2021

Artificial Intelligence and the Rule of Law

Aziz Z. Huq

Artificial Intelligence and the Rule of Law
(forthcoming in *The Routledge Handbook of the Rule of Law*)

Aziz Z. Huq
University of Chicago Law School

This chapter examines an interaction between technological shocks and the "rule of law." It does so by analyzing the implications of a class of loosely related computational technologies termed "machine learning" (ML) or, rather less precisely "artificial intelligence" (AI). These tools are presently employed in the pre-adjudicative phase of enforcing of the laws, for example facilitating the selection of targets for tax and regulatory investigations (Coglianese and Lehr, 2016). They are also increasingly used during adjudication, for example, to facilitate and guide determinations of individual violence risk during pretrial bail determinations (Huq, 2019). Predictions of a general displacement of human judgment by code-driven counterparts abound (Re and Solow-Niedemann, 2019; Volokh, 2019; but see Wu, 2019). But in near equal measure, that prospect is also loudly decried. Anticipated effects on the fairness, transparency, and equity of adjudicative systems are the main grounds for such resistance (Michaels, 2019; O'Neil, 2016). Even if these criticisms are not framed explicitly in terms of the rule of law, they often overlap with, or are closely adjacent to, the normative concerns that ordinarily travel under that rubric.

Two general questions respecting the rule of law arise from these developments. The more immediately apparent one is whether these technologies, when integrated into the legal system, are themselves compatible or in conflict with the rule of law. Depending on which conception of the rule of law is deployed, the substitution of machine decision-making for human judgment can kindle objections based on transparency, predictability, bias, and procedural fairness. A first purpose of this chapter is to examine ways in which this technological shock poses such challenges. The interaction between the normative ambitions of the rule of law and ML technologies, I will suggest, is complex and ambiguous. In many cases, moreover, the more powerful normative objection to technology arises less from the bare fact of its adoption. It is rather a function of the socio-political context in which adoption occurred and the dynamic effect of technology on background disparities of power and resources. ML's adoption likely exacerbates differences of social power and status in ways that place the rule of law under strain. Attending to this dynamic draws useful attention to a topic that has been noticed in theorizations of the rule of law (e.g., Gowder, 2016; Wilmot-Smith, 2019), but not extensively examined: the interaction between social and economic dynamics on the one hand, and the rule of law on the other.

The second question posed by new AI and ML technologies has also not been extensively discussed. Yet it is perhaps of more profound significance. Rather than focusing on the compliance of new technologies with rule-of-law values, it hinges on the implications of ML and AI technologies for how the rule of law itself is conceived or implemented. As Taekema (2020), has recently observed, many of the canonical discussions of the rule of law—including Dicey's and Waldron's—entangle a conceptual definition and a series of institutional entailments. Fuller (1964), Raz (1979), and Waldron (2011), for example, assume that the rule of law requires more or less specific institutional forms, including courts. They presumably also posit human judges exercising discretion and making judgments as necessary rather than optional. For these institutional entailments of the rule of law, a substitution of human for ML technologies likely has destabilizing implications. It sharpens the question whether the abstract concept of the rule of law needs to be realized by a particular institutional form. It raises a question whether instead technological change might demand amendments to the relationship between the concept(s) and the practice of the rule of law. For pre-existing normative concepts and their practical, institutional correlates may no longer hold under conditions of technological change. At the very least then, specification of institutional forms of the rule of law under such conditions raises challenges not just as a practical matter but also in terms of legal theory

The chapter begins by presenting a brief introduction to ML's present and likely future uses in the law's enforcement and adjudication, noting the different considerations that apply to these contexts. (For the balance of the chapter, I focus on machine learning and use the term "ML" because it is the most pertinent new computational technology, and because the term AI is so broad that it risks confusion). I then consider challenges to rule-of-law values presented by this technological shift. In particular, the chapter explores interactions of ML as a new technology with social and economic arrangements, and the implications of this dynamic for the rule of law. Finally, I consider whether the emergence of ML should force a reconsideration of the ways in which the rule of law is conceptualized and implemented, and whether abstract and the practical reflection on the rule of law can be cleanly bifurcated.

ML Uses in Legal Systems

An ML algorithm in general terms solves a "learning problem ... of improving some measure of performance when executing some task through some type of training experience." (Jordan and Mitchell 2015, p.255). It thus "learns rules from data" (Obermeyer and Emanuel 2016, p.1217). This task is analogous to the function performed by familiar tools such as ordinary least squares and logistic regression. These derive an equation from existing data that describes a relationship between different parameters. That equation can then be applied beyond the original sample of data to make what are, in effect, predictions. Generally, ML tools execute this prediction problem to produce results with less bias and lower variance than familiar regression tools. The complexity of an ML model, though, does not reliably index quality. In some contexts, a relatively simple computational model will produce more accurate outputs than a more complex one (Jung et al., 2020).

There are many ways of taxonomizing ML tools. For example, Domingos (2015), identifies five "tribes" in ML research each with their own underlying deep theory. For more practical purposes, the most salient distinction within the class of ML tools is between supervised and unsupervised instruments. Supervised machine learning algorithms define a function f(x) which produces an output y for any given input x. Its outputs hence take the form of a sorting of x onto categories of y. It could sort, for example, images into the classes of "face" and "not face"; suspects into the classes of "dangerous" and "not dangerous," or "cat," "dog," and "mouse" (Alpaydin, 2020). In contrast, an unsupervised machine-learning algorithm begins with unlabeled training data, and then develops classifications based on the data's immanent structure rather than any ex ante guidance by the programmer (Kelleher and Tierney, 2018; Flasch, 2012). Provided a set of online images, for instance, an unsupervised algorithm might sort them into any number of categories that had not been specified a priori: cats v. not cats; people v. objects, etc. These categories can be imagined as clusters of instances in the data that are more similar to each other than to other instances. The algorithm identifies these clusters by constructing multiple layers of representation, with each layer having a different level of abstraction. By iteratively updating the boundaries of different clusters within the data, the algorithm increases within-cluster similarity and between-cluster divergence. The aim of unsupervised machine learning, in colloquial terms, is thus "to see what generally happens and what does not" (Alpaydin, 2020, 11, and Kelleher and Tierney, 2018).

Supervised and unsupervised ML tools have a number of common features. First, both rely on a set of 'training data,' analyzed to gauge how different variables relate to one another. In some instances, this is historical data, such as past medical records, the past crime data for a municipality, or the list of Internet searches typed in by a given population. Second, the ML instrument is tasked with developing a model that can be used to estimate an outcome variable based on a set of inputs. In constructing this model, the instrument will be asked to follow a cost function (sometimes known as a reward function) that defines the sort of inference that the machine should make.. The resulting model can be predictive—in the sense of offering inferences for events that have not happened—or descriptive—in the sense of drawing to human attention correlations or relationships that would have otherwise gone unnoticed. Third, the model is applied to new, 'out-of-sample' data that is not part of the training data-set. Here is the essence of 'learning' being applied.

ML tools are already used widely by private companies. Recommendation systems such as those employed by Netflix and Amazon are common and familiar examples, but not the only ones. Google's Pagerank algorithm and the Facebook feed encountered by millions of social media users also rely upon ML tools. Reflecting on these applications, Kelleher and Tierney (2018, pp.151-80) organize the "real-world" problems amenable to ML solutions into four broad categories: the identification of clusters, or association within a population; the identification of outliers within a population; the development of associational rules; and prediction problems of classification and regression. In contrast, ML tools are of more limited use for causal estimates.

The state is a late, albeit enthusiastic, eager adopter of ML tools. For present purposes, it is useful to distinguish between two sorts of state uses especially pertinent to rule of law questions. The first, more common, use of ML is its deployment in the investigation and targeting stage. The second, so far less common, deployment is as a substitute for the human judge in adjudication. Of course, both together are only a subset of state uses of ML. Defense technologies that use computer vision or audio recognition, state hospitals that use ML diagnostic tools, and state employees who use Google to facilitate their work are all examples of 'state ML' that do not fall within the ambit of the following, adjudication-centered discussion.

To make a theoretical analysis more tractable, it is useful to start by fleshing out these two common deployments with more detail. The state might first employ ML tools for the purpose of deciding who to investigate or how to allocate some scarce resource (whether aid or intrusion). Hence, a recent study of ML adoptions by the U.S. government found some 64 different bodies within the civilian wings of the national government employing 157 "AI/ML" tools. Many of these usages entail the analysis of either self-reported or public data in order to identify potential violations of the law. For example, the American tax agency the Internal Revenue Service analyzes the large corpus of tax filings with an ML tool in order to identify filings that have indicia of fraud (Engstrom et al., 2020). Nor does the federal government have a monopoly on these tools. Many cities use ML tools to allocate fire inspectors or to determine which restaurants to check for health and sanitation violations (Athey, 2017). When private companies adopt ML processes, governments can require certain features to facilitate its own oversight for compliance with legal standards

(Morse, 2019). Starting in Los Angeles, more than 60 American police forces have adopted 'Predpol,' a controversial predictive tool designed to mine historical crime data to identify where crimes will occur in a subsequent day (Huq, 2019). State and local agencies that provide welfare and other services are also turning to ML tools as a means of sorting among clients. Both positive interventions such as (shelter) and also negative ones (such as investigations for child abuse) can be facilitated with ML tools (Eubanks, 2018).

More intrusive possibilities of ML are illustrated by its recent adoptions by the Chinese Communist state. The Chinese state uses "facial recognition and artificial intelligence to identify and track 1.4 billion people" and thereby "assemble a vast and unprecedented national surveillance system" (Mozur, 2018). Chinese police stationed at transportation hubs such as train stations, for instance, use dark glasses with embedded data streams employing facial recognition technology, such that merely looking at a person pulls up their identity and related information. An algorithmic classification tool sorts surveillance data for ethnic Uighur faces, producing a detailed accounting of the precise movements and actions of a single ethnic class.

Second, the state can also employ ML tools in the adjudication context as a partial substitute for human judgment. This has happened to date less frequently than adoptions in the investigative stage. The leading example is still in the proposal stage. It entails the use of ML tools to make predictions as to whether individual criminal defendants are likely to commit acts of violence in the period prior to their trial. Its academic proponents argue that it promises to permit lower rates of pretrial detention with greater gains for public safety (Kleinberg et al., 2015). There is some debate about how extensive ML displacement of the human role in adjudication will be. Wu (2019) argues that it is unlikely that ML will completely displace human judges. He points out that adjudication is a combination of both 'easy' cases and 'hard' cases. Although ML tools are useful for easy cases, they will produce "dangerous or absurd" results when presented with novel constellations of facts. Hence, Wu predicts the emergence of "cyborg" systems "mixing scale and efficacy with human adjudication for hard cases." Volokh (2019) is more optimistic. He predicts that once the technological problem of designing an AI that can produce persuasive legal texts has been solved, there will be "little conceptual reason" to baulk at the prospect of applying that same technology to judges.

For present purposes, it is sufficient to observe that the sort of "cyborg" systems anticipated by Wu already exist, and are characterized by numerically overwhelming uses of ML in "easy cases" and vanishingly few instances of human judgment as applied to "hard cases." Consider two examples. First, a non-ML algorithm called Compas is used in many American jurisdictions to guide judges' bail determinations by generating a risk score from one to ten for defendants. This score that provides guidance to a magistrate charged with setting or denying bail, and hence guides whether or how defendants will be released pending their criminal trial (Huq, 2019). The use of Compas has proved highly controversial. In particular, the instrument has been subject to persistent criticism based on the allegation that it is racially biased. In particular, Compas has been criticized on the ground that the proportion of factually innocent black defendants improperly subject to detention given its predictions tends to be greater than the analogous proportion of factually innocent white defendants.

Second, in 2013, the governor of Michigan Rick Snyder introduced an algorithmic tool called Midas to detect fraudulent applications for unemployment benefits as part of a larger overhaul of the information technology by the state (Charette, 2018). The Midas system had a denial rate of 93 percent, all the while falsely deeming 40,000 Michigan residents to have fraudulently claimed benefits. It also lacked any mechanism to allow individuals to challenge a benefits denial. Until the spike of unemployment claims associated with the Covid-19 pandemic, the state benefits agency also employed only 12 people to resolve and correct fraud allegations. Even as the pandemic accelerated, calls to the agency would result in applicants being connected not with a state employee, but with another benefit claimant who had been denied. Claimants who allegedly have been wrongly denied a benefit report calling a state office more than a thousand times a day, and still not being able to get through. The state thus chose to provide a user interface with relatively limited opportunities for submitting information, and relatively few opportunities for revision or correction after the fact. Whatever the formal status of an instrument's predictions as advisory or only presumptively valid, the institutional context of the Michigan algorithm made its predictions de facto binding for tens of thousands of people. It all but guaranteed an exorbitant false positive rate when it came to fraud detection (Cuéllar and Huq, forthcoming 2021). The Midas system is a good example of why even "cyborg" systems can produce rule-of-law related worries: "easy" cases numerically dominate, and "hard" cases confront high barriers to litigation. In practice, therefore, the Midas systems dominate the decision-space, leaving little room for human judgment in the absence of a legal challenge to the system as a whole. The Midas system also illustrates the potentially porous nature of the distinction between automating enforcement and automating adjudication. In many contexts, automating enforcement discretion squeezes out the possibility of oversight by adjudication. A machine decision such as those produced by Midas may be difficult to understand, and therefore hard to challenge.

State adoptions of ML are only likely to accelerate. Huq and Cuéllar (forthcoming 2021) argue that state adoption of machine learning happens in a dynamic, two-level context. On this political-economy based account, the state is operating within two dynamic contexts: a domestic political environment dominated by firms competing to expand and monetize machine learning capacities, and also in an international environment in which it is competing with other sovereign nations that are cultivating and deploying the same technological capacities for geostrategic ends. How and to what end machine learning instruments are deployed turns on the strategic choices that the national government makes in these two overlapping yet distinct environments. The pressure to adopt ML will depend not just on law, Huq and Cuéllar conclude, but on the dynamic interplay between the geopolitical environment and domestic interest groups. Accordingly, it is useful to think not just about present applications of ML, but to project into the future the possibility of new adoptions that increasingly displace human judgment from the adjudicative process.

ML as Derogation from the Rule of Law

ML's potential autonomation of adjudication poses challenges and opportunities for rule-of-law values. To flesh both those out, it is obviously necessary to start with a definition of the rule of law. Of course, there is wide disagreement about such a definition. But a rough (and intentionally banally noncontroversial) taxonomy of rule-of-law conceptions suggests that whatever definition is adopted, the ouster of human adjudicative judgment by machine decision-making is unlikely to be unproblematic.

The rule of law is can loosely be said to take formal, substantive, and procedural forms. The formal/substantive distinction is a familiar one (Tamanaha, 2004, pp.91-92). Addition of the procedural category reflects recent work that self-consciously takes a more institutional lens to the task of definition. Lon Fuller's (1964) famous catalog of rule-of-law values—that laws must be general, open, prospective, clear, consistent, stable, capable of being obeyed, and upheld by officials—is often labeled "formal" (Craig, 1997; Tamanaha, 2004). As Gardner (2012, pp.199-204) has pointed out, though, Fuller's criteria "pass judgment on the content of the law," and can easily give rise to legal, enforceable rights. Fuller's catalog is better understood, Gardner argued, as "the morality of *how*, not the morality of *why*" (ibid. 206). It concerns law's "mode of generation and application" (Raz, 2019, p.2). I will use the idea of the "formal" rule of law to capture this idea. A second understanding of the rule of law goes to "why," and so warrants the label "substantive." Perhaps the crispest formulation of this possibility is Lord Bingham's argument that the "law must afford adequate protection of fundamental human rights" for the rule of law to obtain (Bingham, 2011, p.75). Similarly, T.R.S. Allen describes the rule of law as "a corpus of basic principles and values, which together lend some stability and coherence to the legal order." He includes with that corpus "traditional ideas about individual liberty and natural justice, and, more generally, ideas about individual justice and fairness in relations between government and governed" (Allen, 1993, pp.21-22).

Finally, Waldron has identified the rule of law with procedural qualities such as the right to a hearing before an impartial tribunal, the right to present evidence and make legal arguments, and the right to a reasoned explanation for a decision. He advances these as necessary to vindicate human dignity, understood as an entitlement to be treated as an "active intelligence" (Waldron, 2011, 23; see also Waldron, 2008). This procedural understanding of the rule of law can be glossed as having been incorporated already within some formal accounts of the rule of law. For example, Waldron's definition can be interpreted as unpacking the eighth and final component of Fuller's formal definition of the rule of law. This requires that officials apply the rules that have been enacted. Yet Waldron's account is also rich enough, and distinct enough in tone and detail from Fuller's terse claim, that it is sensible to treat it as a stand-alone assertion. Unlike Fuller's account, Waldron's centrally embraces and aims to promote "the dignity and agency of persons with their own point of view and arguments worth listening to" in a way that neither the substantive nor procedural formulations of the rule of law can do (Taekema, 2013, 145).

Each of these accounts of the rule of law interact with the emergence of ML and AI in different ways. Tracing these interactions illuminates the normative stakes of these technologies. It also helps clarify technology's most important effect on the rule of law: its preservation and exacerbation of background inequalities of influence, privilege, and status. I will focus here on the formal and procedural understandings of the rule of law. Because substantive understandings entail a specification of "fundamental" or "basic" rights standing that are distinct from the matters of "how" law is applied, their consistency with automated decision-making processes is a contingent matter. To be sure, there are a suite of objections to automated adjudicative systems that are best understood not in terms of impartiality or neutrality, but in terms of basic substantive moral principles. For example, one objection to the Compas system is that its predictions are based on data that reflects and incorporated biased assumptions (e.g.., data that was gathered through discriminatory policing), and that they lend an aura of spurious scientific validity to conceptions of 'dangerousness' that have no good foundation in empirical data. (Koepke & Robinson, 2018). Reliance on training data that reflects historical patterns of racial or gender stratification is also likely to produce outputs that perpetuate such patterns (Huq, 2019). This objection, while important, is not distinct to the ML context. It applies, rather, to any instance in which data distorted by historical bias is used as a basis for future allocations (e.g., the rewards to education in the labor market). But setting aside these quite general objections, there is no reason to assume that the substantive rule of rule will always or necessarily come into conflict with the decision to use an ML tool rather than a human decision-maker.

Consider first the way that ML tools interact with the formal rule of law as Fuller defined it. Depending on the manner in which an automated system is implemented, its effect on the rule of law could be either beneficial or harmful. One of the ways in which Fuller describes the rule of law as failing is through the absence of a "published code declaring the rules to be applied in future disputes." That code, moreover, must be capable of being "understood by an ordinary citizen or a trained lawyer" (Fuller, 1964, pp.35-36). Legislative obscurantism is anathema to the formal rule of law. (It is worth noting that Fuller's formulation seems to allow for a measure of obscurity so long as a trained professional can penetrate it. What about the aid of a computer scientist? Or do guild loyalties preclude that?). A common objection to the use of ML systems, in both adjudication and otherwise, is that they can seem, to use Fuller's terms, a "masterpiece of obscurity" to the person who is on the receiving end of their classifications (Pasquale 2015). The reward function of an automated predictive tool is programmed into its code. It is not necessarily or readily available for examination by those subject to its regime. Even if the code were available, it would not be illuminating for either the "ordinary citizen" or the "trained lawyer." Pace some techno-optimists (Fagan and Levmore, 2019), it seems unlikely that there will many capable of evaluating the quality of a predictive algorithm on the fly. Criminal defendants or benefits claimants subject to non-ML instruments such as Compas or ML instruments such as Midas have no way to effectively understand, or to challenge, the rules under which they are granted or deprived liberty or the means of survival.

In practice, it is likely that the objection from inscrutability will run against many practical instantiations of automated adjudication, including instruments such as Compas and Midas. Yet it is not clear whether the rule-of-law objection to those tools runs applies just to their particular manner of implementation, or whether the objection is one with more general force. Just as there are human-driven adjudicative systems that fall prey to the inscrutability objection so too there are machine-driven systems that have the same flaw. (Think of speed cameras as an example). There is no reason to think that the problems besetting early adopters of a new governance tools will persist in later adoptions. Computer scientists, moreover, have developed a suite of "explainable AIs," which offer a range of different kinds of information to end-users as a means of illuminating a machine decision (Huq, forthcoming 2021). There is no reason why an ML adjudicative tool cannot, for example, offer an ordinary language account of its reward function, the most significant parameters in the determination of an outcome, and an account of what behaviors or factors might be changed to receive a different outcome. (Of course, this might be done either in a way that clarifies or in a way that heaps on obscurity). The literature on transparency in the algorithmic context, moreover, underscores considerations that are somewhat suppressed in the literature on the formal rule of law. Most important, the concept of an "explanation" is neither simple nor singular. To the contrary, explanations can take several different forms, e.g., causal, counterfactuals, motivational, physical, etc. (Miller, 2019). The assumption that disclosure generates trust and facilitates effective human agency has also been challenged on the ground that individuals are likely to have different ex ante epistemic and social resources, and therefore will vary in their ability to respond effectively to explanations (Ananny and Crawford, 2018). Yet the same could be said of a legal code that can be comprehended and applied only by a "trained lawyer."

In other respects, however, an automated decision may be superior to a human decision when viewed through a formal rule-of-law lens. A theme in formal conceptions of the rule of law is the benefit of stability and predictability in the operation of a legal system. Raz emphasizes the predictability of law—or its guidance capacity—as central to his account of the rule of law (Raz, 1979, p.218; see also Barber, 2018, pp.94-95). Shapiro argues that the rule of law gains value "entirely from the benefits that social planning generates and is best served when legal structures maximize those benefits" (2011, p.396). to the extent that the rule of law is understood as instrumental to individual and social planning, the turn from human to ML decision-making has the potential to improve certainty because it squeezes a particular source of uncertainty out of the system. The stochastic effect of individual human judgments, exercised by front-line officials, may well be less predictable than the operation of an automated instrument. This is not to say that the shift from human to machine eliminates all forms of discretion. The problems of factual ignorance and indeterminacy of aim, identified by H.L.A. Hart (2013, p.661) as the springs of discretion, likely persist without regard to the chosen decision tool. But by centralizing the decisional process, automation at least holds the immanent possibility of greater, not lesser, certainty. Law, understood as the facilitation of social and individual planning through guidance, is enabled not handicapped by this shift.

The effect of ML tools on the procedural rule of law seems at first blush less ambiguous. Waldron's account of this conception underscores the relation of judicial process to the vindication of dignity and agency interests. A process staffed by humans treats individuals as "bearers of reason and intelligence" albeit at some context to "law's certainty" (Waldron 2011, p.19). On Waldron's account, individuals within a rule of law regime do not merely passively receive the law in the fashion of Fuller's and Raz's account. They actively contest the law through argument and "elaborate interpretive exercises" (ibid. p.20). On this view of human agency, the loss of human adjudicators capable of listening and responding to the arguments and facts thrown by a person being subject to law would be fatal. A cold and (literally) inhuman indifference of a machine tool is the antipode of the procedural rule of law (O'Neil, 2016).

Yet as forceful as Waldron's argument is, the association of dignity and agency with human-led process is not as obvious nor as strong as first appearances suggest. Notice first that Waldron imagines an individual subject to the law not just as the bearer of "reason and intelligence," but also as possessor of certain dispositions and temperament. Waldron's subject is disputative, with the intellectual and material resources. Just as Fuller's argument against inscrutability rests on the assumption of a certain background level of resources, so Waldron's stipulates away the common condition of individuals who do, in fact, come before the law. They are financially straitened, hassled upon all sides by the daily furies of work or family, and perhaps even beset by the subtle and pervasive indignities of navigating a world in which one's identity as a woman, or non-cis-gendered or trans person, or a racial or ethnic minority, or a disabled person, or a non-citizen, marks one out for regular condescension, contempt, distaste, or pity. The cavalcade of alternatives in that litany underscores how unusual the imagined disputative and empowered subject of Waldron's imaginary really is. Hence, in the United States, for example, it is widely recognized that the criminal justice operates in a massive, effectively industrial scale (Beckett 2018). According to one estimate, some 13.2 million criminal misdemeanor cases are processed in that country every year (Stevenson and Mayson, 2018). In few of these will the defendant arraigned fit Waldron's characterization. Quite the contrary.

Waldron's account may be motivationally as well as institutionally inapt. He assumes that dignity and autonomy are advanced through adversarial process. But for many, an adversarial proceeding can well be a source of distress and even stigmatic harm (Natapoff, 2015). Process, as Feely (1978) said, is itself often punishment. A managerial system of either civil or criminal justice, for example, may well inflict psychological damage upon individuals by the very fact of giving them a hearing. It may often be more respectful of autonomy and dignity to minimize the human interaction that comes with seeking a welfare benefit or contesting a misdemeanor (Huq, 2020). The manner in which dignity and agency are best realized also rests on background expectations of fair and respectful treatment. In the late twentieth century, it was hard to imagine how this could be otherwise. Today, though, the socio-technological landscape is less clear. People at large have likely become far more accustomed to automated interactions with private actors, including over medical issues, personal relations (e.g., in online dating forums), and via social media. The very technological advances that stimulate the emergence of automated adjudication are also catalyzing broader changes in human expectations and behavior. In this context, the question of how interactions with the state are best structured to advance dignity and agency should be recognized as in large part an empirical one. At a minimum, it seems unwise to assume that the substitution of machine for human judgment is necessarily linked to a loss of dignity and autonomy as Waldron does.

The relationship between ML-driven adjudication on the one hand, and the formal and procedural conceptions of the rule of law on the other, in short, is not quite as straightforward as might first appear. Empirical contingencies loom. These brief reflections on these relationships also brings to light a dimension of the rule of law that is easy to gloss over: its relationship to background social and economic conditions, and the way in which changes to the governance choices of the legal system (including the decision whether or not to use ML tools) alter those background conditions. In its formal and procedural formulations, at least, leading theorizations of the rule of law may implicitly lean upon an empirical presupposition of a rough measure of material and intellectual equality of persons before the law. In the absence of such equalities, compliance with the formal or procedural rule of law on its own is likely to have uneven effects across a regulated population. Those with material and intellectual resources are well positioned to leverage the law's "guidance" and to engage in beneficial planning. Those lacking such assets are unlikely to benefit from the rule of law. The net effect of an insistence on the rule of law may well be to increase disparities in entitlements that create the divergence in the first instance. This may be so even if, as E.P. Thompson famously suggested, the rule of law makes all participants absolutely better off. The rule of law, on this account, may be an engine of social inequality under certain (although not all) circumstances. To counteract this effect, it is likely necessary, as Wilmot-Smith (2019, pp.90-91) has argued, that a legal system supply individuals with "the same amount of legal resources" such that legal procedures should promote equal distribution of legal benefits and burdens, including the public good of legality. A related, if more abstract, argument is Gowder's (2016) claim that the rule of law "must be actually justifiable to all on the basis of reasons that are consistent with the equality of all."

New technologies complicate this picture. For if the interaction of formal and procedural conceptions of the rule of law is ambiguous, then perhaps the most important way that new computational technologies matter is their effect upon the background distributions of epistemic, social, and economic attributes that make the rule of law a possible and attractive ambition. ML and AI instruments are not distributionally neutral. They reward access to large accumulations of data—such as those held already by the state and large commercial actors—and the technical resources necessary to extract inferences from that data. The introduction of a new method for corporate actors, whether public or private, to exploit extant resources in a way that private individuals cannot will change the balance of social and economic power between individuals and corporate actors. Such power will likely become more concentrated in corporate actors as a result of new tools for managing and manipulating large populations of individuals. Insulation from the exercise of this new form of power, moreover, requires privacy: the entitlement to withhold information from others. Privacy, of course, is also unevenly distributed; in the United States, it is also often deliberately withheld by state actors on grounds of race and gender, e.g., by conditioning certain welfare benefits used disproportionately by Black women on privacy deprivations (Bridges 2017). The same is true elsewhere. A plausible fear of many technology critics is therefore that new predictive instruments will have the effect of widening social and material inequalities (Zuboff, 2019), even if they fail to identify relevant mechanisms or vulnerable population (Cuéllar and Huq, 2020). Gowder (2018, p.83) has argued that some technological innovations have "a distinctive potential to facilitate egalitarian advances in the rule of law and in access to justice." If the pessimists prove correct, and the equalizing provision of legal resources recommended by Wilmot-Smith remain more aspiration rather than practice, the quality of the rule of law will decline as a practical matter as technological change interacts with, and exacerbates, background hierarchies of wealth and caste.

Theorizing the Rule of Law after ML

A moment of sudden technological change is, finally, an opportunity to reconsider some basic, if typically unexamined, presuppositions in theories of the rule of law. Building on the above reflections upon Waldron's procedural account above, I conclude this chapter by considering whether the emergence of ML tools has implications for the way in which the rule of law is conceptualized. More specifically, inverting the analytic lens in this fashion brings to light the possibility that the normative ambitions and the institutional entailments of the rule of law are more loosely linked than generally supposed. Specifically, the assumption that *courts* as institutions are central to, even necessary for, the rule of law might well require reconsideration.

Many seminal texts on the rule of law explicitly or implicitly assume the centrality of human-managed courts. Raz, for example, admonishes that the "independence of the judiciary must be guaranteed" (Raz 1979, pp. 216-16). As Gardner rightly observes (2012, pp.208-10), Fuller's account is inconsistent with not only the idea of a "legal system without rules" but also with the prospect of a "legal system without rulings, meaning authoritative official decisions in particular cases that purport to be applications of the law's rules." It thus requires a "robustly independent judiciary that does not shy away from decision." And, most obviously, Waldron's procedural account (2011) places courts and the human-driven adjudicative process at its center. Yet as Taekama (2020, 5) observes, theories such as Raz's do not justify the inference from the guidance function of law to the necessity of courts. Her point can be generalized. The formal and procedural conceptions of the rule of law identify normative ambitions (guidance; dignity and agency) of the rule of law. At one point, it may well have been a reasonable assumption that among the institutional entailments of these normative ambitions was an adjudicatory structure such as a court. Whether that was true at one point in time, it is no longer so today. The social and human goods associated with the rule of law can be unbundled from the specific institutional forms with which they have come to be associated. Technological advances invite new ways both to derogate from those values or, alternatively, to achieve them. Either way, the normative, conceptual, and institutional elements of the rule of law—commonly bundled together into a single package by its theorists—are less tightly lashed together than commonly perceived.

The aim of this chapter has not been to offer a positive articulation of how the rule of law could be instantiated under novel technological circumstances. That would be a large undertaking. At the very least, however, the advent of new instruments of decision-making and prediction should not simply be occasions for celebration or despair. They should rather be understood as opportunities for rethinking the elements that we have previously, perhaps prematurely, bundled together into our conception of the rule of law.

References

Allan, T. R. S. (1982). *Law, Liberty, and Justice*. New York: Cambridge University Press.

Alpaydin, E. (2020). *Introduction to machine learning*. Cambridge: MIT Press.

Ananny, M., & Crawford, K. (2018). 'Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability'. *new media & society*, *20*(3), 973-989.

Athey, S. (2017). 'Beyond prediction: Using big data for policy problems'. *Science*, *355*(6324), 483-485.

Barber, N. W. (2018). *The principles of constitutionalism*. New York: Oxford University Press.

Beckett, K. (2018). 'The politics, promise, and peril of criminal justice reform in the context of mass incarceration'. *Annual Review of Criminology*, *1*, 235-259.

Bingham, L. (2007). The rule of law. *Cambridge Law Journal*, 67-85.

Bridges, K. M. (2017). *The poverty of privacy rights*. Palo Alto: Stanford University Press.

Charette, R.N. (2018). 'Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold'. *IEEE Spectrum*.

Coglianese, C., & Lehr, D. (2016). 'Regulating by robot: Administrative decision making in the machine-learning era'. *Georgetown Law Journal*, *105*, 1147.

Cuéllar, M. F., & Huq, A. Z. (forthcoming 2021). 'Toward the Democratic Regulation of AI Systems: A Prolegomenon'. *Available at SSRN 3671011*.

-- (2020). 'Economies of Surveillance'. *Harvard Law Review*, *133*(4), 1280.

Craig, P. (1997). 'Formal and substantive conceptions of the rule of law: an analytical framework'. *Public Law*, (3), 467-487.

Domingos, P. (2015). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. New York: Basic Books.

Engstrom, D. F., Ho, D. E., Sharkey, C. M., & Cuéllar, M. F. (2020). 'Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies'. *Available at SSRN 3551505*.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: St. Martin's Press.

Feeley, M. M. (1979). *The process is the punishment: Handling cases in a lower criminal court*. New York: Russell Sage Foundation.

Fagan, F., & Levmore, S. (2019). 'The Impact of Artificial Intelligence on Rules, Standards, and Judicial Discretion.' *Southern. California Law Review 93*, 1.

Flach, P. (2012). *Machine learning: the art and science of algorithms that make sense of data*. New York: Cambridge University Press.

Fuller, L. L. (1964). *The morality of law*. New Haven: Yale University Press.

Gardner, J. (2012). *Law as a leap of faith: essays on law in general*. New York: Oxford University Press.

Gowder, P. (2018). 'Transformative legal technology and the rule of law'. *University of Toronto Law Journal*, *68*(supplement 1), 82-105.

-- (2016). *The rule of law in the real world*. New York: Cambridge University Press.

Hart, H. L. (2013). 'Discretion'. *Harvard Law Review*, *127*, 652.

Huq, A.Z. (forthcoming 2021). 'Constitutional Rights in the Machine Learning State' *Cornell Law Review* 105.

-- 'A Right to a Human Decision'. *Virginia Law Review* 106 (2020): 611.

-- (2018). 'Racial equity in algorithmic criminal justice'. *Duke Law Journal*, *68*, 1043.

Huq, A. Z., and Cuéllar, M.F. (forthcoming 2019), 'Privacy's Political Economy and the State of Machine Learning.' *NYU Annual Survey of American Law*.

Jordan, M. I., & Mitchell, T. M. (2015). 'Machine learning: Trends, perspectives, and prospects'. *Science*, *349*(6245), 255-260.

Jung, J., Concannon, C., Shroff, R., Goel, S., & Goldstein, D. G. (2020). 'Simple rules to guide expert classifications'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.

Kelleher, J. D., & Tierney, B. (2018). *Data science*. Cambridge: MIT Press.

Kleinberg, Jon, et al. "Prediction Policy Problems." *American Economic Review* 105.5 (2015): 491-495.

Koepke, J. L., & Robinson, D. G. (2018). 'Danger ahead: Risk assessment and the future of bail reform'. *Washington Law Review 93*, 1725.

Michaels, A. C. (2019). 'Artificial Intelligence, Legal Change, and Separation of Powers'. *University of Cincinnati Law Review*, *88*(4), 1083.

Miller, T. (2019). 'Explanation in artificial intelligence: Insights from the social sciences'. *Artificial Intelligence*, *267*, 1-38.

Morse, S. C. (2019). 'Government-to-robot enforcement'. *University of Illinois Law Review*, 1497.

Mozur, P. (2018). 'Inside China's dystopian dreams: AI, shame and lots of cameras'. *The New York Times*, *8*, 2018.

Natapoff, A. (2015). 'Misdemeanor decriminalization'. *Vanderbilt Law Review*, *68*, 1055.

Obermeyer, Z., & Emanuel, E. J. (2016). 'Predicting the future—big data, machine learning, and clinical medicine'. *New England journal of medicine*, *375*(13), 1216.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Broadway Books.

Pasquale, F. (2015). *The black box society*. Cambridge: Harvard University Press.

Raz, J. (2019). 'The Law's own virtue'. *Oxford Journal of Legal Studies*, *39*(1), 1-15.

-- (1979). *The authority of law: essays on law and morality*. New York: Oxford University Press.

Re, R. M., & Solow-Niederman, A. (2019). 'Developing artificially intelligent justice'. *Stanford Technology Law Review*, *22*, 242.

Tamahana, B. Z. (2004). *On the rule of law: History, politics,* theory. New York: Cambridge University Press.

Taekema, S. (2020). 'Methodologies of Rule of Law Research: Why Legal Philosophy Needs Empirical and Doctrinal Scholarship'. *Law and Philosophy*.

Taekema, S. (2013). 'The Procedural Rule of Law: Examining Waldron's Argument on Dignity and Agency'. *JRE*, *21*, 133.

Stevenson, M., & Mayson, S. (2018). 'The scale of misdemeanor justice'. *Boston University Law Review*, *98*, 731.

Volokh, E. (2018). 'Chief Justice Robots'. *Duke Law Journal*, *68*, 1135.

Waldron, J. (2011). 'The rule of law and the importance of procedure'. *Nomos*, *50*, 3-31.

-- (2008). 'The Concept and the Rule of Law'. *Georgia Law Review*, 1, 43.

Wilmot-Smith, F. (2019). *Equal Justice: Fair Legal Systems in an Unfair World*. Cambridge: Harvard University Press.

Wu, T. (2019). 'Will Artificial Intelligence Eat the Law? The Rise of Hybrid Social-Ordering Systems'. *Columbia Law Review*, *119*(7), 2001-2028.

Zuboff, Shoshana. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Profile Books, 2019.