Artificial Intelligence: Responsible AI and the path to long-term growth

October 2022

Théo Kotula
ESG Analyst
AXA IM Core

Key points

- Increased adoption of AI technology and an enhanced understanding of its potential, have raised questions for responsible investors of how businesses might put it to use

- AI systems offer clear opportunities, but regulation is ramping up to address the risks while investors should be conscious of reputational damage from privacy violations or security flaws

- We have seen a lack of awareness of these issues among corporates and have established a framework for engagement to help us actively encourage better disclosure and oversight

- We believe that mitigating the risks associated with AI systems is closely related to long-term value creation with these technologies

Artificial Intelligence (AI) has drastically transformed the business environment and become critical to companies across a broad range of sectors – from finance to healthcare and automotive.

Since their creation, AI systems have come a long way and are now able to combine complex data sets and computer science to mimic human problem-solving and decision-making abilities. AI also has a variety of real-world applications, ranging from customer-service automation to optimising service-operations or product features.

AI adoption continues to grow. According to a McKinsey survey,[i] of the more than 1,800 organisations consulted, 56% had adopted AI in 2021 – compared to 50% in 2020. The survey also found the number of respondents attributing at least 5% of earnings before interest and tax (EBIT) to AI, had risen to 27% from 22% the year before.

A similar study by Gartner in 2022 also recognised the growing applications of AI in business.[ii] It found that 40% of surveyed organisations had deployed thousands of AI models and predicted that the business value of AI would reach $5.1bn by 2025.

However, as the use of AI in business continues to grow, responsible AI practices will be paramount to maintaining this growth and contending with the potential ethical and socio-technical repercussions of AI.

As a global asset manager with extensive investments in the technology sector, we believe we can have a positive role in

encouraging Responsible AI. Our research, therefore, seeks to define the principles behind this and understand how they may be put into practice in the long-term.

AI opportunities and risks

With the real value that businesses can gain from AI systems, they represent potential opportunities to generate returns for clients. AI is notably the bedrock of the metaverse and will enable its scalability and automation[iii] – with deep learning and natural language processing (NLP) enhancing activities and interactions in the Web 3.0.[iv]

We believe that together, AI and the metaverse could drive an expanding and tangible investment opportunity, fuelled by companies with the potential to deliver strong year-on-year growth over the next decade.

From an environmental, social and governance (ESG) standpoint,[v] AI should allow investors to collect and analyse information on risks and opportunities at an unprecedented scale. In particular, we see great potential in using sentiment analysis algorithms and NLP solutions to better integrate ESG considerations into investment strategies.[vi]

At AXA IM, we have developed our own NLP-enabled impact investing technique which scores parts of corporate documents to help assess how companies can support the UN Sustainable Development Goals (SDGs).

Separately, another study has shown that AI may be a net contributor to achieving the 169 more detailed targets within the 17 broad SDGs.[vii] Potential examples include a technology that better identifies recyclable material to reduce ocean-bound plastic or a platform that uses satellite data and AI to reduce water waste.

However, alongside the opportunities and potential applications AI systems can offer, they also bring destructive risks – the most visible being privacy violations and cybersecurity breaches, which can quickly damage a company's reputation.

AI also carries societal and human rights-related risks which, though harder to define and measure, are arguably the most salient.[viii] Most notably, AI systems can be a source of discrimination. When used in financial services, 'network discrimination' can disadvantage certain communities from securing loans; when used in recruitment, AI can reproduce existing patterns of social bias.

Even when used for content moderation, automated content flagging can infringe upon rights to free expression and opinion – not to mention the legal and ethical concerns raised by the ability to identify and track individuals with AI.

Regulation: A nascent environment

Given the significant risks associated with AI adoption, we are witnessing the emergence of a new regulatory landscape. The European Union (EU) is in the process of drafting AI regulation that echoes the General Data Protection Regulation (GDPR), [ix] and which aims to address the risks associated with specific uses of AI while bolstering trust within the AI systems being used.[x]

The regulation classifies AI systems into three risk categories and designates corresponding action:

EU risk categories for AI systems

Unacceptable-risk AI systems
- Subliminal, manipulative, or exploitative techniques causing harm
- Real-time, remote biometric identification systems used in public spaces for law enforcement
- All forms of social scoring

High-risk AI systems
- Systems that evaluate consumer creditworthiness
- Recruiting or employee-management systems
- Systems utilizing biometric identification in nonpublic spaces
- Safety-critical systems (eg, systems that would put the health of citizens at risk due to failure)
- Any systems used in the administration of justice

Limited- and minimal-risk AI systems
- AI chatbots
- AI-enabled video and computer games
- Spam filters
- Inventory-management systems
- Customer- and market-segmentation systems
- Most other AI systems

Source: McKinsey & Co. What the draft European Union AI regulations mean for business, 2021: Relevance of AI risks for companies

This draft regulation would no longer permit the use of *unacceptable-risk systems* within the EU; it would subject *high-risk systems* to various requirements, including human oversight, transparency, cybersecurity, risk management, data quality, monitoring, and mandatory reporting; and it would require transparency within the use of *limited and minimal-risk systems*, so individuals are aware that they are interacting with AI technologies.

The regulation would not only apply to companies and AI providers within the EU but would have extraterritorial jurisdiction over any AI system providing output within the EU. Companies in breach could also face penalties of up to €30m or 6% of their global revenue – whichever is higher – making fines more substantial than those imposed by the GDPR.

As governments and regulators across the globe increasingly consider the governance and regulation of AI, we believe the risks could become very material for companies from a legal and financial perspective.

Responsible approaches to AI

Beyond regulation, the European Commission has produced Ethics Guidelines for Trustworthy AI[xi] which can serve as a useful framework for companies to implement Responsible AI practices. These guidelines seek to support AI that is lawful, ethical and robust in the face of their potential harm.

They are built around four ethical principles which bring concrete requirements for companies:

- **Respect for human autonomy**: Ensuring respect for the freedom and autonomy of human beings; the allocation of functions between humans and AI systems follows human-centric design principles and leaves meaningful opportunity for human choice; human oversight over work processes in AI systems is maintained

- **Prevention of harm**: AI systems should neither cause nor exacerbate harm or otherwise adversely affect human beings; attention must also be paid to situations where AI systems could cause or exacerbate adverse impacts due to asymmetries of power or information; human dignity as well as mental and physical integrity is protected

- **Fairness**: Ensuring equal and just distribution of benefits and costs, and that individuals and groups are free from unfair bias, discrimination, and stigmatisation; the use of AI systems does not lead to people being deceived or unjustifiably impaired in their freedom of choice; the ability to contest and seek effective redress against decisions

made by AI systems and by the humans operating them is preserved

- **Explicability**: Processes should be transparent, the capabilities and purpose of AI systems openly communicated, and decisions explainable to those directly and indirectly affected; "black box" algorithms are avoided

At AXA IM, our approach to Responsible AI draws on a similar structure and requires four clear components. We believe corporate AI systems should be **fair**, **ethical**, and **transparent** while remaining reliant on **human oversight** – both from a technical and high-level governance perspective.

We think that companies will be able to mitigate AI risks and deliver sustainable value creation if they practice Responsible AI policies – benefitting from the deployment of AI systems whilst building trust and improving customer loyalty.

The bottom line is that Responsible AI is not only about avoiding risks but ensuring that these technologies can be harnessed to the advantage of people and businesses alike.

Assessing how Responsible AI is practiced

We have looked at a range of studies and benchmark/assessment tools which deal with how companies address and practice Responsible AI – and they ultimately reveal a large gap in progress.
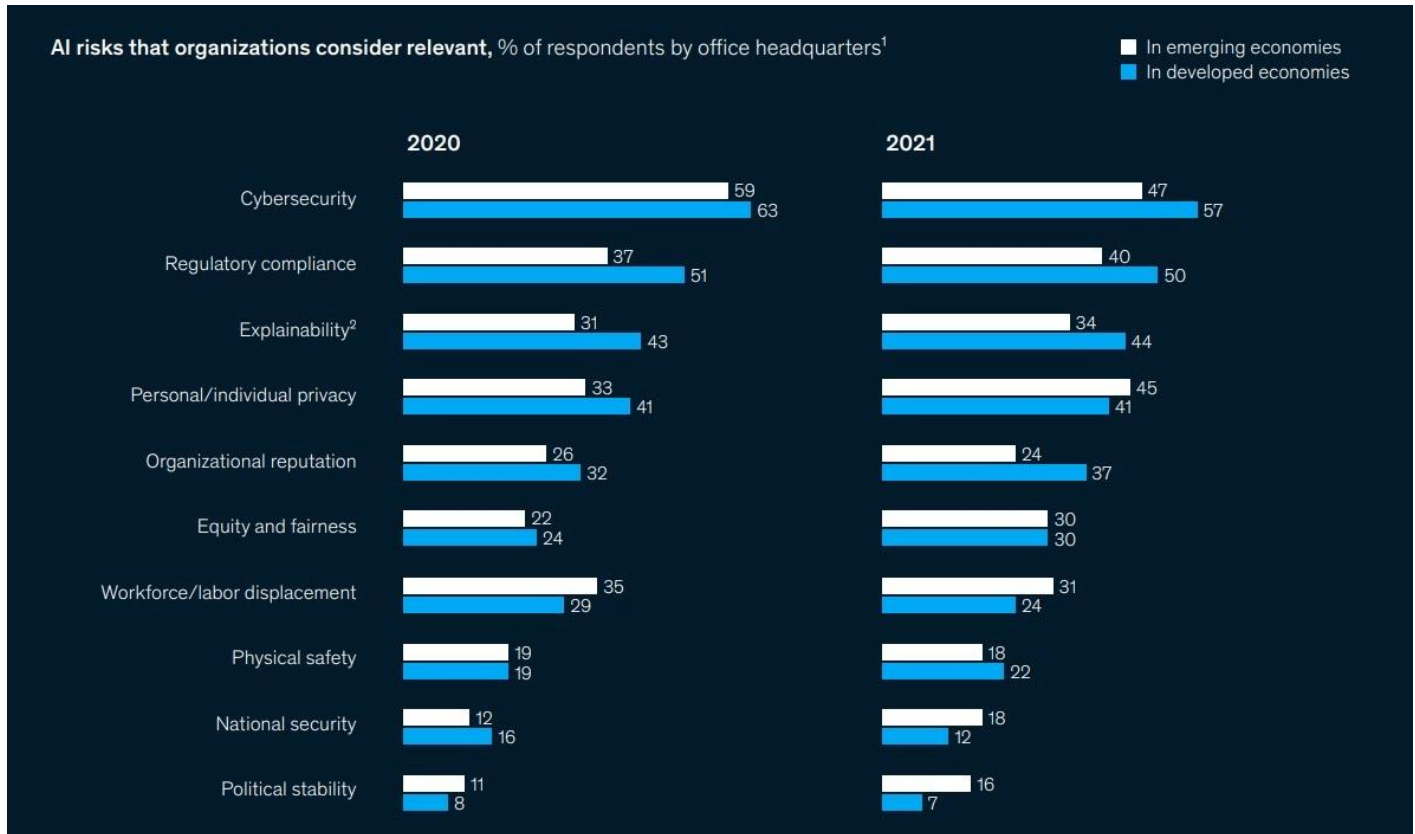
The 2021 McKinsey survey,[xii] for instance, highlights what we see as a lack of awareness amongst the organisations they consulted towards the various risks associated with AI. We think the graphic on the following page is a good illustration of how companies may be underestimating the risks.

Similarly, the 2022 Big Tech Scorecard by Ranking Digital Rights (RDR),[xiii, xiv] an independent research programme, graded tech companies relatively poorly in terms of algorithmic transparency – based on indicators which follow a similar framework to those discussed earlier, including elements relating to human rights policy in the development of algorithmic systems, risk assessments, transparency and privacy.

None of the major tech companies assessed achieved a score above 22%, according to RDR's methodology, with some of the biggest global firms near the bottom of the ranking.

The Digital Inclusion Benchmark from the World Benchmarking Alliance,[xv] meanwhile, ranked companies on "inclusive and

Corporate assessments of AI risks

AI risks that organizations consider relevant, % of respondents by office headquarters[1]

■ In emerging economies
■ In developed economies

2020

Risk	Emerging	Developed
Cybersecurity	59	63
Regulatory compliance	37	51
Explainability[2]	31	43
Personal/individual privacy	33	41
Organizational reputation	26	32
Equity and fairness	22	24
Workforce/labor displacement	35	29
Physical safety	19	19
National security	12	16
Political stability	11	8

2021

Risk	Emerging	Developed
Cybersecurity	47	57
Regulatory compliance	40	50
Explainability[2]	34	44
Personal/individual privacy	45	41
Organizational reputation	24	37
Equity and fairness	30	30
Workforce/labor displacement	31	24
Physical safety	18	22
National security	18	12
Political stability	16	7

Source: McKinsey & Co. The state of AI in 2021, 2021: Relevance of AI risks for companies

ethical research and development" and sought to capture "the degree to which company practices reflect the participation of diverse groups and consideration of ethics in product development". Our assessment of benchmark scores for a range of companies found underwhelming performance, with just 3% of the 77 we studied gaining a top score.

We think these results help demonstrate the need for responsible investors, such as AXA IM, to engage companies in discussions over Responsible AI and indicate how we might put that into practice. As part of our responsible technology research and engagement programme, we focus on firms with significant investment plans in AI development.

Our key engagement recommendations urge companies to:

• Align with the key components of AXA IM's approach to Responsible AI – fair, ethical, transparent, human oversight

• Increase disclosure and transparency around AI system development

• Ensure oversight of Responsible AI by their board and senior executives

We are convinced that mitigating the risks associated with AI systems – and addressing regulatory considerations – are closely related to the ability of companies to deliver long-term value creation with these technologies. We also believe that AI systems can better provide long-term and sustainable opportunities when Responsible AI is practiced.

[i] Global survey: The state of AI in 2021 – McKinsey, 2021
[ii] The Gartner AI Survey Top 4 Findings – Gartner, 2022
[iii] The Metaverse: Driven By AI, Along With The Old Fashioned Kind Of Intelligence – Forbes, 2022
[iv] Web 3.0 is an evolving idea for the third generation of the internet, incorporating concepts of decentralisation and new blockchain technologies to improve web applications. Web 1.0 refers to the early period of the internet to about 2004 while Web 2.0 reflects the move into more user-generated content through centralised hubs.
[v] How can AI help ESG investing? – S&P Global, 2020
[vi] Natural language processing (NLP) is a branch of computer science and AI that allows machines to process large volumes of written documents and rapidly analyse, extract, and use the content. Common uses include auto-correct technology and real-time sentiment analysis.
[vii] The role of artificial intelligence in achieving the Sustainable Development Goals – Nature Communications, 2020
[viii] Artificial Intelligence & Human Rights: Opportunities & Risks – Berkman Klein Center, 2018
[ix] The EU's General Data Protection Regulation (GDPR) is a set of rules on data protection and privacy introduced in 2018 which is aimed at increasing an individual's control over their personal information.
[x] What the draft European Union AI regulations mean for business – McKinsey, 2021
[xi] Ethics guidelines for trustworthy AI – European Commission Report, 2019
[xii] Global survey: The state of AI in 2021 – McKinsey, 2021
[xiii] The 2022 RDR Big Tech Scorecard – Ranking Digital Rights, 2022
[xiv] Who We Are - Ranking Digital Rights
[xv] Digital Inclusion Benchmark – World Benchmarking Alliance, 2021

Ref-27069