Self-Driving Corporations?

Posted by John Armour and Horst Eidenmueller (University of Oxford), on Tuesday, October 7, 2019

In a recent essay, we explore the implications of artificial intelligence (AI) for 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. 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? The standard account is functional in its orientation; that is, it is premised on a social-scientific analysis of what actually happens in a business firm. The starting point for our inquiry is therefore to ask how AI will affect the activities of firms. As a preliminary step, we begin with a clear account of what is technically 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.

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. 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. 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 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. 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 more 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". 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 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.

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

The complete paper is available for download here.