Board Oversight of AI Risk Through an Ethical Lens

In a nascent and fragmented regulatory environment, focusing on the ethical implications of the company's AI activities is a prudent strategy for board oversight of AI risk.

Holly J. Gregory
Partner at Sidley Austin LLP

While advances in generative AI (GenAI) and machine learning present significant potential to solve problems to the benefit of society, these technologies also have the potential as well for significant negative impact. As companies race to develop, deploy, and use AI, boards of directors need to address their responsibility for oversight of the company's AI activities relating to strategic and operational efficiency opportunities and enterprise risk management and compliance. In an environment in which AI-specific regulation is not yet well developed and coherent, board attention to AI ethics and the guidance provided by voluntary ethical frameworks bears emphasis.

This article discusses board oversight of AI risk through an ethical lens and proposes areas of inquiry for boards to consider in discussions with management.

Current AI Concerns

AI's capacity to improve efficiency and enhance predictive capacity appears well accepted, but "[t]rust in AI systems remains a significant challenge," according to a recent global study by KPMG and the University of Melbourne. A majority of respondents expressed skepticism about AI safety, security, and societal impact. According to the study, "[p]eople report experiencing both benefits and negative outcomes from AI use. While many report improved efficiency, accessibility, decision-making and innovation, concerns about cybersecurity, privacy and IP, misinformation, loss of human connection, job loss and de-skilling are widespread." (KPMG International and University of Melbourne, **Trust, Attitudes and Use of Artificial Intelligence: A Global Study 2025**, at 5 (2025) (Global Study).)

Moreover, the data from employees surveyed suggests that company policies and controls for AI use are either not yet in place or are ineffective. While a majority (58%) of employees report that they use AI at work on a regular basis, most employees say they are using AI systems that are free and publicly available rather than employer-provided options. "Many employees report inappropriate, complacent

and non-transparent use of AI in their work, contravening policies and resulting in errors and dependency. Governance and training to support responsible AI use appears to be lagging adoption." (Global Study, at 5.)

The study also reported a strong public mandate for AI regulation, with 70% believing regulation is necessary. However, only 43% believe current laws are adequate. Respondents favor international laws (76%), national government regulation (69%), and co-regulation with industry (71%), with the vast majority (87%) expressing support for laws and fact-checking processes to combat AI-generated misinformation.

The Value of Ethics in a Nascent Regulatory Environment

Laws and regulations to address AI-related activities and related enforcement mechanisms are still developing. It remains unclear whether a more coherent legal and regulatory framework will develop in the US and internationally (and if so, when), adding uncertainty to company efforts to develop, deploy, and use AI and manage risks in a highly competitive business environment. (For more on key developments in AI regulation, see **US AI Law and Regulation: 2025 Tracker** and **EU AI Act** on Practical Law.)

In the absence of a clear set of legal and regulatory boundaries, applying a pointed "soft law" ethical focus to AI activities presents a prudent strategy for board oversight of AI-related risk management. Guidance is available from a variety of voluntary AI ethical frameworks that have been developed to support the responsible development and deployment of AI by promoting fairness, accountability, transparency, safety, human rights, and privacy, including:

- The **Organisation for Economic Co-operation and Development (OECD) AI Principles** (2019, updated 2024).
- The **United Nations Educational, Scientific and Cultural Organization (UNESCO) Recommendation on the Ethics of AI** (2022).

The five core OECD AI Principles focus on:

- Inclusive growth, sustainable development, and well-being.
- Respect for the rule of law, human rights, and democratic values, including fairness and privacy.
- Transparency and explainability, including the ability to understand when one is interacting with AI systems and how AI outcomes are reached.
- Robustness, security, and safety.
- Accountability for the functioning of AI systems and respect for these principles.

Additionally, the OECD AI Principles include guidance for policymakers to support trustworthy AI development.

The UNESCO Recommendation on the Ethics of AI focuses on the protection of human rights and autonomy. It rests on ethical values that include:

- Respect, protection, and promotion of human rights, fundamental freedoms, and human dignity.
- Environmental and ecosystem flourishing.
- Ensuring diversity and inclusiveness.

- Living in peaceful, just, and interconnected societies.

The UNESCO Recommendation promotes the following principles, inspired by these values:

- Proportionality and a “do no harm” approach.
- Safety and security.
- Fairness and non-discrimination.
- Sustainability.
- A right to privacy and data protection.
- Human oversight and determination.
- Transparency and explainability.
- Accountability and responsibility.
- Awareness and literacy.
- Multi-stakeholder and adaptive governance and collaboration.

Other voluntary frameworks include the European Commission’s **Ethics Guidelines for Trustworthy AI** and the **AI Risk Management Framework** issued by National Institute of Standards and Technology (NIST), an agency of the US Department of Commerce. The AI Risk Management Framework was developed with input from the private and public sectors as a voluntary and flexible risk framework. It calls on companies to establish policies that define AI risk management roles and responsibilities, including for the board of directors. Additionally, it identifies characteristics of trustworthy AI systems, such as being “valid and reliable, safe, secure and resilient, accountable and transparent, explainable and interpretable, privacy-enhanced, and fair with harmful bias managed.” (For more information, see **NIST Releases Artificial Intelligence Risk Management Framework** on Practical Law.)

Industry groups have also issued ethical guidance, including:

- The Institute of Electrical and Electronics Engineers’ **General Principles of Ethically Aligned Design**.
- **Ethical Principles and Practices for Inclusive AI** issued by the Partnership on AI, a non-profit partnership of academic, civil society, industry, and media organizations.

Renewed Focus on AI Ethics

The Second Annual Rome Conference on AI, Ethics, and Corporate Governance, held in June 2025 and hosted at the Italian Ministry of Enterprises and the Apostolic Palace in the Vatican, addressed the evolving challenges at the intersection of AI, ethics, and corporate governance. The conference agenda focused on how AI governance frameworks must adapt to the rapid advancement of AI technologies, with a particular emphasis on ethical responsibility, stakeholder impact, the broader societal implications of AI, and the importance of transparency, accountability, and international cooperation in AI regulation. Participants included senior executives from major global technology companies, policymakers, scholars, ethicists, and attorneys, as well as religious representatives.

A **letter to the attendees from Pope Leo XIV**, read at the opening of the second day in the Vatican, identified AI as a tool that should embody the values and intelligence of its human creators. Pope Leo XIV emphasized the need to evaluate AI according to the “integral development of the human person

and society," highlighting the importance of safeguarding human dignity, promoting responsible use, and ensuring that AI serves the well-being of all, especially the young and vulnerable. He called for reflection on the spiritual, cultural, and moral implications of AI, urging that its benefits and risks be weighed against the superior ethical criterion of humanism. In his view, this approach is critical in creating ethical frameworks that effectively manage the long-term issues associated with the speed and scale of this transformative technology.

In a **keynote address**, former Delaware Chief Justice Leo Strine urged business leaders, including independent directors, to confront AI's ethical mirror test. He explained that the starting point for ethical AI is an understanding of the impact of the company's AI-related activities. "Corporate leaders must bear responsibility for the impact of their companies' conduct. Before corporate leaders can act ethically toward others, they must first understand their company's conduct — how it behaves, how it works, and who and what it affects." Strine emphasized that this extends to understanding the potential for misuse of AI that the company has developed for use by third parties (for example, by supporting fraudulent activity or theft of intellectual property (IP)), as well as how AI that the company uses might support fraudulent activity, embed historical discrimination, replace professional judgment, or otherwise lead to erroneous, biased, or detrimental results. "To pass the ethical mirror test, one must accept the human responsibility to use AI only in ways that one understands and that are reasonable and safe for those it stands to affect," Strine noted.

Board Oversight of AI Risk

Board responsibility for managing and directing corporate affairs requires oversight of the authority delegated to management, including with respect to the development, deployment, and use of AI, as well as a fiduciary mindset and attention to internal controls and policies. Directors must:

- Understand AI as a matter of corporate strategy and risk.
- Consider the impact of AI activities on employees, customers, other key stakeholders, and the environment.
- Oversee the company's compliance with laws and regulations that are relevant to AI and the development of related policies, information systems, and internal controls, including corporate ethics. (For more information, see **AI and the Role of the Board of Directors** in *Practical Law The Journal*.)

In a high-pressure race for first-mover advantage with respect to strategic opportunities, the board must ensure that risks to the company, and more broadly to stakeholders and the environment, are given due consideration. (For more on AI risks, see **What's Market: AI Risk Factors** and **AI-Related Disclosure: SEC Comment Letters (2024)** on Practical Law.)

Considering the Impact of AI

AI raises risk and compliance issues that require board consideration. AI systems can incorporate bias and lack transparency, which may lead to concerns about equity and accountability. Additionally, AI systems are heavily reliant on data, often implicate privacy and data protection regulation, and may have implications for IP protections. (For more information, see **Key Legal Issues in Using Generative AI: Overview (US)** on Practical Law.)

AI and GenAI raise considerable concerns about the potential for misuse and unintended consequences. Boards and management teams need to consider the responsible corporate use of AI and how to avoid and mitigate unintended consequences, including through the use of policies and internal controls overseen by the board or an appropriate board committee. The National Association of Corporate Directors (NACD) has issued a **report** outlining a four-pillar framework focused on strategic oversight, capital allocation, AI risk, and technology compliance, which boards can use in organizing their AI oversight.

Renewed Focus on AI Risks

In light of constantly changing AI activities and uncertainties relating to AI use, as well as increasing questions of whether, how, and when AI will provide a return on investment, the NACD has proposed a four-pillar framework for AI governance that includes actions for board oversight and attention in keeping with a board's core fiduciary duties. These four pillars are:

- **AI strategic oversight.** The board should:
  - develop a shared understanding about AI's strategic relevance and importance;
  - set a schedule for AI discussions and updates at each board meeting; and
  - incorporate AI as a topic in the board's annual strategy session.
- **Capital allocation relating to AI.** The board should:
  - include AI expenses in the annual budget approval process; and
  - regularly review the viability and opportunities for M&A and partnerships to acquire AI capabilities, compared to organic in-house development.
- **AI risk oversight.** The board should:
  - integrate AI risks into its enterprise risk management system; and
  - ensure it receives regular briefings from internal and external AI risk experts.
- **AI technology competency.** The board should:
  - develop and maintain the necessary board-level AI and technology proficiency aligned to corporate strategy and governance needs, including through adding directors with technology skills and providing director education;
  - establish clearly designated authority and responsibility for AI within the company;
  - ensure management and workforce readiness for AI transformation; and
  - incorporate AI oversight roles and responsibilities into board committee charters (the framework identifies specific board committee oversight responsibilities).

In addition to the oversight framework, in its report, the NACD identifies "red flags" that directors should look out for and includes practices the board can implement, as well as case studies, examples, and deployment scenarios boards are likely to face.

AI can help users complete routine tasks more efficiently and is improving in its ability to aggregate data, recognize patterns, and create content, with the prospect of freeing up employees to increase their focus on tasks that require more judgment and creativity. AI's growing potential to automate skilled tasks has implications for workforce training, management, and productivity. Boards should consider how the company's use of AI is impacting employees and the talent pipeline, whether employees are being trained to use AI in a manner that leverages their skills and mitigates AI risks, and what types of policies should be implemented to encourage appropriate use of AI by employees for approved use cases — particularly in highly regulated or otherwise high-risk contexts, such as health care, financial

services, or hiring and promotion decisions.

AI's potential for perpetuating bias in the data sets on which it relies raises concerns about the use of AI in employment and promotion decisions. The board should understand whether the company uses AI for these purposes and if so, how AI is used and what policies are in place regarding these uses. (For more information, see **AI in the Workplace (US)** and **AI Bias (US)** on Practical Law.)

Boards should also understand how AI is used by customers and suppliers, as well as its impact on the environment. The use of AI is already well embedded in various industries, such as the automotive, e-commerce, entertainment, financial services, health care, hospitality, insurance, logistics, manufacturing, marketing, retail, and transportation industries. Boards may not understand the ways in which their companies or others in their industries are using AI to interface with customers and suppliers, and the extent to which these uses involve data gathering, with privacy implications and related concerns about bias. Additionally, boards may not appreciate the environmental impact of, for example, training an algorithm to identify reliable patterns, which can require heavy energy use to analyze millions of data sets.

Boards need to understand and stay apprised of regulatory developments and oversee the company's compliance, as well as the development of relevant policies, information systems, and internal controls, to ensure that AI use is consistent with legal, regulatory, and ethical obligations and has appropriate safeguards to protect against potential risks. In doing so, they should be mindful of the variety of ways in which the company may face exposure to AI-related compliance and other risks, including through AI technology that is internally developed, licensed from others, or acquired through M&A activity.

(For resources to assist counsel in identifying potential legal issues concerning AI, see **AI Toolkit (US)** on Practical Law.)

Practice Pointers

The board is responsible for overseeing policies and processes intended to encourage ethical conduct throughout the company. Board oversight is key to ensuring that AI is used in a responsible and ethical manner in alignment with the company's strategic objectives and values. Emphasizing attention to the ethical implications of the company's AI activities is a useful approach for board oversight of enterprise risk management in a nascent, fragmented, and uncertain regulatory environment. (For more on implementing an AI governance framework, see **AI Governance Roadmap (US)** and **AI Governance Checklist** on Practical Law.)

Questions for boards to consider in discussions with management include:

- Do the board and management understand the company's strategy and activities in the development, deployment, and use of AI, including the implications for enterprise risk management, and are these activities consistent with the company's ethical values?
- What AI ethics frameworks has the company committed to, and does the company have a well-defined set of policies, processes, and controls to ensure compliance with its ethical AI principles, as well as with law and regulation? (For a model statement of principles for a company's design, development, and deployment of AI tools, with explanatory notes and drafting tips, see **Principles of Responsible AI Development** on Practical Law.)
- Are there any mission-critical compliance or safety risks related to the company's AI activities

that management has identified? If so, what policies, processes, and controls has management put in place to manage and mitigate these risks? (Mission-critical compliance or safety risks should be mapped to a board committee for more frequent and in-depth attention, and reflected in the committee charter, agenda, and minutes.)

- Has management identified the potential for misuse and unintended consequences from the company's use of AI with respect to employees, customers, and other key stakeholders, and determined how to avoid or mitigate those risks?
- What protections are in place to minimize and mitigate the potential negative impacts of the company's development, deployment, and use of AI, including with respect to the potential for:
    - inaccuracy, including contribution to or amplification of misinformation;
    - creation or exacerbation of bias;
    - violation of IP rights; and
    - breach of data privacy and cybersecurity?
- Does the company have the expertise and resources to develop, deploy, and use AI in a responsible and ethical way, as well as appropriate policies, processes, and controls?
- Who is the senior executive responsible for ensuring that company policies and procedures related to the use of GenAI appropriately address the potential for bias, inaccuracy, breach of privacy, and related issues of consumer protection, cyber and data security, IP protection, and quality control?
- What is the potential impact of the company's AI activities on its workforce? How does the company use AI in tracking and assessing employee performance, and what processes and controls are in place to protect against bias and to foster compliance with relevant regulations?
- What policies and procedures are in place to ensure that the company's AI activities are appropriately transparent and that its disclosures regarding AI activities and risk in public filings are accurate and compliant?
- What policies and procedures are in place to ensure that any AI associated with the company's supply chain is aligned with its commitment to the ethical use of AI?
- What is the potential environmental impact (including energy and water demand) of the company's AI activities, and what efforts is the company taking to mitigate those impacts?

In addition to maintaining an ethical focus on AI activities, boards should ensure that management is keeping up to date on the evolving regulatory environment. (For more on key developments in AI regulation, see **Developments in US AI Law and Regulation: 2025 Tracker** and **EU AI Act** on Practical Law.) The competing interests of promoting an environment in which AI development thrives and the potential harms and risks are controlled may continue to pose challenges for policymakers in developing coherent regulatory approaches.

The views stated above are solely attributable to Ms. Gregory and do not necessarily reflect the views of Sidley Austin LLP or its clients.

https://www.reuters.com/practical-law-the-journal/transactional/board-oversight-ai-risk-through-an-ethical-lens-2025-11-01