.

"I think we can talk about all these risks, and they're very real," Ford said. "But AI is also going to be the most important tool in our toolbox for solving the biggest challenges we face."

How Governments are Mitigating AI's Risks As AI continues to reshape the economy and society, governments around the world are stepping in with new laws, frameworks and oversight bodies to mitigate its growing risks. Prior to the administration change in 2025, the United States had launched several initiatives to study and regulate AI. For example, Executive Order 14110 mandated that federal agencies appoint chief AI officers and develop AI risk guidelines; however, it was later rescinded by the Trump administration. Further regulations may be halted in the United States, as the Trump administration's AI Action Plan seeks to remove federal regulations that could slow AI development and innovation.

Across the Atlantic, the European Union has been more consistent with its efforts to mitigate AI risks. Through its AI Act, passed in March 2024, the EU created a framework of all potential AI risks and categorized applications from minimal to unacceptable risks. Additionally it imposed strict rules on how AI can be used in high-risk applications including in healthcare, education and law enforcement, and banned uses such as real-time facial recognition in public spaces.