Malk
PARTNERS

ESG and Artificial Intelligence

Malk Insights —–July 2023

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.
- When investing in AI, GPs should consider bias mitigation and responsible dataset development, data privacy protection, development team diversity, workforce upskilling, and climate and environmental impact reduction.
- AI has the potential to drive positive impact when applied responsibly; investors should conduct in-depth due diligence to ensure investments are positioned for longevity and scale.

ESG and Artificial Intelligence

As Artificial Intelligence (AI) development accelerates, so too have its potential use cases and public awareness over the past several years. 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. 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 investors should not only consider but also actively protect against in both financial decision-making and ownership.

Why ESG Should Matter to AI Investors

Regardless of industry, the use of AI is more widespread than ever; as of 2022, AI adoption had more than doubled over the past five years, with investments in AI increasing at a similarly rapid rate. Notably, generative AI models—ChatGPT, GPT-3, and DALL-E, for example—can be used to create entirely new content and have soared in popularity over the past year due to their ability to reduce investment into content creation and thereby boost productivity. ChatGPT, the chatbot developed by OpenAI, reached 100 million monthly active users in January 2023 (two months after its launch) as the fastest-growing web platform in history.

The rapid acceleration of AI use and development has brought technology into uncharted territory with regulations, good governance, and legal precedent struggling to keep 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 2021, for example, Zillow shut down its AI-enabled Zillow Offers platform just two years after its launch—the platform was meant to predict housing prices so that the Zillow could buy real estate, flip it, and resell it quickly. In reality, the algorithm was not sufficiently trained to understand the market the way a human broker could—consequently, Zillow suffered losses of more than half a billion dollars on the value of its owned properties. Further, Zillow faced reputational scrutiny through widespread media coverage, and ended up cutting a quarter of its staff. Notably, in response to the mounting number of AI-related incidents such as in the case of Zillow, more than one thousand technology leaders, researchers, and industry experts signed an open letter calling for a pause in AI development to enable a catch-up of AI risk-mitigation measures and regulation; signatories include Apple co-founder Steve Wozniak and engineers from Amazon, DeepMind, Google, Meta, and Microsoft.

In attempts to grapple with the growth of AI, governing bodies have begun to outline frameworks and regulations that will have profound impacts on AI companies. It is therefore critical for investors to recognize their role in the development of this space, as poorly governed businesses can have negative impacts on not only the industry, but also the commercial viability of generative AI products as they exist today. Companies with best practices that can be sustained as they scale can both strengthen their investment proposition and help to advance the positive outcomes AI models can achieve.

Key ESG Considerations

Ethics and Compliance

One of the most prominent ethics risks as it relates to AI is the potential for programs to produce biased, inequitable, or otherwise harmful outputs. These outputs can 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. One of the many drivers of biased outcomes is the presence of bias within datasets used to train AI algorithms. Across the insurance industry, for example, the integration of AI into mortgage underwriting and claims services have the potential to discriminate against protected classes (e.g., by imposing higher insurance premiums or higher rates of debt collection) when models are trained on demographic data, such as zip codes. In this case and many others, locational data, while not inherently biased, is correlated with race and ethnicity due to historical segregation in the US. As such, companies run the risk of proliferating bias through AI models even when the datasets used are seemingly unbiased.

AI use in healthcare can similarly exacerbate inequalities: the World Health Organization recently issued a warning on the potential for AI to generate misleading or inaccurate information when trained on biased or non-representative patient data, particularly when integrated into decision- or diagnosis-making applications. Numerous studies have found that marginalized patient groups, for example, are consistently and selectively under-diagnosed or underserved, with one study in particular finding that Black patients had to be deemed significantly more ill than white patients to be recommended for the same care. Importantly, when applied without proper care or understanding of ethics considerations from the outset, AI can perpetuate serious societal harm.

To mitigate such harm, companies need to practice responsible dataset and algorithm development: data inputs must be accurate, high quality, free of bias (e.g., randomized), and representative of the applications' end subjects. The model itself must also be accurate, which companies can ensure by retraining algorithms at least annually or whenever the application is updated—for reference, AI models should provide accurate results and analysis at least 80 percent of the time. Oversight is key in creating ethical AI and mitigating any inaccuracies in the model; a strong control is to mandate human-in-the-loop reviews to oversee the effectiveness and accuracy of the model to prevent any overreliance on AI models' analysis.

The model itself should be transparent and explainable and should also be complemented by strong corporate governance. Stakeholders should be made aware that they are interacting with an AI system and the analysis and outcome of the model should be traceable and understandable. Importantly, companies should have clear leadership for responsible AI, formal processes to identify and mitigate AI systems biases (which include diverse teams of developers, as they are more likely to identify bias than homogeneous groups), leverage research and education on responsible AI, and maintain a strong practice of pursuing continuous improvement and stakeholder engagement. Business leaders may choose to draw best practices from existing guidelines, 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. To begin implementing these risk mitigation measures, a valuable first step companies can take includes the establishment of a responsible AI committee to oversee forthcoming programmatic developments.

Data Privacy and Security

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. One of the most infamous examples of this risk manifesting itself in recent history is through the Cambridge Analytica scandal in the 2016 US presidential election. The company acquired access to 50 million Facebook users' data without consent and used it to identify, profile, and eventually attempt to influence American voters through algorithmically informed targeted ads.

Another use through which AI can engender data privacy issues is facial recognition technology. Existing facial recognition systems draw from databases that store images from social media, government ID registries, or surveillance cameras. It is not just foreign governments leveraging this technology—Clearview AI (a US company) violated Canadian privacy laws by collecting images through mass surveillance and facial recognition without consent for aggregation and commercial sale to private businesses and the police. As such, the data privacy and security implications of AI are clear: 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.

As the use of large datasets in AI development is critical for ensuring accuracy and enhancing functionality, it is imperative for companies to proactively safeguard data subjects and their information by ensuring proper consent is obtained in a manner that is transparent to the user. In addition to these permissions, companies should de-identify information by randomizing datasets, generalizing identifiable (e.g., geographic or demographic) variables to be unspecific, and preventing the subsequent re-identification of data. To further prevent data privacy and security incidents, companies can similarly use anonymized data that removes personally identifiable variables altogether—techniques for this method include k-anonymization and differential privacy and also necessitate the prevention of re-identification. More broadly, it is critical for companies to align themselves with international data privacy regulations and frameworks, such as the General Data Protection Regulation (GDPR), the EU's policy on trustworthy AI, or the National Institute of Standards and Technology's (NIST) AI Risk Management Framework. By assessing current capabilities against accepted legal guidelines and responsible AI principles, companies can develop tools or policies to bolster their internal data privacy and security programs and compliance postures.

Labor Management

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. In fact, one fourth of current work tasks across all industries can be automated by AI in the US—on the higher end, nearly half of the tasks in the administrative and legal fields are exposed to AI automation. 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.

To meet these workforce challenges, companies can invest in re-education and upskilling of existing workforces. Improved on-the-job training, including basic STEM skills with an emphasis on creativity and critical and systems-based thinking, can help employees adapt to evolving labor markets. At a greater scale, companies need to begin redesigning work; rather than having AI replace humans, jobs can be designed so that technology empowers employees to complete their jobs more effectively, or in a way in which AI can take over tasks workers are overqualified for so that people can focus on valuable, interesting, or more innovative work. Companies can additionally upskill employees to managerial roles that include responsibilities related to AI oversight, which can only be completed by humans. Crucially, business leaders will need to collaborate with governing bodies to contribute to policy and public initiatives to protect people from the most severe labor market disruptions such as mass job loss.

Diversity, Equity, and Inclusion

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, the capacity for AI adoption to dislocate millions of workers across all industries—the lack of diverse representation is particularly concerning. Human biases are reflected throughout the AI lifecycle, including dataset creation, sourcing and labeling, algorithm training and testing, and post-processing performance reviews. When companies lack diverse perspectives, bias issues proliferate; team diversity, on the other hand, allows for not only a variety of perspectives but also for the greater identification of existing/potential biases that would not have been otherwise caught by a homogeneous group. In addition to mitigating risk, the business case for diverse talent in AI development is proved again and again: organizations in which at least 25% of AI development employees are women are 3.2 times more likely than others to be AI high performers. Similarly, when over 25% of AI development employees are racial or ethnic minorities, companies are twice as likely to be high performers.

As such, companies must make increased efforts towards diverse recruiting, with particular attention paid to its engineering and development teams, quality assurance or auditing teams, and any other stakeholders that may be involved in the development, deployment, and monitoring of AI models. Companies can leverage resources such as diverse job sites (e.g., Black Girls Code, Women Who Code), university affinity groups, and diversity career fairs. Importantly, organizations must also foster inclusive environments in which multiple viewpoints are not only celebrated but also encouraged in order to retain and continue to attract employees of all backgrounds.

Climate Change

While AI applications have been lauded as a potential tool to combat climate change (e.g., monitoring weather systems, preempting rainforest destruction), the training of AI models itself has a significant carbon footprint that can worsen climate change if not properly managed. When considering energy consumption, the carbon footprint of training a single big language model is estimated to equal around 300 thousand kilograms of carbon dioxide emissions, or 125 round-trip flights between New York and Beijing. Importantly, the higher the accuracy goal of the model, the higher the emissions from energy used to train the model will be. The environmental cost of training models is just the start, as 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.

The first step to decreasing emissions from AI model training is understanding the scope of emissions generated from these activities; companies can use resources such as the ML CO_2 calculator tool or Microsoft's guidelines on estimating energy use of AI training. Before training a model, companies should consider the importance of incremental increases in accuracy, and balance these with the reduction in emissions that would be achieved by not overtraining the model. Importantly, however, compromises should not be made when it comes to anti-bias controls and data integrity in such training decisions. Companies should also reconsider the sustainability of their data processing centers—location and renewable energy infrastructure can have tangible positive effects on businesses' greenhouse gas emissions. Companies that own on-premises data centers should also consider migrating to the public cloud; cloud data centers typically achieve far higher resource utilization and can therefore save energy and reduce companies' carbon footprints.

Potential Positive Use Cases

There are significant opportunities and use cases in which AI can enable positive impact. It can, for example, enhance crisis response procedures by predicting the progression of natural disasters or helping to locate missing persons. AI models, such as Google's TensorFlow, have been deployed to help combat environmental degradation and injustice by combatting illegal logging of the world's rainforests. To address diversity and inclusion issues, MIT Media Lab and Stanford University used AI to automate the emotion recognition and provide social cues to help individuals on the autism spectrum interact in social environments. As such, investors and businesses can work together to create responsible AI models that have positive social impact. Investors have the opportunity to provide guidance and expertise around AI governance, ethics, and labor management; in turn, companies can reap the benefits of and educate their developers on responsible and inclusive AI.

How Malk Can Support

Public and investor interest in AI has accelerated in recent years, particularly with the rapid development of generative AI models. Investors should continue to make time for in-depth due diligence and remember high market value does not always correlate with a strong investment. Investors should integrate ESG into the entire investment period by continuing to work with scaling companies on managing ongoing and upcoming risks; Malk is well positioned to assist investors in this process. By providing in-depth due diligence and ongoing portfolio monitoring that goes beyond an ESG score to create targeted recommendations, Malk enables investors to proactively manage their ESG risks and identify value creation opportunities across their portfolio.

Authors

Rae Edmunds
Senior Associate
e: redmunds@malk.com

Charlotte Sippel
Project Manager
e: csippel@malk.com

Christina Tedeschi
Project Manager
e: ctedeschi@malk.com

Malk Partners does not make any express or implied representation or warranty on any future realization, outcome or risk associated with the content contained in this material. All recommendations contained herein are made as of the date of circulation and based on current ESG standards. Malk is an ESG advisory firm, and nothing in this material should be construed as, nor a substitute for, legal, technical, scientific, risk management, accounting, financial, or any other type of business advice, as the case may be.

About Malk Partners

Malk Partners is the preeminent advisor to private market investors for creating and protecting value through environmental, social, and governance ("ESG") management and impact investing. Founded in 2009, Malk Partners advises many of the world's leading alternatives managers investing across private equity, growth equity, venture capital, and private credit by helping them define ESG goals, achieve ESG results, and guide their portfolio companies in driving value creation and mitigating risks. The firm is headquartered in La Jolla, California with a second office located in New York. For more information about Malk Partners, please visit www.malk.com.