

# Responsible AI Quick Guide for Asset Owners

## ESG and Artificial Intelligence

As Artificial Intelligence (AI) development has accelerated over the past several years, so too have its potential use cases and public awareness. Recent breakthroughs in generative AI systems such as ChatGPT have only amplified these trends, as these emerging systems can be leveraged by both businesses and individual users. Importantly, the use of AI extends throughout a broad range of industries outside of tech—these include sectors from manufacturing and consumer goods to healthcare and financial services. As such, both companies and the public are exposed to a plethora of risks related to inadequate AI governance—including bias perpetuation, data misuse, job loss, and environmental degradation—which asset owners and General Partners (GPs) alike cannot only consider but also actively protect against in financial decision-making and ownership, respectively.

Notably, despite market downturns, private equity and venture capital investments in AI have continued to grow, with private equity investments in AI and machine learning totaling [\\$5.81B](#) in the first quarter of 2023—up nearly 50% from the previous quarter. The rapid acceleration of AI use and development has brought technology into uncharted territory with regulations, good governance, and legal precedent struggling to keep



IN COLLABORATION WITH:



## KEY TAKEAWAYS

- The development and use of Artificial Intelligence (AI) has outpaced regulatory frameworks and legal precedent, presenting investors with risks related to weak corporate governance and the potential for future noncompliance and broader societal harm if risks are not adequately managed
- As AI becomes more ubiquitous across industries and, to a lesser extent, increasingly embedded in investors' strategies, asset owners must remain abreast of the implications of such risks to their funds, as well as best practice controls among GPs
- Asset owners can establish differing degrees of oversight and ask a range of questions depending on the extent of GPs' investment into or internal use of AI technology

up. Importantly, while AI itself is not a recent invention, its increased exposure in the public sphere, including everyday use by individuals, invites greater risk and public scrutiny.

In attempts to govern the growth of AI, regulatory bodies have begun to outline [legal frameworks](#) to manage risk and establish obligations for both providers and users that will have profound impacts on AI companies. Within private markets, it is critical for investors to recognize their role in the development of this space as poorly governed businesses utilizing AI can have negative impacts on not only the industry, but also the commercial viability and trajectory of generative AI products more broadly. To support the development of responsible AI, asset owners can ensure that they are allocating capital to funds that are governed by GPs promoting thoughtful responsible AI programs across their investments.

## Key ESG Risks

It is critical to recognize AI risks to understand the importance of responsible AI—not only as it relates to dedicated tech/AI strategies, but also where AI applications may be less obvious, such as for investments in the industrial or manufacturing sectors where AI is leveraged for automation. These ESG risks may be direct (i.e., caused by the AI itself) or indirect (i.e., a byproduct of AI adoption).

### Bias Potential

One of the most prominent AI ethics risks is the potential for programs to produce biased, inequitable, or otherwise harmful outputs. These outputs can generate inaccurate or misleading data, or cause resources, information, or opportunities to be allocated unfairly and may infringe on civil liberties by failing to provide the same quality of service to all individuals. When applied without proper ethical controls from the outset (e.g., controlling for bias in machine learning training data), AI can perpetuate serious societal harm.

### Privacy Concerns

The development of AI both necessitates and accelerates the use of big data, as large datasets enhance and increase the speed of analysis. Importantly, the most privacy-sensitive data analysis (i.e., analysis that relies on the collection of personal information) that exists today—recommendation algorithms, facial recognition, advertising technology—are driven by AI and machine learning and trained on big data. The ability to use personal information with such power, speed, and accuracy and in ways that can breach individual privacy is magnified by AI applications.

## Job Loss

While AI has the potential to positively transform business if deployed responsibly, it can also create new societal challenges through its capacity to automate tasks that are currently done by humans—AI models have the capacity to perform tasks more quickly and accurately than humans at a fraction of the cost. As such, just as manual labor-intensive industries have had to adapt to physical automation, businesses prone to intelligent automation will have to grapple with significant workforce transformations and dislocation in the coming years.

## Lack of Diversity

The current state of AI talent is disappointing from a diversity standpoint: on average, only one [quarter](#) of AI teams are women or racial/ethnic minorities. When considering the aforementioned context—the potential for biased AI and the capacity for AI adoption to dislocate millions of workers across all industries—the lack of diverse representation is particularly concerning. When companies lack diverse perspectives, bias issues proliferate, including in dataset creation, sourcing and labeling, algorithm training and testing, and post-processing performance reviews.

## Carbon Emissions

While AI applications can be a great tool in combatting climate change (e.g., monitoring weather systems, preempting rainforest destruction), the training of AI models itself has a [significant](#) carbon footprint that can perpetuate climate change if not properly managed; the carbon footprint of training a single big language model is estimated to equal around [620,000](#) pounds of carbon dioxide emissions, or five times the lifetime emissions of an average American car (including manufacturing the car itself). Further, the data processing and advanced infrastructure needed to sustain AI development is incredibly [energy intensive](#) and typically powered by the public grid and supported by diesel-powered generators—if companies are to deploy AI to help combat climate change, they must first ensure their models do not contribute to environmental degradation.

# Responsible AI Frameworks

It is valuable for asset owners to understand established frameworks outlining responsible AI so that these practices can be incorporated into investment decisions and partnerships with GPs. As a best practice to manage risks presented by AI, GPs can align themselves to recognized responsible or ethical AI frameworks, both as it relates to its development teams and governance structures. The Responsible AI Institute (RAII), for example, has a [certification program](#) aligned with existing ethical

AI laws, regulations, principles, and research. The certification benefits different stakeholders within the AI ecosystem, including senior executives, compliance and procurement officers, investors, and consumers, within a variety of industries, such as financial services, health care, procurement, and human resources. Additionally, Microsoft maintains [six responsible AI principles](#) for AI development and use that focus on operationalizing AI ethics from a software development and programming standpoint, which may be most applicable when leveraged by GPs investing in AI development or training.

For a more comprehensive overview of high-level AI frameworks, outlined below are additional, common frameworks that can be utilized by GPs when developing responsible AI programs.

FRAMEWORK	OVERVIEW
<a href="#">NIST AI Risk Management Framework</a>	Outlines ways to incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems in collaboration with both the private and public sectors
<a href="#">UNESCO Ethics of Artificial Intelligence</a>	Establishes global standards on ethical AI based on the promotion and protection of human rights, human dignity, and environmental sustainability through principles such as transparency, accountability, and the rule of law
<a href="#">OECD AI Principles</a>	Sets practical and flexible AI standards and recommendations from a legal perspective, focused on how to shape a human-centric approach to trustworthy AI
<a href="#">Institute for Ethical AI and Machine Learning</a>	Outlines a framework of eight principles put together by subject matter experts to guide technologists to develop responsible AI/ML systems
<a href="#">EU Ethics Guidelines for Trustworthy AI</a>	Sets out both guidelines and requirements that AI systems should meet to be considered trustworthy, including topics such as human agency, transparency, non-discrimination, and data privacy

# Evaluating GPs' Responsible AI Practices

It is critical to recognize AI risks and what responsible AI looks like in practice—not only as it relates to dedicated tech/AI strategies, but equally across almost all industries as AI adoption continues to grow.

At the fund level, the degree of responsible AI integration will differ depending on a GP's given investment strategy. For example, funds with established AI investment theses may wish to broadly align with a chosen ethical AI framework, while GPs investing in the riskiest industries, such as health tech, or those leveraging AI algorithms in their internal investment decisions, may integrate sophisticated controls both pre- and post-close. Meanwhile, GPs with portfolios in industries that traditionally touch AI less should be aware of the kinds of operational practices that bear AI-related risks (e.g. industrial automation) and be ready to address them on a case-by-case basis. Despite this variation in approach, however, it is valuable for GPs across the board to have a working understanding of what AI means to their portfolio companies to inform not only risk mitigation, but also value creation strategies going forward.

To align with leading practices or manage a higher-risk portfolio, GPs can conduct ESG due diligence that covers responsible AI to gain a thorough understanding of current AI capabilities and any past incidents. Internally, GPs can maintain dedicated expertise on AI ethics and facilitate responsible AI conversations and trainings with management teams. Where feasible, GPs may wish to collect relevant KPIs and monitor portfolio companies' responsible AI performance to ensure adherence to best practices and programmatic improvement over time.

Further, with regard to the promotion of responsible AI best practices across each investment, GPs should make it a practice to ensure portfolio companies maintain some level of developed expertise and oversight of AI risks. For example, GPs investing in companies heavily involved in AI development can encourage management teams to establish clear leadership for responsible AI (i.e., a responsible AI committee) as well as formal processes to identify and mitigate AI systems biases, including engaging diverse teams of developers who are more likely to identify bias than homogeneous groups. Similarly, GPs that utilize AI to make investment decisions can establish these layers of accountability at the firm level. Additionally, GPs of all investment strategies can leverage research and education on responsible AI and maintain a consistent practice of pursuing continuous improvement and stakeholder engagement.

GPs, in partnership with management teams, may choose to draw best practices from the aforementioned AI frameworks, as well as more risk-specific guidance depending on the composition of their portfolios. Risk-specific frameworks, such as the [Mitigating Bias in Artificial Intelligence](#) playbook published by the Center for Equity, Gender, and Leadership at the Haas School of Business or modules offered by [The Alan Turing Institute](#), may provide additional guidance for GPs depending upon their investment thesis and relative risk profile. Further, GPs investing in AI in more non-traditional sectors should ensure that management teams have an understanding of and [clear strategy](#) around advanced analytics, and that employees receive guidance on both how to employ AI models and detect when a model is not working as intended. By understanding expectations for both baseline and risk-specific AI programming for GPs and their respective investments, asset owners can make informed allocations as AI is becoming increasingly ubiquitous across industries.

## Asset Owner Next Steps

To ensure responsible AI is operationalized by funds at the portfolio company level, asset owners can ask GPs any of the following questions depending on the extent or type of their AI investment strategies:

- What is the fund's overall investment thesis as it relates to AI?
- Does the GP maintain internal expertise around AI and best practices?
- To what extent does the fund intend to integrate AI into the sourcing and/or value creation planning of its investments?
- What degree of attention to responsible AI do the industries the GP invests in require?
- What investment diligence process does the fund use as it relates to AI governance?
- How does the GP intend to engage management teams on responsible AI?
- Does the GP maintain policies specifically related to responsible AI, or a responsible investment policy that addresses AI?
- To what extent does the GP utilize AI for making investment decisions, if at all?
- What are the GP's controls around AI use at the fund level?

Identifying best practices and engaging GPs in a high-quality dialogue around the increased relevance of AI, including its approach to risk management, is crucial in advancing a responsible, sustainable approach to new technology.