

Replacing Bureaucrats with Automated Sorcerers?

Bernard W. Bell

Increasingly, federal agencies employ artificial intelligence to help direct their enforcement efforts, adjudicate claims and other matters, and craft regulations or regulatory approaches. Theoretically, artificial intelligence could enable agencies to address endemic problems, most notably 1) the inconsistent decision-making and departure from policy attributable to low-level officials' exercise of discretion; and 2) the imprecise nature of agency rules. But two characteristics of artificial intelligence, its opaqueness and the nonintuitive nature of its correlations, threaten core values of administrative law. Administrative law reflects the principles that 1) persons be judged individually according to announced criteria; 2) administrative regulations reflect some means-end rationality; and 3) administrative decisions be subject to review by external actors and transparent to the public. Artificial intelligence has adverse implications for all three of those critical norms. The resultant tension, at least for now, will constrain administrative agencies' most ambitious potential uses of artificial intelligence.

Artificial intelligence/machine learning (AI/ML) algorithms are widely used in the private sector. We experience the results daily: AI/ML algorithms suggest products for purchase and even finish our sentences. But those uses of AI/ML seem tame and impermanent: we can always reject algorithm-generated suggestions.

Can AI/ML become a resource for government agencies, not just in controlling traffic lights or sorting mail, but in the exercise of the government's coercive powers? The federal government has begun to deploy AI/ML algorithms.¹ The embrace of such technologies will profoundly affect not only the public, but the bureaucracies themselves.² Might AI/ML bring agencies closer to attaining, in Max Weber's words, "the optimum possibility for carrying through the principle of specializing administrative functions according to purely objective considerations"?³

Before exploring such implications, I will discuss AI/ML's capabilities and use by federal agencies, and agencies' functions and environment. Ultimately, we will see that AI/ML is a sort of empirical magic that may assist in coordinating an agency's actions but presents challenges due to its lack of transparency and nonintuitiveness.

The government has long used computers to store and process vast quantities of information.⁴ But human beings fully controlled the computers and wrote their algorithms. Programmers had to do all the work of modeling reality: that is, attempting to ensure that their algorithm reflected the actual world, as well as incorporating the agencies' objectives.⁵

AI/ML is much less dependent on the programmer.⁶ It finds associations and relationships in data, correlations that are both unseen by its programmers and nonintuitive. As to the latter, for example, an AI/ML algorithm might predict a person's preferred style of shoe based upon the type of fruit the person typically purchases for breakfast.⁷ Thus, AI/ML results do not represent cause and effect; correlation does not equal causation. Indeed, as in the example above, AI/ML algorithms may rely upon correlations that defy intuitive expectations about relevance; no one would posit that shoppers consider their breakfast choices when making shoe selections.

The opaque and nonintuitive associations on which AI/ML relies, that is, AI/ML's "black box" quality, have consequences for administrative law.⁸ Even knowing the inputs and the algorithm's results, the algorithm's human creator cannot necessarily fully explain, especially in terms of cause and effect, how the algorithm reached those results. The programmer may also be unable to provide an intuitive rationale for the algorithm's results. While computer experts can describe the algorithm's conclusion that people with a particular combination of attributes generally warrant a particular type of treatment, they cannot claim that the algorithm has established that any particular individual with that combination of attributes deserves such treatment.⁹

AI/ML can be used in either a supervised or unsupervised manner. In *supervised learning*, training data are used to develop a model with features to predict known labels or outcomes. In *unsupervised learning*, a model is trained to identify patterns without such labels.¹⁰

AI/ML is particularly useful in performing four functions: identifying clusters or associations within a population; identifying outliers within a population; developing associational rules; and solving prediction problems of classification and regression.¹¹ AI/ML is currently less useful when a problem requires "estimating the causal effect of an intervention."¹² Nor can such algorithms resolve non-empirical questions, such as normatively inflected ones, like ethical decisions.¹³ Presumably, AI/ML is ill-suited for resolving some empirical questions that frequently arise in administrative and judicial contexts, such as resolving witnesses' differing accounts of past events. In those situations, the data inputs are unclear.

A recent Administrative Conference of the United States (ACUS) study uncovered considerable agency experimentation with or use of AI/ML.¹⁴ Agencies largely employed human-supervised AI/ML algorithms, and

their results were generally used to assist agency decision-makers and agency management in making their own decisions. A few examples follow.

The Securities and Exchange Commission (SEC) uses AI/ML to monitor the securities markets for potential insider trading. The SEC's ARTEMIS system focuses on detecting serial inside traders. A natural language program sifts through 8-K forms submitted by companies to announce important events that occur between their regular securities filings. SEC staff then use a natural language processing algorithm to sift through the forms. Then, a machine learning algorithm identifies trigger events or market changes that warrant investigation. An official reviews the output and decides whether further investigation is justified. If so, SEC staff send a blue sheet request to broker/dealers for relevant trading records. The blue sheet data are analyzed with previously requested blue sheet data by an unsupervised learning model to detect anomalies indicating the presence of insider trading.

The Social Security Administration (SSA) uses several methods to increase the efficiency of its disability benefits claim adjudication process. It has attempted to apply algorithms to claim metadata to create clusters of similar cases it can assign to the same administrative law judge (ALJ). It has also developed an AI/ML analysis of claims to determine the probability of an award of benefits based solely on certain attributes of the claims. Officials use the results in establishing the order in which claims are assigned, moving ones likely to be granted to the head of the line. However, the actual determination of the claim is made by the ALJ.

AI/ML assists adjudicators in preparing disability decisions. The SSA's Insight program allows adjudicators to identify errors in their draft decisions, such as erroneous citations (that is, nonexistent regulation numbers) and misapplication of the vocational grid (the metric used to determine whether sufficient work exists in the national economy for those of a claimant's level of exertional ability, age, and education). Insight also assists the SSA in identifying common errors made by ALJs, outlier ALJs, and areas in which SSA policies need clarification.¹⁵

The ACUS report discusses the use of AI/ML to sift through the massive number of comments made in response to the Federal Communications Commission's proposed rollback of its net neutrality rules and the Consumer Finance Protection Bureau's use of AI/ML to classify the complaints it receives.¹⁶ Algorithms have been deployed to assist agencies in predicting an industry's potential response to various alternative formulations of a contemplated regulation. The Environmental Protection Agency's (EPA) OMEGA model "sift[s] through the multitude of ways . . . automaker[s] could comply with a proposed greenhouse gas emissions standard to identify the most likely compliance decisions."¹⁷ OMEGA thus has helped the EPA set greenhouse gas emissions standards "that [protect] public health and welfare while remaining cognizant of the time and cost burdens imposed on automakers."¹⁸ OMEGA is not an AI/ML algorithm, but we might see it as a forerunner of AI/ML algorithms that would perform a similar function.¹⁹

Administrative agencies perform a wide array of functions. Administrative law scholars tend to focus on three broad categories of agency action that lie at the heart of the government's coercive powers: enforcement, adjudication, and rulemaking. These categories derive from the distinction between legislating, enforcing the law, and adjudicating legal disputes.

Enforcement. Enforcement involves monitoring regulated entities, identifying statutory or regulatory violations, and pursuing sanctions for such violations. Enforcement is largely an executive function.

Moreover, enforcement has heretofore been considered inherently discretionary: agencies' limited resources simply do not allow them to be present everywhere at all times, much less pursue every potential regulatory violation.²⁰ Choosing which regulated entities or activities to investigate can be excluded from the realm of consequential decisions. If the entity or person under investigation has been complying with the law (or if the government cannot amass sufficient evidence to prove otherwise), no adverse consequence will ensue. Generally, the cost of undergoing investigation and defending oneself in an unsuccessful government enforcement action is not considered a harm.²¹

Adjudication. Adjudication involves resolving individuals' rights against, claims of entitlements from, or obligations to the government. Thus, decisions regarding Social Security disability benefits, veterans' benefits, entitlement to a particular immigration status, and the grant or revocation of government licenses or permits, as well as liability for civil fines or injunctive-type relief, are all adjudications. In mass justice agencies, these adjudications differ substantially from traditional judicial determinations. Traditional judicial decisions often involve competing claims of right and frequently require making moral judgments in the course of resolving cases. The specification of rights and obligations is often intertwined with a determination of the applicable facts.²² AI/ML algorithms might make quite good predictions regarding the results in such cases, but we are chary about leaving the *actual* decision to an AI/ML algorithm.

Mass adjudication by administrative agencies can often be much more routinized. Consider insurance companies' resolution of automobile accident claims. The judicially crafted law is complex. Liability turns on each actor's "reasonableness," a judgment based on a mixture of law and fact. The complexity represents an effort to decide whether the injured plaintiff is morally deserving of recovery from the defendant driver. Fully litigating such cases requires questioning all witnesses to the accident closely. But insurance companies seeking to resolve mass claims without litigation use traffic laws to resolve liability issues, as an imperfect but efficient metric.²³

Similarly, the SSA disability determinations could be considered expressions of a societal value judgment regarding which members of society qualify as the *deserving poor*.²⁴ Such a determination could be unstructured and allow significant

room for adjudicators' application of moral judgments and intuition. But the SSA has, of necessity, established a rigid, routinized, five-step process for evaluating disability claims.²⁵ And the final step involves assessing whether sufficient jobs the claimant can perform exist in the national economy. That too was routinized by use of a grid, which provided a yes/no answer for each combination of applicants' age, education, and exertional capacity.²⁶

Another aspect of agency adjudication warrants attention. Much of traditional litigation, particularly suits for damages, involves assessing historical facts, the who, what, when, where, and why of past events. But agency adjudications can involve predictions as well as historical facts. Thus, licensing decisions are grounded on predictions regarding the likelihood that the applicant will comport with professional standards. Likewise, the last step of the SSA disability determination, whether a person with certain age, educational, and exertional limitations could find a sufficient number of jobs available in the national economy, is a prediction. On the other hand, whether an employer committed an unfair labor practice in treating an employee adversely for union activity is a question of historical fact.

AI/ML excels at making predictions – that is its *sine qua non* – and predictions are all we have with regard to future events (or present events we may want to address without taking a wait-and-see approach).²⁷ But for an issue such as whether a particular entity engaged in a specific unfair labor practice, we might want to focus on the witness accounts and documentary evidence relevant to that situation, rather than AI/ML-generated correlations.²⁸ Or to use an example from toxic torts, epidemiological and toxicological studies establishing general causation between a toxin and a toxic harm may be fine for estimating risks to a population exposed to a toxin, but do not prove what courts in toxic torts must determine: namely, whether the harm the *plaintiff* suffered was caused by the *plaintiff's* exposure to a toxin.²⁹

Rulemaking. Rulemaking involves promulgation of imperatives of general applicability akin to statutes. As administrative law scholars Cary Coglianese and David Lehr suggest, AI/ML's use in rulemaking is limited because that process involves normative judgments and requires “overlay[ing] causal interpretations on the relationship between possible regulations and estimated effect.”³⁰ The product of agency rulemaking – regulations – may resemble formal legislation, but the rulemaking process is designed to be far less onerous. Agencies often promulgate such regulations by “notice-and-comment” procedures.³¹ Those procedures seem deceptively simple, but in practice require the agency to identify and categorize assertions made in thousands of comments regarding the rule's propriety. And with the emphasis on the Office of Management and Budget's (OMB) regulatory review of proposed regulatory actions, a significant part of the rulemaking process consists of assessing the overall costs and benefits attendant the rule.³²

These legislative rules differ from the guidance rules used to constrain lower-level officials' discretion, direct their decision-making, or advise the public. Leg-

islatiue rules that are the product of notice-and-comment rulemaking have the “force of law”: violation of the rule itself is unlawful, even if the action does not violate the statutory standard implemented by the rule. Rules in the second sense, guidance rules that merely constrain lower-level officials’ discretion or provide guidance to the public, do not replace the legal standard enunciated in the statute upon which they elaborate. They lack the force of law; an agency’s sanction against violators of such guidance rules can be upheld only if the agency can show that the rule-violator’s conduct has transgressed the underlying statute.

For example, a federal statute grants the Federal Trade Commission (FTC) the power to enjoin unfair and deceptive trade practices. The FTC could issue a guidance rule specifying that gas station operators’ failure to post octane ratings on gas pumps is inherently deceptive. The guidance might well be based on extensive consumer research the FTC has conducted. If the FTC promulgates a guidance rule, each time it goes to court to enforce an order it enters against a rule-violator, it will have to prove that the gas station’s failure to post octane ratings was deceptive. If, however, the FTC promulgates a force of law rule, that is, a “legislative rule,” when it goes to court to enforce an order against a violator, it need merely show that the octane ratings were not posted. The gas station operator can no longer mount a defense asserting that its customers were not confused or deceived by the lack of posted octane ratings.

Legislative rules can be analogized to algorithms. The human lawgiver correlates a trait with a particular mischief the legislative rule is designed to address. The correlation may often be imperfect; but rules are inherently imperfect. However, we would probably not accept laws based on a nonintuitive correlation of traits to the mischief to be prevented, even if the correlation turns out to be a pretty good predictor. Even with respect to legislatures, whose legislative judgments reflected in economic and social legislation are given a particularly wide berth, courts purport to require some “rational basis” for associating the trait that is targeted with the mischief to be prevented.³³ The demands for some intuitive connection, some cause-and-effect relationship between a trait targeted and a harm to be prevented, is even greater when agencies promulgate regulations.³⁴ And to carry the analogy further to guidance rules, it is not clear at all that a nonintuitive connection would be allowed as a guidance rule used to direct the resolution of agency adjudications.

The critical *internal* challenge for government bureaucracies is synchronizing line-level decision-makers, both with the intended agency policy and with each other.³⁵ Internal review processes can serve this function, but such processes still require coordination at the review level and involve duplication of effort. The agency may seek to reduce decision-making metrics to written rules (either legislative rules or guidance rules).³⁶ Agencies also expend resources

on training, and retraining, its lower-level employees. And sometimes agency leadership may encounter bureaucratic resistance, yet another reason some line-level employees' determinations might not comport with the leadership's policy.³⁷

Agencies' internal structures reflect the fundamental tension between rule-like and standard-like decision metrics. Rules are decision metrics that do not vary significantly depending on the circumstances.³⁸ Rules facilitate decisional consistency, assist line officials' efforts to follow agency policy, and allow superiors to more easily detect departures. But rules are invariably over-inclusive or under-inclusive: they sweep within them nonproblematic cases or fail to capture problematic cases, or both.³⁹ And the simpler the rule, the larger the subset of undesirable results.

For instance, due to the increasing heart attack risks as people age, in 1960, the Federal Aviation Administration promulgated the following rule: "No individual who has reached his 60th birthday shall be utilized or serve as a pilot on any aircraft while engaged in air carrier operations."⁴⁰ The rule is over-inclusive: many pilots over sixty have a very low heart attack risk, far lower than that of many pilots under sixty. A case-by-case determination based on medical records would surely have led to a more calibrated response. Even a rule that took into account not only age, but multiple health factors would produce a smaller number of decisions in which relatively risk-free pilots would be grounded.

Some of a rules' inherent limitations can be counteracted by according discretion to line employees. Reintroducing, or retaining, elements of discretion can be particularly important when decisions must be based on circumstances or factors that either: 1) were not envisioned by rule-drafters (rules can quickly be undermined by new scientific, economic, social, or other developments); or 2) cannot be quantified.⁴¹ So agency leadership must accord low-level decision-makers some discretion.⁴² But what if rules could be fine-grained, to take virtually innumerable factors into account? The subset of wrong decisions would become narrower.⁴³

Agencies must also contend with various *external* forces. Agencies' legitimacy rests upon their responsiveness to the elected officials of the executive and legislative branches, namely the president and Congress. The president and Congress must retain the capacity to assert control over agencies, through the exercise of the executive authority and congressional oversight, *inter alia*, and change agency behavior by enactment of statutes modifying the law. But even such legislative and executive oversight is insufficient to ensure agency fidelity to law.⁴⁴ Thus, agency decisions are generally subject to judicial review as well, to ensure that agencies remain faithful to their statutory mandates. Nevertheless, judicial review is generally deferential. On-the-record adjudications need only be based upon "substantial evidence." Less informal adjudications and regulations need only satisfy the "arbitrary and capricious" standard of review.⁴⁵

The public, and both the relevant regulated entities and the beneficiaries, must have notice of their obligations. Regulated entities must be able to predict how agencies will decide cases, and beneficiaries must also be able to determine when a challenge to a regulated entity's actions is warranted. Moreover, no agency can long prosper without the general support of the public, or at least key constituencies.⁴⁶

What are the implications of governments' use of AI/ML? Use of AI/ML algorithms will increase uniformity of adjudicatory and enforcement decisions, and their more fine-grained metrics should minimize the subset of incorrect decisions.⁴⁷ But agencies will face a basic decision: should the algorithms' decisions be binding or nonbinding?

If binding, many fewer line officials, that is, bureaucrats, will be needed to implement the program on the ground, and those that remain may well experience a decline in status within the agency. But in embracing AI/ML algorithms, agency leadership may merely have traded one management problem for another: managing the data specialists assuming a more central role in the agency's implementation of its programs. They will make decisions about the algorithm, the data used to train it, and the tweaks necessary to keep it current. Nonexpert leadership may feel even less capable of managing data scientists than the line officials they replaced.

If the algorithm is nonbinding, the key question will be *when* to permit human intervention. There is reason to believe that permitting overrides will produce no better results than relying on the algorithm itself.⁴⁸ Of course, agency leadership may disagree. In that case, the challenge will be to structure human intervention so as to avoid reintroducing the very problem the AI/ML algorithm was created to solve: unstructured, intuitive discretion leading to discrepant treatment of regulated entities and beneficiaries.

The uniformity wrought by AI/ML algorithms will come at the cost of increasing the opacity of the decision-making criteria and, potentially, the intuitiveness of the decision metric.

Explainability is critical within the agency. It is critical to any attempt to have a line-level, or upper-level, override system. If one does not know the weight the AI/ML algorithm accorded various criteria, how is one supposed to know whether it gave that consideration appropriate weight? At the same time, the algorithms' opacity might lead to staff resistance to such AI/ML decisions.⁴⁹ Lack of explainability poses challenges to agency managers seeking to retain control over policy, because not even the agency head can reliably discern with precision the policy the AI/ML algorithm applies in producing its decisions. At best, agency leadership will be dependent on computer and data processing specialists as critical intermediaries in attempting to manage the algorithm.

AI/ML algorithm's lack of explainability impedes the agency's navigation of its external environment as well. It complicates relationships with Congress and the components of the Executive Office of the President (EOP), like the OMB, with which the agency interacts. The more opaque and less intuitive the explanation of the AI/ML's metrics and decision-making process, the harder it will be to convince members of Congress and the relevant EOP components of the soundness of the agency's decisions. And the more fine-grained the nonintuitive distinctions between applicants for assistance or regulated entities, the more those distinctions will be viewed as literally arbitrary (that is, turning on inexplicable distinctions) and, well, bureaucratic. The reaction of the general public will presumably be even more extreme than that of elected leaders and their staff.

But let us turn to the implications of AI/ML's lack of explainability for judicial review. While judicial review of agency decision-making is deferential, it is hardly perfunctory.⁵⁰ In many circumstances, agency decisions are a type of prediction, even though they may not be framed in that way. Does licensing this pilot pose a risk to public safety? Is this applicant for benefits unable to obtain a job? These are questions in which AI/ML algorithms excel. But, as noted earlier, some agency decisions require a determination regarding past events. Sometimes the facts, one might say "the data," are in dispute. Two people might have a different account of a key conversation between a management official and an employee central to determining whether an unfair labor practice occurred. Current AI/ML algorithms are unlikely to provide much assistance in resolving such a contest.

In addition, if a statute is applicable, an AI/ML algorithm might be incapable of producing a decision *explaining* the result to the satisfaction of a court. The Supreme Court's decision in *Allentown Mack Sales & Service v. NLRB* provides a cautionary tale. There, a company refused to bargain with a union, asserting a "reasonable doubt" that a majority of its workforce continued to support the union. In practice, the National Labor Relations Board (NLRB) required employers making such an assertion to prove the union's loss of majority support. The Court held that an agency's application of a rule of conduct or a standard of proof that diverged from the formally announced rule or standard violated basic principles of adjudication.⁵¹ But that is what an AI/ML algorithm does: it creates a standard different from that announced, which may well be nonintuitive, and then consistently applies it *sub rosa*. AI/ML algorithms reveal that certain data inputs are commonly associated with particular outcomes to which we accord legal significance, but fail to show the basis for believing that the correlation held in a particular circumstance that occurred in the past. In other words, AI/ML can make predictions about the future, but offers little insight into how the record in the particular case leads to particular conclusions with respect to legally significant historical facts.

And often, in close cases, an agency can support either decision open to it. Is the reviewing court to be satisfied with reversing only "clearly erroneous" AI/ML-

produced decisions? Some have suggested that courts review the process for decision-making rather than the outcomes produced by AI/ML algorithms.⁵²

That approach certainly has appeal, but how is the nonexpert (perhaps mostly technophobic) judiciary supposed to review the AI/ML algorithms? Courts faced a similar dilemma when Congress created new regulatory agencies for complex scientific and technological subjects, accorded other agencies more rule-making power, and permitted more pre-enforcement challenges to regulations. The court's response was a "hard-look" approach, ensuring that the relevant factors were considered, irrelevant factors were not, and that public participation was guaranteed.⁵³

Explainability, in another sense, is also important with respect to legislative rules. Let us say that an agency seeks to make explicit what is implicit in an AI/ML algorithm. Assume an AI/ML algorithm finds a correlation between long-haul truck drivers involved in accidents and 1) drivers' credit scores; 2) certain genomic markers; and 3) a family history of alcohol abuse. The agency could license or de-license based on a grid capturing the correlation. How would such a rule fare?

First, correlation does not equal causation. Some additional factor(s) more intuitively relevant to a driver's dangerousness might be propelling the relationship between it and the three variables. Given the basic requirement for some logical relationship between a regulation and its purposes, courts will surely demand either some *intuitive* relationship or *nonintuitive causal* relationship between the variables and truck driver dangerousness. After all, even if there is a fairly high correlation between the variables and truck driver dangerousness, many individuals will be excluded from truck driving due to apparently irrelevant factors. The agency will presumably have to provide the intuitive or causal relationship for the regulation to avoid its invalidation as "arbitrary and capricious."⁵⁴

The example points to another problem. We want to base regulatory limitations (or provision of benefits) on people's conduct, not their traits, either immutable, like genomic markers or family history, or mutable but irrelevant, like credit score. One's reward or punishment by the government should turn on conduct to be encouraged or deterred, not accidents of birth. And to the extent the correlation involves a mutable marker, potential truck drivers will focus on improving their performance on a characteristic that does not improve their driving, like raising their credit score, rather than improving their capabilities as drivers. And, of course, some characteristics, like race and gender, cannot be used, unless the agency can proffer a strong justification that is not based on treating an individual as sharing the characteristic of his or her group.⁵⁵

Third, the function of notice-and-comment requirements would be undermined if the agency can conclude that its process for developing the algorithm is sound, and thus that the correlation is valid, even though the reason the correlation makes sense remains a mystery. Commenters themselves would have to in-

investigate the correlation to either prove it is coincidental (essentially disproving all possible reasons for the existence of the correlation) or identify the underlying causes driving the correlation.

In short, even if an agency reveals its AI/ML algorithms' magic, by attempting to capture an AI/ML-discovered correlation in a legislative rule, the agency's attempt to promulgate a counterintuitive rule will likely fail.

Briefly turning to agency enforcement efforts, courts have recognized, particularly in the Freedom of Information Act (FOIA) context, the inherent tension between making sure there is no "secret law" and preventing circumvention of the law.⁵⁶ Transparency may mean that the enforcement criteria will become the effective rule, replacing the law being enforced. And given the complexity of AI/ML algorithms, transparency could have a disparate effect depending on the wealth and sophistication of the regulated entity.⁵⁷

Nevertheless, to the extent transparency is desirable, it will be more difficult to achieve when the AI/ML algorithm is proprietary, as the FOIA probably allows the agency to withhold such information and the government may feel compelled to do so.⁵⁸

Technology tends to make fools of those who venture predictions. Nevertheless, the potential that AI/ML will reduce the number and status of line-level employees is present. But before AI/ML makes significant inroads, agencies will have to grapple with making AI/ML algorithms' "black box" magic more transparent and intuitive.

ABOUT THE AUTHOR

Bernard W. Bell is Professor of Law and Herbert Hanoach Scholar at Rutgers University. He has published in such journals as *Stanford Law Review*, *Texas Law Review*, and *North Carolina Law Review*.

ENDNOTES

- ¹ See David Freeman Engstrom, Daniel E. Ho, Catherine M. Sharkey, and Mariano-Florentino Cuéllar, *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies* (Washington, D.C.: Administrative Conference of the United States, 2020).
- ² Mark Bovens and Stavros Zourdis, "From Street-Level to System-Level Bureaucracies: How Information and Communication Technology Is Transforming Administrative Discretion and Constitutional Control," *Public Administration Review* 62 (2) (2002): 174.

- ³ Max Weber, *Economy and Society: An Outline of Interpretive Sociology*, ed. Guenther Roth and Claus Wittich (Berkeley: University of California Press, 1978), 975.
- ⁴ The concerns about government data storage and its privacy implications were explored early on by an influential government advisory committee. See U.S. Department of Health, Education, and Welfare, *Records, Computers, and the Rights of Citizens*, Report of the Secretary's Advisory Committee on Automated Personal Data Systems (Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1973).
- ⁵ Paul Schwartz, "Data Processing and Government Information: The Failure of the American Legal Response to the Computer," *Hastings Law Journal* 43 (5) (1992): 1321, 1342.
- ⁶ Even AI/ML algorithms remain dependent on programmers to a certain extent. David Lehr and Paul Ohm, "Playing with the Data: What Legal Scholars Should Learn about Machine Learning," *UC Davis Law Review* 51 (2) (2017): 653, 672–702. This includes unsupervised algorithms, which are dependent upon the choice of data set to train them.
- ⁷ Andrew Selbst and Solon Barocas, "The Intuitive Appeal of Explainable Machines," *Fordham Law Review* 87 (3) (2018): 1085, 1096–1099. This is eerily similar to the derisive caricature of legal realism as suggesting that judges' decisions may turn on the quality of their breakfast. See Alex Kozinski, "What I Ate for Breakfast and Other Mysteries of Judicial Decision Making," *Loyola of Los Angeles Law Review* 26 (4) (1993): 993.
- ⁸ Lehr and Ohm, "Playing with the Data," 706; Cary Coglianese and David Lehr, "Regulating by Robot: Administrative Decision Making in the Machine-Learning Era," *The Georgetown Law Journal* 105 (5) (2017): 1147, 1159–1160; and Engstrom et al., *Government by Algorithm*, 28.
- ⁹ See Lehr and Ohm, "Playing with the Data," 707–710 (describing four potential types of explanations for the results produced by AI/ML algorithms).
- ¹⁰ John D. Kelleher and Brendan Tierney, *Data Science* (Cambridge, Mass.: The MIT Press, 2018), 72–74; and Engstrom et al., *Government by Algorithm*, 12.
- ¹¹ Kelleher and Tierney, *Data Science*, 109.
- ¹² Aziz Z. Huq, "The Right to a Human Decision," *Virginia Law Review* 675 (3) (2020): 611, 634.
- ¹³ *Ibid.*
- ¹⁴ Engstrom et al., *Government by Algorithm*.
- ¹⁵ Felix F. Bajandas and Gerald K. Ray, "Implementation and Use of Electronic Case Management Systems in Federal Agency Adjudication," Administrative Conference Recommendation 2018-3 (Washington, D.C.: Administrative Conference of the United States, 2018), 44–49.
- ¹⁶ Engstrom et al., *Government by Algorithm*, 59–64.
- ¹⁷ *Natural Resources Defense Council v. EPA*, 954 F.3d 150, 153 (2d Cir. 2020).
- ¹⁸ *Ibid.*
- ¹⁹ Coglianese and Lehr, "Regulating by Robot," 1174–1175.
- ²⁰ *Heckler v. Chaney*, 470 U.S. 821 (1985); and *Morrison v. Olson*, 487 U.S. 654, 727–728, 731–732 (1988) (Scalia dissenting).

- ²¹ See *Environmental Defense Fund v. Ruckelshaus*, 439 F.2d 584, 592 (D.C. Cir. 1971); *Nor-Am Agri Prods. v. Hardin*, 435 F.2d 1151 (7th Cir. 1970) (*en banc*); and *Dow Chemical v. Ruckelshaus*, 477 F.2d 1317 (8th Cir. 1973).
- ²² Jerry L. Mashaw, *Bureaucratic Justice: Managing Social Security Disability Claims* (New Haven, Conn.: Yale University Press, 1983).
- ²³ H. Laurence Ross, *Settled Out of Court: The Social Process of Insurance Claims Adjustment*, 2nd ed. (New York: Routledge, 1980).
- ²⁴ Matthew Diller, “Entitlement and Exclusion: The Role of Disability in the Social Welfare System,” *UCLA Law Review* 44 (1998): 361, 372–374.
- ²⁵ See *Bowen v. Yuckert*, 482 U.S. 137, 140–142 (1987).
- ²⁶ *Heckler v. Campbell*, 461 U.S. 458, 460–462 (1983).
- ²⁷ Lehr and Ohm, “Playing with the Data,” 670–672.
- ²⁸ Say A’s computer screensaver is blue 90 percent of the time. But yesterday, B had a view of the computer screen at a particular time, say 9:33 pm. If we wanted to predict the color of A’s screen in the future, reliance on the 90 percent algorithm is fine. But it would seem odd to do so if the issue is whether the screen was blue at 9:33 pm yesterday, given the availability of a witness. In short, reliance on the rule can never tell one when the rule is incorrect in a particular situation. See Huq, “The Right to a Human Decision,” 679.
- ²⁹ *Restatement of the Law, Third, Torts: Liability for Physical and Emotional Harm, Volume 1* (Philadelphia: American Law Institute, 2010), §28, comm. c(3) & c(4).
- ³⁰ Coglianese and Lehr, “Regulating by Robot,” 1172–1173.
- ³¹ Some can be promulgated even less informally. See 5 U.S.C §553.
- ³² Jeffrey S. Lubbers, *A Guide to Federal Agency Rulemaking*, 6th ed. (Chicago: American Bar Association, 2018), 252–263.
- ³³ Bernard W. Bell, “Legislative History without Legislative Intent: The Public Justification Approach to Statutory Interpretation,” *Ohio State Law Journal* 60 (1) (1999): 1, 35 (discussing the rational basis test). Granted, such rationality rarely results in invalidation of federal statutes because of the federal judiciary’s concerns about its institutional position in a democracy. *Ibid.*, 28–30.
- ³⁴ *Motor Vehicle Manufacturers Association v. State Farm Insurance*, 463 U.S. 29, 43, n. 9 (1983).
- ³⁵ *Sun Ray Drive-In Dairy, Inc. v. Oregon Liquor Control Commission*, 517 P.2d 289, 293 (Or. Ct. App. 1973).
- ³⁶ *Sun Ray Drive-In Dairy*, 517 P.2d, 293 (discussing necessity of “written standards and policies”).
- ³⁷ Jennifer Nou, “Civil Servant Disobedience,” *Chicago Kent Law Review* 94 (2) (2019): 349, 349–350, n. 1–4. Bureaucratic resistance is hardly new. President Truman reportedly quipped that if General Eisenhower became president: “He’ll sit here, and he’ll say, ‘Do this! Do that!’ And *nothing will happen*. Poor Ike—it won’t be a bit like the Army.” Richard E. Neustadt, *Presidential Power: The Politics of Leadership* (New York: John Wiley & Sons, 1960).
- ³⁸ See Frederick Schauer, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* (Oxford: Clarendon Press, 1993), 48–52, 86–87, 104, n. 35.

- ³⁹ *Ibid.*, 31–34.
- ⁴⁰ *Air Line Pilots v. Quesada*, 276 F.2d 892 (2d Cir. 1961), *cert. denied*, 366 U.S. 962 (1961).
- ⁴¹ *Yetman v. Garvey*, 261 F.3d 664, 679 (7th Cir. 2001) (observing that “a strict age sixty cutoff, without exceptions, [might be] better suited to 1959 than to 2001”).
- ⁴² Rule-like adjudicatory criteria may need to have exceptions—*Heckler v. Campbell*, 461 U.S., 467, n. 11; and *U.S. v. Storer Broadcasting*, 351 U.S. 192 (1956)—or a robust regime permitting waivers for situations in which the rule is just not appropriate. See Alfred Aman, “Administrative Equity: An Analysis of Exceptions to Administrative Rules,” *Duke Law Journal* 31 (2) (1982): 277; and Aaron L. Nielson, “Waivers, Exemptions, and Prosecutorial Discretion: An Examination of Agency Nonenforcement Practices,” J. Reuben Clark Law School, Brigham Young University Research Paper No. 17-27 (Washington, D.C.: Administrative Conference of the United States, 2017).
- ⁴³ See Schauer, *Playing by the Rules*, 83, 155.
- ⁴⁴ The political branches of government may not be able to respond to every agency action or may desire or even encourage agencies’ unfaithfulness to the intent of the legislature that enacted the statute.
- ⁴⁵ 5 U.S.C. §706(2).
- ⁴⁶ For a general discussion of the external environment of agencies and its implications for agency decision-making, see Rachel Augustine Potter, *Bending the Rules: Procedural Politicking in the Bureaucracy* (Chicago: University of Chicago Press, 2019); and *Sun Ray Drive-In Dairy v. Oregon Liquor Control Commission*, 517 P.2d at 293–294.
- ⁴⁷ This all assumes that circumstances do not fundamentally change. Ali Alkhatib and Michael Bernstein, “Street-Level Algorithms: A Theory at the Gaps Between Policy and Discretion,” CHI ’19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems Paper No. 530, May 2019. However, AI/ML algorithms are dynamic; they can continue to revise themselves in light of later data. Coglianese and Lehr, “Regulating by Robot,” 1159. Legislative rules are notorious for their “ossification.” See, for example, Lubbers, *A Guide to Federal Agency Rulemaking*, 404–405. The 1978 grid is still used by the SSA. While AI/ML algorithms offer the prospect of better modifying decision rules in light of changes in the data, one can anticipate judicial discomfort with nontransparent, nonintuitive rules changes in decision rules outside the normal “notice-and-comment” procedures applicable to the revision of rules. See, for example, *Appalachian Power Co. v. EPA*, 208 F.3d 1015, 1019 (D.C. Cir. 2000).
- ⁴⁸ Huq, “The Right to a Human Decision,” 665–667.
- ⁴⁹ Engstrom et al., *Government by Algorithm*, 28 (reporting SEC line-level enforcement staff seeking more information regarding AI/ML designation of certain investment advisors as “high risk”).
- ⁵⁰ Michael Asimow, ed., *A Guide to Federal Agency Adjudication* (Chicago: American Bar Association, 2003), 85, n. 9–10.
- ⁵¹ The Court noted that “the consistent repetition of [such a] breach can hardly mend it.” *Allentown Mack Sales & Service v. NLRB*, 522 U.S. 359, 374 (1998).
- ⁵² Huq, “The Right to a Human Decision,” 675; and Lehr and Ohm, “Playing with the Data,” 710–716.

- ⁵³ See *Natural Resources Defense Council v. U.S. Nuclear Regulatory Commission*, 547 F.2d 633, 655–657 (D.C. Cir. 1976) (Bazelon concurring) (“in highly technical areas, where judges are institutionally incompetent to weigh evidence for themselves, a focus on agency procedures will prove less intrusive, and more likely to improve the quality of decisionmaking, than judges ‘steeping’ themselves ‘in technical matters’”).
- ⁵⁴ Here I disagree with Coglianesi and Lehr, “Regulating by Robot,” 1202.
- ⁵⁵ Coglianesi and Lehr believe this can be overcome, if only the AI/ML algorithm is sufficiently fine-tuned and takes multiple factors into account. *Ibid.*, 1201. I am far more skeptical. For instance, considering life expectancy charts that use race as a variable would be unconstitutional, even if an AI/ML algorithm were used to make the tables more complex and fine-tuned by taking numerous other variables into account.
- ⁵⁶ See *Dirksen v. HHS*, 803 F.2d 1456, 1458–1459, 1461–1462 (9th Cir. 1986) (Ferguson dissenting) (the majority’s approach makes “the risk of circumvention . . . indistinguishable from the prospect of enhanced compliance”). The high-2 exemption that *Dirksen* applied was overturned in *Milner v. Department of the Navy*, 562 U.S. 562 (2011). The FOIA exemption for law enforcement records that poses a risk of circumvention remains.
- ⁵⁷ Engstrom et al., *Government by Algorithm*, 86–87.
- ⁵⁸ Coglianesi and Lehr, “Regulating by Robot,” 1210–1211. See also *Food Marketing Institute v. Argus Leader Media*, 139 S.Ct. 2356 (June 24, 2019) (expanding FOIA exemption 4). However, the ACUS study reported that a majority of the AI/ML algorithms in their survey were developed by the government in-house. Engstrom et al., *Government by Algorithm*, 18.