

## DIRECTOR ADVISORY

### DIGITAL TRANSFORMATION

# Promote Trust in AI to Enhance Long-Term Value

By Irfan Saif

Humans are working with machines in more ways than ever to harness the power of data and artificial intelligence (AI). These technologies are a critical part of many companies' strategies, as organizations of all types and sizes are transforming their businesses with innovations aimed at enriching customer experiences, reimagining processes, and improving efficiencies, insights, and precision.

AI can help deliver exponential benefits to companies that can effectively leverage its power. However, as with other emerging technologies, there is a great deal for boards and senior leaders to understand about the nuanced risks and other potential implications of AI.

Keith Darcy, president of Darcy Partners and a career senior executive and corporate director, says it is important for boards to confirm that their organizations are proactively governing AI to protect their trust with their stakeholders. "Technology can inform but never substitute for judgment," he said. "In particular, as it pertains to AI, there is nothing artificial about ethics."

Directors should challenge their management teams to explain how they are mitigating risk as they integrate AI into products and operations. If implemented hastily or in an uncoordinated manner, AI could harm a company's stakeholders and its reputation, which could bring adverse brand, financial, operational, and regulatory consequences. Addressing AI risks and trust early not only helps safeguard against potential downside risks, but can result in several positive benefits, such as accelerated adoption and operational efficiency.

#### A FRAMEWORK APPROACH

In a time of heightened corporate responsibility and scrutiny, the board should exercise its fiduciary duty by asking C-suite leaders if and why they are comfortable that the organization's uses of AI are trustworthy and ethical. Leading practices for promoting trustworthy uses of AI that protect consumers and organizations are commonly built on a framework such as Deloitte's Trustworthy AI™ Framework. This framework is built on six critical dimensions: fairness and impartiality, transparency and explainability, privacy,

responsibility and accountability, reliability, and safety and security.

Each of these dimensions is important for organizations to consider as they implement AI that can affect people in a multitude of ways. Not every dimension will be relevant to every AI use case, but a framework that guides implementation through each of these dimensions promotes a comprehensive consideration of possible risks. An AI governance framework can help organizations implement a taxonomy, guardrails, and consistent, repeatable practices to promote a common culture and understanding of AI uses. This type of governance provides tactics for aligning people, processes, and technology toward risk-informed uses of AI with trustworthiness at the core.

A trustworthy approach to AI should consider a broad range of risk factors, some of which may be similar to those in the adoption of other technologies. Each organization must determine which risk dimensions are important, who will own them, and how they will intersect to mitigate a challenge or access greater value.

Governance driven by a framework approach is also valuable for promoting robust, reliable, and repeatable outcomes. Consider the importance of such outcomes in health care, for example. Drug development companies are increasingly using AI in compound synthesis, biological modeling, and clinical testing. Health-care companies might also use AI techniques to identify brain scan abnormalities and assist in prescribing treatment recommendations. These types of uses demonstrate the extent to which AI is playing a key role in potentially life-saving interventions, where the consequences for ineffective risk management could have dire consequences.

#### THE FRAMEWORK IN PRACTICE

To understand how such a holistic approach might work in practice, consider the thought process that might be involved in promoting one dimension: fairness and impartiality. Many financial institutions have developed models that are designed to identify anomalous transactions, such as credit card transactions, that might be fraudulent. In so doing, it's important that organizations establish checks throughout the AI life cycle to achieve

equitable outcomes across participants.

Data analyzed to assess fraudulent transactions may include prior purchase history as well as the size, nature, and location of a current purchase. Inputs based on factors such as a cardholder's gender, age, or level of education, however, could lead to an unintended potential for discriminatory bias. It's important for organizations to understand how these factors are contributing to outcomes, which promotes transparency.

In another example, consider how a framework can promote uses of AI that are consistent with another dimension of trustworthiness: transparency and explainability. To enhance and protect trust, participants should be able to understand what data are being used and how AI systems make decisions, potentially with algorithms, attributes, and correlations that are open to inspection. Retail companies increasingly use AI to make product recommendations to customers online. In so doing, those companies should be able to explain what data were collected and used to make these recommendations. Individuals should be able to access their own data, to opt out of its use, and to make inquiries and receive feedback.

Transparency and explainability can help build accuracy in decisions and trust in the model, both for consumers and an organization. If an organization can't clearly articulate how the data are being used and how decisions are being made, that should prompt further inquiry about whether the algorithm should be modified or the use continued.

The framework approach can continue guiding the thought process through further dimensions, such as privacy. Consumers may expect to be able to limit the use of their data beyond a stated purpose and to opt in or out of the sharing of their data with other parties. As such, organizations should consider the increased proliferation of data for AI purposes. To promote responsibility, organizations should consider policies that promote accountability over AI use cases, including the data, algorithm, and output of AI systems. With so many people and processes involved, there are many places at which one might point a finger in the event of a lapse or a loss. AI systems must also be protected from cyber risks. Security measures such as encryption, anonymization (where possible), and privacy policies are important for people who rely on virtual home assistants, for example, to help protect them from potential hackers.

## A PATH FORWARD

There are no global standards for appropriate implementations of AI. The legal and regulatory framework, such as privacy requirements that are relevant to uses of technology, is evolving. This does not mean, however, that organizations can take a wait-and-see approach and respond to new legal or regulatory requirements


as they emerge. To do so risks mishaps that can erode consumer trust, which can lead to financial implications and impede long-term value.

What's more, an effective approach to AI governance can act as a proactive means of extracting value from AI programs, with risks and trust considered and managed across the AI life cycle. Trustworthy uses of AI can increase brand equity and consumer loyalty, which can lead to customer growth and employee retention.

A trustworthy approach to AI can also increase the likelihood that customers opt in to share their data, creating a more valuable dataset. A data advantage enables improved decision-making that translates to tangible outcomes in product performance, customer experience, and operational efficiency, which can help increase revenue and reduce costs.

## QUESTIONS FOR BOARDS TO CONSIDER

1. What governance framework does the organization leverage to implement AI? Is there a strategy and program for the implementation of AI systems, both for back-office and front-of-house operations, as well as for the associated data and outcomes from those systems?
2. Does the organization have effective leadership and adequate talent to execute its AI strategy?
3. How is management monitoring the ongoing risks, including regulatory and contractual risks, associated with the use of AI in the business and the outcomes from those systems?
4. What metrics should the board expect to see periodically to understand how this risk is evolving, and who is responsible for providing such information?

Boards that can answer these questions may put their companies in a better position to enhance the long-term value of AI. 



Irfan Saif is a principal and chief strategy officer for Deloitte Risk & Financial Advisory, Deloitte & Touche LLP, and a member of Deloitte US's board of directors.

As used above, Deloitte refers to a US member firm of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee (DTTL). This article contains general information only and Deloitte is not, by means of this article, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This article is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte shall not be responsible for any loss sustained by any person who relies on this article. Copyright © 2022 Deloitte Development LLC.