

INVESTOR RECOMMENDATIONS FOR AN EFFECTIVE RIGHTS-BASED APPROACH TO AI

With the rapid and widespread uptake of artificial intelligence (AI), it is increasingly necessary to ensure that AI-driven products and services spur technological innovation in a manner that respects the rights of consumers and the public. To address the novel opportunities and risks presented by these technologies, both companies and their shareholders can benefit from a responsible AI program or approach comprised by robust governance structures, policies, and practices. Proactive companies will be better positioned to deliver high-quality products and mitigate financial and reputational risk, thereby fostering a pro-competitive environment that contributes to long-term shareholder value.

It is with this goal in mind that independent investors have participated in the creation of the recommendations set forth below. Each is an independent fiduciary responsible for the investment and voting decisions of its customers. There is no collective decision-making with respect to acquiring, holding, disposing, and/or voting of securities, and all communications relating to the development of these recommendations comply with applicable laws.

Key elements of a responsible AI approach:

- **GOVERNANCE:** Establish a framework to direct, oversee, and monitor activities related to the development and use of AI in products and services.
- **POLICIES AND PRACTICES:** Develop rights-based policies and implement a risk management approach, including but not limited to adequate human rights due diligence and remediation in the event of harm to address salient risks and impacts of AI systems.
- **TRANSPARENCY AND ACCOUNTABILITY:** Provide timely, consistent, and transparent disclosure on the application and efficacy of relevant policies and practices.

These elements are further detailed below and are informed by globally accepted norms regarding the protection of human and labor rights – such as the United Nations Guiding Principles on Business and Human Rights (UNGPs), Organization for Economic Cooperation and Development (OECD) Guidelines for Multinational Enterprises on Responsible Business Conduct, and International Labor Organization (ILO) Conventions – and by current best practice standards and regulation regarding the development and deployment of AI such as the U.S.’s National Institute of Standards and Technology (NIST) AI Risk Management Framework and the European Union (EU) AI Act. Together, these frameworks address AI from both an enterprise risk and rights-preserving perspective.

GOVERNANCE: Establish a framework to direct, oversee, and monitor activities related to the development and use of AI in products and services.

As long-term investors, we believe robust governance structures are critical for strategy setting, effective implementation, resource allocation, accountability, and organizational buy-in. Companies should provide transparency in this area and demonstrate how these governance structures are integrated across product life cycles. Components of an effective governance structure include:

- Clear allocation and delineation of decision-making, oversight, advisory, and reporting responsibilities at the board, executive, and business unit levels.
- Board-level oversight that is embedded in the charter of the relevant committee, ensures board-level expertise and training, includes performance metrics related to a company’s impact on society, and draws on the perspectives of independent experts.
- Adequate resourcing and appropriate staffing of committees and/or functions responsible for developing, implementing, and governing the company’s responsible AI approach.
- Appropriate training and auditing for employees, teams, supply chain and business partners involved in the development and/or deployment of AI-driven products and services.
- Alignment of management, employee, and team incentives with responsible AI objectives and performance.

Companies may retain advisory councils to provide domain expertise and external perspectives on emerging topics and concerns. However, these are not sufficient substitutes for clear, standing, internal governance structures as these bodies generally do not wield decision-making authority and are not integrated into core business operations.

POLICIES AND PRACTICES: Develop rights-based policies and implement a risk management approach, including but not limited to adequate human rights due diligence and remediation in the event of harm to address salient risks and impacts of AI systems.

To provide the necessary transparency on how rights-based risks are being proactively addressed, companies should develop and make public a set of principles and/or commitments—grounded in a respect for human rights, ensuring safe, accurate and reliable technologies and including robust due diligence, post-deployment monitoring, and

comprehensive remedy—that guide their strategy and decision-making with regards to AI development, commercialization, and use. To be effective and ensure integration into long-term business strategy, it is important for such policies to align with the company’s mission and stated values, define acceptable (and unacceptable) use, outline specific processes and how they enhance or intersect with other aspects of operations (e.g. data privacy and governance, human rights, supply chain, customer due diligence, etc.), and allocate responsibility and oversight.

Translating Principles and Policy into Practice

To be effective, it is important for responsible AI guidelines and policies to be integrated into product lifecycles, including risk identification and assessment, product design, development, testing, validation, post-deployment monitoring, and sales and marketing practices, and they should ensure the development of high quality, safe, and accountable AI systems.

Due to heightened risks associated with the use of AI in sensitive domains, including military uses, the application of human rights due diligence (HRDD) processes is critical to ensure that fundamental rights are protected and upheld. Human rights impact assessments (HRIAs) are an important aspect of HRDD, especially for identifying and addressing products or features that pose elevated risks to protected rights. Notably, HRIAs differ from traditional enterprise risk management assessments as they consider both risks to the business and risks from the business on key stakeholders, including individuals and communities affected by the end use or application of products and services. Such assessments should consider the realized and potential impacts on human, civil, and labor rights; they should be conducted regularly, and the results should inform the company’s risk management policies and practices; and ideally, they should be conducted by independent experts and supplemented and updated with ongoing internal assessments.

Effective processes and procedures include the following elements:

- Requirements for testing, documentation, performance, and review throughout product lifecycles to identify and address high-risk products and features, and where appropriate, subject them to additional safeguards and performance criteria.
- Meaningful channels for stakeholder consultation, including representatives from civil society and affected communities, accompanied by robust remediation and grievance policies and whistleblower protections.
- Strong data governance to ensure the lawful collection, use, and protection of high quality, relevant data sets that are representative of the intended use population, including security and privacy controls for personal/sensitive data.
- Customer and end-use due diligence to monitor and mitigate risks of product abuse or misuse.
- Independent audits to assess efficacy of policies and processes.
- Guidelines for lobbying and political advocacy activities to uphold protections for human, labor, and civil rights.

TRANSPARENCY AND ACCOUNTABILITY: Provide timely, consistent, and transparent disclosure on the application and efficacy of relevant policies and practices.

Companies should report publicly on their efforts to demonstrate implementation and effectiveness of responsible AI practices and build trust and accountability with stakeholders. Innovation, regulation, and best practices of emerging technologies evolve rapidly. Thus, companies should demonstrate how their risk management processes prepare their businesses to mitigate financial, regulatory, and reputational risks.

It is important for companies to comprehensively describe their AI governance structure and supporting practices, as described in the preceding sections. This reporting should be easily accessible and regularly updated – for example, integrated into annual sustainability reporting or published in standalone AI transparency reports. Similar to other salient ESG and sustainability issues, AI governance and risk management reporting should set relevant qualitative and quantitative metrics to assess its societal impact and adherence to international human rights norms, disclose its progress towards fulfilling them, and evaluate the efficacy of such measures.

Human Rights Considerations

The UNGPs, adopted by the UN Human Rights Council in 2011, lay out the responsibility of private sector actors to respect human rights through their operations and business relationships. The scope of business responsibility extends to all internationally recognized human rights standards, including the International Bill of Human Rights and the International Labor Organization (ILO) core conventions, regardless of whether such standards are codified into law in all contexts where a company operates.

In carrying out their business functions, companies are responsible for fulfilling their human rights obligations by conducting human rights due diligence (HRDD) to identify and manage impacts that they may cause or contribute to. Human rights impact assessments (HRIAs) are a critical component of HRDD as they help companies identify and prioritize their most salient impacts, and in turn develop effective policies, processes, and procedures to address them. HRIAs are often conducted in conjunction with an independent third-party, such as a consultant or advisory firm with expertise in human rights, and are generally applied to specific products, processes, or contexts.

Importantly, HRDD and HRIA focus on salient human rights issues. Saliency concerns risks to people, not the company— however, frequently the areas that pose the most significant risk to human rights are also the ones that pose the most significant risk to the business. For this reason, companies that only utilize traditional enterprise risk management processes and general materiality assessments often fail to identify and adequately prioritize human rights in their business and sustainability strategies. Robust HRDD and HRIAs should be incorporated into companies' risk management frameworks and processes for all AI systems that pose meaningful risks to human rights.

Workplace Considerations

Workers are integral to the success of companies' AI strategies, and concerns around job displacement, wage impacts, and workplace surveillance and monitoring are important issues for management teams to address. Developing and deploying AI responsibly and effectively necessitates meaningful worker engagement such that new technologies may benefit both companies and workforces. Where AI systems will directly impact workers, companies should seek to protect workers' dignity, safety, privacy, and autonomy, invest in necessary upskilling or reskilling resources, and proactively seek feedback and input. Doing so will position companies to attract, hire, and retain an engaged workforce and uphold their commitments to inclusion, equity, and diversity, contributing to sustainable business success.

Standards, Norms, and Resources

- [NIST AI Risk Management Framework](#)
- [EU Artificial Intelligence Act](#)
 - and [General-Purpose AI Code of Practice](#)
- [OECD Guidelines for Multinational Enterprises](#)
 - and [AI Principles](#)
- [UN Guiding Principles on Business and Human Rights](#)
 - and [B-Tech Project](#)
- [ISO/EIC 42001 - AI Management System \(AIMs\)](#)
 - and [ISO/IEC 23894 - AI risk management](#)
 - and [ISO/IEC TR 24027 - Bias in AI systems](#)
- [Hiroshima Process International Guiding AI Principles](#)
- [International Labor Organization \(ILO\) Conventions](#)
- [MIT AI Risk Repository](#)

For additional information, refer to the National Institute of Standards and Technology (NIST) Risk Management Framework, the EU Artificial Intelligence Act, the OECD Guidelines for Multinational Enterprises and AI Principles, the UN Guiding Principles on Business and Human Rights and the B-Tech Project, the ISO/EIC 42001 – AI management System (AIMs), the Hiroshima Process International Guiding AI Principles, the International Labor Organization (ILO) Conventions, and the MIT AI Risk Repository.