# DIRECTORS IN THE LOOP? RESPONSIBLE CORPORATE GOVERNANCE FOR THE ERA OF AI

Deliversity Lynn Warneke The Australian National University

#### **ABSTRACT**

This article examines how the relationship between corporate success and technological progress has become more evident in the era of Artificial Intelligence (AI), highlighting its disruptive impact on the economy, law, and society. As AI becomes a key driver of profitability and competitive differentiation, it also generates socioeconomic externalities that pose significant challenges for corporate governance and the interplay between private value and public interest. This paper assesses the effectiveness of corporate law and governance in the context of AI, arguing that directors are not sufficiently prepared to govern AI in a way that promotes longterm corporate value. The article proposes reforms for "responsible AI governance," indicating that substantial legal normative changes are necessary to address the risks and benefits associated with AI. In conclusion, it is suggested that directors must adopt principles of "corporate techno-social responsibility" to establish a new model of responsible governance that redefines corporate value in this disruptive era.

#### **RESUMEN**

Este artículo analiza cómo la relación entre el éxito corporativo y el progreso tecnológico se ha vuelto más evidente en la era de la Inteligencia Artificial (IA), destacando su impacto disruptivo en la economía, el derecho y la sociedad. A medida que la IA se convierte en un motor clave de rentabilidad y diferenciación competitiva, también genera externalidades socioeconómicas que plantean importantes desafíos para la gobernanza corporativa y la interacción entre el valor privado y el interés público. Este trabajo evalúa la efectividad del derecho y la gobernanza corporativa en el contexto de la IA, argumentando que los directores no están suficientemente preparados para gobernar la IA de manera que promueva el valor corporativo a largo plazo. El artículo sugiere reformas para la "gobernanza responsable de la IA", indicando que son necesarios cambios legales y normativos sustanciales para enfrentar los riesgos y beneficios asociados con la IA. En conclusión, se plantea que los directores deben adoptar principios de "responsabilidad tecno-social corporativa" para establecer un nuevo modelo de gobernanza responsable que redefina el valor corporativo en esta era disruptiva.

**KEYWORDS:** Corporate Governance, Artificial Intelligence, Techno-Social Responsibility, Corporate Value, Regulatory Reforms, Socioeconomic Risks

PALABRAS CLAVE: Gobernanza Corporativa, Inteligencia Artificial, Responsabilidad

Tecno-Social, Valor Corporativo, Reformas Regulatorias, Riesgos Socioeconómicos.

**JEL CODE:** G34; K22

**RECIBIDO:** 25/03/2024 **ACEPTADO:** 04/09/2024 **DOI:** 10.26807/rfj.v1i15.507

#### INTRODUCTION

The relationship between corporate success and technological progress has never been more overt, with digital businesses and products proliferating at extraordinary pace and scale. Of the many innovations to have emerged, Artificial Intelligence (AI) entails the greatest disruption to the corporation, economy, law, and society, and thus represents a singular challenge for the company director. AI is becoming pervasive, driving profitability and competitive differentiation (Williams, 2022), (Commonwealth of Australia, 2022), yet its profound socioeconomic externalities are provoking attention on corporate governance and the nexus between private value and public interest (Cihon, Schuett, & Baum, 2021), (Dignam, 2020) (Ford, 2021), (Land, 2020). At the same time, a lack of contextualization for AI in contemporary regulatory frameworks creates ongoing legal uncertainties for industry and society. This paper critically assesses the effectiveness of corporate law and governance in this context and argues—primarily from the Australian perspective—that directors are not adequately prepared to govern AI for long-term corporate value. Part 1 descriptively examines the distinctive challenge of AI governance, contextualizing the subsequent normative arguments. Part 2 critically analyses present-day board and legal effectiveness in the governance of AI for shareholder and stakeholder benefit. Part 3 explores "responsible AI" governance reforms, contending substantial legal and normative changes are required in future. AI will effect a momentous socioeconomic transformation, promising great benefits but carrying equally profound risks; therefore, the paper concludes that while an appropriate regulatory framework for AI is now essential, the director has a critical role to pre-emptively adopt "corporate techno-social responsibility" principles and establish a new model of responsible governance to redefine corporate value for this most disruptive era.

# Part 1 AI — A Distinctive Governance Challenge

If a precondition of effective governance is clarity on what is being governed, AI can be challenging even at the definitional level. While acknowledging that reductive characterizations can obscure governance complexities, AI is commonly defined simply as computing systems and techniques that simulate human cognition.

The *Macquarie Dictionary* defines Artificial Intelligence (AI) as "the ability of a computer or other device or application to function as if possessing human intelligence [and] the branch of computer science which deals with the design and use of machines that have this ability" (Macquiere Dictionary, 2022). The term is frequently used in an "umbrella" sense to describe a field that comprises a diverse range of methodologies and practices which leverage technologies, algorithms and datasets, and the application of human expertise at points in the AI lifecycle, to achieve functionality that could be described as being on an "intelligence spectrum": from relatively simple automated data analysis and decision-making, often with high levels of human intervention or controls; to complex and sophisticated machine learning solutions to achieve defined objectives with limited human intervention or controls; through to autonomous machine problem-solving, which may in some cases produce original determinations, new knowledge or discoveries with minimal or no explicit human involvement.

Distinction may be made between "narrow AI" (also "assisted" and "augmented" AI), comprising technologies and techniques that are currently achievable or feasible in the near term, and "general AI" or "artificial general intelligence" (also "strong AI" and "the singularity") which, thus far, is a speculative concept and contested by experts in the field. This paper refers only to narrow AI. The term will be used broadly herein, in an "umbrella" sense and without further specification, given the focus of the paper is on broad normative and regulatory implications for corporate governance of "narrow AI" (Macquiere Dictionary, 2022).

Increasingly *surpassing* human-level achievement (Brynjolfsson, Rock, & Syverson, 2016), offering "new business capabilities with significant potential for value creation" (Fuhrman & Mooney, 2021) and material financial returns (McKinsey & Company, 2019). AI is now sufficiently pervasive, powerful and productive to matter to the board; but its myriad opportunities and risks, non-exhaustively examined in this Part, constitute a distinctive governance challenge.

# Opportunities and Benefits

The director's statutory role to act in "the best interests of the corporation" (Corporations Act (Cth), 2001, s 181) includes, inter alia, strategic decisionmaking in pursuit of shareholder profit (Australian Institute of Corporate Directors (AICD), 2020). With AI progressively underpinning corporate performance, boards must determine how to extract value from, and maintain competitive advantage in, a rapidly evolving landscape. Opportunities abound, with myriad automated and machine-learning solutions operating in diverse industries and contexts (Fuhrman & Mooney, 2021): AI powers mortgage approvals (Eyers, 2022), investment services (Featherstone, 2017), predictive medicine, agricultural and environmental applications (McKinsey Global Institute, 2019), the creation of original artwork (Roose, 2022) (Perrigo, 2021), and remarkable scientific discoveries (Callaway, 2022). Corporations are driving substantial AI research and development: One study "identified 4403 AI-related companies that received a total of USD 55.7 billion in funding in the year ending July 2019" (Cihon, Schuett, & Baum, 2021). New AI companies are multiplying (Cihon, Schuett, & Baum, 2021). The director faces a key challenge to identify strategic or operational opportunities within this plethora that will meet company objectives and achieve return-on-investment and shareholder value (Board Agenda, 2021). Inaction, short-termism, "dashboard myopia" (Armour & Eidenmüller, 2019) or ill-informed decision-making are obverse challenges—the "opportunity cost" of AI. Many once-leading corporations have suffered value erosion by failing to keep up with technology opportunity (Valentine et al., 2020, pp. 225, 228), hence if the board is "slow to embrace technology, compared to its rivals [...] activists will be all over them" (Featherstone, 2017). Sophisticated investors may consider strategies that merely replicate those of competitors a "massive opportunity lost" (Governance Institute of Australia, 2022), driving their agents to pursue competitive differentiation from AI. Additionally, many scholars have conjectured opportunities for AI to strengthen governance while reducing agency costs (see generally: (Möslein, 2018), (Fenwick & Vermeulen, 2018), (Picciau, 2021), (Kalmanath, 2019), (Enriques & Zetzsche, 2020), (Hilb, 2020) and (Gramitto Ricci, 2020)); therefore principals could

conceivably challenge agents to augment their own capability, extending the director role beyond corporate governance of AI, to governance with AI (Hilb, 2020, p. 867). Instances of AI-augmented boardrooms are limited, for example, "Edison" at Salesforce and "Vital" at a Hong Kong investment firm were two early examples, enthusiastically reported on by media at the time, but apparently not yet replicated to a material degree in the intervening years and therefore possibly more marketing hype than currently feasible governance innovation (Burridge, 2017) (Hickey, 2018). however theoretical opportunities for AI to benefit corporate performance and conformance include investor profiling (Armour & Eidenmüeller, 2019), selecting directors and remunerating officers (Featherstone, 2017), (Laptev & Feyzrakhmanova, 2021) and (Fenwick & Vermeulen, 2018) reducing information asymmetries between actors (Picciau, 2021, p. 106), enhancing director independence and minimizing 'groupthink', and mitigating corporate liability by pre-emptively identifying potential non-compliance (Kalmanath, 2019, pp. 6, 7-8, 12-13), (Enriques & Zetzsche, 2020, pp. 7, 66).

AI opportunity is a complex, multifactorial, and dynamic governance challenge, but it is claimed that "for any organization that wants to leap forward [...] meeting that challenge will determine their future". (Board Agenda, 2021, p. 4) Thus, maintaining long-term corporate value in the era of AI is emerging as a singular director role and responsibility.

#### Risks and Harms

Notwithstanding potential rewards, AI's inherent risks are currently acute: Acemoğlu contends that current AI technologies "are more likely to generate various adverse social consequences, rather than the promised gains" (2021). High failure or error rates persist, Gartner predicts that through 2022, "85 percent of AI projects will deliver erroneous outcomes due to bias in data, algorithms, or the teams responsible for managing them". (Gartner, 2018) (Nimdzi Insights, 2019) and it is claimed that compliance failures are "expected to multiply in the near future" (World Economic Forum [WEF], 2022). Duties of care and diligence (Corporations Act (Cth), 2001, s 180 (1)) demand that the director pays close regard to AI's endogenous and exogenous regulatory, economic, and reputational risks.

Non-legal risks in AI adoption include potential reputational and financial impacts (which of course can develop into legal issues). High-quality, contextually accurate AI models are costly (Dignam, 2020), but lower-quality models that suffer well-known accuracy and bias problems (Dignam, 2020) risk causing economic harm to the corporation—for example, compensatory settlements or loss of revenue arising from erroneous, algorithmically-biased exclusion of customers. Reputational risk subsists in deficient or defective AI datasets, algorithms and human expertise, with extensive evidence that AI continues to misdiagnose patients, discriminate against minorities, systematically impinge upon human and consumer rights, and injure even kill—citizens (Dignam, 2020). Commercially compelling but ethically ambiguous AI adoption risks employee and investor activism, negative media coverage and related damage to the corporation (Cihon, Schuett, & Baum, 2021), (Sim, 2019). Legal and regulatory risk can originate ex ante in flawed AI designs or arise *ex post* in unanticipated results that infringe existing laws: examples of unlawful AI-facilitated outcomes include discriminatory hiring and credit approval practices, profit-optimizing distortion of share markets, and algorithmic collusion on pricing (Diamantis, 2020). AI-generated collection and use of personal data risks non-compliance with privacy, cybersecurity, anti-discrimination and consumer laws (Armour & Eidenmüller, 2019, p. 18), (Chiu & Lim, 2021) Conversely, lacunae in Australian law represents regulatory risk, as return-on-investment and shareholder value could be impaired if future legislation were to render an existing AI product unlawful. Supply-chain risk can manifest in opaque, "black-box" AI procured from third parties. The "tech nirvana fallacy" (Enriques & Zetzsche, 2020), risk of overconfidence in AI could result in poor governance decisions and adverse results. Critically, exogenous to the corporation at the intersection of business and society, growing public awareness of AI harms—from widespread workforce displacement to privacy infringements and discrimination—is creating deep societal distrust and mounting expectations of corporate transparency, fairness and accountability (Williams, 2022). Critically, the board must therefore guard against the "moral hazard" risk of creating externalities that damage consumer trust and "business-society relations" (Chiu & Lim, 2021), and impair the firm's market value. AI risks are profuse, multifactorial, and dynamic, with implications for corporate social responsibility.

AI opportunities and risks are not neatly divisible into "benefit" and "harm" respectively but represent a complex admixture of corporate incentives and disincentives, with potential for immense social externalities. Many predict "this is just *the tip of the iceberg*, with the vast majority of digitization yet to occur" (Commonwealth Scientific and Industrial Research Organization [CSIRO], 2022), and the distinctive governance challenge for directors primed to grow:

we can expect shareholders to point to those who were in a position to act during this window when the harms are increasingly visible, especially as regulators clarify the rules of the AI road. (Eccles & Vogel, 2022)

Part 2 will therefore examine the effectiveness of current normative and legal modalities in the governance of AI for short- and long-term shareholder value.

# Part 2 Contemporary Governance of AI

Actions by corporate governance actors today will have long-term impact "through path dependence in governance regimes" (Cihon, Schuett, & Baum, 2021, p. 21). This invites normative assessment of AI governance skills and practices "in the boardroom" and the application of corporate and related laws to AI "in the courtroom".

#### In the Boardroom

As the apex corporate governance body, the board is claimed to have "the greatest potential impact on organizational performance and behavior" (Bankewitz, Åberg, & Teuchert, 2016, p. 58-59). Extensive studies (Board Agenda, 2021), (Governance Institute of Australia, 2022), (Australian Institute of Company Directors [AICD], 2019), (Valentine, 2016), (Voogt & Verreynne, 2018), (Watermark Search International, 2021) and (Weill et al., 2019) have therefore researched the preparedness of directors today to "create value for organizations and society tomorrow" (Bankewitz, Åberg, & Teuchert, 2016, p. 58). Globally, many corporate actors believe a "lack of skills and knowledge at the top of organizations about [AI's] transformative

capacity" is inhibiting adoption (Board Agenda, 2021), (Governance Institute of Australia, 2022). Relatedly, studies confirm a growing gap between "AI power users and adoption laggards" (McKinsey & Company, 2019): empirical research, utilizing machine-learning analysis, demonstrated that companies governed by boards comprising a minimum of three directors with a specific digital skill-set (Weill et al., 2019, p. 41) outperformed competitors on all key valuation metrics: "We found that among companies with over \$ 1 billion in revenues, 24 % had digitally savvy boards, and those businesses significantly outperformed others on key metrics—such as revenue growth, return on assets, and market cap growth [...] it takes three members to have a statistically significant impact" (Weill et al., 2019, p. 41-42). Earlier research found similar correlations between the technical/digital capability of board directors and firm performance, and "current and future value creation through digital transformation was driven from the top. These results occurred across all industry sectors, without exception" (Valentine et al., 2020, p. 227). In Australia, only three percent of directors have technology expertise, (Australian Institute of Company Directors [AICD], 2019, p. 28) Findings are corroborated by the GIA survey which "uncovered a distinct lack of digital skills in the boardroom" (Governance Institute of Australia, 2022, p. 10) rising to just under seven percent in the top 300 public companies (Watermark Search International, 2021 p. 15-16). Despite findings that Australian boards do not "prioritize innovation or disruption risks to the extent seen in overseas boardrooms" (Australian Institute of Company Directors [AICD], 2019, p. 10), and directors admitting minimal ability to assess "both the *ethical and practical implications* of using modern technologies. emphasis added" (Australian Institute of Company Directors [AICD], 2019, p. 30) the imperative to add this expertise to the board remains contentious, "The push for more technical experts on boards – technology, cybersecurity, human resources or scientific experts – is being resisted" (Durkin, 2021), suggesting problematic "over-confidence' in the status quo" (Enriques & Zetzsche, 2020, p. 55). Noting that corporate governance codes have strong normative and indirect legal effects on the development of directors' duties, a multi-jurisdictional academic study found none currently "refer to technology skills, digital literacy or cyber fluency as important [and only one...] includes

significant benchmarks to deal with the effects of technology" (Voogt & Verreynne, 2018). Because AI will have substantial impacts on industry and society, inadequate expertise on boards creates constraints and significant risks for corporations (Board Agenda, 2021, p. 4); however the evidence suggests Australian boards overall may be deficient in expertise correlated with financial performance in the digital economy, complacent, and lacking optimal governance frameworks and norms. Currently, directors appear illequipped to fulfil their fiduciary and statutory duties to maintain long-term corporate value in the era of AI (Evans, 2020, pp. 210-217), (Valentine et al., 2020, p. 227).

#### In the Courtroom

Absent contextualized or specific laws, only a few AI-related cases have come before Australian courts and regulators; however, determinations that existing laws were breached, and customers were harmed led to severe financial and reputational damage to the companies involved. The Australian Federal Court imposed major pecuniary penalties on Trivago for breaching Consumer Law by falsely claiming its pricing algorithms advantaged customers (Australian Competition and Consumer Commission [ACCC], 2021), (ACCC v Trivago NV [2020] FCA 16), (Trivago N.V. v ACCC [2020] FCAFC 185).

The Office of the Australian Information Commissioner found 7-Eleven (Office of the Australian Information Commissioner [OAIC], 2021) and Clearview AI (Office of the Australian Information Commissioner [OAIC], 2021) breached the Privacy Act (1998) by unlawfully collecting sensitive personal data for AI-enabled facial recognition applications (implicating Clearview AI's customers in illegal activity, and exemplifying supply-chain risk). As faceprint technologies are not explicitly regulated (Davis, Perry, & Santow, 2022), these breaches were of privacy *consent* law: three retailers therefore recently argued that entry signage informing customers about in-store use of facial recognition constituted the necessary consent, and suspended their practices only after an investigation was announced and reputational harm became acute (Blakkarly, 2022). These preliminary cases

illuminate risks for corporate AI adoption within Australia's extant legal framework, together with emergent gaps; as firms deploy AI, the board's care and diligence role must include assessing regulatory compliance of novel AI applications. While these judgments may not have clearly implicated the companies' directors, they suggest a failure to appropriately inform themselves (Corporations Act 2001 (Cth) s 180(2)(c)) or to prevent foreseeable harms—with significant consequences for both corporation and customer.

# Governance Gaps

AI's practical and legal novelty and related regulatory gaps may represent a particularly significant governance risk given the trend in Australian law towards imposing greater accountability for discharge of director duties under sections 180 and 181 of the Corporations Act 2001 (Cth) (Lowry, 2012), and the uncertain defense the business judgment rule provides for board oversight failure (Nettle, 2018). Courts have determined directors owe "a core, irreducible requirement of involvement in the management of the company" (Deputy Commissioner of Taxation v Clark, 2003), hence those who fail to make informed decisions, adequately manage AI risks, or undertake prudent oversight of management when adopting AI, could potentially breach their duties of care and diligence (Petrin, 2019). Legal scholars trace increasing strictness in court interpretation of the standard of care expected of the modern director (Lowry, 2012, p. 257), and some posit that boards will only comply with non-delegable statutory duties in future by demonstrating expertise in data and AI governance (Armour & Eidenmüeller, 2019), (Möslein, 2018, pp. 660-662), (Picciau, 2021, p. 130) —much as high— profile corporate failures led to courts establishing the objective standard for a director's financial literacy (Australian Securities and Investments Commission v Healey, 2011), AI may provoke a similar clarification regarding technical literacy. Theoretically, demonstrably inadequate board expertise to convert AI to capital value could constitute statutory failure to act in the company's best long-term interests.

Australian corporate law operates within a broader civil enforcement regime, in which courts increasingly regard director duties as "public

obligations bearing an important social function" (Hill, 2020). Breach actions "usually brought by ASIC, [have an] 'extraordinarily high success rate" (Hill, 2020, p. 27), hence deficiencies in the board's ability to assess the impact of AI risk on strategy, to the detriment of shareholder and stakeholder interests, could fall short of the standard required, leading to breach of duties and possible liability (Voogt & Verreynne, 2018, p. 1354). With corporate regulators now seeking to understand AI use and risk mitigations in banking (Eyers, 2022) and declaring plans for AI-enabled compliance innovations (Australian Securities and Investments Commission, 2022), it is clear that algorithmic scrutiny and assurance will escalate. Current laws cannot hold AI directly liable (Hilb, 2020, p. 859), (Kalmanath, 2019, p. 12) and therefore the corporation and its directors could become liable for harms arising from the autonomous algorithms they create or deploy (Abbott & Sarch, 2019), (Armour & Eidenmüeller, 2019), (Chiu & Lim, 2021), (Diamantis, 2020), (Hilb, 2020), (Laptev & Feyzrakhmanova, 2021), (Selbst, 2021), (European Commission, 2022). The possibility of corporations and the natural persons who govern them becoming a "liability sponge" (Johnson, 2020) may have a chilling effect on AI adoption and innovation in an era when "embracing technology is becoming increasingly a matter of survival" (Picciau, 2021), reducing long-term shareholder and societal benefits. Equally however, AI exacerbates existing socioeconomic disparities and generates harms at scale, therefore regulatory oversight and sanctions should be expected.

The well-documented "pacing problem"—in which technology rapidly outpaces the law, creating gaps and ambiguities in its wake—is evident and to what extent director duties and corporate obligations will be normatively and legally prescribed in the era of AI remains uncertain. Part 3 will examine potential reforms to "hard" and "soft" law and the director role that could ensure the corporation, as the locus of private decision-making with acute public impact, adopts AI in accordance with shareholder and stakeholder interests.

# PART 3 Future Governance of 'Responsible AI'

Broad consensus is emerging that corporate self-regulation of AI is unsustainable, and laws are needed to control AI risks, cultivate industry and public confidence, and secure national prosperity and international competitiveness (Edelman, 2019). Enabling AI use and innovation by industry, while prescribing stakeholder protections and societal obligations to prevent harms, necessitates a systemic "responsible AI" (Gillis, 2021) (Ford, 2021) modality in the private and public interest, employing "hard laws" and sanctions, "soft law" fiduciary standards and ethical governance practices.

# Hard(er) Legislative and Regulatory Reforms

Internationally, AI-related laws are under development in several districts and de facto or de jure extra-territorial effects are anticipated. (Siegmann & Anderljung, 2022), (Townshend, 2022) Observing this "strong global competition", the Australian government is calling for views on AI regulation, aiming to position "Australia as a leader in digital economy regulation" to enhance public trust and encourage uptake. The Australian Human Rights Commission has proposed comprehensive human rightsrespecting AI laws and an independent AI Commissioner to oversee compliance (Australian Human Rights Commission, 2021). Applicationand sector-specific AI laws could form part of Australia's regulatory mix: for example, a model facial recognition technology law has been proposed (Davis, Perry, & Santow, 2022), and targeted regulation for AI-enabled policing or medical applications, an international example being the United States' proposed regulation of AI solutions that constitute a medical device: US Food and Drug Administration (FDA) (U.S. Food and Drug Administration, 2021) could balance industry innovation and public safeguards in high-risk contexts. Comprehensive *sui generis* law may prove impractical for a mutable technology like AI. Moses, writing on emerging technology and legal problems, outlines a model design for "a legal system that treats different technologies fairly and is resistant to difficulties associated with technological change" (2007) therefore contextualizing existing statute, such as anti-discrimination, employment, and competition and consumer laws, could complement and minimize reliance on dedicated AI legislation. The Australian Competition and Consumer Commission has, for example, recommended reform of extant law for the digital era (Australian Competition and Consumer Commission, 2019), particularly the Privacy Act which is widely regarded as inadequate to protect consumers from emerging technology harms (Attorney-General's Department, n.d.).

To the extent that "the current problems of AI are problems of unregulated AI" (Acemoğlu, 2021) a proportionate hard law framework that prescribes acceptable, and proscribes unacceptable, AI use, with enforcement and redress provisions, should be broadly welcomed by industry, regulators, and civil society. Notably, however, corporate law remains "extraordinarily complex, imprecise, confusing, imperfect and very much in need of reform and clarification" (Voogt & Verreynne, 2018, p. 1342) —acutely so, given director duties were codified before extensive corporate uptake of AI. Although the Australian Law Reform Commission is undertaking a multi-year inquiry into corporate law, both the terms of reference and recently released interim submission focus principally on simplifying financial services regulation and make no reference to mounting technology-related gaps and imperatives (Australian Law Reform Commission, 2020). Therefore, absent reform of corporate law that addresses AI governance, soft law reform becomes critical to further define and drive responsible corporate governance norms for the era of AJ.

#### Soft Law Governance Reforms

Soft law, in the form of non-binding AI guidelines or standards, is recognized by various experts (Cihon, Schuett, & Baum, 2021, pp. 12-13) for its flexibility and utility in supporting responsible development AI, technically and normatively, while hard law reform progresses slowly. An already extensive voluntary AI soft law apparatus, ranging from ethical principles (Australian Government, 2019), to technical standards and certification frameworks (Boza & Evgeniou, 2021), has recently been complemented by governance-specific instruments (International Standards Organization, 2022). At present, all are discretionary and of questionable prominence (Eyers, 2022).

Extensive scholarship therefore argues that AI soft law should progress from voluntary, to a "comply or explain" model, or even become binding as AI becomes more pervasive and powerful (Enriques & Zetzsche, 2020), (Cihon, Schuett, & Baum, 2021), (Picciau, 2021), (Chiu & Lim, 2021) and (Voogt & Verreynne, 2018, pp. 1359-1360). For example, corporate governance codes such as the "ASX Corporate Governance Principles" (ASX Corporate Governance Council, 2019) could initially provoke directors to engage with AI-related governance duties by requiring technology-related reporting and AI-specific governance disclosures, particularly where material to the firm's strategy and risk profile. Amending Australian corporate regulatory standards to mandate a specific responsible AI framework would establish an objective yardstick against which board effectiveness could be evaluated, and provide investors, regulators, and civil society with access to information that will be essential to evaluate corporate compliance, market value and social impact in the era of AI. Such transparency and accountability mechanisms could in turn help to develop public trust in AI necessary to realize its full potential (Commonwealth Scientific and Industrial Research Organization [CSIRO], 2022). Together, normative soft law incentives and deterrents would ideally foster a virtuous "race to the top" toward responsible AI and, in combination with hard laws, effect adaptation of director duties and corporate governance practices for responsible AI.

# Corporate Governance of Responsible AI: Directorsin-the-Loop

Regulatory conditions are currently uncertain, and it is becoming accepted that legal clarity or reforms in relation to AI will be essential to ensure responsible and safe adoption. Nonetheless, corporate agents must always govern beyond the minimum standard required by law (Australian Institute of Corporate Directors (AICD), 2020) and therefore the board should engage critically, now, with this most disruptive of technologies. With the corporation leading AI research and development, and AI adoption necessitating a "boardroom-led strategy" that will have near- and long-term impacts (Board Agenda, 2021, p. 9), the director is a critical human "in-the-loop" of AI governance. As research implies many Australian boards are

currently ill-equipped for this role, creating attendant risks for companies and society, the director role must include making immediate precautionary and tactical changes to address AI governance. Interdisciplinary sub-committees, inclusive of strategic, technical, legal and ethics skillsets, could assist the board in the interim to manage the breadth and complexity of the responsible AI governance agenda (Picciau, 2021, p. 130) (Enriques & Zetzsche, 2020, p. 94). Arguably, the most pressing reform required is recognizing digital skillsets are now a 'crucial' capability within the board's "universal skills" (Voogt & Verreynne, 2018, pp. 1349) and adjusting board composition and practices accordingly—such as elevating technology and innovation on the agenda (Evans, 2020, p. 213), (Australian Institute of Company Directors [AICD], 2019, p. 10-11) and developing director capabilities that will ensure holistic governance of AI threats and opportunities (Bankewitz, Åberg, & Teuchert, 2016).

However, boards should look beyond mere tactical reforms. Compelling arguments made by many stakeholders assert that AI represents a unique socioeconomic paradigm, requiring responsible governance in the private and public interest, and therefore that "stakeholderism" must become the dominant modality of corporate governance (Korinek & Balwit, 2022). AI may therefore compel boards to undertake a genuine transformation of corporate governance that cohesively integrates harms-based and benefitsbased approaches. Firstly, by adopting a "forward compliance" strategy that does not "merely wait for or rely on regulatory parameters" (Chiu & Lim, 2021), directors could establish a firm-wide, rules-oriented AI methodology that pre-emptively forecasts and mitigates risks and prevents harms to both corporation and society. Secondly, and relatedly, by building technical and ethical expertise consciously aligned with responsible AI values and the corporation's strategic purpose, directors could ensure the firm delivers private and public value from AI, for mutual business and societal benefit. Deriving social legitimacy from fair, accountable and transparent AI governance additionally creates an opportunity for competitive differentiation and market advantage. This implicates an adaptation of corporate social responsibility principles for AI—a new paradigm of "Corporate Technosocial Responsibility" (CTR) (Bughin & Hazan, 2019), potentially reflected

in a firm-specific code or covenant, in which the creation of shareholder value from AI is consciously aligned with long-term societal needs and harms prevention—for example, purposefully adopting AI for organizational growth over cost-reduction and prioritizing worker reskilling over redundancy, could realize shareholder value and, simultaneously, broader multi-stakeholder and societal benefits. By reporting on the corporation's responsible AI code, under a fourth pillar within an integrated Environment Social *Technology* and Governance (ESTG) disclosure framework, the board could capitalize on governance as a competitive differentiator, while also delineating a new fiduciary yardstick and positive role for the director as creator and trustee of shareholder profit and societal purpose and redefining the scope and import of corporate value in the era of AI.

### **CONCLUSION**

AI-enabled transformation of industry, the economy, law, and society are nascent, but progressing rapidly. This paper has argued that industry is currently adopting AI at pace, ahead of effective corporate governance capabilities, norms, and laws, and thus risks both short-term shareholder value and long-term societal well-being. As AI threatens immense socioeconomic and citizen harms, its equally immense potential benefits for industry and humanity can only be assured within a comprehensive hard and soft corporate law and governance framework—one that enables the creation of corporate value and economic prosperity, while simultaneously prescribing AI risk-management and harms-prevention, in a modality commensurate with societal expectations, interests and needs. As legislators and regulators inevitably seek to create a public-private regulatory framework for AI, the director remains a vital governance actor overseeing responsible AI adoption at the apex of the corporation. Effectively fulfilling fiduciary and statutory duties to act in the company's best interests and realize long-term value from AI will likely require the board to undertake an ambitious and far-reaching transformation of private corporate governance in the public interest, predicated on Corporate Techno-social Responsibility principles and a responsible AI covenant. How effectively the director accepts and acquits the

# **RFJ**

critical role of governing AI responsibly for profit and purpose will materially impact not only the corporation's shareholders but employees, consumers, and citizens in the era of AI.

#### **REFERENCES**

- Abbott, R., & Sarch, A. (2019). Punishing artificial intelligence: Legal fiction or science fiction. *University of California Davis Law Review*, *53*, 323. https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3327485
- Acemoğlu, D. (2021, December 1). What are the dangers of unregulated AI? An expert explains. World Economic Forum (WEF). https://www.weforum.org/agenda/2021/12/unregulated-artificial-intelligence-aitech
- Armour, J., & Eidenmüeller, H. (2019). Self-driving corporations? (Law Working Paper No. 475/2019). European Corporate Governance Institute (ECGI). https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3442447
- Armour, J., & Eidenmüeller, H. (2019, August). Self-driving corporations? (Law Working Paper No. 475/2019). European Corporate Governance Institute (ECGI). https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3442447
- ASX Corporate Governance Council. (2019). ASX corporate governance principles and recommendations (4th ed.). ASX.
- Attorney-General's Department. (n.d.). Privacy Act review: Discussion paper—Published responses. Attorney-General's Department. https://consultations.ag.gov.au/rights-and-protections/privacy-act-review-discussion-paper/consultation/published\_select\_respondent
- Australian Competition and Consumer Commission (ACCC) v Trivago NV [2020] FCA 16.
- Australian Competition and Consumer Commission (ACCC). (2021, April 22). Trivago to pay \$44.7 million in penalties for misleading consumers over hotel room rates (Media release). https://www.accc.gov.au/media-release/trivago-to-pay-447-million-in-penalties-for-misleading-consumers-over-hotel-room-rates

- Australian Competition and Consumer Commission. (2019). Digital platforms inquiry: Final report. Australian Competition and Consumer Commission. https://www.accc.gov.au/system/files/Digital%20platforms%20inquiry%20-%20final%20report.pdf
- Australian Government. (2019, November 7). Australia's artificial intelligence ethics framework. Department of Industry, Science and Resources. https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework
- Australian Human Rights Commission. (2021). Human rights and technology: Final report 2021. Australian Human Rights Commission. https://humanrights.gov.au/our-work/rights-and-freedoms/publications/human-rights-and-technology-final-report-2021
- Australian Institute of Company Directors (AICD). (2019, September). Driving innovation: The boardroom gap (AICD and University of Sydney Business School). https://www.aicd.com.au/content/dam/aicd/pdf/news-media/research/2019/Driving-Innovation-The-Boardroom-Gap.pdf
- Australian Institute of Corporate Directors (AICD). (2020, January 1). Director Tools: Role of the Board. Retrieved from https://www.aicd.com.au/content/dam/aicd/pdf/tools-resources/director-tools/board/role-of-board-director-tool.pdf
- Australian Law Reform Commission. (2020, September 11). Review of the legislative framework for corporations and financial services regulation. Australian Law Reform Commission. https://www.alrc.gov.au/inquiry/review-of-the-legislative-framework-for-corporations-and-financial-services-regulation/
- Australian Securities and Investments Commission v Healey [2011] FCA 717 (Austl.).

- Australian Securities and Investments Commission. (2022). ASIC corporate plan 2022–26: Focus 2022–23 (August 2022). ASIC. https://asic.gov.au/about-asic/corporate-governance/asic-corporate-plan/
- Bankewitz, M., Åberg, C., & Teuchert, C. (2016). Digitalization and boards of directors: A new era of corporate governance? *Business and Management Research*, *5*(2), 58-59. https://doi.org/10.5430/bmr.v5n2p58
- Blakkarly, J. (2022, July 12). Kmart, Bunnings and The Good Guys using facial recognition technology in stores. *Choice*. https://www.choice.com.au/consumers-and-data/data-collection-and-use/how-your-data-is-used/articles/kmart-bunnings-and-the-good-guys-using-facial-recognition-technology-in-store
- Board Agenda. (2021). Leadership in AI 2021: Boards, barriers and new beginnings (Research report, p. 3). Mazars & INSEAD Corporate Governance Centre. https://www.insead.edu/sites/default/files/assets/dept/centres/icgc/docs/leadership-in-ai-2021.pdf
- Boza, P., & Evgeniou, T. (2021, April). Implementing AI principles: Frameworks, processes, and tools (Working Paper). *INSEAD*. https://doi.org/10.2139/ssrn.3824551
- Brynjolfsson, E., Rock, D., & Syverson, C. (16 de January de 2016). Unpacking the AI-Productivity Paradox. Retrieved from MITSloan Management Review: https://sloanreview.mit.edu/article/unpacking-the-ai-productivity-paradox/
- Bughin, J., & Hazan, E. (2019, August 6). Can artificial intelligence help society as much as it helps business? *McKinsey Quarterly*. https://www.mckinsey.com/business-functions/quantumblack/our-insights/can-artificial-intelligence-help-society-as-much-as-it-helps-business
- Burridge, N. (2017, May 25). AI takes its place in the boardroom: Directors' tasks are becoming increasingly automated: Are fully autonomous companies next? *Nikkei Asia*. https://asia.nikkei.com/Business/AI-takes-its-place-in-the-boardroom

- Callaway, E. (28 de July de 2022). "The entire protein universe": AI predicts shape of nearly every known protein. Retrieved from *Nature:* https://www.nature.com/articles/d41586-022-02083-2
- Chiu, I. H.-Y., & Lim, E. W. K. (2021). Managing corporations' risk in adopting artificial intelligence: A corporate responsibility paradigm. Washington University Global Studies Law Review, 20(2), 347-367. https://ssrn.com/abstract=3780586
- Chiu, I. H.-Y., & Lim, E. W. K. (2021). Managing corporations' risk in adopting artificial intelligence: A corporate responsibility paradigm. *Washington University Global Studies Law Review*, 20(2), 347–367. https://ssrn.com/abstract=3780586
- Cihon, P., Schuett, J., & Baum, S. (2021). Corporate Governance of Artificial Intelligence in the Public Interest. Information, 1 30.
- Commonwealth of Australia. (2022). Positioning Australia as a leader in digital economy regulation Automated decision making and AI regulation. Department of the Prime Minister and Cabinet. Retrieved from https://storage.googleapis.com/converlens-au-industry/industry/p/prj211c4e81fb27d147ec9c1/public\_assets/automated-decision-making-ai-regulation-issues-paper.pdf
- Commonwealth Scientific and Industrial Research Organization (CSIRO). (2022, July). Our future world: Global megatrends impacting the way we live over coming decades (p. 29). CSIRO. https://www.csiro.au/en/about/facilities-operations-our-sites/our-future-world
- Corporations Act 2001 (Cth) s 180(2)(c) (Austl.).
- Davis, N., Perry, L., & Santow, E. (2022, September). Facial recognition technology: Towards a model law (Human Technology Institute Report, pp. 7-8). Human Technology Institute. https://www.uts.edu.au/sites/default/files/2022-09/Facial%20recognition%20model%20 law%20report.pdf

- Deputy Commissioner of Taxation v Clark (2003) 57 NSWLR 113, 140 [108] (Austl.).
- Diamantis, M. (2020). Who pays for AI injury? Oxford Business Law Blog. https://www.law.ox.ac.uk/business-law-blog/blog/2020/05/who-pays-ai-injury
- Diamantis, M. (2020, mayo 4). Who pays for AI injury? Oxford Business Law Blog. https://www.law.ox.ac.uk/business-law-blog/blog/2020/05/who-pays-ai-injury
- Dignam, A. (2020). Artificial intelligence, tech corporate governance and the public interest regulatory response. *Cambridge Journal of Regions, Economy and Society, 13*(1), 37-54. https://ideas.repec.org/a/oup/cjrecs/v13y2020i1p37-54..html
- Durkin, P. (2021, September 21). Six business leaders predict the future for boards. *Australian Financial Review*. https://www.afr.com/work-and-careers/leaders/six-business-leaders-predict-the-future-for-boards-20210909-p58q4t
- Eccles, R. G., & Vogel, M. (2022, January 5). Board responsibility for artificial intelligence oversight. Harvard Law School Forum on Corporate Governance. https://corpgov.law.harvard.edu/2022/01/05/board-responsibility-for-artificial-intelligence-oversight/
- Edelman. (2019). 2019 Edelman AI survey (Survey Results Report, pp. 28-29). Edelman. https://www.edelman.com/research/2019-edelman-ai-survey
- Enriques, L., & Zetzsche, D. A. (2020). Corporate technologies and the tech nirvana fallacy. *Hastings Law Journal*, 72(1), 55–102. https://repository.uchastings.edu/hastings\_law\_journal/vol72/iss1/2

- European Commission. (2022). Proposal for a directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) COM (2022) 496 Final (28 September 2022). https://ec.europa.eu/info/sites/default/files/1\_1\_197605\_prop\_dir\_ai\_en.pdf
- Evans, G. L. (2020). Technology and the corporate board 2020 and beyond. In R. Leblanc (Ed.), The handbook of board governance: A comprehensive guide for public, private and not-for-profit board members (2nd ed., pp. 210-217). Wiley.
- Eyers, J. (21 de September de 2022). Why Westpac's board is obsessed with artificial intelligence. Retrieved from Financial Review: https://www.afr.com/companies/financial-services/westpac-wants-an-ai-law-to-guide-adoption-and-build-trust-20220920-p5bjgc
- Featherstone, T. (24 de March de 2017). Governance in the new machine age.

  Retrieved from Australian Institute of Company Directors: https://www.aicd.com.au/innovative-technology/disruptive-innovation/examples/governance-in-the-new-machine-age.html
- Fenwick, M., & Vermeulen, E. P. M. (2018). Technology and corporate governance: Blockchain, crypto, and artificial intelligence (Law Working Paper No. 424/2018). ECGI. https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3263222
- Ford, J. (2021). Ford, Jolyon, Ethical AI: The Role of Law and Regulation. SSRN 3831211, 1-13.
- Fuhrman, P., & Mooney, J. (2021). Business Adoption of Artificial Intelligence: An Analysis of Scope, Intent and Realized Business Benefits. *Graziadio Business Review*, 24(1).
- Gartner. (2018, February 13). Gartner says nearly half of CIOs are planning to deploy artificial intelligence. *Gartner Newsroom*. https://www.gartner.com/en/newsroom/press-releases/2018-02-13-gartner-says-nearly-half-of-cios-are-planning-to-deploy-artificial-intelligence

- Governance Institute of Australia. (2022, May). Driving the digital revolution: A guide for boards. Governance Institute of Australia. https://www.governanceinstitute.com.au/media/887017/driving-digital-revolution.pdf
- Gramitto Ricci, S. A. (2020). Artificial agents in corporate boardrooms. *Cornell Law Review*, 105(3), 869–914. https://ssrn.com/abstract=3677627
- Hickey, A. (2018, January 26). AI sits in on Salesforce's board meetings and has something to say. CIO Dive. https://www.ciodive.com/news/ai-sits-in-on-salesforces-board-meetings-and-has-something-to-say/515606/
- Hilb, M. (2020). Corporate governance and artificial intelligence. *Journal of Business Ethics*, 162(4), 849–864.
- Hilb, M. (2020). Toward artificial governance? The role of artificial intelligence in shaping the future of corporate governance. *Journal of Management and Governance*, 24(2), 851–873. https://doi.org/10.1007/s10997-020-09519-9
- Hill, J. G. (2020). Legal personhood and liability for flawed corporate cultures (Law Working Paper No. 431/2018, ECGI, pp. 25-26). *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3309697
- International Standards Organization. (2022, April). ISO/IEC 38507:2022: Information technology—Governance of IT: Governance implications of the use of artificial intelligence by organizations. ISO. https://www.iso.org/standard/56641.html
- Johnson, K. (2020, October 12). When AI hurts people, who is held responsible? *VentureBeat*. https://venturebeat.com/ai/when-ai-hurts-people-who-is-held-responsible/
- Kalmanath, A. (2019, May 16). The perennial quest for board independence: Artificial intelligence to the rescue? *Albany Law Review* (Forthcoming). https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3360349

- Korinek, A., & Balwit, A. (2022, May). Aligned with whom? Direct and social goals for AI systems (Center on Regulation and Markets Working Paper #2). Brookings. https://www.brookings.edu/wp-content/uploads/2022/05/Aligned-with-whom-1.pdf
- Land, M. K.-l.-0.-0. (2020). Human Rights and Technology: New Challenges for Justice and Accountability. *Annual Review of Law and Social Science*.
- Laptev, V. A., & Feyzrakhmanova, D. R. (2021). Digitalization of institutions of corporate law: Current trends and future prospects. *Laws*, *10*(93), 1-16. https://doi.org/10.3390/laws10040093
- Lowry, J. (2012). The irreducible core of the duty of care, skill and diligence of company directors: Australian Securities and Investments Commission v Healey. *The Modern Law Review*, 75(2), 249–249. https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=2014563
- Macquiere Dictionary. (15 de march de 2022). Search word: Artificial Intelligence. Retrieved from Macquiere Dictionary: https://www-macquariedictionary-com-au.virtual.anu.edu.au/features/word/search/?search\_word\_type=Dictionary&word=artificial+intelligence
- McKinsey & Company. (22 de November de 2019). Global AI Survey: AI proves its worth, but few scale impact. Retrieved from https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact#/
- McKinsey Global Institute. (15 de May de 2019). "Tech for Good": Using technology to smooth disruption and improve well-being. Retrieved from https://www.mckinsey.com/featured-insights/future-of-work/tech-for-good-using-technology-to-smooth-disruption-and-improve-well-being
- Moses, L. B. (2007). Recurring dilemmas: The law's race to keep up with technological change. *University of Illinois Journal of Law, Technology & Policy, 2007*(2), 243. https://doi.org/10.2139/ssrn.980045

- Möslein, F. (2018). Robots in the boardroom: Artificial intelligence and corporate law. In W. Barfield & U. Pagallo (Eds.), *Research handbook on the law of artificial intelligence* (p. 649). Edward Elgar Publishing.
- Nettle, G. (2018). The changing position and duties of company directors. *Melbourne University Law Review*, 41, 1402–1422.
- Nimdzi Insights. (2019). Artificial intelligence, localization, winners, losers, heroes, spectators, and you (Whitepaper, pp. 13-15). Nimdzi Insights & Pactera EDGE. https://www.nimdzi.com/wp-content/uploads/2019/06/Nimdzi-AI-whitepaper.pdf
- Office of the Australian Information Commissioner (OAIC). (2021, 14 de October). Decision: Commissioner initiated investigation into Clearview AI, Inc. (Privacy) [2021] AICmr 54. AustLII. https://www.austlii.edu.au/cgi-bin/viewdoc/au/cases/cth/AICmr/2021/54.html
- Office of the Australian Information Commissioner (OAIC). (2021, 29 de September). Decision: Commissioner initiated investigation into 7-Eleven Stores Pty Ltd (Privacy) (Corrigendum dated 12 October 2021) [2021] AICmr 50. AustLII. https://www.austlii.edu.au/cgi-bin/viewdoc/au/cases/cth/AICmr/2021/50.html
- Perrigo, B. (23 de August de 2021). An Artificial Intelligence Helped Write This Play. It May Contain Racism. Retrieved from TIME: https://time.com/6092078/artificial-intelligence-play/
- Petrin, M. (2019). Corporate management in the age of AI (University College London (UCL) Working Paper Series No. 3/2019, p. 37). SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3346722
- Picciau, C. (2021). The (un)predictable impact of technology on corporate governance. *Hastings Business Law Journal*, 17(1), 67–96. https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3643500
- Privacy Act 1988 (Cth).

- Roose, K. (24 de August de 2022). We Need to Talk About How Good A.I. Is Getting. *The New York Times*. https://www.nytimes.com/2022/08/24/technology/ai-technology-progress.html
- Selbst, A. D. (2021). An institutional view of algorithmic impact assessments. *Harvard Journal of Law & Technology*, 35(1), 117. https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3867634
- Siegmann, C., & Anderljung, M. (2022). The Brussels effect and artificial intelligence: How EU regulation will impact the global AI market (Report). Centre for the Governance of AI.
- Sim, B. (2019, June 18). Shareholders quiz Google on AI risks. Financial News. https://www.fnlondon.com/articles/shareholders-quiz-google-on-ai-risks-20190618?mod=hp\_LATEST
- Townshend, P. (2022, April 28). How is AI regulated around the world? SmartFrame Blog. https://smartframe.io/blog/how-is-ai-regulated-around-the-world/
- Trivago N.V. v Australian Competition and Consumer Commission [2020] FCAFC 185.
- U.S. Food and Drug Administration. (2021, January). Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) action plan. U.S. Food and Drug Administration. https://www.fda.gov/media/145022/download
- Valentine, E. L. H. (2016). Enterprise technology governance: New information and technology core competencies for boards of directors (PhD thesis). Queensland University of Technology. https://doi.org/10.13140/RG.2.2.34027.95529
- Voogt, T., & Verreynne, M.-L. (2018). Director appointments: Expressing board care and diligence. *University of New South Wales Law Journal*, 41(4), 1335-1359. https://www.unswlawjournal.unsw.edu.au/wp-content/uploads/2018/12/Voogt-and-Verreynne.pdf

- Watermark Search International. (2021). 2021 board diversity index (Watermark Search International & Governance Institute of Australia). https://www.watermarksearch.com.au/2021-board-diversity-index
- Weill, P., Ross, J. W., & others. (2019). It pays to have a digitally savvy board. MIT Sloan Management Review, 60(3), 41-48.
- Williams, S. (17 de March de 2022). Only 22% of Australians trust AI implementation study. Retrieved from IT Brief Australia Technology news for CIOs & IT decision-makers: https://itbrief.com.au/story/only-22-of-australians-trust-ai-implementation-study
- World Economic Forum (WEF). (2022, January). Empowering AI leadership: AI C-suite toolkit (p. 63). https://www3.weforum.org/docs/WEF\_Empowering\_AI\_Leadership\_2022.pdf