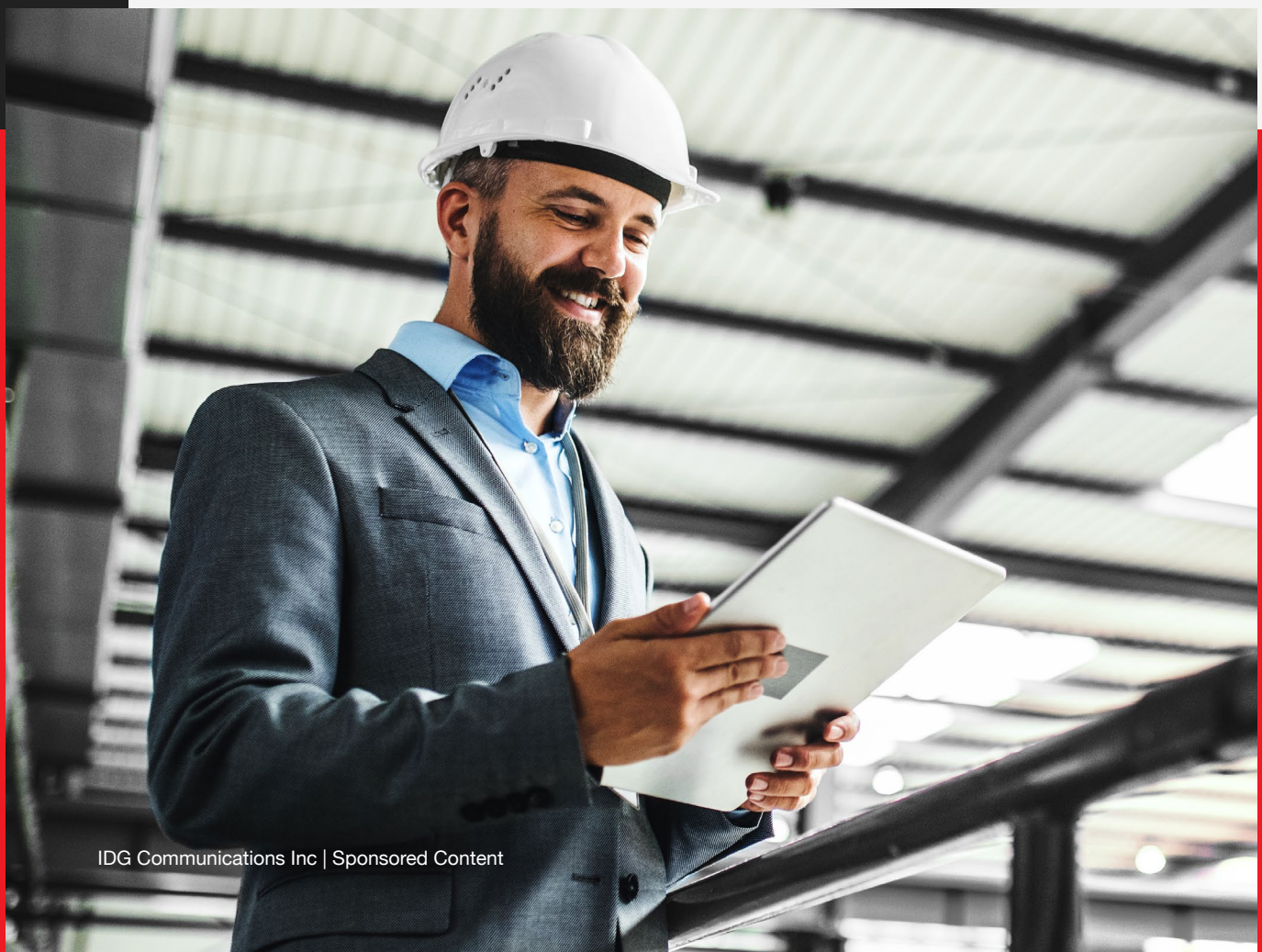
CIO FROM IDG

SPONSORED BY Adobe × Microsoft

Ethical AI is paramount: understanding your responsibilities

Organizations are defining best practices for ethical AI.

CIO 5M READ TIME



Behind all the technologies that emerge in an apparent breakthrough moment, there's typically a long trail of theory, experimentation, and discovery.

If we trace back along the lineage of artificial intelligence (AI), we encounter the German mathematician Carl Friedrich Gauss who devised the "least squares regression" technique in 1809, to work out where the asteroid Ceres would emerge after travelling behind the sun.

Part of what we might call the canon of mathematical concepts, the same mathematical technique still plays a significant role in AI computation routines today.

The rise of AI: opportunities and challenges

While the underpinnings of AI have been around for a long time, society is often forced to react to powerful new technologies rapidly. At best, lawmakers, academics, and scientists devise new frameworks of governance — as in the case of human-embryo research in the 1990s. At worst, the result could be societal, or even global dissonance.

> **The opportunity to learn from high-profile failures like these is invaluable.**

So far, society's response to the deployment of AI in commercial and governmental settings has been largely permissive, with a few important reservations in areas close to what former Google Chairman Eric Schmidt once described as the "creepy line" (for example, facial recognition).

However, many deployments exist on something of a knife edge. AI, for example, has the potential to radically enhance what humans are capable of. But biased data or poor system design can easily remove human input from aspects of decision-making where it really ought to be present.

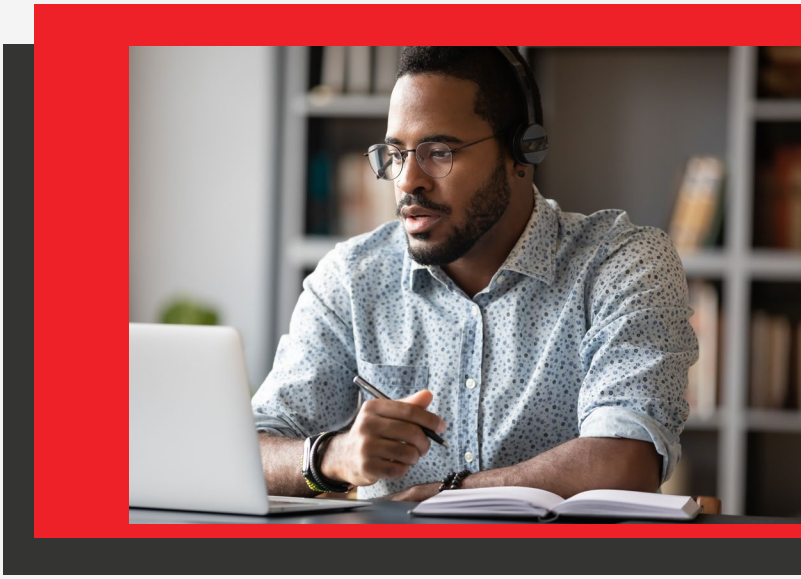
We have seen an increasing number of problematic AI deployments. These include CV assessment applications that favor male over female job applicants, AI systems in the US courts that overestimate the likelihood of black defendants reoffending, and banking systems that appear to discriminate against women by granting larger credit limits to men with poorer credit scores.

Why AI needs ethical risk assessment

The opportunity to learn from high-profile failures like these is invaluable. As Bill Gates once noted, success is a lousy teacher because "it seduces smart people into thinking that they can't lose."

But there are different categories of success and failure. In the world of technology, we're accustomed to being judged on the quality of technical execution, or on technology's ability to generate a return on investment. With AI, however, there's an additional category of judgment, one with which CIOs are not always intimately familiar: the ethical dimension.

With AI, ethics matter. Trust takes a long time to build and can be undone in a moment. Those who hope to benefit from AI deployment need to keep demonstrating their ability to think about, and act upon, the technology's potential weaknesses. As Luciano Floridi, Professor of Philosophy and Ethics of Information at Oxford University, points out: "Public acceptance and adoption of AI technologies will occur only if the benefits are seen as meaningful and risks as potential, yet preventable, minimizable, or at least something against which one can be protected, through risk management (e.g. insurance) or [other] redressing."



This no doubt explains why the UK's intelligence agency GCHQ – the equivalent of the National Security Agency in the US — published its first ever white paper for public consumption in early 2021, on the theme of the ethics of artificial intelligence.

Ethics allow us to strike the balance, to leverage new opportunities that are socially beneficial, while simultaneously anticipating and avoiding costly errors. Informed judgment of this kind also allows us to tread with more certainty. As Floridi puts it, ethics also "lowers the opportunity costs of choices not made or options not grabbed for fear of mistakes."

Emerging best practice for ethical AI

CIOs commissioning AI projects may find themselves and their teams operating beyond their comfort zone when dealing with the ethics of AI. That's understandable. Nevertheless, a body of best practice thinking is starting to emerge. Here are a few suggestions for mitigating the ethical risks.

- **Build a diverse workforce**

  Time and again, diversity proves its value. Among developers, architects, and data scientists working for your organization, diversity can act as an initial defense against ethical risk. Bias that may not be apparent to one member of the team may be obvious to another. Different individuals bring a stronger, healthier mix of perspectives to the table.

Ethics allow us to strike the balance, to leverage new opportunities that are socially beneficial, while simultaneously anticipating and avoiding costly errors.

- **Develop internal review boards**

  The notion of an internal review board comes from healthcare, where committees are designed to mitigate the ethical risks involved in research on human subjects. Here again, diversity is key: potential participants

include corporate lawyers, data scientists, data privacy officers, experts in markets, products, applications, and ethics. At Microsoft, for example, three committees oversee AI ethics risks: one sets company-wide governance policies, a second advises the company's leadership on the challenges and opportunities of AI, and a third is dedicated to raising awareness of ethical risk among the company's engineers. Adobe has a structure that looks similar: its Ethics Committee includes a diverse selection of expert voices from around the world, while the AI Ethics Review Board examines new features and guides development teams.

- **Create a framework for making judgements**

  Consider devising a policy that builds ethical risk mitigation into operations and products. What standards are product developers and product managers expected to live up to? At what point should ethical concerns be escalated to senior leadership or an in-house ethics committee? What processes can be used to identify biased algorithms, privacy violations, or incorrect AI outputs that need explanation? Ideally, your teams will have preset procedures for these scenarios, and more. Microsoft's policies in this area are laid out in detail here and in this blog post.

- **Consider explainability**

  When AI makes decisions in a regulated industry that affect people's lives in significant ways — for example, by approving or denying a mortgage application — it becomes important that these decisions are explainable. But, in some cases, it can be particularly challenging. Indeed, not all outputs need to be explainable (for example, in the context of complex pattern recognition undertaken by machine learning). Product managers need guidance and tools to address these challenges.

- **Raise awareness and incentivize good behavior**

  AI applications can be designed and deployed in good and bad ways. This broadens the scope of concern beyond developers and architects to product managers and even line of business managers. In organizations that don't invest resources into AI ethics or reward employees for thinking about ethics, it's clear that revenue and profit growth remains the uncontested priority. This clearly opens the potential for adverse outcomes.

Among developers, architects, and data scientists working for your organization, diversity can act as an initial defense against ethical risk.

The technology behind AI may feel sufficiently complex without adding on an additional layer of activity related to ethical risk. But the roll call of ethical failures in AI is already lengthy, and the accompanying risk of reputational damage is real. If your organization is developing AI, and if those applications have the potential to affect outcomes for customers, it's time to start thinking about the ethics of AI.

To find out more about why trust is becoming a key business challenge, and how CIOs can respond, click here to read the white paper.