european corporate governance institute

ecgi

Self-Driving Corporations?

Law Working Paper N° 475/2019
August 2019

John Armour
University of Oxford and ECGI
Horst Eidenmueller
University of Oxford and ECGI

© John Armour and Horst Eidenmueller 2019. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This paper can be downloaded without charge from:
https://doi.org/10.67929/ecgi-law-475-2019

european corporate governance institute

ECGI Working Paper Series in Law

Self-Driving Corporations?

Working Paper N° 475/2019

August 2019

John Armour
Horst Eidenmueller

We have benefited from helpful comments from participants following presentations of prior versions of this work at an ECGI Roundtable on *Technology and Corporate Governance*, a ZHR Symposium on Corporate and Capital Markets Law in Glashütten, an Oxford University Business Law Workshop, a seminar at Texas Law School and a conference on *The Future of the Firm* at University College London. We are grateful for comments from Chris Brummer, Jens Dammann, Jeff Gordon, Vanessa Knapp and Mari Sako. We also thank Jordan Gifford-Moore and Moritz Spenke for outstanding research assistance.

© John Armour and Horst Eidenmueller 2019. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Abstract

What are the implications of artificial intelligence (AI) for corporate law? In this essay, we consider the trajectory of AI's evolution, analyze the effects of its application on business practice, and investigate the impact of these developments for corporate law. Overall, we claim that the increasing use of AI in corporations implies a shift from viewing the enterprise as primarily private and facilitative, towards a more public, and regulatory, conception of the law governing corporate activity. Today's AI is dominated by machine learning applications which assist and augment human decision-making. These raise multiple challenges for business organization, the management of which we collectively term 'data governance'. The impact of today's AI on corporate law is coming to be felt along two margins. First, we expect a reduction across many standard dimensions of internal agency and coordination costs. Second, the oversight challenges—and liability risks—at the top of the firm will rise significantly. Tomorrow's AI may permit humans to be replaced even at the apex of corporate decision-making. This is likely to happen first in what we call 'self-driving subsidiaries' performing very limited corporate functions. Replacing humans on corporate boards by machines implies a fundamental shift in focus: from controlling internal costs to the design of appropriate strategies for controlling 'algorithmic failure', i.e. unlawful acts by an algorithm with potentially severe negative effects (physical or financial harm) on external third parties. We discuss corporate goal-setting, which in the medium term is likely to become the center of gravity for debate on AI and corporate law. This will only intensify as technical progress moves toward the possibility of fully self-driving corporations. We outline potential regulatory strategies for their control. The potential for regulatory competition weakens lawmakers' ability to respond, and so even though the self-driving corporation is not yet a reality, we believe the regulatory issues deserve attention well before tomorrow's AI becomes today's.

Keywords: Corporate Law, Corporate Governance, Artificial Intelligence, Machine Learning, Business Organization, Data Governance, Self-Driving Subsidiaries, Regulation, Corporate Objectives, Control and Liability Regimes, Regulatory Competition

JEL Classifications: K22, M1, O3

John Armour
Professor of Law and Finance
University of Oxford, Faculty of Law
St.Cross Building, St. Cross Road
Oxford, OX1 3UL, United Kingdom
phone: +44 186 528 161 6
e-mail: john.armour@law.ox.ac.uk

Horst Eidenmueller*
Statutory Professor for Commercial Law
University of Oxford, Faculty of Law
St Margaret's Road
Oxford, OX2 6LE, United Kingdom
phone: +44 1865 284 472
e-mail: horst.eidenmueller@law.ox.ac.uk

*Corresponding Author

Self-Driving Corporations?

John Armour[*]
Horst Eidenmüller[**]

ABSTRACT

What are the implications of artificial intelligence (AI) for corporate law? In this essay, we consider the trajectory of AI's evolution, analyze the effects of its application on business practice, and investigate the impact of these developments for corporate law. Overall, we claim that the increasing use of AI in corporations implies a shift from viewing the enterprise as primarily private and facilitative, towards a more public, and regulatory, conception of the law governing corporate activity. Today's AI is dominated by machine learning applications which assist and augment human decision-making. These raise multiple challenges for business organization, the management of which we collectively term 'data governance'. The impact of today's AI on corporate law is coming to be felt along two margins. First, we expect a reduction across many standard dimensions of internal agency and coordination costs. Second, the oversight challenges—and liability risks—at the top of the firm will rise significantly. Tomorrow's AI may permit humans to be replaced even at the apex of corporate decision-making. This is likely to happen first in what we call 'self-driving subsidiaries' performing very limited corporate functions. Replacing humans on corporate boards by machines implies a fundamental shift in focus: from controlling internal costs to the design of appropriate strategies for controlling 'algorithmic failure', i.e. unlawful acts by an algorithm with potentially severe negative effects (physical or financial harm) on external third parties. We discuss corporate goal-setting, which in the medium term is likely to become the center of gravity for debate on AI and corporate law. This will only intensify as technical progress moves toward the possibility of fully self-driving corporations. We outline potential regulatory strategies for their control. The potential for regulatory competition weakens lawmakers' ability to respond, and so even though the self-driving corporation is not yet a reality, we believe the regulatory issues deserve attention well before tomorrow's AI becomes today's.

[*] Professor of Law and Finance, University of Oxford; Fellow of the British Academy and of the ECGI.

[**] Professor of Commercial Law, University of Oxford; Member of the BBAW and ECGI Research Associate.

We have benefited from helpful comments from participants following presentations of prior versions of this work at an ECGI Roundtable on *Technology and Corporate Governance,* a ZHR Symposium on Corporate and Capital Markets Law in Glashütten, an Oxford University Business Law Workshop, a seminar at Texas Law School and a conference on *The Future of the Firm* at University College London. We are grateful for comments from Chris Brummer, Jens Dammann, Jeff Gordon, Vanessa Knapp and Mari Sako. We also thank Jordan Gifford-Moore and Moritz Spenke for outstanding research assistance. The usual disclaimers apply.

Electronic copy available at: https://ssrn.com/abstract=3442447

TABLE OF CONTENTS

INTRODUCTION

I. THE TRAJECTORY OF AI

A. The History of AI

B. Today's AI

C. Tomorrow's AI?

II. TODAY'S AI AND CORPORATE LAW

A. Impact of Today's AI on Business Organization

B. 'Data Governance': How Today's AI Impacts Corporate Governance

1. Scope of available data

2. Model selection and training

3. Model predictions and wider corporate goals

4. Deployment and Organizational Structure

C. Data Governance and Corporate Law

III. TOMORROW'S AI AND CORPORATE LAW

A. Self-Driving Subsidiaries

B. Calibrating Corporate Objectives for Self-Driving Corporations

C. Alternative Control and Liability Regimes

D. Regulatory Competition

CONCLUSION

Electronic copy available at: https://ssrn.com/abstract=3442447

Self-Driving Corporations?

INTRODUCTION

In this essay, we explore the implications of artificial intelligence (AI) for corporate law. The field of AI has seen dramatic technical advances in recent years, in turn spurring much speculative media attention, [1] to which corporate boardrooms have not been immune. Five years ago, a Hong Kong-based venture capital firm made a big splash by announcing that a robot named "Vital" had been appointed as a board member.[2] Even discounting for the intense hype surrounding the topic, the opportunities for productivity gains and welfare increases appear to be considerable. At the same time, AI brings new economic and legal risks. These too feature hyperbole—with some claiming algorithmic entities are a "threat to humanity". [3] We seek here to offer a considered assessment of how this new technology may come to impact corporate law.

Today, corporate law is primarily understood as a means of facilitating productive activity in business firms. On this view, it is a predominantly private endeavor, concerned with helping parties to lower the costs they encounter. Much of 'core corporate law' can hence be explained as responses to agency and coordination problems arising between investors and managers. [4] That is,

---

[1] On AI applications for businesses see, in particular, THOMAS H. DAVENPORT, THE AI ADVANTAGE: HOW TO PUT THE ARTIFICIAL INTELLIGENCE REVOLUTION TO WORK (2018); Robert J. Thomas, Rouven Fuchs & Yaarit Silverstone, *A machine in the C-suite* (2016) https://www.accenture.com/t00010101T000000Z__w__/br-pt/_acnmedia/PDF-13/Accenture-Strategy-WotF-Machine-CSuite.pdf.

[2] Rob Wile, *A Venture Capital Firm Just Named An Algorithm To Its Board Of Directors — Here's What It Actually Does*, https://www.businessinsider.com.au/vital-named-to-board-2014-5 ("Just like other members of the board, the algorithm gets to vote on whether the firm makes an investment in a specific company or not."). In fact, Vital did not actually become a board member in law, but technically only had the status of "observer". See Florian Möslein, *Robots in the Boardroom: Artificial Intelligence and Corporate Law*, Oxford Business Law Blog, Nov. 16, 2017, https://www.law.ox.ac.uk/business-law-blog/blog/2017/11/robots-boardroom-artificial-intelligence-and-corporate-law.

[3] Lynn M. LoPucki, *Algorithmic Entities*, 95 WASH. UNIV. L. REV. 887, 951 (2018): "Algorithmic Entities constitute a threat to humanity …".

[4] FRANK H. EASTERBROOK & DANIEL R. FISCHEL, THE ECONOMIC STRUCTURE OF CORPORATE LAW (1991); REINIER KRAAKMAN, JOHN ARMOUR, PAUL DAVIES, LUCA ENRIQUES, HENRY HANSMANN, GÉRARD HERTIG, KLAUS HOPT, HIDEKI KANDA, MARIANA PARGENDLER, WOLF-GEORG RINGE & EDWARD ROCK, THE ANATOMY OF CORPORATE LAW: A COMPARATIVE AND FUNCTIONAL APPROACH 29–

Electronic copy available at: https://ssrn.com/abstract=3442447

numerous shareholders who each have only a small stake in a firm cannot readily coordinate to monitor the actions of managers, to whom power to make decisions about corporate assets has been delegated; managers may therefore be tempted to further their own interests at the expense of shareholders'. This account is typically framed as a bilateral dynamic between managers and shareholders. As a corollary, the impact of business activity on society at large is typically treated as outside the remit of core corporate law, in line with the theory that regulatory norms should apply equally to all actors, corporations or otherwise.

To what extent will AI change the regnant account of corporate law?[5] The standard account is functional in its orientation; that is, it is premised on a social-scientific account of what actually happens in a business firm.[6] The starting point for our analysis is therefore to ask how the activities of firms will be affected by AI. As a preliminary step, however, we need to begin with a clear account of what is technically possible. In contrast, much of the current legal literature on AI makes strong assumptions about technical capabilities without a well-informed account of the limits of what is currently possible. The next step is then to apply standard analytic tools from social science—the economics of business organization—to explore the likely impact of these innovations on business practice. With a model of business practice in mind, we can begin to visualize how corporate law may be affected.

---

47 (3rd ed. 2017); GREGOR BACHMANN, HORST EIDENMÜLLER, ANDREAS ENGERT, HOLGER FLEISCHER & WOLFGANG. SCHÖN, REGULATING THE CLOSED CORPORATION 8–25 (2014).

[5] In our analysis, we focus on developments in the field of AI and do not deal with blockchain technology. AI applications are already extensively used in businesses today. By contrast, blockchain applications in the field of corporate governance are more of a future topic. At any rate, blockchain technology is currently not changing the types and forms of corporate activity, or does so only to a limited extent. This is because every implementation or execution requires an action/decision that should be implemented or executed. Apart from very simple transactions (contracts), the discretion of the actors who make decisions on the allocation of resources always plays a role here, and that is the domain of AI.

[6] This is based on the assumption that *firms*—as a means to organize productive activities different from *markets*—exist in the first place. On this issue, see the seminal analysis by Ronald H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386 (1937). For an assessment against the background of AI/blockchain technology, see VIKTOR MAYER-SCHÖNBERGER & THOMAS RAMGE, REINVENTING CAPITALISM IN THE AGE OF BIG DATA (2018); John Armour et al., *Putting technology to good use for society: the role of corporate, competition and tax law*, 6 J. BRIT. ACAD. 285, 293 (2018).

Electronic copy available at: https://ssrn.com/abstract=3442447

We consequently start in Section I with a detailed overview of what AI is, and what it can and cannot do. Two major points emerge from this enquiry. First, the current state of AI development, while impressive, is far from providing a general human-level intelligence (so-called 'artificial general intelligence' or 'AGI'). Rather, the effective deployment of today's AI depends on sufficient quantities of relevant data and problems that can usefully be answered by predictive analytics based on this. Today's AI is not going to replace humans in the C-suite.

Second, the path of technological development is a trajectory, not a simple on-off switch. For AI, this trajectory is evolving continuously in a way that is not linear. A recent survey of scientific leaders in AI suggests a wide range of estimates of the time horizon until the advent of AGI – from a decade to two centuries.[7] Over the past few decades, rapid and sudden change has occurred unpredictably, and there is no reason to think that this *lack* of pattern will change any time soon. Consequently, while we can articulate the 'deployment conditions' for current AI applications, we would likely be negligent to fail also to consider the potential implications of future advances.

This suggests two distinct approaches to the subsequent stages in our enquiry. In Section II, we draw on standard social science tools to explore the implications for business activity of *today's* AI. Indeed, much research activity is currently investigating these issues.[8] The emerging picture is one in which human decision-making can efficiently be assisted and augmented by AI applications. The impact on corporate law is coming to be felt along two margins. First, we expect a reduction across many standard dimensions of internal agency and coordination costs. Augmenting human decision-making typically means that fewer human

[7] MARTIN FORD, ARCHITECTS OF INTELLIGENCE 528–29 (2018).

[8] See, e.g., Erik Brynjolfsson & Andrew McAfee, *How AI Fits into Your Science Team: What it can do—and cannot do—for your organization,* HARV. BUS. REV. (2017); Sam Ransbotham, David Kiron, Philipp Gerbert & Martin Reeves, *Reshaping Business with Artificial Intelligence: Closing the Gap Between Ambition and Action*, MIT SLOAN MANAGEMENT REVIEW RESEARCH REPORT (2017); DAVENPORT, *supra* note 1; Daren Acemoglu & Pascual Restrepo, *Artificial Intelligence, Automation, and Work*, MIT DEPARTMENT OF ECONOMICS WORKING PAPER 18-01 (2018); Ajay Agrawal, Jishua S. Gans & Avi Goldfarb (eds.) THE ECONOMICS OF ARTIFICIAL INTELLIGENCE (2019); INSTITUTE OF CHARTERED ACCOUNTANTS OF ENGLAND & WALES, AI IN CORPORATE ADVISORY: INVESTMENT, M&A AND TRANSACTION SERVICES (2019); Tim Fontaine, Brian McCarthy & Tamim Saleb, *Building the AI-Powered Organization*, HARV. BUS. REV. 62 (2019).

Electronic copy available at: https://ssrn.com/abstract=3442447

beings are needed to deliver the same results. Automated decision processes do not lack fidelity, and so agency costs associated with automated decisions should reduce. At the same time, along a second margin, new types of discretionary decision-making become important—the decisions involved in establishing and testing the automated systems themselves.[9] One may be tempted to see this simply as a substitution of agency costs from one domain to another. Yet if the overall level of human participation decreases, the impact of the new agency costs will be increasingly 'strategic' in their reach—that is, having potentially far-reaching consequences for corporate performance. Identifying where to monitor these, and how best to do it, will be a progressively complex and important task. This will necessitate increasing energy being devoted to the mapping and governance of these risks, an endeavor which we term '*data governance*'. The high-level implication is that this will place increasing demands on oversight at the top of the firm. For corporate law, this means that the duties of directors, who are ultimately responsible for oversight of firms' performance, will increasingly come to recognize the significance of data governance for corporate success.

In Section III, we pursue a different tack with respect to *future* AI. Here, we wish to envisage the consequences of replacing humans with AI in the apex of corporate decision-making—giving rise to what might be termed 'self-driving corporations'. While the technology to implement this is not here yet, it seems plausible that it could arrive sooner than AGI is achieved. To show this, we begin with a thought experiment framed around the most likely early use-cases: what we term 'self-driving subsidiaries'.[10] Subsidiaries are currently used by many corporations to achieve specific limited functions. It is conceivable that using little more than today's technologies, entities performing very limited functions could be fully automated. In systems of corporate law that permit firms to be organized without human directors, the self-driving subsidiary could soon be a reality. Yet for such firms, the principal product of our discussion in Section II—the increasing

[9] See generally, David H. Autor, *Why Are There Still So Many Jobs? The History and Future of Workplace Automation*, 29 J. Econ. Persp. 3 (2015); Acemoglu & Restrepo, *supra* note 8.

[10] The term 'self-driving' in the title of our article does not aim to draw a direct analogy with self-driving cars. Rather, we are proposing it as a metaphor with a certain heuristic usefulness.

Electronic copy available at: https://ssrn.com/abstract=3442447

importance of oversight for directors—is abruptly neutered. Without directors, oversight liability for board members can have no traction.

Of course, investors in a parent company who are unhappy with the decisions made by a self-driving subsidiary may still have recourse against those charged with oversight of the parent company for their decision to establish the subsidiary as 'self-driving'. The analysis in Section II will continue to hold, in attenuated form, for such internal concerns. Yet with respect to external liabilities of the subsidiary—for example, torts and crimes—there is no longer a point of human contact. Hence, the deployment of automated subsidiaries appears to have an appealing application for the avoidance of regulatory or tortious liabilities. This implies a more fundamental shift in focus, from controlling internal costs—which in a fully self-driving context are automated from the outset—to the design of appropriate strategies for controlling the costs that corporate activity imposes on persons external to the endeavor. That is, a shift from viewing the enterprise as primarily private and facilitative, towards a more public, and regulatory, conception of the law governing corporate activity.

There is a strong hint in this direction from the analytic significance of whether corporate law mandates human directors. Mandating such directors in an era of self-driving corporations will not be a means of facilitating the lowering of costs of organizing activities in firms. Rather, it will primarily be a means of regulating such firms to ensure that humans are charged with oversight of their activities.

If corporate law does not mandate that companies have human directors when fully self-driving corporations become a reality, it must deploy other regulatory devices to protect investors and third parties from what we refer to as 'algorithmic failure': unlawful acts triggered by an algorithm, which cause physical or financial harm. We discuss the issue of corporate goal-setting which is likely to become more and more the focus of debate on AI and corporate law in the medium term. We also explore further regulatory implications in an environment characterized by regulatory competition. Fully self-driving corporations might be subject to an *ex ante* assessment of controlling algorithms as well as to strict liability for algorithmic failure, combined with compulsory corporate liability

Electronic copy available at: https://ssrn.com/abstract=3442447

insurance. As a regulatory alternative, we consider an unlimited pro rata liability of shareholders for corporate torts.

I. THE TRAJECTORY OF AI

A. The History of AI

AI is not a new concept, the first usage generally being attributed to computer scientists *John McCarthy* and *Myron Minsky* in the mid-1950s.[11] The convention is to use the term in a functional sense, meaning that an artificial system functions as well as, or better than, a human. Clearly, machines can perform many tasks better than humans that involve strength or endurance as opposed to intelligence. The 'intelligence' qualifier therefore can usefully be understood as restricting the comparison to activities for which a human would use their brain—most obviously, processing and analyzing information.[12]

The classic assessment of whether a system functions as well as a human is the so-called 'Turing test', in which a human is asked to engage in conversation with messages sent through a mechanism that does not reveal whether the party on the other side is human or not.[13] If a human participant cannot distinguish the communications of an artificial system from a human, then the test is passed by that system. To pass a Turing test without any constraints around the type of conversation that could be had, the machine would need to exhibit AGI; that is, as good as human in every dimension of intelligence. Modern AI systems do not come anywhere near AGI. This is—according to experts—anywhere between a decade

[11] John McCarthy et al., *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence (Aug. 31, 1955)*, 27 AI MAG. 12 (2006) (proposing a "2 month, 10 man study of artificial intelligence", organized around "the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.").

[12] See, for example, John McCarthy, *What is AI? / Basic Questions*, http://jmc.stanford.edu/artificial-intelligence/ what-is-ai/index.html ("the science and engineering of making intelligent machines, especially intelligent computer programs.")

[13] Alan M. Turing, *Computing Machinery and Intelligence*, 49 MIND 433, 434 (1950). A test so formulated "has the advantage of drawing a fairly sharp line between the physical and the intellectual capacities of a man" (ibid).

Electronic copy available at: https://ssrn.com/abstract=3442447

and two centuries away.[14] Rather, the AI deployed today only has (super)human-level capability in respect of narrowly-defined functions, such as image recognition, driving vehicles in straightforward surroundings, or the classification of documents.

Defining AI in this functional way means that no particular restrictions are put on the nature of the computing system used. Indeed, there have been at least three distinct technical approaches to AI since the birth of the field.[15] The first, popular in the 1960s, involved logical rules: the idea was to develop a general-purpose system capable of deriving answers to problems through formal logical reasoning.[16] This approach fell afoul of the problem that deterministic calculation of outcomes becomes exceedingly complex for even moderately challenging real-world problems.

A different tack was taken in the 1980s and 90s, with the advent of so-called 'expert systems'.[17] These were designed to give human users the benefit of expert answers to problems in an area of practice. The problems and answers—what computer scientists call 'domain knowledge'—were characterized with the help of relevant human experts. These were then coded into a system designed to answer questions related to that particular body of knowledge. Expert systems in turn proved quite brittle, however. If the question asked by a user fell outside the system's expertise, it could not give an answer. And moreover, the framing of questions had to be done in terms of the specific syntax of the system. If the user was unable to express themselves in terms the system could understand, then again it would fail. These limitations created roadblocks to the roll-out of such systems.

---

[14] See *supra* note 7 and accompanying text.

[15] See STUART RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH 16–28 (3rd ed. 2010). See also MIKE WOOLDRIDGE, ARTIFICIAL INTELLIGENCE (2018).

[16] See generally Jack Minker, *Introduction to Logic-Based Artificial Intelligence, in* LOGIC-BASED ARTIFICIAL INTELLIGENCE 3 (Jack Minker ed., 2000).

[17] See, for example, PAUL HARMON & DAVID KING, EXPERT SYSTEMS: ARTIFICIAL INTELLIGENCE IN BUSINESS (1985). For a classic account of their application to law, see RICHARD E. SUSSKIND, EXPERT SYSTEMS IN LAW: A JURISPRUDENTIAL ENQUIRY (1987).

Electronic copy available at: https://ssrn.com/abstract=3442447

B. Today's AI

The third and most recent development in AI has related primarily to *machine learning* (ML). This is an approach to computing in which the solution to an optimization problem is not coded in advance, but is derived inductively by reference to data.[18] In a sense, ML turns the logic of expert systems upside down: instead of deriving answers from rules and data, rules are developed from data and answers. The technique relies on applying computing power to very large amounts of data, the availability of which of both has blossomed in recent years.[19]

Progress since 2012 has largely been in a particular type of ML known as *deep learning*, which involves running multiple layers of representation of the data in series.[20] A typical deep learning setup consists of an input and an output layer, with multiple hidden layers in between that lie at different levels of abstraction and are linked to each other (see Fig. 1).[21]

[18] Alan Turing anticipated this in 1950 as follows: "Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's?" (Turing, *supra* note 13, at 456). For a comprehensive treatment see, for example, ETHEM ALPAYDIN, MACHINE LEARNING (2016).

[19] On the significance of hardware for AI performance, see Tim Hwang, *Computational Power and the Social Impact of Artificial Intelligence*, WORKING PAPER, MIT MEDIA LAB (2018); on the importance of data for ML performance, see Alon Halevy, Peter Norvig & Fernando Pereira, *The Unreasonable Effectiveness of Data*, 24 IEEE INTELLIGENT SYSTEMS 8 (2009); Chen Sun et al., *Revisiting Unreasonable Effectiveness of Data in Deep Learning Era,* arXiv:1707.02968v2 (2017).

[20] See, for example, FRANÇOIS CHOLLET, DEEP LEARNING WITH PYTHON 8–11 (2018).

[21] These processes are often called 'neural networks', a term drawn from neurobiology, reflecting the fact that some of these processes draw inspiration from contemporary understanding of the functioning of the brain. However, their actual operation is quite different from the human brain: Chollet, *supra* note 20, at 8.

Electronic copy available at: https://ssrn.com/abstract=3442447

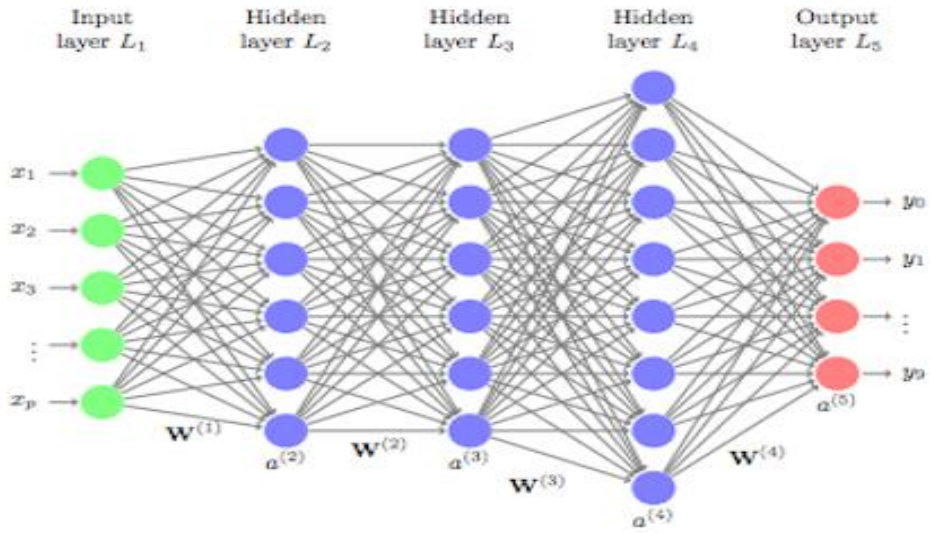


Fig. 1: Deep 'neural network'[22]

The various (hidden) layers represent different stages of abstraction of a thought process. For example, if the question is to identify at the output layer whether an image is a door, the first level would be pixels, the second would be edges, then corners, and so on up to the output layer: door (or not).

The learning process of the algorithm takes place via so-called *back-propagation*: In the course of training the algorithm, new information is fed back from the output layer over the various hidden levels and recalibrates the settings or weights of the individual neurons with the aim of improving the accuracy of results.[23]

The greatest practical successes with ML to date have been in the use of *supervised* learning techniques.[24] This refers to a process that begins with a dataset that is classified or labelled by humans according to the dimension of interest, known as the *training data*. The system analyses this dataset and determines the best way to predict the relevant outcome variable (also, if necessary, classified by the experts) by reference to the other available features of the data. The nature of the features, and the relationships between them, relevant for predicting the outcomes can be exceedingly complex: the power of ML lies in

[22] http://uc-r.github.io/feedforward_DNN.

[23] See, for example, Russel & Norvig, *supra* note 15, at 733–736.

[24] Approximately 95% of all ML applications today are based on this method: Ford, *supra* note 7, at 186.

Electronic copy available at: https://ssrn.com/abstract=3442447

identifying the optimal mix of input variables. The trained model—that is, the algorithm with the set of parameters that optimized performance on the training dataset—is then put to work on a new *test* dataset, to see how effective it is at predicting outside the original training sample. These results must now be checked by human experts.

C. Tomorrow's AI?

Two other approaches to ML that are subjects of much current research activity have yet to see the same level of practical application. These are *unsupervised learning* and *reinforcement learning*.[25] Unsupervised learning relies on the model itself to identify patterns in the data. In contrast to supervised learning, this does not necessitate a labelled training data set. However, the setup of the model parameters becomes correspondingly more important in order to ensure that the resulting patterns are open to meaningful interpretation.

In a reinforcement learning setup, the algorithm learns by trial and error. It receives a 'reward' by finding the correct answer to a specified problem. A practical example is the training process relating to board games such as Go. AlphaGo by DeepMind has achieved superhuman performance within a very short time by playing very many games against itself. While reinforcement learning has much promise as a technique, its applications are currently limited to contexts in which the entire payoff structure can be specified with clarity. In most real-world settings, as opposed to game-playing, the complexity of the payoff structure rapidly exceeds the capability of hardware for the application of reinforcement learning techniques.

Another promising research strategy is to blend together the structured knowledge frameworks used in expert systems (known as 'ontologies') with ML approaches. These seek to combine the structured reasoning of the expert systems approach with the inductive engagement with data of ML, and may offer a way to move beyond the 'black box' character of current applications. [26]

---

[25] See Alpaydin, *supra* note 18, at 111–23, 125–39.

[26] See, for example, Patrick Hohenecker & Thomas Lukasiewicz, *Ontology Reasoning with Deep Neural Networks*, arXiv:1808.07980 (2018).

Electronic copy available at: https://ssrn.com/abstract=3442447

Considering the evolution of AI in historical perspective, as we have done, illustrates that the course of its development has been neither smooth nor linear. Rather, there have been a series of periods of rapid progress, each of which has (so far) engendered more excitement than was justified by the capabilities of the technology, leading inevitably to disappointment and withdrawal of funding—the so-called 'AI winter' effect. The stop-start nature of AI's development to date makes it very difficult to predict a forward trajectory towards AGI. However, the very unpredictability of the process, coupled with the speed at which development has occurred, means it is appropriate to reflect on the likely consequences of future developments, including—as we do in Section III—for corporate law.

II. TODAY'S AI AND CORPORATE LAW

Our review of AI's development suggests it is useful to consider its implications for corporate law from two perspectives on a trajectory: first, from the perspective of current technology—for which performance is a known quantity, and the implications for business organization can be derived with a reasonable degree of predictability; and second, from the perspective of future AI: a more speculative exercise justified by the significance of the potential impact. In this section, we focus on the implementation, and consequences for corporate law, of today's AI.

A. Impact of Today's AI on Business Organization

As we have discussed, the vast majority of current AI applications are supervised ML models. Now and in the near future, the focus is on applications that support and improve human decision-making (assisted AI) or enable humans to do new things (augmented AI). The development is driven by the goal of achieving productivity gains. There are many practical applications, relating to almost all aspects of corporate functions: information and communication management, investor profiling, product development, monitoring (profitability, risks, compliance etc.) and decision-making with respect to corporate strategy, investment, personnel selection, procurement, production planning and control, pricing, sales and marketing.

Electronic copy available at: https://ssrn.com/abstract=3442447

Successful preparation of a ML application depends on the availability of appropriate data.[27] There must be a large dataset, which for supervised learning models needs to have a significant component that is labelled in the dimensions of interest. As discussed, the ML model is then trained using the labelled dataset to identify variables of interest by reference to other variables correlated with them in the training data. Assembling a sufficiently large dataset, and ensuring it is labelled appropriately, is a costly exercise, meaning that there are network effects and economies of scale to the application of ML.[28] For developmental purposes, ML developers often resort to readily available records and data. However, the training data must not only be *available*, but also *appropriate*. That is, the data used to train the model must be sufficiently similar to the data the trained model is deployed to analyze to be able to give meaningful results.[29] A well-known example illustrates the way in which these challenges operate in practice. Perhaps the largest and best-known publicly available dataset of email communications is the Enron email record published by the Federal Energy Regulation Commission (FERC).[30] This has consequently been used by many developers working on a diverse range of applications. Yet for many users and their application data, this is not or only partially helpful today: the Enron dataset is almost two decades old, and the emails were written in the US business idiom of that time.

Together, these data demands of current ML technology have implications both for the kinds of questions it can be used to analyze, and how datasets are assembled. First, the feasible set of applications: if a sufficiently large dataset that is relevant for the question in issue cannot be accessed, then ML cannot be applied usefully. This implies that questions that are idiosyncratic, or arise only infrequently, are less likely to be good candidates for ML analysis. In contrast, generic and frequently-recurring issues will be more readily-susceptible.

---

[27] *Supra*, Section I.B.

[28] Hal Varian, *Artificial Intelligence, Economics, and Industrial Organization, in* Agrawal, Gans & Goldfarb, *supra* note 8. More users mean more data, smarter algorithms, i.e. a better product, again more users etc. ('data flywheel').

[29] Another way of putting this is to say that the trained model must have sufficient "external validity" with respect to the data it is used to analyze.

[30] https://www.cs.cmu.edu/~./enron/.

Electronic copy available at: https://ssrn.com/abstract=3442447

There are also implications for the way in which datasets for training ML are assembled. In most business contexts, there are unlikely to be publicly-available datasets. Firms that offer AI analytics products typically seek to scale their datasets by agglomerating the training benefits of many users' data. Such firms use feedback from the results from individual clients to improve the quality of the analysis, such that the product can achieve increasing returns to scale. This strategy is well-known in the consumer application context, where firms such as Amazon, Facebook and Google have offered users products that are free at the point of use in return for permission to harvest personal data for training purposes. However, this only works where the issues to be analyzed are ones that are comparable across clients. For example, decisions about which sales leads to follow can be made on the basis of analysis of past customer behavior and the sentiment of email traffic. Predictions based on these sorts of indicator are likely to be generic across businesses (although sentiment analysis is likely to vary regionally and intertemporally). Real-world examples include the Customer Relationship Management tool 'Einstein' by Salesforce,[31] the 'Ideal' platform for personnel selection[32] and various applications in the field of robot-controlled process automation.[33] Similarly, for the assessment of litigation risks, tools are now available that predict the outcome of disputes based on prior precedents and the records of the assigned judges and litigators.[34]

[31] https://www.salesforce.com/de/products/einstein/overview/.

[32] https://ideal.com/. On selecting corporate managers with the help of ML applications see Isil Erel et al., *Selecting Directors Using Machine Learning* (Oct. 6, 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3144080.

[33] See, for example, https://www.uipath.com/de/; https://www.hyperscience.com/; https://www.blueprism.com/; https://www.automationanywhere.com/; https://vidado.ai/; https://www.appzen.com/; https://www.workfusion.com/.

[34] See, for example, 'Legal Analytics', https://premonition.ai/legal_analytics/. Prior tools were based on decision theory ('litigation risk analysis') and did not feature a ML component. See, for example, 'TreeAge', https://www.treeage.com/, or 'Litigation Risk analysis', https://www.litigationrisk.com/. On litigation risk analysis see Horst Eidenmüller, *Prozessrisikoanalyse*, 113 ZEITSCHRIFT FÜR ZIVILPROZEß [ZZP] 5 (2000). ML tools replace human probability assessments at the chance nodes with machine knowledge.

Electronic copy available at: https://ssrn.com/abstract=3442447

Larger businesses, in particular, may have an alternative option: training an instance of an AI model on their proprietary data.[35] This may generate a more appropriate dataset, especially where the activities sought to be analyzed are idiosyncratic to the firm. It may also permit the firm to retain a source of competitive advantage, by preventing the benefits of the trained model being shared with other users. For small businesses, however, the scale of their internal data may be insufficient to train an ML model, meaning that their only option would be to provide the training benefits of their data to an AI vendor.

We set out the implications of this for firms' decisions whether to make or buy their AI analysis in Table 1:

Table 1: Dataset size and viability of AI procurement strategies

Dataset size *AI Procurement*	*Proprietary*	*Third Party*
Large	Viable/Idiosyncratic	Viable/Generic
Small	Not viable	Viable/Generic

As can be seen, where firms have sufficiently large quantities of internal data, idiosyncratic matters may also be available for ML analysis. This could be a significant source of competitive advantage.

In turn, this provokes reflection on the organization of corporate processes so as to maximize the capture of internal data. Restructuring a firm's data architecture so as to permit this is a complementary, but often non-trivial, investment that must be made to deliver maximum value from today's AI. It will be especially difficult if it involves midstream changes from 'entrenched' legacy systems, requiring sophisticated change and conflict management.[36]

[35] This would not necessitate the firm developing its own AI software. Some vendors, such as IBM, permit users to license their AI products such that the results of training are proprietary to the user.

[36] On this see CHRISTIAN DUVE, HORST EIDENMÜLLER, ANDREAS HACKE & MARTIN FRIES, MEDIATION IN DER WIRTSCHAFT: WEGE ZUM EFFEKTIVEN KONFLIKTMANAGEMENT (3rd ed. 2019).

Electronic copy available at: https://ssrn.com/abstract=3442447

Idiosyncratic matters for which insufficient internal data are available to justify ML analysis fall outside the scope of questions for which human decision-making can be directly augmented by today's AI. Into this category likely fall most strategic questions considered at the C-suite level, which will arise relatively infrequently and will typically be idiosyncratic to the firm. However, external generic data can be used to assist in scenario planning. A good example is the 'IBM Research Scenario Planning Advisor (SPA)'.[37] The factors analyzed as 'main forces' (economic environment, technology, currency, social order/unrest, corruption, natural disaster, market disruptors, government stability) are all non-firm specific. These are used to derive accounts of likely future possible trajectories, against which strategic planning can be stress-tested.

B. 'Data Governance': How Today's AI Impacts Corporate Governance

We now turn to the implications of today's AI for corporate governance. The application of AI within the firm, in the ways described in Section II A, gives rise to a number of challenging oversight issues, which collectively we describe as 'data governance'.

The oversight challenge can usefully be framed in terms of the data-dependent limitations of ML applications outlined in the previous section. A core question for those wishing to apply ML to deliver productive gains is: *are the available training data capable of providing answers to the relevant questions*? This in turn can be broken down into three distinct components: first, what is the scope of the 'available' data? Second, how can the efficacy of ML applications be maximized, and their success benchmarked? Third, how important or relevant are the predictions made by the model? And fourth, what are the implications for a firm's organizational structure? We shall consider each of these aspects in turn.

[37] https://researcher.watson.ibm.com/researcher/view_group.php?id=9444.

Electronic copy available at: https://ssrn.com/abstract=3442447

1. Scope of available data

Firms wishing to maximize their capture of internal data may need to engage in restructuring of their data architecture, as discussed above. Beyond this purely economic consideration, however, there are significant legal and ethical constraints that may also restrict the volume of data that may be captured. Legal limitations include data protection (is data acquisition compatible with the applicable rules, in particular the EU's General Data Protection Regulation?),[38] confidentiality, copyright issues, and concerns about potential liability in respect of data breaches.[39] Going beyond purely legal concerns, ethical and cultural considerations are also relevant. These may play out internally through corporate culture—effect on employee morale, productivity and retention—and externally through reputation—impacting on market participants' willingness to purchase a firm's products.[40]

2. Model selection and training

Second, what are the right ML applications for each company, and how can their performance be optimized? Answering this question requires expertise in data science.[41] For example, consider the various ways in which the success of a model's predictions against a test data set may be benchmarked. 'Accuracy' is an overall evaluation metric that takes into account both false positives and false negatives. However, 'precision' is a measure that focuses solely on false positives, whereas 'recall' focuses solely on false negatives. Which of these is more important depends on the context in which the application is being deployed. If the desire is to pre-screen materials so as to ensure human decision-makers focus their

[38] 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 2016, ("General Data Protection Regulation" or "GDPR").

[39] See e.g., William McGeveran, *The Duty of Data Security*, 103 MINN. L. REV. 1135 (2018).

[40] For an illuminating account of the situation with Uber see ALEX ROSENBLAT, UBERLAND: HOW ALGORITHMS ARE REWRITING THE RULES OF WORK (2018).

[41] On this see, for example, JOHN D. KELLEHER & BRENDAN TIERNEY, DATA SCIENCE (2018); David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653 (2017).

Electronic copy available at: https://ssrn.com/abstract=3442447

attention on cases that are relevant, then precision might seem a good metric to maximize. A high-precision pre-screening ensures that human decision-makers only consider cases that are likely to be relevant, thus making effective use of their time. However, this may be at the expense of low recall, meaning that some potentially relevant cases are (wrongly) discarded by the pre-screening process. How much this matters depends on the context. Nevertheless, appropriately benchmarking the performance of AI for the task in hand is relevant regardless of whether an AI model is trained in-house or by a third party vendor.

3. Model predictions and wider corporate goals

Third, what predictions or claims does the model make, and how interesting or important are these? In particular: what does the model say about relevant corporate goals? Within discussions on the future of AI, the issues of 'goal content' and 'goal alignment' play a central role:[42] the challenge is to determine the objective function of an application and to ensure its correspondence with desirable human values.

Above all, for our purposes, the problem is that in applying ML analytics, the focus must necessarily be on measurable outcome variables. This may exacerbate a more general tendency to focus excessively on factors it is possible to quantify such as stock prices, turnover or short-term risks—a problem of 'dashboard myopia'. Important but difficult-to-measure variables such as the external effects of corporate policies, ethical implications or long-term risks are easily ignored. This is already a problem with regard to jurisdictions in which shareholder value maximization is the only corporate goal. It becomes quite problematic when multifactorial goals are pursued or must be pursued, as in the case of the stakeholder value approach. A focus on what can easily be measured may lead to overlooking what is important.

[42] See, for example, NICK BOSTROM, SUPERINTELLIGENCE: PATHS, DANGERS, STRATEGIES 132–34 (2014).

Electronic copy available at: https://ssrn.com/abstract=3442447

4. Deployment and Organizational Structure

In addition to the issues of data ingestion, data analytics and system specification considered above, successful deployment of AI at scale requires the realignment of the firm's existing business processes to enable implementation of AI effectively in appropriate use-cases (see also Section II A *supra*).[43] This necessitates an iterative process of identifying appropriate use cases, data marshalling, model preparation and integration with on-the-ground decision-making. Such decisions are strategic, and imply a centralized, high-level function within the firm for delivering these tasks.[44] However, this needs to be done in cooperation with individual business units to identify use-cases, align data capture, and prepare employees for AI roll-out. For example, employees whose decisions will be augmented by AI will need to be trained to make use of the system, and incentivized to want to do so—which in turn may require rethinking compensation and promotion incentives.

In summary, data governance presents business executives with significant challenges in data management, the selection and training of ML applications and the alignment of ML predictions with key corporate goals. Obviously, these challenges can only be met through adjustments to business organization and processes. In particular, they imply changes in the skills and training of managers: relevant technical and analytic expertise will become increasingly important.

This point applies to all levels of the corporate hierarchy. Not only must relevant expertise be deployed for decision-making about data governance; it must also be overseen—at the most senior level, by the board of directors. This raises questions about board expertise. In a typical US public company, almost all members of the board—bar the CEO—are independent directors.[45] To what extent, though, are independent directors capable of asking meaningful questions regarding executives' management of data governance programs? Our analysis

[43] See Davenport, *supra* note 1; Ransbotham et al, *supra* note 8; Fontaine et al, *supra* note 8.

[44] Fontaine et al, *supra* note 8, at 69–71.

[45] Jeffrey N. Gordon, *The Rise of Independent Directors in the United States, 1950-2005: of Shareholder Value and Stock Market Prices*, 59 STAN. L. REV. 1465, 1476 (2007).

Electronic copy available at: https://ssrn.com/abstract=3442447

suggests that boards will increasingly find it useful to add expertise in technology to the list of capabilities they seek to retain.

The challenge of data governance also generates questions about board structure. Most US company boards have three committees: Nomination, Compensation, and Audit. Each of these deals with a particular area of conflict of interest for management, and allows for a focused and technical discussion by independent directors with appropriate expertise. This logic might be thought to imply a strong case for establishing a distinct committee for data governance, staffed by independent directors recruited for their expertise in technology. However, other committees are rare in US public company boards.[46] As of 2013, only three per cent had established a Technology Committee,[47] and as of 2017, less than seven per cent had established a Compliance Committee.[48] In many instances, compliance-related issues are often added to the (already long) list of tasks overseen by the Audit Committee.[49] This seems an unsatisfactory way to deal with the data governance questions we have considered.

Moreover, there are reasons for thinking that encouraging independent directors to engage in the oversight of data governance may be capable of adding value to corporate decision-making, rather than simply the addition of further costly compliance processes. It is well-known that empirical corporate governance literature struggles to find a relationship between director independence and firm performance. This is primarily due to problems of measurement: board structure is endogenous to firm performance—that is, causality may run both ways.[50] However, there are theoretical challenges as well: independent directors must

---

[46] See Kevin D. Chen & Andy Wu, *The Structure of Board Committees*, HARVARD BUSINESS SCHOOL WORKING PAPER 17-032, 24 (2016), http://www.hbs.edu/faculty/Publication%20Files/17-032_22ea9e7a-4f26-4645-af3d-042f2b4e058c.pdf; John Armour et al., *Board Compliance*, 104 MINN. L. REV. (forthcoming 2019).

[47] *Chen & Wu, supra* note 46.

[48] *Armour et al*., *supra* note 46.

[49] See John Armour, Jeffrey Gordon & Geeyoung Min, *Taking Compliance Seriously*, 36 YALE J. REG. (forthcoming 2019).

[50] See e.g. Renée B. Adams, Benjamin E. Hermalin and Michael S. Weisbach, *The Role of Boards of Directors in Corporate Governance: A Conceptual Framework and Survey*, 48 J. ECON. LIT. 58 (2010).

Electronic copy available at: https://ssrn.com/abstract=3442447

walk a tightrope between being sufficiently disinterested in a firm's business to be classed as "independent" while, at the same time, sufficiently engaged to be capable of adding value.[51] However, we have argued that the data demands of successful ML models mean that the kinds of issue these models will be used to decide are likely to be more generic than idiosyncratic.[52] Consequently, independent directors with generic capabilities may be well-placed to provide effective oversight of these processes.

C. Data Governance and Corporate Law

As we have seen, data governance raises significant opportunities for productivity gains and value-creation, provided the issues are appropriately handled. This in turn raises questions regarding oversight. What implications, if any, does this have for corporate law?

Augmenting human decision-making means that fewer human beings are needed to deliver the same results. This may be expected to reduce some types of agency costs within firms; at the same time, new types of discretionary decision-making become important—the decisions involved in establishing and testing the automated systems themselves.[53] To the extent that decisions about the deployment of AI result in a decrease in a firm's human workforce, the remaining agency costs of human decision-making will increasingly relate to 'strategic' choices—such as the data governance issues discussed above. These will have potentially far-reaching consequences for corporate performance, such that the aggregate agency costs may not decrease; rather just change shape.[54] Identifying where to monitor these, and how best to do it, will be an increasingly complex and

[51] See Kraakman et al., *supra* note 4, at 65 with further references.

[52] *Supra*, Section II A.

[53] See also Luca Enriques and Dirk Zetzsche, *Corporate Technologies and the Tech Nirvana Fallacy* (May 31, 2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3392321.

[54] To illustrate this point, managers/board members might (ab)use AI applications to *entrench* their position. Alternatively, managers/board members might use a surveillance algorithm in clear breach of applicable data privacy rules.

Electronic copy available at: https://ssrn.com/abstract=3442447

important task. The high-level implication is that this will place increasing demands on 'oversight' at the top of the firm.

Boards are formally responsible for oversight of corporations,[55] and directors owe their firms fiduciary duties of loyalty.[56] Where decision-making has been delegated—as is typical in US public company boards—the board retains residual duties of oversight. The duty was set out by Chancellor Allen in his well-known opinion in *Caremark*,[57] in which he explained that boards needed to assure themselves that their firm had:

> [I]nformation and reporting systems … that are reasonably designed to provide to senior management and to the board itself timely, accurate information sufficient to allow management and the board, each within its scope, to reach informed judgments concerning both the corporation's compliance with law and its business performance.[58]

Should such a system of oversight give an indication of problems—a so-called 'red flag'—then the board are expected to take steps to investigate and take remedial action.[59] However, all aspects of this 'oversight duty'—both to ensure some system of oversight exists, and to take action if it flags a problem—impose only obligations of good faith. This means that liability is only triggered by a failure so egregious

---

[55] See, for example, DGCL § 141(a) ("The business and affairs of every corporation organized under this chapter shall be managed by *or under the direction of* a board of directors…") (emphasis added).

[56] Stone ex rel. AmSouth Bancorporation v. Ritter*,* 911 A.2d 362, 370 (Del. 2006); Accord Mills Acq. Co. v. Macmillan, Inc.*,* 559 A.2d 1261, 1280 (Del. 1989) ("[D]irectors owe fiduciary duties of care and loyalty to the corporation and its shareholders."); Polk v. Good*,* 507 A.2d 531, 536 (Del. 1986) ("In performing their duties the directors owe fundamental fiduciary duties of loyalty and care to the corporation and its shareholders.").

[57] In re Caremark International Inc., Derivative Litigation, 698 A.2d 959 (Del. Ch., 1996). See generally Hillary A. Sale, *Monitoring Caremark's Good Faith,* 32 DEL. J. CORP. L. 719 (2007) (detailing the history of *Caremark*); Jennifer Arlen, *The Story of Allis-Chalmers, Caremark and Stone: Directors' Evolving Duty to Monitor, in* CORPORATE LAW STORIES 323 (J. Mark Ramseyer ed., 2009) (same); Donald C. Langevoort, *Caremark and Compliance: A Twenty Year Lookback*, 90 TEMP. L. REV. 727 (2018) (review of subsequent developments).

[58] *Caremark, supra* note 57, 698 A.2d at 970.

[59] Graham v. Allis-Chalmers Mfg. Co., 188 A.2d 125 (Del. 1963); In re Massey Energy Company Derivative & Class Action Litigation, 2011 WL 2176479 (Del. Ch. 2011); Melbourne Municipal Firefighters' Pension Trust Fund v. Jacobs, 2016 WL 4076369 (Del. Ch. 2016); Oklahoma Firefighters Pension & Retirement System v. Corbat 2017 WL 6452240 (Del. Ch. 2017); In re Wells Fargo & Co. Shareholder Derivative Litigation (N.D. Cal., 2017); In the matter of Wells Fargo & Company, Federal Reserve Docket No 18-007-B-HC (2018).

Electronic copy available at: https://ssrn.com/abstract=3442447

as to call into question the board's good faith.[60] The rationale is that the board, not the court, knows best how to pursue the firm's internal compliance activities.[61]

This rationale is premised on the idea that the decisions being reviewed are matters of business judgment. This is true for some aspects of data governance—for example, the question whether a firm should buy ML analytics from third party providers, or make its own in-house, is a classic business strategy decision. However, to the extent that data governance involves the management of potential liabilities for the firm—as with privacy or copyright violations—then it encompasses issues of compliance with obligations imposed by law to protect the interests of persons outside the firm. Similarly, when it comes to assessment of whether the goals being pursued by the firm's data-driven decisions are the right ones. One of us has argued elsewhere that a duty framed solely in terms of good faith does little to incentivize directors to take compliance seriously.[62] To the extent that poor data governance results in externalities, a similar analysis also applies here.

The Delaware Supreme Court has recently emphasized that directors' oversight duty is non-delegable: directors must ensure that there is some *board-level* mechanism for overseeing the operation and functioning of internal compliance programs.[63] In our view, directors can properly perform their oversight duty only if they are familiar with the main issues of data governance in relation to ML applications discussed above in Section II B. In other words, what we have suggested top executives need to know in the future in order to deploy AI and especially ML applications productively in the firm is also relevant to determining due care for boards in the selection, instruction and monitoring of specialist management personnel. The non-delegable nature of the oversight duty implies

[60] The necessary degree of oversight failure to trigger liability was later characterized by the Delaware Supreme Court as an "utter fail[ure] to implement any reporting or information controls", Stone v. Ritter, 911 A.2d 362, 370 (Del. 2006).

[61] As Chancellor Allen put it in *Caremark:* "Obviously the level of detail that is appropriate for such an information system is a question of business judgment" (*Caremark, supra* note 57, 698 A.2d at 970).

[62] Armour et al., *supra* note 49.

[63] *Marchand* v. *Barnhill* (Del. 2019) 2019 WL 2509617.

Electronic copy available at: https://ssrn.com/abstract=3442447

that, over time, boards will increasingly be expected to demonstrate engagement with data governance issues in order to comply with it.

III. TOMORROW'S AI AND CORPORATE LAW

The focus of Sections I and II was on the implications of AI-assisted and AI-augmented decision-making processes for corporate governance and corporate law. However, AI functionality is continuously evolving. Going beyond assisted and augmented intelligence, autonomous AI decision-making ("autonomous intelligence") for certain limited functions is already a distinct possibility today. In this section, we shift focus to the future of the corporation with increased AI functionality.

At some point, we may see humans on boards of corporations being replaced by algorithms. This may happen sooner than the technical breakthroughs necessary for AGI: if the business activity of the company in question is sufficiently limited, then an AI application that is sub-human level in a number of areas of performance but super-human in others might nevertheless be valuable, provided that the necessary decisions harness the super-human capacity. We argue that this could occur first—indeed, sooner rather than later—in the context of subsidiaries with very narrowly-defined domains of operation. Hence, in a first step, we consider the 'self-driving subsidiary'.

The analysis will demonstrate that with further advances in technology, the function of corporate law in AI-controlled corporations will also change. Instead of facilitating productive exchange, it will increasingly be concerned with the regulation of corporate activities. In particular, we suggest that calibrating corporate objectives will become a central issue, which will have important regulatory implications. We explore potential regulatory responses to what we describe as 'algorithmic failure' in an environment likely characterized by intense regulatory competition.

Electronic copy available at: https://ssrn.com/abstract=3442447

A. Self-Driving Subsidiaries

Self-driving subsidiaries can be regarded as an intermediate step on the road towards a fully autonomous or self-driving (independent) corporation. Many jurisdictions around the world today do not demand that a corporation is run by human directors.[64] For example, the Cayman Islands Companies Law contains no requirement that any company directors be a natural person.[65] While Delaware General Corporation Law (DGCL) §141(b) provides that members of a Delaware corporation's board of directors must all be natural persons, § 141(a), which enshrines the management power of the board of directors, may be waived in a corporation's charter.[66] On one reading, it is permissible to waive the requirement that the company have a board altogether.[67]

To date, AI applications—based mainly on ML—relate to specific entrepreneurial functions (see Section I B *supra*). Integrated applications for an overall corporate management do not currently exist. It is therefore to be expected that automated corporate management might be tried for the first time for companies with limited functions. One can think, for example, of special purpose vehicles (SPVs) in the financial sector ('self-driving securitization'). In the transport sector, too, opportunities for such vehicles exist. Autonomous driving is already a reality today. A self-driving taxi that can be ordered with an application could be operated as an incorporated enterprise. It is then not only the motor vehicle that is 'autonomous', but also the corporation which steers the taxi's deployment and the associated entrepreneurial functions—such as order processing, billing etc.—and thus itself with an algorithm.

---

[64] For an overview see Shawn Bayern et al., *Company Law and Autonomous Systems: A Blueprint for Lawyers, Entrepreneurs, and Regulators*, 9 HASTINGS SCI. & TECH. L.J. 135 (2017).

[65] See Cayman Islands Company Law, §55.

[66] The exact wording of the provision is as follows: "§ 141 Board of directors; powers; number, qualifications, terms and quorum; committees; classes of directors; nonstock corporations; reliance upon books; action without meeting; removal. (a) The business and affairs of every corporation organized under this chapter shall be managed by or under the direction of a board of directors, *except as may be otherwise provided in this chapter or in its certificate of incorporation*. …" (emphasis added).

[67] See LoPucki, *supra* note 3, at 907–10.

Electronic copy available at: https://ssrn.com/abstract=3442447

The 'business case' for this organizational set-up is straightforward: the deployment of autonomous AI—if it works—brings significant productivity gains. At the same time, incorporating a specific business function—such as transportation by a vehicle—in a separate corporation increases the transparency of the activity while also containing liability risks for the entrepreneur.

If one imagines that taxi companies will in future organize their self-driving taxi fleet as a group of incorporated individual taxis operating as self-driving subsidiaries, then the core corporate law issue can be clearly identified. It is about the asymmetry of risks and liability substrate: costs are partially externalized if and to the extent the corporation as liability addressee is 'judgment proof' due to possibly insufficient assets. Given limited liability, shareholders—and managers acting on their behalf—have an incentive to engage in inefficient risk-taking (agency costs of debt).[68] Of course, this (liability) problem is not new. But it could become supercharged in corporations with autonomous AI decision-making.

In a corporation managed by humans, managerial liability is one key mechanism available to control externalities caused by inefficient risk-taking. This device is removed once a corporation runs autonomously, controlled by an algorithm.[69] A fully self-driving corporation can cause high damage without humans being involved in the decision-making chain, and without human managers that could be held accountable for the firm's actions. Creditors will then usually attempt to go beyond the assets of the corporation and target its shareholders. In the case of abusive schemes, courts may allow them to do just that under doctrines such as piercing of the corporate veil.[70] However, applying this doctrine is not an efficient solution in 'ordinary' cases of algorithmic failure, i.e. if no element of abuse of the corporate form is present. Some other regulatory approach is needed to address such failure. For a meaningful discussion of such an approach, we first need to better understand the regulatory problem.

---

[68] See, for example, Bachmann et al., *supra* note 4, at 11–13.

[69] See Martin Petrin, *Corporate Management in the Age of AI,* forthcoming COLUM. BUS. L. REV. (2019).

[70] See Kraakman et al., *supra* note 4, at 131–34; *Hisaei Ito & Hiroyuki Watanabe (eds.)*, *Piercing the Corporate Veil, in* COMPARATIVE COMPANY LAW: A CASE-BASED APPROACH 211–49 (Mathias Siems & David Cabrelli eds., 2nd ed. 2018).

Electronic copy available at: https://ssrn.com/abstract=3442447

B. Calibrating Corporate Objectives for Self-Driving Corporations

In the context of a self-driving subsidiary, but also beyond, i.e. independent of the use of AI/ML applications for managing subsidiaries in groups of companies, calibrating the goals of a corporation will become a central corporate governance question with autonomous AI decision-making.

We addressed the problem already in the introduction to this article and then in the context of the discussion of important corporate law problems of data governance (see Section II C *supra*). As corporate managers, human actors pursue their own interests, which may differ from those of other stakeholders (shareholders, creditors, third parties, etc.). The consequences are agency issues and costs. By contrast, algorithms pursue set goals, precisely and thoroughly. Algorithms also do not act in bad or good faith.[71] Rather, they simply optimize a given goal function.

In the long run, we can expect algorithms to also manage corporations outside the group context and substitute human directors (if this is permitted by the applicable corporate law rules): if and to the extent that algorithms perform better than humans, it would be inefficient to continue using the latter.[72] Competition pushes businesses to maximize profits and will force them, at some point, to operate without human decision-makers.[73]

---

[71] Florian Möslein, *Robots in the boardroom: artificial intelligence and corporate law, in* RESEARCH HANDBOOK ON THE LAW OF ARTIFICIAL INTELLIGENCE 649, 666 (Woodrow Barfield & Ugo Pagallo eds., 2018) is too imprecise on this point ("In the same vein, a breach of [robots'] fiduciary duty is *unlikely*" [emphasis added]).

[72] Hence, we do not expect that market forces will push firms to 'outsorce the board' to professional (corporate) service providers (but see STEPHEN M. BAINBRIDGE & M. TODD HENDERSON, OUTSOURCING THE BOARD (2018)). Rather, we expect the board to be substituted by machines.

[73] The *factual* substitution of human decision-makers because of the competitive dynamic described in the text should not be confused with a *normative* substitution based on legislative fiat. Such normative substitution will become relevant for many areas of our daily lives and AI applications at some point in time (by way of example: When will we no longer allow humans to drive cars?), see Horst Eidenmüller, *Machine Performance and Human Failure, in* AUTONOMOUS SYSTEMS AND THE LAW 75 (Nikita Aggarwal et al. eds., 2019); Horst Eidenmüller, *Machine Performance and Human Failure: How Shall We Regulate Autonomous Machines?* (July 14, 2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3414602, 15 J. OF BUS. & TECHN. LAW (forthcoming 2020). With respect to using AI as a corporate management tool the normative question loses much of its relevance because of the factual development described.

Electronic copy available at: https://ssrn.com/abstract=3442447

Instead of agency problems, it is then goal specification and goal setting that assume the greatest importance. Which goal or goals may or should the algorithm pursue, and how should any goal conflicts be resolved? It is obvious that these questions are relatively easier to answer in jurisdictions with a one-dimensional financial target such as shareholder value maximization than in jurisdictions with more complex benchmarks for corporate managers (stakeholder value maximization or similar, see above Section II B).

For example, under current German law the management board of a stock corporation must aim to act in the 'best interests of the company' (§ 93 (1) sentence 2 AktG). This is a vague goal which gives the board some leeway when taking specific actions. This leeway is further increased by the business judgment rule for the choice of means. In the future, it will be possible to determine an algorithmic goal function specifically and precisely and thus the exact content of the 'best interests of the company': to what extent may/should the development of the share price be the goal of entrepreneurial action? Which time horizon should be used? What other goals should be pursued, and how should the achievement of goals be measured? How are goal conflicts to be resolved? Which additional constraints must be observed?

It is not obvious that legislators will or should allow company-controlling algorithms the same discretion as human decision-makers today. If the (complex) goal function of an algorithm can be precisely defined *ex ante*, we can expect an intensive and controversial policy discussion on goals to take place. This is especially true due to the rigidity with which algorithms pursue set goals. It not only carries the risk that extremely one-sided goals will be pursued with utmost effectiveness. Above all, it raises the risk of 'algorithmic failure', which can never be completely ruled out (see below). For example, do we want to allow a self-driving corporation to operate a nuclear power plant and thereby primarily push for a short-term share price increase? The problem of the correct calibration of a corporation's goals and, most importantly, corporate control and liability regimes to address 'algorithmic failure' pose major regulatory challenges to (corporate) law, which we turn to in the final sections of this article.

Electronic copy available at: https://ssrn.com/abstract=3442447

C. Alternative Control and Liability Regimes

'From the control of managerial action to the control of algorithms'—this is how, in a nutshell, the regulatory problem can be described. If the exact *ex ante* specification of the algorithmic goal function comes to the foreground of the corporate governance relevant questions, the question of liability for deficient/incorrect specifications will inevitably arise. In a first step, we will therefore first consider alternative control and liability regimes, assuming that the applicable corporate law does no longer mandate human directors.[74] Next, we look at the expected global regulatory competition in the field of self-driving corporations.

According to today's corporation laws in many jurisdictions, setting up a corporation is an act of right if a set of clearly defined stipulations are met.[75] The conferral of legal capacity is therefore not an act at the discretion of the regulatory authority; rather, it is the automatic consequence of the absence of clearly defined legal impediments.

This system does not have to be shelved with respect to a fully self-driving corporation. In particular, it would be wrong to grant the registrar the discretionary competence to admit only 'socially acceptable' algorithmic entities and to monitor them accordingly in terms of their business activity. Giving registrars such a discretionary competence would create uncertainty and, as a consequence, unnecessary transactions and opportunity costs (because it would deter risky but profitable investments). It might also be considered problematic from a human rights perspective as it could lead to arbitrary restrictions of entrepreneurial activity.

---

[74] This is not an unreasonable assumption. The 'traditional' case for mandating human directors rests on grounds of transparency and accountability. This case is weak, see Bainbridge & Henderson, *supra* note 72, at 137-152. It is also inconsistent with allowing corporate entities to act as general partner in limited liability partnerships—a common feature of most corporate law systems in the world. Jurisdictions might nevertheless continue to mandate human directors—who fit, for example, certain gender or ethnicity criteria—for reasons other than transparency or accountability.

[75] Under German law, this has come to be known as *System der Normativbestimmungen* ('system of normative stipulations'). On the 'concession theory' and how it has developed over time see Susan M. Watson, *The corporate legal person*, 19 J. CORP. L. STUD. 137, 142–45, 162–64 (2019).

Electronic copy available at: https://ssrn.com/abstract=3442447

Consideration should, however, be given to adapting the registration requirements such that an algorithmically controlled corporation must act within the framework of a given corridor of permissible corporate objectives and goal functions.[76] Conformity with this corridor would have to be declared by the founders when the company is set up, and it would have to be maintained during the life of the company. Otherwise, the company would have to be compulsorily wound up and struck off the company register.

Even an algorithm that is correctly geared to a socially desirable target within the framework of a set of permissible corporate objectives can make mistakes in individual decisions and possibly even act unlawfully. Such unlawful action may be defined as 'algorithmic failure'.[77] Algorithmic failure can, in particular due to the rigidity with which the algorithm pursues its objectives, have serious negative effects—physical harm or financial loss—on business partners of the company, but also on third parties. Who should then be liable, and how should this liability be designed?

As a starting point, responsibility for algorithmic failure should, in our opinion, lie with the entity using it, i.e. the corporation itself, and not with any third party vendors or an AI platform. This is obvious if the AI is a proprietary model developed by the corporation. But even with external input, it is ultimately the corporation itself that is responsible for its own data governance (see above Section II B/C). The corporation decides on the design/specification of algorithms, their deployment, their interaction etc., and it benefits from them. Consequently, the corporation is likely to be the 'least-cost avoider';[78] that is, able to prevent

[76] Given that we are talking about corporate entities that are fully controlled by algorithms, this requirement is equivalent to stipulating that the controlling algorithm must pursue (an) objective(s) within a range of permissible corporate objectives and goal functions.

[77] It does of course depend on the laws applicable in a particular jurisdiction whether and when algorithmic action(s) could be characterized as 'unlawful', whether such action(s) caused damage, what the appropriate measure of damages is, who bears the burden of proof etc. We do not express an opinion on these issues.

[78] See, for example, EJAN MACKAAY, LAW AND ECONOMICS FOR CIVIL LAW SYSTEMS 218–21 (2013).

Electronic copy available at: https://ssrn.com/abstract=3442447

algorithmic failure at lowest cost.[79] Allocating liability to the least-cost avoider is generally efficient.

If the corporation were to be held strictly liable for any damage caused by algorithmic failure it would be forced to internalize all costs of algorithm-powered activities. By contrast, a regime based on liability (only) for negligence would in many cases leave an injured party without compensation.[80] Further, determining the (efficient) standard of care can be exceedingly difficult, and a negligence regime does not control the 'activity level' of a dangerous or risky activity.[81] On the other hand, strict liability undoubtedly reduces incentives for companies to innovate.[82] Here, jurisdictions will have to make a critical (policy) decision against the background of the respective willingness to innovate or to take risks. At least in the European Union (EU), the regulatory tradition suggests strict liability for hazardous assets or activities.[83]

Even strict liability of the corporation for damage caused by algorithmic failure would not be a satisfactory solution to the regulatory problem described above without changing the applicable legal framework. This is because of the already mentioned 'judgment-proofing' issue: the corporation may well lack sufficient funds to compensate all those damaged by algorithmic failure. Per the assumptions in this section, a fully self-driving corporation also does not have human directors who could be held liable. This leaves us essentially with two

---

[79] The situation is similar to the liability problem with respect to self-driving cars. Here it is the manufacturer who is the least-cost avoider. On this see Horst Eidenmüller, *The Rise of Robots and the Law of Humans*, 25 ZEITSCHRIFT FÜR EUROPÄISCHES PRIVATRECHT [ZEUP] 765, 772 (2017); Gerhard Wagner, *Robot Liability* (July 14, 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3198764.

[80] For a general discussion, see Steven Shavell, *Strict Liability versus Negligence*, 9 J. LEG. STUD. 1 (1980).

[81] For these reasons, it has been argued that liability for accidents involving self-driving cars should be strict, see Eidenmüller, *supra* note 79, at 772.

[82] See Armour et al., *supra* note 6, at 294–95.

[83] Council Directive 85/374/EEC of Jul. 25, 1985 establishes a (strict) liability regime for defective products. However, under Article 7(e) the producer shall not be liable as a result of this Directive if he proves "… that the state of scientific and technical knowledge at the time when he put the product into circulation was not such as to enable the existence of the defect to be discovered." We do not think that it would be wise if this defense were allowed under the liability regime proposed in the text. It transforms a *de iure* strict liability regime into a *de facto* negligence regime.

Electronic copy available at: https://ssrn.com/abstract=3442447

regulatory options: improving the corporation's loss-absorbing capacity in cases of algorithmic failure or some form of shareholder liability under a veil-piercing doctrine or similar mechanism.[84]

In many jurisdictions worldwide, notably in the EU, rules on legal capital of a corporation (minimum capital requirements, distribution restrictions) are considered to be an important creditor protection device.[85] However, these rules do not provide sufficient protection for creditors in the event of a crisis as losses will have diminished or even eliminated the capital cushion at that point in time.[86] A tightening of legal capital rules by, for example, stipulating a very high minimum capital requirement, is neither realistic nor would it be sensible as it would erect a significant barrier to entrepreneurship and technological innovation.

On the other hand, it would be realistic and sensible to require self-driving corporations to purchase some form of liability insurance with a legally prescribed minimum amount of coverage, possibly graduated according to a corporation's line of business.[87] This would allow private insurance companies to price the company-specific risk of algorithmic failure of fully self-driving corporations. Given their market knowledge, these private insurance companies are typically better placed to accomplish this task than governmental agencies. 'Safe' corporations would benefit from lower premia and thus receive a subsidy for their business model. At the same time, such mandatory insurance should not constitute a significant market entry barrier.

---

[84] A third option would be to focus liability on vendors of AI decision-making software used in the corporate context: see Petrin, *supra* note 69. However, vendor liability is not an efficient solution, see *supra* in the text. Further, it would also necessitate a form of veil-piercing, because imposing liability on the vendor would require courts to disregard limitations of liability established in the contract between the vendor and the corporation deploying the AI.

[85] The key statutory instrument in the EU is the Directive 2017/1132 of Jun. 14, 2017. For a comparative perspective on EU, German, UK and US law see ANDREAS CAHN & DAVID C. DONALD, COMPARATIVE COMPANY LAW 254–85 (2nd ed. 2018). For the US, see BAYLESS MANNING & JAMES J. HANKS, LEGAL CAPITAL (4th ed. 2013).

[86] On the lack of functionality of legal capital as a creditor protection instrument, see the contributions in THE LAW AND ECONOMICS OF CREDITOR PROTECTION: A TRANSATLANTIC PERSPECTIVE (Horst Eidenmüller & Wolfgang Schön eds., 2008); Kraakman et al., *supra* note 4, at 124–27.

[87] For this regulatory proposal regarding self-driving cars, see Eidenmüller, *supra* note 79, at 772–73.

Electronic copy available at: https://ssrn.com/abstract=3442447

As a regulatory alternative, a general liability of the shareholders of fully self-driving corporations for corporate debts should at least be considered. Some caution is warranted: the principle of limited liability is economically significant. It supports capital collection and risk diversification and thus risky investments in innovative businesses. This is especially important in the context of the development of AI. On the other hand, if some form of abuse—of limited liability—is present, piercing the corporate veil is also economically justified (see above Section III A). Serious algorithmic failure would lead to mass tort claims. It has long been argued convincingly in the scholarly literature that an unlimited pro rata liability of shareholders vis-à-vis tort creditors is well founded.[88] On the other hand, despite attempts to increase transparency it will probably become more and more difficult to identify sufficiently capitalized shareholders: by means of cross-shareholdings and pyramid constructs involving foreign legal entities, the traces of initiators and beneficial owners can easily be blurred.[89] The compulsory insurance regime discussed above does not suffer from this problem. Registration of a self-driving corporation without human directors could be made dependent on proof of the legally required insurance coverage.

D. Regulatory Competition

It is highly unlikely that the regulatory measures discussed in the previous section for controlling fully self-driving corporations will receive immediate worldwide support. Rather, we expect stiff regulatory competition with a tendency towards lax regulations (race to the bottom), at least in the short run. Many countries have made AI and related investment a strategic priority.[90] Investors are not attracted by harsh regulation. For example, Switzerland fears that the status

[88] Henry Hansmann & Reinier Kraakman, *Toward Unlimited Shareholder Liability for Corporate Torts*, 100 YALE L. J. 1879 (1991).

[89] See LoPucki, *supra* note 3, at 919–24.

[90] This includes China in particular. If it is true that supervised learning-applications will dominate the ML landscape in the near future, then China has a clear competitive advantage because of the wealth of available data and its lax data protection rules. On this, see generally KAI-FU LEE, AI SUPERPOWERS: CHINA, SILICON VALLEY, AND THE NEW WORLD ORDER (2018).

Electronic copy available at: https://ssrn.com/abstract=3442447

of the industrial district in Zug as 'Swiss Crypto Valley' could lose out to new, even laxer laws in other countries.[91]

Private international law rules are not functional to counteract this development. More specifically, the so-called 'real seat theory' does not offer meaningful protection to those who suffer loss as a result of self-driving corporations. According to this theory, a corporation is subject to the laws and regulations of the jurisdiction in which it has its real seat. This is the location where the day-to-day management of the corporation takes place. However, the real seat of a self-driving corporation can strategically be arranged to coincide with the place of registration. As a consequence, lax registration rules of an incorporating state can easily be exported. Whether recognition obligations dictate the application of the laws of the incorporating state—which is often the case under supranational rules (such as in the EU) or bilateral/multilateral trade agreements—therefore is only of secondary importance.[92]

To rely on voluntary regulatory cooperation between states, on the grounds that algorithmically controlled corporations are said to possess a global threat potential, is naive.[93] Coordinated regulation is realistic only in supranational legal systems such as in the EU, which can pass harmonization regulations with binding effect for all Member States. Corresponding EU regulatory initiatives in the area of AI have been launched.[94] Apart from that, the development of model laws or rules that reflect best practice would certainly be a sensible undertaking within the framework of organizations such as UNCITRAL or UNIDROIT. Such projects should be initiated.

A key question is how popular incorporation states such as Delaware will react to this development. Joining the regulatory race to the bottom would be strategically unwise. Rather, given Delaware's status as premier supplier of

---

[91] See, for example, https://www.blick.ch/news/wirtschaft/experte-zur-schweiz-als-zentrum-der-krypto-welt-blockchain-koennte-100000-arbeitsplaetze-schaffen-id15063854.html; Andrew Edgecliffe-Johnson, *The Canadian brain gain*, FINANCIAL TIMES, Feb. 27, 2019, at 9.

[92] For a comprehensive treatment of the subject matter, see the contributions in AUSLÄNDISCHE KAPITALGESELLSCHAFTEN IM DEUTSCHEN RECHT (Horst Eidenmüller ed., 2004).

[93] But see JACOB TURNER, ROBOT RULES: REGULATING ARTIFICIAL INTELLIGENCE 237–47 (2019).

[94] For an overview, see https://ec.europa.eu/digital-single-market/en/artificial-intelligence.

Electronic copy available at: https://ssrn.com/abstract=3442447

corporate laws for businesses, the state might undertake to relaunch the facilitative model of corporate law in an age of AI by, for example, offering the 'legal operating system' for AI applications that sit on corporate laws, work with AI developers to create Delaware specific applications etc. The aim could be to preempt domestic or coordinated international regulation of self-driving corporations by demonstrating that the market is capable of producing more functional solutions for businesses in an age of AI than regulators could. Whether this strategy works—for Delaware or any other state—will probably much depend on the regulatory environment and traditions, especially on the ease of federal or international regulatory intervention.

CONCLUSION

The use of artificial intelligence (AI) is changing the way we live and work. It will also strongly influence corporate governance and corporate law over the years and decades to come. In this essay, we have analyzed in detail the likely effects starting from assisted and augmented intelligence to autonomous intelligence: the self-driving corporation is visible on the horizon.

Of fundamental relevance is the fact that 'traditional' agency problems in corporations will increasingly lose importance as an analytical issue: algorithms optimize a given goal function. At the same time, new oversight challenges and, as a consequence, 'strategic' agency costs are created at the top of a firm's hierarchy. Further, with increasing AI functionality, a shift from viewing the enterprise as primarily private and facilitative, towards a more public, and regulatory, conception of the law governing corporate activity is likely. Calibrating corporate goals and providing redress for 'algorithmic failure' will emerge as key regulatory challenges. The most important results of our contribution can be summarized as follows.

Overall, the implications of AI for corporate governance and corporate law are scientifically and practically significant. Our goal has been to analyze and assess these implications. This in turn necessitated opening the black box of AI applications and to study these applications in detail.

Electronic copy available at: https://ssrn.com/abstract=3442447

In the foreseeable future, machine learning (ML), in particular so-called supervised learning, will dominate AI applications for businesses. Such applications require extensive, labelled datasets—a continuous 'data pipeline'. This raises key 'data governance' challenges for corporate directors and managers. These relate in particular to the scope and quality of the (factually and legally) available data, the selection of suitable ML applications and the training of these applications, as well as the alignment of model forecasts with the (further) objectives of the company.

Meeting these challenges—and delivering the productivity gains made possible by AI applications—will require adjustments to a firm's data architecture. It will also pose new oversight tasks for board members. A case can be made for establishing a distinct committee for data governance, staffed by independent directors recruited for their expertise in technology. Directors' liability for deficient oversight should not be curtailed by the business judgment rule if and to the extent poor data governance results in externalities.

In the medium term, full corporate management by algorithms is most likely to be possible in the first instance for group-affiliated corporations with a narrowly limited business function. The 'self-driving subsidiary' raises well-known legal problems such as inefficient risk-taking by shareholders/managers, 'judgment proofing' and shareholder liability under a veil piercing doctrine for abuses of the corporate form. However, these problems could become supercharged in corporations with autonomous AI decision-making.

With a further increase in AI functionality, determination of the goal function of self-driving corporations will become the central corporate governance question in the medium to long term, even outside the group context. 'Traditional' agency problems will, on the other hand, take a back seat.

Adequate protection against algorithmic failure of a self-driving corporation could be ensured by an *ex ante* control of the (eligible) AI applications (corporate objectives) and strict liability of the corporation combined with firm-specific mandatory liability insurance. The main regulatory alternative is unlimited pro rata liability of the shareholders vis-à-vis tort creditors.

Electronic copy available at: https://ssrn.com/abstract=3442447

The regulation of corporate governance related AI applications is expected to be marked by global regulatory competition. Only in supranational legal systems such as the EU are certain harmonization projects to be expected. However, the development of model laws or best practice rules within the framework of institutions such as UNCITRAL or UNIDROIT should also be initiated. Leading providers of corporate laws such as Delaware might seek to preempt regulatory intervention by reinventing the facilitative model of corporate law in an age of AI.

While the state of current technology means there is no immediate need for regulatory action, the development of AI in general and of applications for corporate governance is progressing rapidly. In a few years, things are likely to look different. Given the extent of the likely challenges, it is important for lawmakers to begin to consider these issues well in advance, before tomorrow's AI becomes today's.

Just as companies will only achieve productivity gains from AI if business models are adapted and the people in charge improve their technical understanding and ability, lawyers will only be able to make good AI-related laws if they are at least familiar with AI basics and especially with ML applications. Here lies the most pressing task: you cannot meaningfully regulate AI, if you do not understand it.

In 1897, Oliver Wendell Holmes famously suggested that "[f]or the rational study of the law the black-letter man may be the man of the present, but the man of the future is the man of statistics and the master of economics."[95] When the economic analysis of law began to influence legal scholarship and legal education in the 1970s, we were reminded of his dictum—decades later. Holmes did not mention technology. It should be mentioned today, and we think it should not take decades before its impact on the study of the law becomes visible.

[95] Oliver W. Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 469 (1897).

Electronic copy available at: https://ssrn.com/abstract=3442447

european corporate governance institute

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.global

Electronic copy available at: https://ssrn.com/abstract=3442447

european corporate governance institute

ECGI Working Paper Series in Law

Editorial Board

Editor	Amir Licht, Professor of Law, Radzyner Law School, Interdisciplinary Center Herzliya
Consulting Editors	John Coates, John F. Cogan, Jr. Professor of Law and Economics, Harvard Law School
	Horst Eidenmüller, Freshfields Professor of Commercial Law, University of Oxford
	Curtis Milhaupt, Professor of Law, Stanford Law School
	Niamh Moloney, Professor of Law, Department of Law, London School of Economics and Political Science
Editorial Assistants	Alison Schultz, University of Mannheim
	Julian Hanf, University of Mannheim

www.ecgi.global/content/working-papers

Electronic copy available at: https://ssrn.com/abstract=3442447

european corporate governance institute

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html

www.ecgi.global/content/working-papers

Electronic copy available at: https://ssrn.com/abstract=3442447