

**Is the challenge of governing AI any different to previous scenarios that have required boards to increase their business oversight? Compare the introduction of AI to a previous situation that has precipitated significant governance changes.**

---

**Introduction:**

Board oversight, as we know it, is the oversight of *processes*. A board receives reports on how a company is managing its risks and is assured by the knowledge that robust processes are in place. The process behind artificial intelligence (AI), however, can be so opaque, even to its designers, that it cannot be overseen in the same way. Governing AI requires boards to oversee something entirely different: *outcomes*. And this makes the challenge unique.

From the Enron Scandal to the 2008 Financial Crisis, the turn of the twenty-first century ushered in a series of events that catalysed – and in some cases instigated – significant changes in how corporate governance is regulated and practised. Such scenarios required heightened stability and predictability from corporate governance, however, navigating the multifaceted challenges and opportunities of technological advancement demands a more dynamic approach. A more applicable comparison to the introduction of AI is the digital age; in demonstrating that these two stages in technological evolution pose distinct governance challenges, this essay will establish that boards are faced with an entirely new challenge in governing AI.

In serving this comparison, this essay will draw upon technological theories and philosophical concepts, alongside governance analyses. Rather than contribute to an ever-growing fearmongering discourse, it will offer some practical steps for boards to consider in mitigating AI risks – above all, reorientating focus towards reviewing outcomes, rather than processes. It will challenge previous governance approaches that have focused on maximising shareholder value above a broader range of stakeholders and that have been overly rigid in their nature – endorsing a more agile approach.

**i. The Nature of These Two Challenges:**

In many ways, the internet and the proliferation of computers confronted boards with a similar challenge to that posed by AI. They had the capacity to transform business in ways that could not be easily predicted, creating new industries and revolutionising existing ones. To ensure their effectiveness, boards had to evaluate whether they had the necessary technical expertise, operating model, data and technology architecture, and risk management processes. Companies' exposure to third-party and data privacy risks intensified to an unprecedented degree and it became essential to ensure that staff had digital literacy. New jobs emerged and old ones became redundant: boards were faced with managing the structural and social implications of these changes on their workforce, as the prospect of job displacement became very real for many people.

The internet is focused on communication and information sharing. Its introduction revolutionised logistics and supply chain management in an unprecedented advancement of globalisation.<sup>1</sup> It enhanced collaboration between suppliers and manufacturers, introduced live

---

<sup>1</sup> Stenger, Alan J. "Advances in Information Technology Applications for Supply Chain Management." *Transportation Journal*, vol. 50, no. 1, 2011, pp. 37–52

tracking and monitoring, and enhanced customer experience.<sup>2</sup> Efficiency peaked, but this came at the cost of resilience, as inherent vulnerabilities halted global manufacturing and shipping and a worldwide overreliance on China developed.<sup>3</sup> Similarly, AI promises to further maximise efficiency, however, its focus is not on connecting systems and transmitting information, but instead on automation and decision making – therefore, something else is at stake: *transparency*.

Previous technologies, such as computers and the internet, have had clear processes with a chain of cause and effect that boards could easily oversee and receive assurance on. However, there is no audit trail for the process behind AI. Particularly with machine learning and deep learning models, AI systems can have a ‘black box’ process. Even those with the relevant expertise can struggle to account for a machine’s output. From a technological perspective, it may be somewhat reductive to describe a technology solely in terms of its inputs and outputs, and obfuscate the extent to which these algorithms can be understood, nevertheless, this granularity is not compatible with the high-level view of board oversight.<sup>4</sup> Moreover, the very nature of these systems is adaptive and autonomous; models are trained once or continually to infer trends and connections in data that are not easily discernible to humans, and they can continue to develop new ways of inference that were not envisioned by their creators. This opacity makes the ability of boards to assess risks, receive assurance, and ensure accountability infinitely more difficult.

It is also worth acknowledging the ways in which the digital age introduced new governance challenges, distinct from those posed by AI. With the internet, information, previously limited to professionals in insiders, became readily available to all stakeholders who were now able to analyse corporate institutions far more closely and subject their actions to far greater scrutiny.<sup>5</sup> This has developed beyond shareholder interests, and companies are now expected to present sustainability stories to broader societal stakeholders.<sup>6</sup>

## ii. The Specifics of the AI Challenge:

AI systems are trained with past data, and this forms the basis of their assumptions. Their performance is therefore dependent on the quality of this input data. If the data is unrepresentative of the reality with which the system is interacting or reflective of existing prejudices this will result either in inaccuracy or in **perpetuating biases**.<sup>7</sup> Google recently came under fire for its Gemini AI tool generating racially diverse Nazi-era German soldiers, with CEO Sundar Pichai acknowledging

---

<sup>2</sup> Jackson, Rob ‘The Revolutionary Impact of the Internet of Things on Supply Chain’, *IT Supply Chain*, 5, April, 2023

<sup>3</sup> Runde, Daniel F., and Sundar R. Ramanujam. *Recovery with Resilience: Diversifying Supply Chains to Reduce Risk in the Global Economy*. Centre for Strategic and International Studies (CSIS), 2020

<sup>4</sup> See Kroll Joshua A., ‘The Fallacy of Inscrutability’, *Philosophical Transactions of the Royal Society: Mathematical, Physical and Engineering Sciences*, 2018; Jarke, Juliane, and Hendrik Heuer. “Reassembling the Black Box of Machine Learning: Of Monsters and the Reversibility of Foldings”, *Algorithmic Regimes: Methods, Interactions, and Politics*, Amsterdam University Press, 2024, pp. 103–26

<sup>5</sup> See Raymond, Mark, et al. “MULTI-STAKEHOLDERISM: ANATOMY OF AN INCHOATE GLOBAL INSTITUTION.” *Who Runs the Internet?: The Global Multi-Stakeholder Model of Internet Governance*, Centre for International Governance Innovation, 2017, pp. 19–44

<sup>6</sup> See Kashmanian, Richard M., et al. “Corporate Environmental Sustainability Strategy: Key Elements.” *The Journal of Corporate Citizenship*, no. 44, 2011, pp. 107–30

<sup>7</sup> Hao, Karen, ‘This is how AI bias really happens—and why it’s so hard to fix’, *MIT Technology Review*, 4 Feb. 2019

and condemning the bias it displayed.<sup>8</sup> And Amazon has previously faced criticism for developing an AI hiring algorithm that systematically discriminated against women applying for technical jobs.<sup>9</sup> These biases can affect both employees and customers and therefore carry enormous reputational risks. Boards should consider deferring the use of AI in sensitive areas, such as recruitment, until these issues have been assuredly resolved. Or at least ensure a human is in place to check their outputs. Given the amount of bias already baked into our systems, if this is to change a new approach will be required.<sup>10</sup>

Perhaps the biggest risk, particularly of Generative AI, is that it is prone to simply **make things up**. The danger here is that it appears to display logic and reasoning without actually understanding the inputted prompt or the answer it provides; they simply assemble a string of words probabilistically. Philosophically, this is known as competence without comprehension. And while Daniel Dennett may obscure the importance of this distinction, the lack of comprehension as it relates to board oversight significantly obscures transparency and accountability.<sup>11</sup> Large Language Models (LLMs) routinely and confidently produce, what the industry terms, ‘hallucinations’. Some believe this problem will only get worse, as LMM-generated material increasingly floods the internet, and these models feed on their own falsehoods in subsequent training, gradually corrupting all corners of the internet.<sup>12</sup>

With the creation, storage, and transmission of digital information in the digital age, cyber security emerged as a critical risk facing all organisations. However, AI ‘**deepfake**’ technology is revolutionising the sophistication of fraud, and the threat now facing boards is of an entirely different scale. These highly realistic synthetic creations are becoming increasingly challenging to discern from reality. To illustrate, a multinational company was recently defrauded of HK\$200 million after one of its employees fell victim to deepfake technology: the victim joined a conference call and was ordered to transfer money by the chief financial officer, in the presence of many other familiar employees – all of whom turned out to be entirely fake digital reconstructions.<sup>13</sup> ‘Deepfake’ fraud has evolved from one-to-one calls using voice cloning software to convincing multi-person conference calls in an alarmingly short time frame; they may soon become indistinguishable and boards will have to adopt a completely new approach to defend against these threats. In the meantime, boards must commend vigilance, provide regular and engaging training, and constantly monitor emerging threats to ensure controls are sufficiently resilient.

All technologies evolve; the internet developed incredibly fast, often in unpredictable ways, but its basic principles remained relatively stable, and this allowed boards to develop governance frameworks at a measured pace. However, AI is **hyper-evolutionary**: it is advancing at an unprecedented pace, far surpassing Moore’s Law, which has so far successfully predicted the

---

<sup>8</sup> Heath, Alex, ‘Google CEO Says Gemini AI Diversity Errors Completely Unacceptable’, *The Verge*, 28 Feb. 2024

<sup>9</sup> Goodman, Rachel ‘Why Amazon’s Automated Hiring Tool Discriminated against Women’, *American Civil Liberties Union*, Oct. 12, 2018

<sup>10</sup> See Pooley, Jefferson. *Large Language Publishing*, 2024, for potential solutions

<sup>11</sup> See Dennett, C., Daniel “‘A Perfect and Beautiful Machine’: What Darwin’s Theory of Evolution Reveals about Artificial Intelligence”, *The Atlantic*, 2012

<sup>12</sup> Pooley, Jefferson. *Large Language Publishing*, 2024

<sup>13</sup> Kong, Harvey, “‘Everyone looked real’: multinational firm’s Hong Kong office loses HK\$200 million after scammers stage deepfake video meeting”, *South China Morning Post*, 4 Feb. 2024

doubling of computer power every two years.<sup>14</sup> Instead, the computation used to train the most powerful AI models has grown tenfold every year for the past 10 years.<sup>15</sup> Not only is computing capacity expanding exponentially, but AI may likely become the first technology capable of improving itself. These challenges will only grow as the foundation technologies that facilitate AI will continue to become smaller, cheaper, and more accessible.

### **iii. Looking Forward: What this Means for Boards**

Governing AI is among the most difficult challenges facing boards over the coming decades. They are tasked with overseeing something way beyond their competencies, changing very fast, and not just unknown, but unknowable. Responses to previous scenarios precipitating governance changes have favoured predictability – with Sarbanes-Oxley driving mandated financial reporting practices – and stability – with tightening capital requirements following 2008. However, if boards are to oversee AI effectively, their governance approach must reflect its inherently unpredictable nature.

Boards demonstrated commendable adaptability in generally handling the COVID-19 pandemic effectively, but AI will require boards – as well as regulators – to adopt an entirely new level of agility. Given the ever-evolving risks, opportunities, and ethical dilemmas AI poses, approaches and policies in place today may be ineffectual or restrictive in a matter of months – not years. And this requires a precautionary – but at the same time, responsive and self-corrective approach.

In managing previous governance challenges, regulation has been reactive. However, given the speed and scale of potential AI catastrophes, this response may be inadequate. Regulating the use of AI will require a more proactive approach that identifies and mitigates potential crises before these systems are deployed and it is too late. In light of this, boards must be vigilant in managing heightened legal and regulatory risks. It will be essential to closely monitor regulatory evolution, keep track of your organisation's use of AI tools, and consider assigning someone to oversee AI development.

Given the speed at which AI is developing, there is also immense pressure on boards to champion innovation and not get left behind. This poses the risk of companies rushing into AI, driven by earnings target pressures, fears of competitors gaining advantages, or simply wanting to jump on the trend. Potentially resulting in it being used in the wrong cases, poorly implemented, or creating additional cybersecurity and reputational risks. Boards will need to ensure all stakeholders are involved in deciding new use cases, including the relevant representatives in customer, sustainability, and compliance. It will also be crucial to ensure that organisations have the technical capabilities and capacity to execute AI projects. Given the complexity of the technology behind AI, incorporating external advice into boardroom discussions will be essential.

Boards must always keep the risks front of mind when using AI or sanctioning its use. They must decide where AI is being used and whether the lack of accountability is legally and morally permissible. All outputs should be subject to a vigilant review process that includes human decisions

---

<sup>14</sup> Bremmer, Ian, Suleyman and Suleyman, Mustafa, 'Building Blocks for AI Governance by Bremmer and Suleyman', *International Monetary Fund*, 2023

<sup>15</sup> Ibid

and considers all AI outputs inherently untrustworthy. The effectiveness of this process should be routinely assessed by the Board. However, with more than half of workers using AI at work doing so unofficially, review processes will struggle without knowing exactly what work has used AI.<sup>16</sup> It is therefore crucial for boards to vet its use in all cases. Clear internal disciplinary rules around unsanctioned use are also essential, but given the accessibility of Generative AI, boards must understand that its use is somewhat inevitable. Therefore, fostering a culture of transparency and disclosure around the use of AI will be instrumental.

### **Conclusion:**

The governance challenge posed by the rise of AI is unlike any other boards have faced in the past. Previous technological advancements, such as the internet, posed similar challenges in requiring boards to evaluate the extent of their technical expertise, the capabilities of their risk management functions, and their resilience to third-party and data-privacy risks, however, AI demands a fundamentally different approach. Unlike the internet, which is focused on communication and sharing information, AI introduces a level of opacity that challenges traditional methods of oversight. The opacity of these algorithms presents a significant obstacle for boards in overseeing risk management and ensuring accountability. Previous technology presents clear chains of cause-and-effect, while the autonomy and adaptability – not to mention biases and hallucinations – with which AI operates, makes it near impossible for boards to oversee the process.

Managing this challenge has significant implications for board oversight. It requires a fundamental shift from process-oriented, to outcome-focused oversight. This requires boards to adopt an agile and proactive approach to governance that emphasises transparency, accountability, and ethical considerations amidst a broader stakeholder universe. Boards must look to establish robust review processes and foster a culture of transparency and disclosure around the use of AI. A more proactive regulatory approach is necessary to anticipate emerging risks before it is already too late, and Boards must be closely monitoring regulatory developments to ensure their compliance with ever-evolving standards.

Overall, governing AI requires boards to navigate uncharted waters with agility, foresight, and ethical integrity. By embracing these principles and reorientating their assurance focus, boards can effectively mitigate the risks and at the same time capitalise on the opportunities that AI promises.

---

<sup>16</sup> 'More than Half of Generative AI Adopters Use Unapproved Tools at Work', *Salesforce*, 2023

## Bibliography

---

- Bremmer, Ian, Suleyman and Suleyman, Mustafa, 'Building Blocks for AI Governance by Bremmer and Suleyman', *International Monetary Fund*, 2023
- Broadbent, Meredith. *What's Ahead for a Cooperative Regulatory Agenda on Artificial Intelligence?* Center for Strategic and International Studies (CSIS), 2021.
- Cataleta, Maria Stefania. *Humane Artificial Intelligence: The Fragility of Human Rights Facing AI*. East-West Center, 2020
- Cath, Corinne. "Governing Artificial Intelligence: Ethical, Legal and Technical Opportunities and Challenges." *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2133, 2018
- Csernaton, Raluca. "To Innovate or to Regulate?: That Is the Misleading Question." *Charting the Geopolitics and European Governance of Artificial Intelligence*, Carnegie Endowment for International Peace, 2024
- Dennett, C., Daniel "'A Perfect and Beautiful Machine': What Darwin's Theory of Evolution Reveals about Artificial Intelligence", *The Atlantic*, 2012
- Engelke, Peter, *AI, Society, and Governance: An Introduction*. Atlantic Council, 2020.
- Goodman, Rachel 'Why Amazon's Automated Hiring Tool Discriminated against Women', *American Civil Liberties Union*, Oct. 12, 2018
- Hao, Karen, 'This is how AI bias really happens—and why it's so hard to fix', *MIT Technology Review*, 4 Feb. 2019
- Heath, Alex, 'Google CEO Says Gemini AI Diversity Errors Completely Unacceptable', *The Verge*, 28 Feb. 2024
- Jackson, Rob 'The Revolutionary Impact of the Internet of Things on Supply Chain', *IT Supply Chain*, 5, April, 2023
- Jarke, Juliane, and Hendrik Heuer. "Reassembling the Black Box of Machine Learning: Of Monsters and the Reversibility of Foldings", *Algorithmic Regimes: Methods, Interactions, and Politics*, Amsterdam University Press, 2024
- Kashmanian, Richard M., et al. "Corporate Environmental Sustainability Strategy: Key Elements." *The Journal of Corporate Citizenship*, no. 44, 2011
- Kong, Harvey, "Everyone looked real": multinational firm's Hong Kong office loses HK\$200 million after scammers stage deepfake video meeting', *South China Morning Post*, 4 Feb. 2024
- Kroll Joshua A., 'The Fallacy of Inscrutability', *Philosophical Transactions of the Royal Society: Mathematical, Physical and Engineering Sciences*, 2018
- Pooley, Jefferson. *Large Language Publishing*, 2024

Raymond, Mark, et al. "MULTI-STAKEHOLDERISM: ANATOMY OF AN INCHOATE GLOBAL INSTITUTION." *Who Runs the Internet?: The Global Multi-Stakeholder Model of Internet Governance*, Centre for International Governance Innovation, 2017

Reed, Chris. "How Should We Regulate Artificial Intelligence?" *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2128, 2018

Runde, Daniel F., and Sundar R. Ramanujam. *Recovery with Resilience: Diversifying Supply Chains to Reduce Risk in the Global Economy*. Centre for Strategic and International Studies (CSIS), 2020

Salesforce, 'More than Half of Generative AI Adopters Use Unapproved Tools at Work', 2023

Stenger, Alan J. "Advances in Information Technology Applications for Supply Chain Management." *Transportation Journal*, vol. 50, no. 1, 2011

Willim, Robert. *Mundania: How and Where Technologies Are Made Ordinary*. 1st ed., Bristol University Press, 2024