European Business Organization Law Review
https://doi.org/10.1007/s40804-025-00356-7

ARTICLE

AI and Strategic Decisions: Facing the Incompleteness

Michele Mozzarelli[1] · **Giulia Schneider**[2]

Accepted: 12 November 2025

Abstract

This study examines the role of artificial intelligence (AI) systems—and the related legal requirements—in strategic corporate governance decisions, characterised by a high degree of incompleteness and, consequently, discretion. It illustrates the inadequacy of current European regulatory developments in the information technology field, primarily based on a product liability approach, to govern the complexities underlying the introduction of AI tools in the strategic choices made by directors. Given the core elements of corporate governance and its specific accountability needs, the study demonstrates how algorithmic decision-making cannot be an exclusive strategy for strategic decisions characterised by unpredictability and incompleteness. Therefore, the ultimate objective of the analysis is to underscore the persisting gap in legislation regarding the need to shift from a product-based and risk-minimisation approach in AI-driven decision-making to an integrated approach wherein the decision-making process is shaped by the complex, yet balanced, interaction between two distinct decision-making components, namely humans and the algorithm. To this end, the paper advocates for a stronger emphasis on human enhancement, algorithm ergonomics, and a legal strategy-driven design of AI.

Keywords AI · Corporate governance · Strategic decisions · Human in the loop

✉ Michele Mozzarelli
michele.mozzarelli@unicatt.it

✉ Giulia Schneider
giulia.schneider@santannapisa.it

[1] Full Professor of Business Law, School of Banking, Finance and Insurance, Università Cattolica del Sacro Cuore, Milan, Italy

[2] Assistant Professor of Business Law, DIRPOLIS Institute, Sant'Anna School of Advanced Studies, Pisa, Italy

Published online: 12 January 2026

1 Introduction

Is it possible to entrust an algorithm with defining a company's industrial policy, making a significant decision such as a merger, or identifying members for an optimal composition of the board of directors? The answer to this question appears increasingly disconnected from the technological sphere, especially since the advent of artificial intelligence (AI), whose impact is now pervasive in commercial activities and beyond. The distinctive features of AI systems lie in their ability to achieve a given set of objectives autonomously, thereby generating predictions, recommendations, and decisions. Artificial intelligence's significant revolution is precisely related to its expanding scope of application, which increasingly covers the entire decision-making process, thereby leading to optimisation and efficiency gains.

The source of concern when entrusting a significant decision to an algorithm is not the algorithm's technical ability to make decisions, but rather the legal ramifications associated with such an ability. As is widely recognised, optimising algorithmic decision-making comes with inherent risks.

European regulators have variously addressed the risks stemming from the adoption of new technologies by businesses. Regulations such as the General Data Protection Regulation (GDPR), the Artificial Intelligence Act (AI Act), the Digital Services Act (DSA), the Digital Markets Act (DMA), and, lately, the Digital Operational Resilience Act (DORA) for the financial sector have introduced different tools to control some of the risks associated with these new technologies. This continuously evolving legislative framework within the EU aims to guide the efficiency-driven advancement of digital transformation in both the private and public sectors along the lines of consolidating legitimacy paradigms.

Specifically, our understanding is that the chosen approach to maintaining an anthropocentric conception of ongoing digital transformation focuses, at the procedural level, on the 'human in the loop'. As discussed later, AI—broadly understood as automated systems—is treated by the regulations above as a product, the safety of which, in terms of preventing harm to third parties, must be ensured through human monitoring. Little attention is paid to the whole decision-making process, including the human decision-making component and its effective interaction with the AI systems employed.

Against this backdrop, this study examines the role of AI systems—and the related legal requirements—in strategic corporate governance decisions, characterised by a high degree of incompleteness and, consequently, discretion. It illustrates the inadequacy of current European regulatory developments in the IT field, primarily based on a product liability approach, to govern the complexities underlying the introduction of AI tools in the strategic choices made by directors.

Given the core elements of corporate governance and its specific accountability needs, the study demonstrates how algorithmic decision-making cannot be an exclusive strategy for strategic decisions characterised by unpredictability.

Therefore, the ultimate objective of the analysis is to underscore the persisting gap in legislation regarding the need to shift from a product-based and risk-minimisation approach in AI-driven decision-making to an integrated approach wherein

the decision-making process is shaped by the complex, yet balanced, interaction between two distinct decision-making components, namely humans and the algorithm. This presents an ambitious, albeit necessary, challenge.

To this end, the paper is structured as follows: Sect. 2 outlines the regulatory perspective on AI accountability in accordance with the 'human in the loop' and thus a product-based approach.

Section 3 contextualises the presently dominant approach of a 'human in the loop' as a risk-mitigating factor against potential harms arising from the AI models deployed in the corporate technology (CorpTech) domain. It illustrates how the polarisation characterising corporate scholarship around CorpTech solutions can be viewed as a direct by-product of the endeavour to emphasise the distinctions between corporate mandated procedural (protocol-based) decisions and discretionary decisions. Then, in Sect. 4, we contend that, in strategic corporate decisions, the 'human in the loop' approach is inadequate due to the combination of the incompatibility of AI and corporate governance legal strategies and the risk of overloading the human controller. In the final Sect. 5, we propose a change of perspective in AI regulation aimed at keeping humans in command and focusing on the interaction between humans and machines rather than solely on the algorithms. To this end, we advocate for a stronger emphasis on human enhancement, algorithm ergonomics, and a legal strategy-driven design of AI. Conclusions follow.

2 AI as a Product

2.1 The State of the Art of EU Digital Regulations

The employment of AI tools for decision-making purposes has raised some non-neglectable practical problems directly stemming from the specific technological features of AI systems. The most common and widely documented concerns relate to the perpetuation of various types of biases (historical, social and statistical) and the structural obscurity of AI models.[1] These features have resulted in explainability and understandability gaps in AI-driven decisions, which have quickly climbed to the forefront of the European regulatory agenda. Indeed, the focus of emerging IT regulations has been directed mainly at achieving transparency of the automated technologies used for decision-making, as a means to control the threats to the protection of fundamental rights in accordance with a risk-based approach.

Overall, the European IT framework is meant to control the risks directly arising from the impact of the employment of technological tools regarding: the protection of personal fundamental rights and the adverse effects on end-users as the

[1] Zuboff (2019).

addressees of automated decision-making processes; the businesses making use of these technological tools, which amplify in particular their legal and operational risks; and, finally, the broader economic system, in terms of market integrity and stability.[2]

This risk-based approach targeting the protection of the personal fundamental rights enshrined in the Charter of Fundamental Rights of the European Union, is one of the distinctive features of the emerging European 'digital constitutionalism',[3] as opposed to the American and Chinese models of technology regulation.[4]

Although with substantial differences, the IT regulations issued so far at the European level share the common target of focusing on the structural features of technologies enabling the pursuit of accountability objectives.[5]

At the general level, the accountability of technological means relates to the ability to establish whether a technological system works in conformity with substantive and procedural standards. Where these standards are not accomplished, the controller's liability arises as a result of the failure to control the various risks stemming from the departure from applicable standards.[6] One of the primary means used to hold an automated system accountable is the establishment of transparency requirements.

The first steps regarding automated decision-making accountability were taken in the General Data Protection Regulation (GDPR)[7] which, in its shift from a data protection to a data management regulatory paradigm,[8] obliged businesses processing personal data to ensure that the processing operations they perform are externally traceable and justifiable,[9] through the production of 'meaningful information'[10] for the ultimate purpose of protecting data subjects' (personal) fundamental right to data protection.

Very similarly to the GDPR, the AI Act[11] defines AI transparency from a functional standpoint and links this structural feature to the achievement of 'compliance with the relevant obligations of the deployer and of the provider' set out in the Regulation for those systems posing a high risk to personal fundamental rights.[12]

[2] With regard to the consideration of economic interests affected by automated tools, see Regulatory Obstacles to Financial Innovation Experts Group—ROFIEG (2019).

[3] De Gregorio and Dunn (2022), pp 479–483.

[4] Bradford (2023).

[5] Carnat (2023), pp 104–110.

[6] Schneider (2022a), p 7 ff.

[7] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[8] For a recent analysis see Balboni and Francis (2023), in particular pp 22–47.

[9] See principles of transparency and accountability under Art. 5 GDPR.

[10] Arts. 13-15 GDPR.

[11] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (hereinafter AI Act), 12 July 2024.

[12] See Art. 13 AI Act.

In the (difficult) effort to regulate the highly changing landscape of AI systems, the Regulation has opted for establishing only programmatic objectives for those systems that fall under the category of high-risk systems.[13]

Interestingly, in accordance with the AI Act's objective of protecting users' personal fundamental rights, the systems listed as high risk under Annex III of the Regulation do not encompass many applications of AI in the business sector.

2.2 The EU Regulatory Technique: Human Subjects as Monitoring Gatekeepers

From a functional standpoint, the EU IT transparency-based framework requires that the systems employed for autonomous decision-making have some structural features enabling the activation of human-machine interaction. Enacted rules regarding the IT systems indeed set the first limits on ongoing processes of dehumanisation of decision-making triggered by optimisation promises. These limits are established in accordance with an anthropocentric conception, which is based, at the procedural/methodological level, on the 'human in the loop' command. The resulting regulatory paradigm focuses on the functional control relationship between humans and machines: with regard to the automated decision-making loop, human subjects are in charge of controlling and managing the risks associated with the technological infrastructure.

This standpoint is clearly expressed by Recital 1 of the AI Act, which recalls a paradigm of human-centric and trustworthy AI. Recital 6 of the AI Act confirms this stance, highlighting that 'as a pre-requisite, artificial intelligence should be a human-centric technology', which should 'serve as a tool for people, with the ultimate aim of increasing human well-being'. The human-centric approach to AI is thus declared to be at the very core of the European legislative intervention in the field of AI. Nonetheless, this European manifesto is only partly addressed in the provisions laid down in the AI Act. The core problem of keeping AI-driven decisions under the control of human subjects remains largely unanswered, as if it were taken for granted by European regulators.

As shaped by legislative requirements, indeed, the role of human subjects in an AI-driven decision-making process is that of supervising and, in accordance with the risk-based approach adopted by both the GDPR and the AI Act, minimising the possible risks attached to entirely technologically driven outputs.

Ultimately, both transparency and thus accountability requirements are meant to ensure human scrutiny over the automated decision-making loop so as to preserve the fairness of the processing, mainly ensured by the absence of biases affecting the automated decision-making loop to the detriment of vulnerable parties.[14]

An example of a technology-related accountability requirement under the GDPR is found in the supervisory obligations attributed to data controllers and processors regarding their means of processing. For example, for businesses processing personal data, Article 35 of the GDPR imposes the burden of measuring the threats to

[13] See Annex III AI Act.

[14] Schneider (2022b), pp 168 ff., 171–173.

data subjects' fundamental rights, arising from the processing activities carried out, in the form of a Data Protection Impact Assessment (DPIA).[15]

With specific regard to automated decision-making processes, Article 22 of the GDPR establishes a right not to be subject 'to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her'. The provision has been recently interpreted by the Court of Justice of the European Union (CJEU) in the *Schufa* judgment as an outright prohibition on fully automated decisions, echoing an interpretation found in a strand of the literature.

This prohibition should first be interpreted as banning automated decisions unless human monitoring of the system's features, potential biases, and ultimate risks is properly conducted. Along these lines, Article 22 of the GDPR lays down a user's 'right to intervention', that is, a right to challenge machine-generated outputs, which is strictly related to the right to object to algorithmic determinations.

Along similar regulatory lines, the latest version of the AI Act requires high-risk systems to be subject to human oversight. This supervisory function, however, as revealed by the analysis of the same provision under Article 14 of the AI Act, is merely monitoring support for the output of the automated decision and not much of the algorithmic process leading to the decision.

Indeed, despite the first paragraph demanding high-risk AI systems to be designed and developed in such a way, including with appropriate human-machine interface tools, 'that they can be effectively overseen by natural persons during the period in which the AI system is in use', the second paragraph acknowledges that this human oversight shall be ensured so as to prevent or minimise the

> risks to health, safety or fundamental rights that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse, in particular where such risks persist despite the application of other requirements set out in this section.

As this wording suggests, human oversight shall be functional just to detect potential risks to fundamental rights, but it is not meant to touch on or interfere with a decision-making process that could also be of a fully automated nature.[16]

If it does not regard the algorithmic process directly, human oversight cannot but refer to the output of the (also fully automated) decision, as confirmed by Article 14(4) of the AI Act, requiring the human overseer to 'correctly interpret the high-risk system's output' (lett. c), as well as 'to decide, in any particular situation, not to use the high-risk AI system or otherwise disregard, override or reverse the output of the high-risk AI system' (lett. d).

In accordance with the oversight functions under Article 14 of the AI Act, the human subject can either decide to 'intervene in the operation of the high-risk AI

[15] Kaminski and Malgieri (2021), pp 129–131.

[16] This perspective appears to be further confirmed by the erasure in the latest circulated version of the AI Act of the overarching principle of non-exclusivity of AI decisions under a previously proposed Art. 4a AI Act, the introduction of which had been suggested by the European Parliament as a direct expression of the human-centric approach to AI articulated in Recitals 1 and 4 AI Act.

system', mainly by interrupting the system 'through a "stop" button or a similar procedure' (lett. e)—thus, refusing to use the output of the automated system—or internalise this output in the final decision outcome with a mere warning to be aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system ('automation bias') (lett. b).

The regulatory perspective is thus one of an AI-driven decision-making 'loop' in which the human component serves as an external watchdog over the technological decision-making outcome and over the associated risks to health, safety or fundamental rights recalled under Article 14(2) of the AI Act (the 'human in the loop' approach). In other words, the cited Regulations treat AI as a product, the safety of which—in terms of preventing harm to third parties—must be ensured by human monitoring and evaluation. Nothing is said about the decision-making process considered as a whole, also possibly including the human decision-making component.

It is no coincidence that recent legal literature[17] has identified various critical issues concerning the human oversight remedy at the heart of the European Regulation on artificial intelligence. Among these, it should first be noted that the monitoring duties envisaged by the European legislator are mainly enabled by the initial design options of the machine as determined by the manufacturer: the business user is indeed forced to circumscribe its monitoring capabilities within the instructions given by the manufacturer.

The problem of the human chain of responsibility, from the manufacturer to the user of the system, has been commented on in various studies in the field of tort law.[18] The complexity of the human network—and thus of responsibility—has been posited by some as the basis of a necessarily complex liability system that must accompany the life cycle of an automated system. However, as suggested by the proposed framework on AI liability, lastly proposed by the European Commission in September 2022, and given a proposal for a revision of the Directive on product liability and a proposal for a directive on AI liability,[19] the complexity of a possible liability system regarding AI risks has been simplified by positioning the ultimate liability burden on the provider of a given AI system.

Contrary to the stance initially proposed by the European Parliament,[20] the proposed revision of the Product Liability Directive[21] tackles only the liability of manufacturers, governed by a strict liability regime. The proposed AI Liability Directive[22] also only focuses narrowly on the liability of manufacturers and users in the event of damage arising from high-risk systems under the AI Act, without addressing the more complex issue of the user's (in our case, 'a business') extra-contractual

[17] See, in particular, Crootof et al. (2023), pp 503–505.
[18] Comandé (2019), pp 179–180.
[19] European Commission (2021b).
[20] European Parliament (2020a, b).
[21] European Commission (2022a).
[22] European Commission (2022b). See further European Parliament (2023).

(fault-based) liability for low-risk AI systems under the AI-related regulatory framework. Hence, the proposed liability regimes for harm arising from AI systems confirm the product-centred viewpoint taken by the European regulator in matters of AI. They indeed ensure the coverage of torts arising from AI systems by identifying the manufacturers as the main source of liability and the primary point of deterrence.[23]

2.3 The Latest Generation of 'Human in the Loop' Rules: DORA

In line with the approach identified above, financial sector regulations have also adopted 'human in the loop-styled' requirements. This is well reflected in the Regulation on the digital operational resilience of the financial sector (DORA),[24] which builds on the technological resilience requirements first proposed in the MiFID II framework[25] for the protection of the fundamental economic interests of financial stability and market integrity.[26]

Referring to the governance of ICT risks in the financial sector, DORA appears to be the latest generation of 'human in the loop' regulation, laying down external limits on the employment of new technologies, if not properly supervised through a complex (human) organisation.

Similarly to what occurs under the general IT regulations cited above, DORA, too, remains loyal to a product-based approach to technologies: it indeed calls for the adaptation of financial entities' corporate governance structures to the early prevention, detection and handling of ICT risks. The risks targeted by this piece of regulation are those potentially impacting financial institutions' digital operational resilience, which Article 3(1) of DORA defines as the

> ability of a financial entity to build, assure and review its operational integrity and reliability by ensuring, either directly or indirectly through the use of services provided by ICT third-party service providers, the full range of ICT-related capabilities needed to address the security of the network and informa-

[23] Wagner (2023), p 191 ff., and in particular pp 196–197. As has been suggested by this literature, the option of leaving the liability of users of AI unaddressed (at least through specific tort rules) reflects the European regulators' intention not to excessively freeze the market of AI systems and thus their use by businesses. However, the implicit appeal to the more general technology-neutral liability rules for fault in regulating users' liability raises a series of unanswered questions regarding how to tailor the general precepts of extra-contractual liability with regard to users' monitoring duties required by the AI Act: when is a monitoring activity conducted negligently? What is the monitoring standard that the user has to meet in order to demonstrate their diligence? How should this liability be matched with the manufacturers' liability? The 'human in the loop' paradigm affirmed by the AI Act triggers these queries, and no response is to be found in the neighbouring emerging AI liability framework.

[24] Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector and amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014, (EU) No 909/2014 and (EU) 2016/1011 (hereinafter DORA).

[25] See Art. 47 MiFID II (Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU).

[26] European Commission (2020).

tion systems which a financial entity uses, and which support the continued provision of financial services and their quality, including throughout disruptions.

Moving beyond a mere fundamental rights-based conception of IT accountability, DORA thus targets the control of the economic interests potentially affected by the digitalisation of financial business structures.[27]

With regard to these ICT risks, human subjects, and in particular directors, serve as a liability site, so as to remedy any eventual, legally relevant automation-driven 'adverse effects'. From this perspective, the turning point of DORA, compared to the GDPR as well as the AI Act, lies in the fact that the supervision of technologies becomes a corporate governance matter.[28]

To these ends, assuming that, in line with strict ICT regulations, technologies have to be designed in a manner that enables human supervision, DORA specifies *how* technologies in the financial sector have to be supervised.

By laying down specific procedural and organisational requirements regarding how financial entities are to manage ICT tools, the Regulation determines the management body's 'ultimate responsibility' for the prevention and mitigation of ICT risk.[29] The management body, in the continuous interaction with ICT and the risk management function,[30] is required to take a 'pivotal and active role in steering and adapting the ICT risk management framework and the overall digital operational resilience strategy'.[31] In this way, it expects directors of financial entities to take a result-oriented and functional approach that involves close monitoring of employed technologies and that needs to be sensitive to the mitigation of ICT risks beyond the dictates of specific ICT requirements.[32]

In contrast to the GDPR and the AI Act, DORA places a greater focus on the human behaviour component vis-à-vis technological tools through the lenses of corporate governance structures. However, the methodological approach to the regulation of (financial) businesses' digital assets—whether personal data (under the GDPR), AI systems (under the AI Act) or ICT infrastructure (as addressed by DORA)—remains, also under DORA, that of entrusting the oversight and control of the functioning, as well as the final output of the employed technologies to human subjects.

[27] ICT risks are largely defined as 'any reasonably identifiable circumstance in relation to the use of network and information systems which, if materialised, may compromise the security of the network and information systems, of any technology dependent tool or process, of operations and processes, or of the provision of services by producing adverse effects in the digital or physical environment', Art. 3(5) DORA.

[28] Suggesting the opportunity for an entity-based 'human in the loop approach' for the supervision of AI in the financial sector, Buckley et al. (2021), p 66 ff.

[29] Recital 45 DORA.

[30] Art. 5 DORA.

[31] Recital 45 DORA.

[32] Advocating for a functional role of boards' oversight of businesses' compliance choices, Armour et al. (2020), pp 31–39.

Under these premises, it can be concluded that the objective pursued by contemporary IT regulations is the understanding and control by humans of the risks to either personal or economic interests arising from the employment of automated systems. With regard to these risks, humans handling automated tools have the functions of preventing harm and of monitoring the expected risks through the available regulatory tools (in terms of transparency, explainability and auditing), so as to pull the emergency brake when needed and either remove or substitute a faulty device.

However, the monitoring role attributed to the human 'controller' and the remedies it provides to them implicitly suggest that the legislator admits the possibility of an entirely automated decision-making process, provided that the human actor oversees this same process. In other words, the regulations issued so far appear to focus on use cases where the option of a 100% algorithmic decision-making process may be permitted, on the condition that a 'human in the loop' defence is in effect. From this specific standpoint, it seems that the considered regulatory framework implicitly legitimises AI's replacement of human decision-makers in technology-driven processes.

While the human monitoring paradigm inflecting European IT regulations is surely to be welcomed as a first direct response by the European regulator to curb fast-changing and highly intensive innovation scenarios, it offers just 'one view of the cathedral'[33] regarding the implication of employing AI in the decision-making process. A slightly bigger piece of our AI-related regulatory cathedral can be glimpsed if the described approach is tested against the backdrop of the distinction between organisational or routine business decisions and strategic governance decisions in the corporate realm.

Indeed, with regard to the varied landscape of businesses' decisions, the specific use cases to which each of the aforementioned IT regulations refer appear to relate to low-discretion administrative structures of contemporary corporations.

Certainly, data processing activities addressed by the GDPR or the ICT infrastructure under DORA may also possibly be employed in the context of businesses' strategic decisions. However, in light of the above analysis, it can be said that the European legislator's intent is that of addressing the assessed 'human in the loop-styled' requirements mainly for the purposes of enacting external control checkpoints in businesses' organisational routine decisions.

None of the available regulatory tools in the traced framework tackle, to date, the specific risks related to the integration of technologies in the dynamics of corporate governance in the case of directors' strategic decisions,[34] such as those regarding fundamental changes.

Interestingly, the legal debate that has arisen around the phenomenon of the digitalisation of businesses, known as 'CorpTech', has failed to adequately consider the (potentially) diverse nature of businesses' automated decisions and the different

[33] The quote is taken from the famous article Calabresi and Melamed (1972), p 1089 ff.

[34] The distinction between organisational and administrative decisions is also mentioned by Petrin (2019), pp 966, 983. See also Kolbjørnsrud et al. (2016).

governance demands stemming from digitalised organisational and strategic corporate structures.

3 The CorpTech Dilemma

3.1 Setting the Focus: The CorpTech Dilemma

In parallel with the evolving regulatory baselines, there is the market reality of businesses boosting their employment of AI models within their organisational and managerial structures.

As increasingly acknowledged by empirical studies at both the European[35] and the domestic levels,[36] autonomous computational models are creeping into the most sensitive corporate decision-making areas. This is well illustrated by the uptake of legal tech tools, predicting, for example, the convenience and success rate of litigation,[37] as well as by the growing employment of AI systems for evaluating and forecasting scenarios in the context of mergers.[38] As is apparent, litigation or merger decisions are not routine decisions but rather strategic ones.

The application of technological tools to the corporate realm is generally referred to as the 'CorpTech' domain, so as to mirror the developments occurring in the 'FinTech' sector.

Precisely because of the growing application of AI-driven decisions in strategic corporate contexts, the literature has started to consider—also under the pressure of evolving practices, especially in Asian countries[39]—the extreme hypothesis of the substitution of directors by machines, in the form of a 'robo-director',[40] or even of an entirely automated board, as shown in the case of two different hypotheses regarding a 'robo-board'[41] and the Decentralised Autonomous Organisation (DAO).[42]

Against the backdrop of these possible—and in some cases actual—transformations of the board of directors, the doctrinal debate has evolved around the two major trajectories of enthusiasm on the one hand, and scepticism and rejection on the other.

A first strand of the literature[43] has warmly welcomed the integration of technology into corporate structures and has considered technologies as a panacea for the 'evil' of corporations' internal functioning, above all conflicts of (human) interests. It has been said that the new digital means can reduce the risks of conflicts of interest in corporate relationships, in particular between shareholders and directors or

[35] European Commission (2021a).
[36] See, for example, Assonime (2023), pp 60–69.
[37] Armour et al. (2022), pp 107–198.
[38] Lehot (2023).
[39] Wile (2014).
[40] Abriani and Schneider (2021), p 197 ff.
[41] Mosco (2020), pp 89 and 91–92.
[42] Borgogno (2022), especially pp 7–14.
[43] See, lately, Li (2024), pp 1320–1336, Yermack (2017), pp 17–20.

between different categories of shareholders, while enabling more direct interaction between the different parties and thus greater alignment between them. In accordance with these views, reducing agency problems would make redundant, at their very roots, some of the corporate tools that have been introduced as a remedy for the said agency problems, such as independent directors.

Following these lines of reasoning, both blockchain and artificial intelligence systems are said to have the potential to more efficiently link the different groups of interests orbiting around a corporation, minimising the risks of self-serving behaviour while reaching the optimal allocation of financial resources within the new corporate 'machine'.[44]

In line with these views, the advancements brought about by new technologies should be promoted just for their ability to mitigate those behavioural externalities of the traditional hierarchical corporate structures. Automated tools thus pave the way for horizontal (or platform-based) governance models, which are expected to be more efficient than the corporate governance solutions known to date.

This first approach of 'CorpTech-enthusiasts' thus supports radical evolutions of traditional corporate structures into forms that either strongly depart from the corporate anatomy, as seen with decentralised autonomous organisations (DAOs),[45] or that propose a maximum integration of technology in corporate dynamics, up to the point of admitting the presence of outright 'robo-directors'.[46]

A second strand of the literature accuses CorpTech proponents of being blinded by a 'tech nirvana fallacy'.[47] In the views of this strand, technologies are merely new instruments through which traditional corporate relationships may unfold: technology is not going to change anything in the corporate structure, not even the common conflicts of interest situations among parties. Rather, the contrary is more probable, namely that traditional agency problems are propagated, and in some cases even worsened, through algorithmic codes.

Technology 'sceptics' are thus suspicious (to say the least) of the employment of technology—in its fully autonomous forms—in the internal corporate governance sphere, given the various technology-related risks that are often not fully tackled by the more slowly evolving regulatory framework. The costs of digitalisation are considered to be higher than its expected benefits, thus offering a ground for discouraging corporations from light-heartedly embracing at least the most AI-intensive solutions, leading to the replacement of humans by algorithms.

As summarised, the current debate on CorpTech tends to be polarised around two different reactions to the digital evolution of corporate governance: a welcoming and a more suspicious one. Upon closer consideration, the current debate analyses the possibilities of integrating technologies in the corporate governance domain, in light

[44] Erel et al. (2021), p 3227, Hamdani et al. (2018), pp 223–233.

[45] DAOs are not corporations; they are new organisational models for conducting business that are different from corporations. Borgogno and Martino (2024), especially pp 24–32.

[46] This happened in the case of the 'Vital' algorithm. This solution would not be possible in the European Union, where AI systems lack legal personality. See Teubner (2018), pp 113–115.

[47] Enriques and Zetzsche (2020), pp 71–74.

of the traditional corporate governance problems (agency problems) and remedies (the tools restoring these same problems).

Although insightful, the level of the debate so far appears to remain confined to overly general terms. It indeed mostly seems to consider AI as an alter ego of a human being, adding fuel to a subsequent polarisation: to put it in extremely radical terms, either AI is to be largely admitted as a player in corporate structures, or, although it may bring about efficiencies in corporate decisions, it will not, fundamentally, change anything with regard to traditional corporate governance mechanisms.

In the end, this framing of the relationship between AI and corporate structures does not seem to provide concrete answers regarding the degree to which a corporation's internal digital transformation should be allowed under the present IT and corporate governance framework.

From this more specific standpoint, the major flaws in the present CorpTech debate relate to the fact that it assumes that CorpTech is a unique technological product and does not generally delve into the heterogeneity of CorpTech tools, which not only arises from the variety of possibly applicable technologies, but—what is relevant for the purposes of our study—also stems from the different roles CorpTech tools can have at the various levels of the corporate governance structure.

In line with the conception of corporate governance as an organisational underpinning for more efficient decision-making processes in corporate matters,[48] these different levels are to be, first of all, identified in light of the different types of decisions that the board of directors takes in a corporation. Hence, the role of CorpTech tools must be assessed based on the *type* of corporate decisions for which they are employed.

3.2 Corporate Decisions Between Procedures and Discretion

In our view, AI tools should be assigned a more specific—and more realistic—role, given the diversity of corporate decisions in which AI can be employed.

We propose a different framework, starting from the variables of completeness and discretion involved in a given decision-making process. For the sake of clarity, we can consider two opposing cases.

A first hypothesis relates to a constrained decision, as occurs in the case of the already mentioned high-frequency trading system. In this case, the machine applies a protocol (that is, a set of criteria decided *ex ante*), which provides a precise formalisation of the decision-making process and which is thus validated by established rules. Precisely because a protocol exists, these decisions are bounded in their patterns and outcomes, because the protocols are normally based on the same (quantitative) variables.

This is why it can be argued that decisions regarding businesses' organisational structures are performed in a 'complete' environment: indeed, the elements upon

[48] Armour et al. (2017a), pp 22–24.

which the decision is based are known and only have to be assembled so as to identify the decision's outcome. For these decisions, automated systems appear to perform better in the execution of the protocol. Minor discretion is left to the machine because a (relatively low) discretion has already been exercised by the human in the definition of the protocol/procedure. If the score is above a certain threshold, then an action is performed. These decisions are recurring and continuous; indeed, they concern the processing of multiple data in accordance with a pre-determined protocol, which is sufficient to fully substantiate the decision-making outcome. This is what occurs, for example, in the case of high-frequency trading, the processing of information for anti-money laundering (AML) enquiries, or the distribution and surveillance of employees.

These examples relate to limited-scope decisions, which we may call organisationally 'routine' or 'complete' decisions, defined elsewhere as high-frequency, low-impact decisions.[49] They involve little discretion because the outcome is bound by a given protocol. In these cases, the AI device only substitutes the human decision-maker at the secondary stage of executing a protocol—a task that the device can do faster, at a cheaper cost and with greater mathematical precision.

This type of recurring decision can be taken in a nearly fully automated way,[50] provided the human factor is capable of intervening as a supervisor or controller of the enacted system, mainly in terms of monitoring and reviewing the system's output, just as the legislative requirements demand.

In these cases, accountability for the decision-making processes can—or, better said, should—be reached through the product-based approach underlying the 'human in the loop' approach: the decision loop here is mainly determined by the machine's output, and the role of the human decision-maker is to ascertain whether the machine-driven outputs are suitable or fair (e.g., free of biases) in relation to the framing developed *ex ante* in the protocol. Ultimately, with regard to routine and complete decisions, the 'human in the loop' paradigm may be sufficient to guarantee the legitimacy of the decision; given the protocol, it is easy to check whether the machine has applied it correctly. The human subject is thus capable of exercising oversight functions on the model, while the decision-making loop is maintained for efficiency reasons.

From the perspective of corporate structures, it can thus be said that, with regard to routine decisions, directors may employ fully automated decision-making processes and thus rely on fully automated outputs, provided that corporate structures enable, in addition to an efficient automated process, a risk-based monitoring function.

In this respect, however, it must be admitted that the oversight function is not always effective with regard to the risk-minimisation objectives in the context of organisational decisions; as some recent empirical studies have demonstrated, the human monitoring function is better suited to detecting smaller, machine-driven

[49] Camuffo et al. (2023), pp 127 ff., 131–133. See also Petrin (2019), pp 966 ff., 983, Kolbjørnsrud et al. (2016).

[50] Camuffo et al. (2023), pp 131–133.

mistakes than to fixing larger and more severe ones, thus rendering the human oversight function not always 'secure', including with regard to highly repetitive, and thus, predictable decisions.[51]

3.3 The Case of Strategic Decisions in Incomplete Scenarios

At the other end of the spectrum, we find strategic decisions, such as those involving a corporation's fundamental changes (e.g., a decision regarding a merger with another company or the appointment of a board member). Such decisions are made in an incomplete environment: the protocol is not precise, by hypothesis.

From the decision-maker's standpoint, this type of decision involves a number of potentially infinite framing options, which, respectively, might lead to different outcomes.[52] Incompleteness makes the formalisation of the decision at stake difficult and, consequently constructs a one-sided protocol.[53]

Strategic decisions can indeed be framed differently and, as a result, may lead to different outcomes. It all depends on the functional standpoint and thus the sensitivity with which they are taken. In these decisions, there is no right mathematical answer, but the decision output will depend on how different qualitative factors are evaluated and 'assembled'.

Against this backdrop, for the purposes of this analysis, a strategic decision assumes incompleteness, and such incompleteness underpins the discretionary nature of a decision-making process.

In the field of corporate action, incompleteness appears to be, first of all, directly related to the complexity of the interests at stake in corporate governance; the so-called compositive nature of corporate governance is becoming increasingly relevant in times when corporate efficiency goals are the subject of a heated debate and of reform projects requiring boards to take into consideration different and 'diverse' elements for the purposes of their decision-making.

The incompleteness of the setting in which strategic corporate decisions are made is exacerbated by the unpredictability and uncertainty of business relations, which can be subject to sudden health or geopolitical crises, thus requiring an additional effort in corporate-governance flexibility.

As a result of the above, strategic decisions structurally involve an area of discretion. The degree of discretion involved in a decision can clearly vary, along the lines of a continuum in which low-discretion decisions are more similar to routine tasks and high-discretion decisions correspond to truly strategic decisions. The greater the degree of discretion, the more difficult it will be to properly allocate decision tasks between the human decision-maker and the machine. In the absence of a defined protocol, the machine evidently cannot perform the execution task outlined in the previous paragraph.

[51] Sele and Chugunova (2022), p 14.
[52] Alvarez and Porac (2020), p 735 ff.
[53] Erikson and Knockaert (2021), p 153 ff.

Nonetheless, also in these decisions, AI tools may facilitate the development of an efficient decision-making process. This can occur by delegating only specific sub-tasks of the overall decision-making process to the machine.

Ideally, algorithms can also frame a protocol for strategic decisions, as in the case of an M&A plan defined through automated means.

The developments in the field of machine learning clearly show the creative potential of the latest-generation AI models to create a decision-making process from the start of a given problem. The range of problems addressed by machine learning typically involves tasks for which other approaches fail, either due to a lack of a suitable formalisation of the problem, or because the resolution of the problem is intractable using non-learning approaches.

Indeed, machine learning typically refers to algorithms that learn to complete tasks by identifying statistical patterns in data, rather than by following instructions provided by humans. Machine-learning approaches focus on the development of systems capable of learning and inferring from data, to solve an application problem without being explicitly programmed with a set of step-by-step instructions from input to output. Learning refers to the computational process of optimising the parameters of the model, which is a mathematical construct generating an output based on input data.[32]

However, although feasible from a technical standpoint, delegating a strategic decision to an AI system needs to be better assessed from a legal standpoint in light of the very features of such a decision, namely incompleteness and discretion. The challenge is to justify the employment of AI in strategic decisions not only based on an efficiency rationale but also by identifying its legitimacy in the specific context of corporate decisions; this justification must be considered *in addition* to basic human monitoring requirements, but also *provided that* those requirements are fulfilled.

From this different perspective, when it comes to strategic corporate decisions, the factual feature of incompleteness appears to challenge the effectiveness of the current regulatory approach based on a human monitoring function, such as that outlined in the AI Act.

The role of the human decision-maker as a mere monitoring principal of the machine-agent appears to be highly risky in the context of strategic decisions. Precisely because these decisions are taken in incomplete environments, they may be conducted without any direct prejudice to a legally relevant interest, such as a fundamental right, but they could still be erroneous in the sense of not indicating the 'better' decisional outcome for the corporation that addresses it.

Last but not least, the monitoring paradigm informing current IT-related regulations is not capable of ensuring the accountability of a human actor's strategic decisions conducted with the support of AI. In these cases, there is, indeed, not only the AI monitoring problem (an issue with which the law struggles), but also the problem of the accountability regarding how AI is used in the decision-making process (a problem that the law has not yet addressed).

Under these premises, the following paragraphs will examine the *internal* limits that corporate governance imposes on the integration of its traditional structures—and particularly management—with digital components, specifically concerning strategic decisions. Consideration will ultimately be given to the impact

on the directors' duties when AI tools are employed for the purposes of strategic decision-making.

Hence, taking into consideration a strategic corporate decision, characterised by its high impact and its large degree of incompleteness, the following analysis will be divided into two major parts: first, consideration will be given to a strategic corporate decision based solely on AI and its decrease in accuracy; in the second part, the human monitoring capabilities regarding this type of decision will be discussed.

4 Strategic Decisions Based Solely on AI

4.1 Strategic Decisions in Incomplete Environments and AI: Losing Accuracy

As mentioned earlier, the integration of AI into decision-making processes is primarily driven by efficiency considerations.[54]

In routine decisions, the accuracy of the AI tool seems an implicit but powerful source of its legitimacy. The algorithm serves as a self-sufficient and potentially exclusive decision-making strategy for routine corporate decisions. Under the 'human in the loop' paradigm, human involvement is limited to overseeing the AI system in use. Intervention is necessary only in emergencies to protect fundamental rights as defined by applicable regulations.

The implicit assumption behind these assertions is that humans may refrain from intervening in the algorithmic decision-making process because, as long as the results are highly reliable, the high costs associated with ongoing human intervention may not be justified. If the process produces accurate results, there may be no need to structure monitoring instruments. Instead, one could rely on remedies designed to pre-emptively address any negative externalities related to health, safety, human rights and the environment rather than scrutinising the decision-making process. In essence, the machine is entrusted with finding the optimised and most efficient solution, while humans manage any side effects impacting critical areas.

However, this equilibrium is no longer applicable when dealing with strategic decisions, which are structurally incomplete and whose exact outcome is thus not entirely within the control of the decision-maker, whether human or algorithmic.[55] Unlike routine tasks, strategic decisions inherently entail a margin of error due to the incomplete nature of the available information. Even the most high-performing AI cannot comprehend everything, and the broader the incompleteness, the less accurate a decision based solely on AI may become.[56]

The issue extends beyond less accurate calculations. As noted, incompleteness not only results in an inevitable reduction in accuracy but also introduces the potential

[54] Malgieri and Pasquale (2022), p 12, Kroll et al. (2017), p 636, who state that 'the efficiency and accuracy of automated decision-making ensures that its domain will continue to expand'.

[55] See Enriques and Zetzsche (2020), p 27.

[56] See Matsumi and Solove (2025), p 20 ff.

misplacement of the core idea on which the use of the algorithm is based—identifying the right decision with the accurately optimised decision (Sect. 3.2). Incompleteness may signify not just a scarcity of observations within a given variable (such as the number of sick days for director X in the next three years) but also, and more fundamentally, the inadequate consideration of certain variables (e.g., the algorithm fails to properly weigh the ability of director X and director Y to work together, thereby neglecting or significantly underestimating that variable).[57] In the latter case, incompleteness means that an algorithmic optimisation, even if computationally exact, would not necessarily lead to the 'right' outcome.[58]

Therefore, in decisions taken in structurally incomplete settings, the genetically diminished statistical accuracy is compounded by the compression of the margins of 'legal' accuracy, which stems from an organised hierarchy and systematisation of legally relevant interests and regulatory objectives impacting a given decision-making framework. Consequently, the more a decision depends on an incomplete factual matrix, the more challenging it becomes to reconcile individual and collective interests with the 'system' (specifically, an algorithmic system) that is impossible to identify *ex ante* in a protocol.

Hence, a fundamental tension arises between statistical prediction and 'legal' accuracy in scenarios with incomplete risks, as is often the case in business activities.

This inevitable expansion of uncertainty renders the algorithmic result doubly unreliable, and effectively managing the incompleteness associated with algorithmic decision-making in a strategic context becomes crucial.

4.2 'A Soul to Damn and a Body to Kick': AI Decisions and the Legal Strategies

With its innate confrontation with an incomplete reality, the law has developed multiple tools and strategies to deal with this issue.

The challenges associated with managing strategic decision-making are particularly acute in the context of corporate regulation. The complexities of overseeing organisations entrusted with long-term productive activities necessitate that corporate law addresses the management of strategic decisions due to their inherent uncertainty.

A widely held and compelling interpretation of corporate law posits that it has two primary objectives: first, to establish the structure of the corporate form, and second, to minimise or eliminate conflicts of interest among various corporate stakeholders.[59]

[57] Ibid., p 1928.

[58] Ibid., p 1928: '[P]roblems emerge when too much quantitative data is relied upon to the exclusion of qualitative data. Not everything is readily quantifiable'.

[59] See, *ex multis*, Armour et al. (2017b), p 29 ff.

In an ideal world of perfect decision-making, market logic might achieve the second objective through an efficient negotiation between principals and agents.[60] However, in the real world, subject to market failures, developed corporate systems employ sophisticated strategies to allocate decision-making discretion. Dealing with future events, most of these legal strategies refrain from pursuing the exact outcome of a particular decision. Instead, they focus mainly on the decision-making process as a critical area to mitigate agency problems, enhancing the principal's ability to control or structure the agent's decisions (through agent constraint, incentive alignment, decision rights, etc.).[61]

Given these legal strategies and assuming their effectiveness, the question arises as to whether they can be applied to decisions based solely on AI to compensate for the mentioned loss of algorithmic accuracy.

An interesting vantage point from which to investigate this issue is provided by the Vital case, where—according to media reports—an algorithm was granted voting rights on the board of a Hong Kong venture capital firm when it came to deciding 'whether the firm invests in a specific company or not'.[62] Legally speaking, Vital was not technically a corporate director under Hong Kong's corporate law. It was, however, considered 'a member of [the] board with observer status' by the other human board members and, in that capacity, 'contributed' to approving some investment decisions.[63]

Despite the AI's involvement in decision-making, the complex duties and responsibilities associated with the office remained with the human directors.[64] The AI functioned as a 'narrow-gauge' director, while the human directors played a role akin to that of caretakers for the digital prodigy, being liable for its limits and shortfalls.

In the perspective considered here, of entrusting a strategic decision to AI, this example highlights two critical issues: AI is indeed (a) unable to be the legal recipient of either civil or criminal liability, and, particularly relevant in this context, (b) impervious to the set of duties typical of a director of a corporation.

It is well known that a machine cannot be held legally accountable or be subject to the same set of duties as a corporate director, and there is no easy solution. Simply changing the law to allow for the appointment of a self-learning algorithm as a director would not be enough, as the role of a director is rooted in the legal concept of accountability.[65]

[60] On the difference between efficient and desirable allocation of resources, see Armour et al. (2016), p 54 ff.

[61] Davies (2020), p 30.

[62] Wile (2014).

[63] Burridge (2017). In the legal literature on this topic, see, *ex multis*, Möslein (2018), p 649 ff, Gramitto Ricci (2019), p 871 f., Mosco (2020), p 89, Eroğlu and Karatepe Kaya (2022), p 541 f.

[64] Goyal (2014). On this topic, see also Mosco (2020), p 92, noting that 'not only is it currently undesirable for AI-based directors to sit in on the board, but times are also unripe for even the most preliminary discussion as to AI-related subjectivity and legal capacity. This, in turn, makes it impossible to conceive of AI as a truly independent entity—one that could be distinguished from human directors'.

[65] See Gramitto Ricci (2019), p 886.

In this sense, starting from the practical observation that AI lacks 'a soul to damn and a body to kick',[66] it has been correctly observed that 'governance relies on accountability, accountability presupposes a conscience, a conscience might presuppose consciousness'.[67] Although it cannot be ruled out that one day AI may become advanced enough to develop or at least imitate such human characteristics, this is not yet the case.[68]

Since 'accountability requires more than legal capacity',[69] simply allowing today's AI into the boardroom would ultimately produce a shift of liability onto humans, as seems to have happened in the case of Vital.[70]

However, it is unnecessary to go that far (i.e., to search for AI's consciousness) to highlight the tensions between current corporate regulation and AI decisions.

The passage regarding the inability of an algorithm to fulfil all the duties imposed by the legal system on every director is also, if not more, significant.

The issue concerns not only the constraints imposed on directors but, more broadly, many of the legal strategies used by the most sophisticated legal systems to manage agency problems and, more generally, market failures. When the point of application of current corporate legal strategies shifts from the human decision-maker to the algorithmic one, the effectiveness of these strategies cannot be taken for granted.

The main corporate law strategies are structurally based on the specific characteristics of human beings, which cannot consistently be replicated or even exhausted in an automatic optimisation process. This does not mean appealing to the irrational or instinctual side of humans, simply because it remains beyond the reach of machines. To do so, it would first be necessary to demonstrate that human irrationality solves more problems than it creates.[71] Instead, the current corporate law system has built legal strategies to exploit dynamics that are (not good or bad in themselves but) quintessentially human. In this sense, these strategies are designed to be structurally immanent and functional to the human nature of their recipients. Therefore, their immediate applicability to algorithms appears to be anything but obvious, given the radical difference in decision-making processes between algorithms and humans.[72]

[66] The original quotation from Lord Thurlow was about corporations: 'Corporations have neither bodies to be punished, nor souls to be condemned, they therefore do as they like'.

[67] Gramitto Ricci (2019), p 894.

[68] Ibid., pp 894, 906, also noting that such a situation, once occurred, would likely pose a series of more existential problems than those—albeit important—encountered in regulating a board.

[69] Ibid.

[70] Correctly, Gramitto Ricci (2019), p 906, noted that 'proposals that emphasized the role of insurance in order to repair damages caused by artificial agents in boardrooms exclusively consider *ex-post* remedies that aim to repair already caused damages. Such proposals would fail to address or enhance accountability itself'.

[71] After the global endorsement of behavioural economics and its conclusions, it seems increasingly difficult to imagine that we can rely directly on our own biases. See, *ex multis*, Gilovich et al. (2002), Kahneman (2011), Sunstein (2011), p 1349 ff.

[72] See Solove and Matsumi (2024), p 1927, according to whom 'machine decisions are fundamentally different from human ones. Comparing human to machine decision-making is akin to comparing apples and oranges, not rotten apples to fresh ones'.

Two examples can be helpful to highlight this difference concerning the board of directors.

4.2.1 Human Versatility and the Limited Operational Scope of AI

Human directors are inherently versatile subjects, capable of interacting with and reasoning about problems with different characteristics. Therefore, they can apply their sensitivity to a wide range of decision-making processes. While hyper-specialisation appears to be a growing trend in the selection of directors, versatility is increasingly seen as a distinctive strength in the board room.[73]

On the other hand, well-known operational examples, such as the aforementioned Vital case, the DAO experiment, and algorithms used in mergers and acquisitions,[74] as well as legal scholars converge on the narrow 'operational domain' of the algorithm.[75] It is worth noting that this assertion is not contradicted by Large Language Models (LLMs) today, whose progress in the language domain has attracted public attention in the past year. Still, LLM-generated content is not as reliable as that generated by humans.[76]

Even when considering cases of autonomous governance intelligence, the current conclusion seems to be that 'the goals of the AI system and the company are restricted to a narrow operational domain, in addition to being closely linked and intertwined'.[77]

The underlying implications are significant. It has been correctly observed in legal contexts that discussions about algorithmic decision-making often imply a division of labour between AI and humans, based on an ideal division of tasks according to their respective competencies (machines are better at this/humans are better at that).[78]

If not properly managed, this division risks losing sight of the complexity of actual cases, focusing only on the algorithm's advantageous aspect and leaving all other aspects to an overwhelmed human director. In essence, it echoes the Vital case, highlighting the gap between the machine's narrow domain of operation and human versatility.

[73] See Shapira and Nili (2024).

[74] Having proposed at least six types of different functions performed by algorithms, Gal and Rubinfeld (2024), p 691 noted that 'this range of uses also exemplifies the fact that there is no one-type-fits-all algorithm. Rather, different types of algorithms are needed to perform different tasks'.

[75] Mertens (2023), p 12.

[76] At present, the testing of LLMs in total autonomy in the face of complex problems has made headlines more for extravagant results than for robust returns. However, it must be emphasised that the ability to handle language—the medium through which humans share their logic—appears to be one of the most interesting frontiers in the development of AI towards greater versatility.

[77] Mertens (2023), p 13.

[78] Crootof et al. (2023), p 460 ff.

4.2.2 Collegiality and Diversity

Corporate governance utilises the versatility of human directors also in monitoring, balancing, and transparency, for example, by requiring certain decisions to be made not by the individual CEO but by the board as a whole. This introduces a significant difference compared to algorithmic decision-making, where the algorithm optimises, but the board weighs. The decision to shift responsibility to the board level allows for solutions that result from reconciling different positions and weighing diverse perspectives. Replicating this dynamic is complex, not only when entrusting the decision to a single AI but also when imagining a plurality of algorithms interacting. Indeed, a collegial decision is not simply the average of each board member's optimisations but rather a creative process through which the decision takes shape.[79]

It has been observed that simply replicating the current strategy for human gender diversity through an algorithm without requiring actual representation of both genders on the board would be incorrect.[80] This approach would fail to acknowledge that diverse and original perspectives are shaped by the same incentives and disincentives, which are inherently different in the case of algorithms.[81] Algorithmic logic operates within a framework of paradigms that are not directly comparable to human ones and is not motivated by the same incentives that affect human decision-making. Equating algorithmic opinions with those of a human director is an erroneous approach due to the fundamental differences between their decision-making processes.

4.3 The Contraction of Space for Strategies and the Loss of Control over the Decision-making Process

The core of the argument lies in the explicit design of current corporate law strategies for a human decision-making process aimed at mitigating risks arising from agency relationships and the consequent threat of opportunism by the agent side. CorpTech itself was initially proposed as a potential tool to alleviate agency problems.[82]

For the reasons outlined above—ranging from the absence of a 'soul to damn and a body to kick' to the distinct and 'foreign' logic of AI—some existing corporate law strategies prove less effective in addressing the challenges that arise when AI is used for strategic decision-making.

Consequently, reintroducing these strategies in a context where discretion is assigned to the algorithm appears to miss the final goal, i.e., the control over the

[79] Kolbjørnsrud et al. (2016), p 13, Enriques and Zetzsche (2020), p 29.

[80] On this topic, see Eroğlu and Karatepe Kaya (2022), pp 541 ff., 565.

[81] But *contra*, see Petrin (2019), p 1003, according to whom 'the combined knowledge and skills, benefits of group-decision making, and characteristics such as diversity and independence, which previously could only be offered by a collective, will be replicated in fused boards through an algorithm's coding features'.

[82] See nn. 44–45 above.

discretion of the AI's decision-making process. In fact, for each decision-making process, corporate law typically provides a certain number of strategies, usually proportionate to the significance of the decision-making process in terms of its impact on society, agency risks, effects on stakeholders, etc. However, to implement such legal strategies, there is a need for legal 'space'. An eloquent example is the separation of powers, a foundation of every liberal democracy.[83] More pragmatically, within the corporate realm, one could highlight the difference between a decision made by the board of directors and one made by a sole director.

While, in the former case, activities like gatekeeping can be implemented using legal strategies such as a trusteeship or diversity within the board or through a reconstruction of how the decision was reached via an analytical record of the board meeting, in the latter case, the legal 'space' for the preparation of these strategies is unavailable, as the entire process remains within the mind of the sole director, residing in a region impervious to legal strategies.

Similarly, suppose the algorithm proves to be impervious to legal strategies designed for human directors. In that case, the consequence is that the fraction of the decision-making process entrusted to the algorithm lacks safeguards, reducing the overall legal space available to mitigate potential issues.

In summary, the case of strategic corporate decisions taken in incomplete scenarios highlights how the ineffectiveness of many current legal strategies compounds the inevitable reduction in the algorithm's accuracy. The outcome is a loss of reliability in the algorithmic result, unbalanced by remedies effectively addressing the process leading to that result.

To draw a comparison, the situation is not markedly different from the scenario in which, in a self-driving car, the algorithmic guidance system disengages just before impact with an obstacle, leaving the human driver to manage an irreversibly compromised situation.[84] At that point, even if the person behind the wheel were the best driver in the world (i.e., a plastic embodiment of all human strategies), the outcome could not change: the available space would be insufficient.

4.4 The Overloaded 'Human in the Loop'

Drawing on the example of self-driving cars, it might be argued that driving a vehicle, which involves a continuous flow of actions and real-time corrections, fundamentally differs from making decisions relevant to the community. In such decisions, humans are expected to have ample time to make the final decision. Unfortunately, when it comes to human-machine interaction, reality seems to offer less optimistic insights.

This is shown by the 'human in the loop' approach, conceived as a human safety valve to 'prevent or minimise the risks to health, safety or fundamental rights that may emerge when a high-risk AI system is used' (Article 14 of the AI Act). Despite

[83] The modern theory of distribution of powers is commonly ascribed to Baron de Montesquieu (1749).

[84] The case is far from theorical: see Crootof et al. (2023), p 438.

being the preferred strategy for decisions based solely on AI, the text enshrined in the Regulation raises concerns, even without considering strategic decisions.

Criticism has been directed at the inadequate attention provided by Article 14 to the human user's role, skills, and duties. The AI Act's human oversight requirements for providers have been said to 'focus less on understanding and mitigating known human frailties or on designing an effective human-machine system than on increasing the agency and power of the human in the loop'.[85] The absence of clear procedural safeguards related to the specific monitoring functions assigned to the human controller opens the door to a significant risk of overloading the human controller. Such potential overloading, from a factual standpoint, diminishes monitoring effectiveness and, from a legal perspective, makes it exceedingly challenging to verify when and how such functions have been diligently performed. Consequently, the supervising natural person would be endowed with broad theoretical powers and subject to duties beyond their actual capabilities.

The situation becomes even more complicated in the context of strategic decision-making. In these decisions, human oversight, as described in Article 14 of the AI Act, appears insufficient to address the increased margin of error.

While the outlined regulatory developments provide a crucial baseline for delineating corporations' efficiency choices in business and digitalisation transformation matters, limited analysis exists on how the consolidating 'human in the loop' paradigm tackles the risks arising from the structural incompleteness features of AI-driven decision-making in a strategic context, beyond oversight tasks focused on the protection of fundamental rights.

In this scenario, merely adding inaccuracy to the list of risks to be monitored according to Article 14 of the AI Act does not resolve the issue. Although monitoring can help mitigate negative externalities, achieving complete control over the potentially infinite variables involved in discretionary decision-making is a much more complex task. In other words, the issues raised in Article 14 of the AI Act concerning human oversight seem destined to worsen when monitoring shifts from mitigating negative externalities related to fundamental rights to reviewing the intrinsic correctness of the whole decision-making process. The human compensating for the loss of reliability in the algorithmic outcome results in an overextension of human oversight and, inevitably, an (even more severe) overload of the strategy, transforming it into something more akin to strict liability.[86]

That said, we do not imply that the 'human in the loop' approach is inherently ineffective, but solely that, by its nature, its scope cannot be excessively extended, as would happen if oversight were imposed on the intrinsic correctness of the algorithmic process.

[85] Ibid., p 504.

[86] See Sele and Chugunova (2022), p 16, who state that 'in a more sophisticated system, the number of features incorporated into the automated recommendation may exceed human capacity. If this is the case, human monitors may have to rely on inferior (or at least limited) information when deciding on the adjustment. This imbalance may increase the risk of decreasing decision accuracy due to human intervention further'.

In simpler terms, it is evident that the 'human in the loop' approach, as laid out in Article 14 of the AI Act, does not serve as a valid safeguard against the loss of reliability in the algorithmic result in the case of strategic decisions.

This assertion is further reinforced by automation bias in human-algorithm interactions. Automation bias refers to the human tendency to let the machine's dictates take over the decision-making process or, in other words, to make the wrong decision based on what the machine has (wrongly) indicated as the correct outcome.[87]

Automation bias has proven to be a pervasive phenomenon that is not necessarily mitigated by the significance of the decision. Uncritical reliance on algorithmic decisions has occurred even in highly sensitive circumstances, such as criminal judgments. The reference is notable to the *Compas* case, decided by the Supreme Court of Wisconsin,[88] where the algorithm's decision on the duration of the sentence in proportion to the risk of recidivism was accepted without real counterbalances or monitoring activities regarding the decision-making process. Implicitly, it formed the basis for a human decision justified *ex post* by applying typical human decision-making rules. The result was an opaque algorithmic decision, uncritically legitimised by the human decision-maker who accepted and justified the outcome *ex post*.[89]

Coordinating these considerations with Article 14 of the AI Act, it should be noted that paragraph 4 of the provision highlights the risk of automation bias but does not address it. As correctly noted, such a choice shifts the management of the automation bias risk onto human oversight, becoming an additional element of overload for this strategy.

4.5 Strategic Decisions and a Human-centric and Trustworthy AI

The approach chosen by the AI Act for decisions based solely on AI becomes problematic as one moves from decisions dealing with low incompleteness to those coping with increasing incompleteness. Specifically, within the realm of AI-human interaction, the 'human in the loop' strategy can be considered a simplified approach applicable to particular cases. In these cases, the substantial human replacement resulting from decisions based on AI, with human oversight to protect fundamental rights, seems justified by the minimal or limited degree of discretion required for the decision and the high level of accuracy achieved. In this sense, it would appear that current European regulations based on the 'human in the loop' approach envision

[87] On empirical evidence of automation bias, see ibid., p 4 ff.

[88] *State v. Loomis,* 881 N.W.2d 749 (Wis. 2016). In early 2013, Wisconsin charged Eric Loomis with five criminal counts, and in its sentencing, the trial court referred to the assessment made by COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) with regard to the risk of recidivism. COMPAS is a case management and decision support tool developed by a private company and has been used by some US jurisdictions to assess the risk of recidivism.

[89] The ruling has been intensively criticised from different perspectives. See, *ex multis*, Mattu et al. (2016), Freeman (2016), p 75 ff., De Miguel Beriain (2018), p 45 ff., Mayson (2018), p 2218 ff.

a situation characterised by low discretion and high reliability in the algorithmic outcome.

When delving into the domain of strategic decisions, the regulatory architecture established by the AI Act for decisions based solely on AI appears inadequate to meet the aforementioned principal prerequisite of the European project, i.e., that 'AI should be a human-centric technology. It should serve as a tool for people, with the ultimate aim of increasing human well-being'.[90]

The analysis undertaken here calls for a radical shift in the regulatory perspective. In the incomplete environment of strategic decisions, the normative goal cannot solely be an algorithm that does not endanger the safety of its users. Instead, the normative goal should be a human who can harness, while remaining in command, a tool that enhances their decision-making capabilities.

The final part of this work contemplates some guidelines for implementing what has been defined at the European level as the 'human in command' approach within strategic decisions. According to this approach, 'machines remain machines and people retain control over these machines at all times'.[91]

5 AI as a Tool

5.1 A Ban of the Algorithm: A Wrong Turn

It is worth immediately clearing the field of what seems to be the simplest solution to restore the 'human in command', namely a preventive and absolute ban on using the algorithm for strategic decisions.

This approach can lead to undesirable outcomes. It is a losing proposition at the macroeconomic level, since it is evident that not embracing the race towards innovation today means suffering from the decisions of competing jurisdictions tomorrow.[92]

Even without dwelling on the systemic opportunities that might be lost by refraining from embracing disruptive AI innovation, at the microeconomic level, a ban on using AI risks turning out to be only a formal prohibition, which is impossible or extremely challenging to enforce in practice. A recent investigation has shown that AI's ability to save time and resources is significant enough to prompt a considerable percentage of generative AI users to use the tool at work, even without authorisation or in defiance of the employer's ban, while recognising that the safe and ethical use

[90] Recital 6 AI Act.

[91] European Economic and Social Committee (2017).

[92] The very text of the AI Act stresses the importance of developing a human-centric as well as trustworthy AI. According to Recital 4 AI Act, 'AI is a fast evolving family of technologies that contributes to a wide array of economic, environmental and societal benefits across the entire spectrum of industries and social activities. By improving prediction, optimising operations and resource allocation, and personalising digital solutions available for individuals and organisations, the use of AI can provide key competitive advantages to undertakings and support socially and environmentally beneficial outcomes'. See also Mozzarelli (2022), p 281.

of this tool is through company-approved programmes.[93] They may also pass off the machine's work as their own.[94] The probable outcome of such a ban on the algorithm could easily encourage its hidden use, probably without valid strategies (or even proper training) for managing the risks associated with such use.

5.2 AI and the Human Decision-maker: From Human Replacement to Human Enhancement

If prohibiting AI usage is not a viable solution, then the answer may be to better regulate the interactions between the algorithm and the human. Considering a future comprehensive exploration of the subject, we can still attempt to recapitulate the considerations made so far to identify three schematic coordinates to deepen future research.

The first coordinate arising from the above analysis is that the legal perspective on the role of AI in strategic decisions calls for a significant shift: the algorithm cannot be managed as a mere product aimed at autonomously replacing human skills. Still, it must be considered a tool, an enhancer or augmenter of available resources, with which humans interact.[95] Though human enhancement and human replacement have often been placed on a sort of evolutionary scale regarding the AI role,[96] the preceding sections advocate a strong discontinuity in terms of regulation between enhancement and replacement.

In this perspective, the regulatory objective of AI as a tool is no longer solely to minimise risks stemming from AI as a 'product', but it also and primarily focuses on the interaction between AI and the human decision-maker, with the perspective of keeping humans in command.

The shift in perspective towards interaction implies a particular focus on the integrated decision-making process between the human decision-maker and the algorithm and the two further coordinates: on the one side, the actual control by the human decision-maker must not be questioned, and on the other side, the process should be permeable to currently available legal strategies.

5.2.1 The Ergonomics of the Algorithm

The second coordinate is all about the risk that the 'human in command' is reduced to a mere formal statement: human overreliance on algorithmic output and the inability, or serious difficulty, in understanding the algorithmic path and the reasons for one particular outcome rather than another are typical examples. To structure legal

[93] A recent empirical research by Salesforce regarding 14,000 workers across 14 countries revealed that over a quarter of workers are currently using AI at work and 55% of them without the formal approval of their employers or (40%) even against an explicit ban, Salesforce (2023).

[94] According to the cited research, 64% of workers have passed off work done by AI as their own.

[95] Petrin (2019), p 982, proposes the concept of advisory AI or augmented intelligence, referring 'to a combination of artificial and human intelligence, in which AI does not replace human intelligence, but leverages or improves it by, for example, giving information and advice that would otherwise be unavailable or more difficult and time consuming to obtain'.

[96] See Petrin (2019), p 980, Armour and Eidenmüller (2020), p 87 ff.

precepts in a way that provides the appropriate tools to enable the human decision-maker not only to interpret the algorithmic outcome but also to remain the true centre of the decision-making process, it is crucial to reconsider the interaction between humans and machines from an interdisciplinary perspective, potentially drawing on behavioural approaches.

We could define this novel approach as the ergonomics of the algorithm. Its aim would be a human-machine interaction that should avert human mechanisation in favour of genuine human empowerment.

5.2.2 A Legal Strategy-driven AI

The third, and probably most crucial from a legal perspective, coordinate is the role of legal strategy in the AI design.

The loss of reliability in the algorithmic result for strategic decisions requires intervention in the decision-making process, including the legal tools necessary to ensure (not the exact result but) the correct process.

Recently, it has been observed that 'a good blueprint' for integrating the respective human and algorithmic decision-making processes is still lacking. Drafting this blueprint is primarily the task of the legal scholar: the correctness of the decision-making process, unlike the accuracy of the result, does not respond to mathematical logic but to essentially legal parameters conveyed through legal strategies.

Designing the algorithm based on corporate legal strategies may limit its use: as observed, it would be wrong to assume that all corporate legal strategies can be effectively implemented through a decision-making process algorithm. Some strategies, particularly those based on collegiality and diversity within a board of directors, appear to be far from the algorithmic logic (at least at the current state of the art), and their actual effectiveness within an algorithmic decision-making process should be demonstrated with particular rigour. Otherwise, the use of AI in that circumstance should be excluded.

Instead, when legal strategies and AI are not structurally incompatible, defining minimum AI admissibility requirements is the priority. A clear definition is essential, as it serves as the entry point for an algorithm's legitimate use in strategic corporate decision-making.

Our first impression is that these requirements must be conceived in such a way that the decision taken with the help of AI is legally equivalent to the decision taken without such assistance. The critical regulatory consequence is that the focus should not be on the algorithm and on how to make it equivalent to a (decent) human being (fair, transparent, accurate, etc.), but on the interaction between human and machine—and on how to make it equivalent to an (adequate) interaction between humans.

Compared to an approach focused solely on the algorithm, aimed, for example, at increasing its accuracy or transparency,[97] we believe that taking the interaction between humans and machines as a reference, and therefore focusing on the (possible) conditions of equivalence between the two processes (with and without the algorithm), will help achieve some significant results. On the one hand, the integrated approach appears better suited to mitigate the risk of human overload and cases of automation bias. On the other hand, comparing the integrated conduct (humans assisted by AI) with solely human conduct allows for a more precise definition of the integrated conduct, linking the minimum admissibility criteria to the rule's final goal.

Examples might include identifying the minimum requirements for using AI to activate the business judgment rule and consequently protecting the (human, assisted by AI) decision from being second-guessed in a court review,[98] or defining the corporate AI system requirements, if any, to preserve the independence of an independent director who is monitoring the executives' conduct through that AI system.

The legal answer to this question will, in our opinion, make it possible to provide much more comprehensive and precise legal guidance, not only on the design of the algorithm but also on the entire interaction between human and machine, making it possible to weigh up the various elements involved (from the literacy of the human user to the possibility of intervening in the algorithmic patterns by inserting company-specific conditions and qualitative assessments, from the transparency of the algorithmic outcome to the degree of approximation of the result, etc.).

6 Conclusions

The current regulatory framework is insufficient for managing decisions based solely on AI, as evidenced by strategic decisions that reveal its inherent limitations. This indicates first and foremost that, from the perspective of the board of directors, relying on AI for strategic decision-making is currently particularly risky, even when operating within the bounds of existing regulations. If, as has been argued,

[97] See Rudin (2019), p 212, who stated that 'since the definition of what constitutes a viable explanation is unclear, even strong regulations such as "right to explanation" can be undermined with less-than-satisfactory explanations'. In the same way, Guidotti et al. (2018), p 36, noted that 'one of the most important open problems is that, until now, there is no agreement on what an explanation is. Indeed, some works provide as explanation a set of rules, others a decision tree, others a prototype (especially in the context of images). It is evident that the research activity in this field is not providing yet a sufficient level of importance in the study of a general and common formalism for defining an explanation, identifying which are the properties that an explanation should guarantee, e.g., soundness, completeness, compactness and comprehensibility. Concerning this last property, there is no work that seriously addresses the problem of quantifying the grade of comprehensibility of an explanation for humans, although it is of fundamental importance', while Sele and Chugunova (2022), p 16, stated that 'the simple inclusion of a human in the loop is unlikely to prevent inaccurate predictions based on algorithmic recommendations'.

[98] See, on that subject, Mosco (2020), p 93, and recently the interesting proposal of Langenbucher (2024), who focuses on the dimensions of ownership and trust to build a framework for understanding how corporate law shapes board decision-making.

the regulatory framework is inadequate for managing the risks associated with high levels of incompleteness, then the use of AI by companies for strategic decision-making may result in decisions lacking the level of accuracy that AI can achieve in low(er)-discretion environments, while also being structurally detached from the legal strategies designed for human agents.

The solution proposed here to address the issue of incompleteness therefore necessitates a change in perspective, with a focus on empowering the human decision-maker and designing a more ergonomic AI that operates within the framework of existing legal strategies. This shift in perspective provides a robust foundation for holding the human decision-maker accountable rather than the algorithm, thus placing the decision-maker in a position to fully comprehend the algorithm's potential and assume complete responsibility for its use.

The path to achieving such a change in perspective entails the establishment of novel standards for integrated decision-making processes. This is a relatively unexplored legal territory that, in our view, requires significant effort from legal scholars to gain a deeper understanding of the operational realities involved. This will enable them to determine whether it is possible to identify the criteria that would allow a strategic decision made with the assistance of AI to be legally equivalent to the same decision made without such assistance. It is essential to clearly define these criteria, as they serve as the entry point for the legitimate use of algorithms in strategic corporate decision-making.

It is then incumbent upon policymakers to apply the most promising theoretical solutions to draft guidelines and engage in dialogue with companies to ascertain their true effectiveness in practice. While the initial practical applications may be limited in scope, this starting point is essential for the sustainable development of human-algorithm interaction, capable of managing any decisions, even strategic ones, despite their inherent incompleteness and high levels of discretion.

Acknowledgments We thank Mario Libertini, Gaetano Presti and the seminar participants at the conference 'Orizzonti del diritto commerciale' in Rome on 24 February 2024 for helpful conversations.

Author contributions Although this paper is the result of joint work, Sect. 2 and 3 should be attributed to Giulia Schneider and Sect. 4 and 5 to Michele Mozzarelli.

Funding Open access funding provided by Università Cattolica del Sacro Cuore within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors have no competing interests or funding to disclose.

References

Abriani N, Schneider G (2021) Diritto delle imprese e intelligenza artificiale. Il Mulino, Bologna

Alvarez SA, Porac J (2020) Imagination, indeterminacy, and managerial choice at the limit of knowledge. AMR 45:735–744. https://doi.org/10.5465/amr.2020.0366

Armour J, Eidenmüller H (2020) Self-driving corporations? Harv Bus Law Rev 10:87–116

Armour J, Hansmann H, Kraakman RH (2017a) Agency problems and legal strategies. In: Armour J, Enriques L, Davies P et al (eds) The anatomy of corporate law: a comparative and functional approach, 3rd edn. Oxford University Press, Oxford, pp 29–48

Armour J, Gordon J, Min G (2020) Taking compliance seriously. Yale J Regul 37:3–66

Armour J, Parnham R, Sako M (2022) Augmented lawyering. U Ill L Rev 71–134

Armour J, Awrey D, Davies P et al (2016) Principles of financial regulation. Oxford University Press, Oxford

Armour J, Enriques L, Davies P et al (2017b) The anatomy of corporate law: a comparative and functional approach, 3rd edn. Oxford University Press, Oxford

Assonime (2023) L'evoluzione dell'organo amministrativo tra sostenibilità e trasformazione digitale. https://www.assonime.it/_layouts/15/Assonime.CustomAction/GetPdfToUrl.aspx?PathPdf=https://www.assonime.it/attivita-editoriale/studi/Documents/Note%20e%20Studi%201%20-%202023.pdf

Balboni P, Francis K (2023) Data protection as a corporate social responsibility. Edward Elgar Publishing, Cheltenham

Baron de Montesquieu C de S (1749) De l'esprit des loix. Aux dépens de la Compagnie

Borgogno O (2022) Making decentralized autonomous organizations (DAOs) fit for legal life: mind the gap. Bank of Italy Occasional Paper No. 718

Bradford A (2023) Digital empires: the global battle to regulate technology. Oxford University Press, Oxford

Borgogno O, Martino ED (2024) Decentralised autonomous organizations: targeting the potential beyond the hype. European Banking Institute Working Paper Series 161, Amsterdam Law School Research Paper No. 2024-02, Amsterdam Center for Law & Economics Working Paper No. 2024-01. https://doi.org/10.2139/ssrn.4692754

Buckley RP, Zetzsche DA, Arner DW, Tang B (2021) Regulating artificial intelligence in finance: putting the human in the loop. Syd Law Rev 43(1):43–81

Burridge N (2017 Artificial intelligence gets a seat in the boardroom. Nikkei Asia, 10 May 2017. https://asia.nikkei.com/business/artificial-intelligence-gets-a-seat-in-the-boardroom. Accessed 29 Jan 2024

Calabresi G, Melamed AD (1972) Property rules, liability rules, and inalienability: one view of the cathedral. Harv Law Rev 85:1089–1128

Camuffo A, Gambardella A, Pignataro A (2023) Framing strategic decisions in the digital world. Strateg Manag Rev 4:127–160

Carnat I (2023) Ethics lost in translation: trustworthy AI from governance to regulation. Opin Juris Comp 4:90

Comandé G (2019) Multilayered (accountable) liability for artificial intelligence. In: Lohsse S, Schulze R, Staudenmayer D (eds) Liability for artificial intelligence and the internet of things. Nomos Verlagsgesellschaft mbH & Co KG, pp 165–184

Crootof R, Kaminski ME, Price WN II (2023) Humans in the loop. Vand Law Rev 76:429–510. https://doi.org/10.2139/ssrn.4066781

Davies P (2020) Introduction to company law. Oxford University Press, Oxford

De Gregorio G, Dunn P (2022) The European risk-based approaches: connecting constitutional dots in the digital age. Common Market Law Rev 59(2):473. https://doi.org/10.54648/COLA2022032

De Miguel Beriain I (2018) Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the *Wisconsin v. Loomis* ruling. Law Probab Risk 17(1):45–53. https://doi.org/10.1093/lpr/mgy001

Enriques L, Zetzsche DA (2020) Corporate technologies and the tech nirvana fallacy. Hastings Law J 72:55

Erel I, Stern LH, Tan C, Weisbach MS (2021) Selecting directors using machine learning. Rev Financ Stud 34:3226–3264

Erikson T, Knockaert M (2021) Negotiating incomplete contracts. J Strateg Contract Negot 5:153–169. https://doi.org/10.1177/20555636211042093

Eroğlu M, Karatepe Kaya M (2022) Impact of artificial intelligence on corporate board diversity policies and regulations. EBOR 23:541–572. https://doi.org/10.1007/s40804-022-00251-5
European Economic and Social Committee (2017) Artificial Intelligence: Europe needs to take a human-in-command approach. https://www.eesc.europa.eu/en/news-media/press-releases/artificial-intelligence-europe-needs-take-human-command-approach-says-eesc
European Commission (2020) Digital finance package. https://finance.ec.europa.eu/publications/digital-finance-package_en. Accessed 17 March 2024
European Parliament (2020a) Civil liability regime for artificial intelligence. September 2020. https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654178/EPRS_STU(2020)654178_EN.pdf. Accessed 17 Mar 2024
European Parliament (2020b) European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence. 2020/2014(INL)
European Commission (2021a) Study on the relevance and impact of artificial intelligence for company law and corporate governance. https://op.europa.eu/en/publication-detail/-/publication/13e6a212-6181-11ec-9c6c-01aa75ed71a1/language-en. Accessed 17 Mar 2024
European Commission (2021b) Liability Rules for Artificial Intelligence. https://commission.europa.eu/business-economy-euro/ng-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en. Accessed 17 Mar 2024
European Commission (2022a) Proposal for a Directive of the European Parliament and of the Council on liability for defective products. COM/2022/495 final
European Commission (2022b) 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
European Parliament (2023) Artificial intelligence liability directive. February 2023. https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI(2023)739342_EN.pdf. Accessed 17 Mar 2024
Freeman K (2016) Algorithmic injustice: how the Wisconsin Supreme Court failed to protect due process rights in *State v. Loomis*. N C J Law Technol 18:75
Gal M, Rubinfeld DL (2024) Algorithms, AI and mergers. Antitrust Law J 85:683
Gilovich T, Griffin D, Kahneman D (2002) Heuristics and biases: the psychology of intuitive judgment. Cambridge University Press, London
Goyal M (2014) Hong Kong VC firm appoints AI to board of directors. IT Business Blog, itbusiness.ca. https://www.itbusiness.ca/blog/hong-kong-vc-firm-appoints-ai-to-board-of-directors/48815. Accessed 29 Jan 2024
Gramitto Ricci SA (2019) Artificial agents in corporate boardrooms. Cornell Law Rev 105:869
Guidotti R, Monreale A, Ruggieri S et al (2018) A survey of methods for explaining black box models. ACM Comput Surv CSUR 51:1–42
Hamdani A, Hashai N, Kandel E, Yafeh Y (2018) Technological progress and the future of the corporation. J Br Acad 6:215–225
Kahneman D (2011) Thinking, fast and slow. Farrar, Straus and Giroux, New York
Kaminski ME, Malgieri G (2021) Algorithmic impact assessments under the GDPR: producing multi-layered explanations. IDPL 11:125–144
Kolbjørnsrud V, Amico R, Thomas R (2016) The promise of artificial intelligence: redefining management in the workforce of the future. Accenture Institute for High Performance
Kroll JA, Huey J, Barocas S et al (2017) Accountable algorithms. Univ Pa Law Rev 165:633–705
Langenbucher KC (2024) Ownership and trust—a corporate law framework for board decision-making in the age of AI. European Corporate Governance Institute—Law Working Paper No. 758/2024. https://doi.org/10.2139/ssrn.4746241
Lehot L (2023) How artificial intelligence is disrupting the dealmaking process. Medium, 18 September 2023. https://lehotlouis.medium.com/how-artificial-intelligence-is-disrupting-the-dealmaking-process-eb6a2ae3bf4c. Accessed 18 Mar 2024
Li Z (2024) Artificial fiduciaries. Wash Lee L Rev 81:1299
Malgieri G, Pasquale FA (2022) From transparency to justification: toward ex ante accountability for AI. Brooklyn Law School, Legal studies paper no. 712
Matsumi H, Solove DJ (2025) The prediction society: algorithms and the problems of forecasting the future. Ill Law Rev 2025:1

Mattu J, Larson J, Angwin L, Kirchner S (2016) How we analyzed the COMPAS recidivism algorithm. ProPublica, 23 May 2016. https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. Accessed 2 Feb 2024

Mayson SG (2018) Bias in, bias out. Yale LJ 128:2218

Mertens F (2023) The use of artificial intelligence in corporate decision-making at board level: a preliminary legal analysis. Financial Law Institute Working Paper Series 2023-01. https://doi.org/10.2139/ssrn.4339413

Mosco GD (2020) AI and the board within Italian corporate law: preliminary notes. Eur Company Law 17:87–96

Möslein F (2018) Robots in the boardroom: artificial intelligence and corporate law. In: Barfield W, Pagallo U (eds) Research handbook on the law of artificial intelligence. Edward Elgar, pp 649–670

Mozzarelli M (2022) Digital compliance: the case for algorithmic transparency. In: Manacorda S, Centonze F (eds) Corporate compliance on a global scale. Springer, Cham, pp 259–284

Petrin M (2019) Corporate management in the age of AI. Colum Bus L Rev 2019:965–1030

Regulatory Obstacles to Financial Innovation Experts Group—ROFIEG (2019) 30 recommendations on regulation, innovation and finance (13 December 2019). https://finance.ec.ef.uropa.eu/publications/final-report-expert-group-regulatory-obstacles-financial-innovation-30-recommendations-regulation_en. Accessed 17 Mar 2024

Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 1:206–215

Salesforce (2023) More than half of generative AI adopters use unapproved tools at work. 15 November 2023. https://www.salesforce.com/news/stories/ai-at-work-research/?utm_campaign=amer_cbaw&utm_content=Salesforce_World+Tour&utm_medium=organic_social&utm_source=linkedin. Accessed 29 Jan 2024

Schneider G (2022a) Accountability. In: Comandé G (ed) Encyclopedia of law and data science. Edward Elgar, pp 7–12

Schneider G (2022b) Fairness. In: Comandé G (ed) Encyclopedia of law and data science. Edward Elgar, pp 168–175

Sele D, Chugunova M (2022) Putting a human in the loop: increasing uptake, but decreasing accuracy of automated decision-making. Max Planck Institute for Innovation & Competition Research Paper No. 22-20. https://papers.ssrn.com/abstract=4285645

Shapira R, Nili Y (2024) Specialist directors. Yale J Regul 41:652

Solove DJ, Matsumi H (2024) AI, algorithms, and awful humans. Fordham Law Rev 92:1923

Sunstein CR (2011) Empirically informed regulation. U Chi L Rev 78:1349–1429

Teubner G (2018) Digital personhood? The status of autonomous software agents in private law. Ancilla Iuris, pp 106–130

Wagner G (2023) Liability rules for the digital age—aiming for the Brussels Effect. JETL 13:191–243. https://doi.org/10.1515/jetl-2022-0012

Wile R (2014) A venture capital firm just named an algorithm to its board of directors—here's what it actually does. Business Insider, 13 May 2014

Yermack D (2017) Corporate governance and blockchains. Rev Finance 21(1):7–31. https://doi.org/10.1093/rof/rfw074

Zuboff S (2019) The age of surveillance capitalism: the fight for a human future at the new frontier of power. Public Affairs, New York