A GOVERNMENT OF LAWS AND NOT OF MACHINES

EMILY BERMAN[*]

[*] Assistant Professor of Law, University of Houston Law Center. Thanks to Seth Chandler, Ashley Deeks, Victor Flatt, Tracy Hester, Aziz Huq, David Kwok, Peter Margulies, Doug Moll, James Nelson, D. Theodore Rave, Joe Sanders, and participants in the 2018 American Association of Law Schools session on the intersection of technology and civil rights as well as a South Texas School of Law workshop for helpful comments. Thanks also go to Seth Chandler and James Winkle for their patience in explaining some of the more technical aspects of machine learning.

Each week brings another story touting the miracle of "machine learning"—a strand of artificial intelligence that uses mathematical algorithms to construct computer models that analyze enormous data sets, often for the purpose of making predictions about the future. Machine learning is all around us—it is used for spam filters, facial recognition, detecting bank fraud, calculating credit risk, and much more—and it is immensely powerful. This Article considers the government's use of machine learning in the context of law enforcement and national security decision-making, taking a step back from the nuts and bolts questions surrounding the implementation of predictive analytics, on which most scholarly commentary has focused, to assess their use from a more conceptual perspective. The question I seek to answer is this: whether reliance on the output of machine-learning models—even if highly accurate—is consistent with the goal to maintain "a government of laws" and not of machines. I conclude that government officials operating in contexts where they enjoy broad decision-making discretion should embrace machine-learning predictions as a valuable tool. By contrast, when government discretion is highly constrained by existing constitutional, statutory, or regulatory rules, the use of machine-learning predictions represents a threat to the rule-of-law.

INTRODUCTION

"Machine learning" is a strand of artificial intelligence that sits at the intersection of computer science, statistics, and mathematics, and it is changing the world.[1] The applications of machine learning in modern society are nearly

[1] *See, e.g.*, Steve Barrett, *AI is Changing the World, but Will It End in Utopia or Dystopia*?, PR WEEK (Feb. 9, 2018), https://www.prweek.com/article/1456842/ai-changing-world-will-end-utopia-dystopia (highlighting benefits and dangers of improved artificial intelligence technology); Bernard Marr, *5 Key Artificial Intelligence Predictions for 2018: How Machine Learning Will Change Everything*, FORBES (Dec. 18, 2017, 12:28 AM), https://www.forbes.com/sites/bernardmarr/2017/12/18/5-key-artificial-intelligence-predictions-for-2018-how-machine-learning-will-change-everything/#7a1c79c56545 ("I expect 2018 to provide a continuous stream of small but sure steps forward, as machine learning and neural network technology takes on more routine tasks."); Roger Parloff, *Why Deep Learning Is Suddenly Changing Your Life*, FORTUNE (Sept. 28, 2016, 5:00 PM), http://fortune.com/ai-artificial-

endless: search engines, spam filters, Amazon and Netflix recommendations, voice and facial recognition, self-driving cars, spotting bank fraud, creditworthiness determinations, medical diagnoses, apps that transform your photos into the style of your favorite painter, robotic vacuum cleaners, and automated weapons.[2] Machine learning is thus an immensely powerful tool that already has transformed society's ability to exploit data.

Machine learning is, in essence, a particularly powerful version of data mining. The value of data mining stems from its capacity to make sense of so-called "big data"—enormous databases full of various bits of information. These data sets are too complex for humans to understand because of the volume of the information, because there are too many variables for humans to process, or because the meaningful relationships among the data are not self-evident. What sets machine learning apart from other forms of data mining is that the mathematical algorithms on which it relies to identify the meaningful patterns within a data set are able to learn from experience and become more accurate over time.[3] These algorithms make inferences from the data to generate

intelligence-deep-machine-learning/ [https://perma.cc/T7UU-CZ9G] (stating breakthroughs in voice recognition, image recognition, and machine translation are all due to artificial intelligence); Tom Simonite, *The Wired Guide to Artificial Intelligence*, WIRED (Feb. 1, 2018, 9:22 AM), https://www.wired.com/story/guide-artificial-intelligence/ ("The current boom in all things AI was catalyzed by breakthroughs in an area known as machine learning.").

[2] *See, e.g.*, Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 4 (2017) ("A credit card company uses behavioral-scoring algorithms to rate consumers' credit risk . . . ."); Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 88 (2014) ("In the last few decades, researchers have successfully used machine learning to automate . . . autonomous (i.e., self-driving) cars . . . ."); John Brandon, *Why the iRobot Roomba 980 Is a Great Lesson on the State of AI*, VENTUREBEAT (Nov. 3, 2016, 4:10 PM), https://venturebeat.com/2016/11/03/why-the-irobot-roomba-980-is-a-great-lesson-on-the-state-of-ai/ [https://perma.cc/D22G-6APR] ("[T]he latest Roomba . . . uses true AI."); Kyle Mizokami, *Kalashnikov Will Make an A.I.-Powered Killer Robot*, POPULAR MECHANICS (July 19, 2017), https://www.popularmechanics.com/military/weapons/news/a27393/kalashnikov-to-make-ai-directed-machine-guns/ [https://perma.cc/QX92-7F6G] ("Russian weapons maker Kalashnikov is working on an automated gun system that uses artificial intelligence to make 'shoot/no shoot' decisions."); Kumba Sennaar, *Machine Learning for Medical Diagnostics—4 Current Applications*, TECHEMERGENCE (Jan. 11, 2018), https://www.techemergence.com/machine-learning-medical-diagnostics-4-current-applications/ [https://perma.cc/NL4C-E4NX] ("Today, AI is playing an integral role in the evolution of the field of medical diagnostics."); PIKAZO, INC., http://www.pikazoapp.com/ [https://perma.cc/T7MK-AYLF] (last visited Sept. 11, 2018).

[3] *See* Surden, supra note 2, at 89 (explaining that machine-learning algorithms are "capable of changing their behavior to enhance their performance on some task through experience").

computer models that expose new insights about the world and, in many instances, make predictions about the future.[4]

Given its utility, it is not surprising that government decision-makers seek to harness machine learning's predictive power for public-sector use. These tools already have made significant inroads in the contexts of national security and law enforcement. In these areas, predictive computer models promise to allocate government resources more efficiently, reduce the impact of conscious or unconscious bias in decisionmaking, pinpoint criminal activity and threats that would otherwise go undetected, and provide unexpected insights into the behavior of those who pose threats to national security and the security of our communities.[5] The use of machine-learning predictions has been integral, for example, to bail reform efforts across the country, where computer models that generate "risk assessments"—which purport to identify which defendants pose a flight risk or a danger to the community—have resulted in vastly reduced numbers of pretrial detainees.[6]

Despite machine learning's promise, there is significant and growing literature highlighting challenges that arise when using machine-learning predictions.[7] Use of machine learning for risk assessments in pretrial

[4] *See* JOHN D. KELLEHER, BRIAN MAC NAMEE & AOIFE D'ARCY, FUNDAMENTALS OF MACHINE LEARNING FOR PREDICTIVE DATA ANALYTICS 1 (2015) (describing predictive data analytics). Systems that use machine learning to make predictions can also be labeled "predictive analytics," "predictive algorithms," or "predictive models." *See id.*; David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653, 666, 671 (2017).

[5] *See* Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327, 389-95 (2015) (explaining benefits of law enforcement use of big data as indicator of suspicion).

[6] *See* Lauryn P. Gouldin, *Disentangling Flight Risk from Dangerousness*, 2016 BYU L. REV. 837, 867-71 (2016) (describing data-driven risk-assessment tools used by judges at pretrial proceedings). For many more examples, see *infra* Section I.B.

[7] *See, e.g.*, BRUCE SCHNEIER, DATA AND GOLIATH: THE HIDDEN BATTLES TO COLLECT YOUR DATA AND CONTROL YOUR WORLD 273 (2015) (noting balancing of benefits to society versus costs to individuals involved in use of big data "is done for us by governments and corporations with their own agendas"); Steven M. Bellovin et al., *When Enough Is Enough: Location Tracking, Mosaic Theory, and Machine Learning*, 8 N.Y.U. J.L. & LIBERTY 556, 621-24 (2014) (describing possible violations of reasonable expectation of privacy from law enforcement use of machine learning in location tracking); Kiel Brennan-Marquez, *"Plausible Cause": Explanatory Standards in the Age of Powerful Machines*, 70 VAND. L. REV. 1249, 1255-57 (2017) (arguing use of algorithms in law enforcement decisionmaking threatens traditional criminal justice system values); Citron & Pasquale, *supra* note 2, at 4 ("Because human beings program predictive algorithms, their biases and values are embedded

proceedings, to take an example from the law enforcement context, has not yielded unanimous praise. A 2016 report by the nonprofit investigative journalism organization ProPublica studied one risk assessment tool, called Correctional Offender Management Profiling for Alternative Sanctions ("COMPAS"), and determined that it was racially biased.[8] It misidentified African Americans as risks at double the rate of white people.[9] That is to say,

into the software's instructions, known as the source code and predictive algorithms."); Ferguson, *supra* note 5, *passim* (exploring "whether a Fourth Amendment stop can be predicated on the aggregation of specific and individualized, but otherwise noncriminal, factors"); Andrew Guthrie Ferguson, *Predictive Policing and Reasonable Suspicion*, 62 EMORY L.J. 259, 313-25 (2017) (discussing primary constitutional concerns implicated by predictive policing); Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*, 10 HARV. L. & POL'Y REV. 15, 17 (2016) (describing expansion of surveillance discretion made possible by big data tools as implicating legal questions about police oversight); Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 695-96 (2017) (emphasizing need for transparency and accountability in public policy decisionmaking based on algorithms); Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 U. PA. L. REV. 871, 925 (2016) (describing inaccurate data and human error as two possible causes of error in using algorithms to automate findings of suspicion); Ric Simmons, *Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in Our Criminal Justice System*, 2016 MICH. ST. L. REV. 947, 950 (2016) (presenting obstacles to effective incorporation of big data tools into criminal justice system); Surden, *supra* note 2, at 105-07 (discussing limitations of machine-learning legal predictive models); Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1503, 1506 (2013) (stating use of predictive practices based on analysis of personal information and data mining by law enforcement may result in biased, discriminatory processes threatening privacy and autonomy). Cautionary calls have emerged in other contexts as well. *See, e.g.*, Citron & Pasquale, *supra* note 2, at 10-13 (explaining concerns surrounding use of machine-learning predictions in credit scores); Dana A. Remus, *The Uncertain Promise of Predictive Coding*, 99 IOWA L. REV. 1691, 1706-17 (2014) (discussing concerns surrounding use of machine-learning predictions in discovery in legal cases).

[8] Julia Angwin et al., *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [https://perma.cc/4KQV-QC7K] (concluding that when isolating effect of race from other variables, statistical test of COMPAS showed black defendants were seventy-seven percent more likely to be identified as higher risk of committing future violent crime and forty-five percent more likely to be predicted to commit any future crime); *see also Algorithms in the Criminal Justice System*, ELECTRONIC PRIVACY INFORMATION CENTER, https://epic.org/ algorithmic-transparency/crim-justice/ [https://perma.cc/M34D-9FE9] (last visited Sept. 11, 2018) (warning that algorithms may use information on demographics, socioeconomic background, and family characteristics as proxies for race).

[9] Angwin et. al., *supra* note 8.

when it came to Type I errors, or false positives—a prediction of future criminal activity that turns out to be incorrect—the algorithm determined erroneously that black people would transgress twice as often as it made the same error with respect to white people.[10] Existing critiques of machine learning not only point to concerns about bias but also identify possible pitfalls surrounding data quality, algorithm selection, and model verification that might result in inaccurate models.[11] In addition, this scholarly commentary has explored whether the use of machine learning is consistent with norms such as transparency, accountability, and due process.[12] These accounts adeptly identify important challenges posed by certain uses of machine-learning predictions and merit significant attention.

This Article steps back from these nuts and bolts questions surrounding the implementation of predictive analytics, however, to assess their use from a more conceptual perspective. The question this Article seeks to answer is this: even assuming algorithmic models yield an accurate result in ninety-nine percent of national security and law enforcement decisions, can the use of these models to make such decisions *ever* conform to the fundamental values underlying our legal framework? That is to say, whether reliance on the output of machine-learning models—even if highly accurate—is in tension with the goal to maintain "a government of laws" and not of machines.[13]

This Article contends that certain characteristics of predictive analytics inevitably bring them into tension with rule-of-law principles. As a result, while machine-learning predictions can be valuable instruments in some decision-making contexts, they constitute a threat to fundamental values in others. Three

[10] *Algorithms in the Criminal Justice System*, *supra* note 8 ("[T]he [COMPAS] formula was particularly likely to flag black defendants as future criminals, labeling them as such at almost twice the rate as white defendants."). *But see* WILLIAM DIETERICH, CHRISTINA MENDOZA & TIM BRENNAN, COMPAS RISK SCALES: DEMONSTRATING ACCURACY EQUITY AND PREDICTIVE PARITY 30-32 (2016), https://university.pretrial.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=b31d4b9f-9ba8-6357-4c08-9839963679df&forceDialog=0 [https://perma.cc/K7T6-QJPZ] (critiquing ProPublica's findings).

[11] *See infra* Section II.A (discussing "concrete obstacles inherent in building and implementing accurate, effective machine-learning programs").

[12] *See* sources cited *supra* note 7.

[13] JOHN ADAMS, *Essay No. 7*, *in* NOVANGLUS 84 (Hews & Goss ed. 1819) (noting that Aristotle and Livy "define a republic to be a government of laws, and not of men" (emphasis omitted)). In President Adams's formulation, a government of men ruled according to the unpredictable whim of those in power, whereas a government of laws was based on universally applicable rules.

particular characteristics of machine learning, as it currently exists,[14] are relevant to the discussion. First, many computer models that emerge from the machine-learning process cannot be explained in terms intelligible to humans—they are unavoidably opaque.[15] Second, the models and their predictions are based on identified correlations within a data set rather than proven causal relationships. The result is that inevitable errors—even if small in number—will be in some sense arbitrary.[16] Third and finally, models generated through machine learning inescapably reflect the values, biases, and judgment calls of their programmers, often in ways that are invisible on the face of the algorithm.[17]

Each of these characteristics threatens crucial elements of a rule-of-law framework—the predictability that makes it possible for individuals to plan their affairs, the need for constraints on arbitrary exercises of government power, and ultimately government legitimacy.[18] Now, it is not the case that the "rule of law" is either present or absent in any given system. Rather, "rule of law" is a question of degree, and in some contexts, the idea does less work than in others. When it comes to use of machine-learning predictions, there is a continuum on which rule-of-law concerns are less salient where executive decision-making enjoys a great deal of legal discretion, the costs of false positives are relatively modest, and efficiency gains are potentially significant. By contrast, the more highly constrained the decision—whether those constraints originate in constitutional doctrine, statutory rules, or regulatory strictures—and the higher the costs of false positives, the more important it becomes to satisfy rule-of-law principles and, as a result, the more problematic the use of machine-learning predictions becomes. Ultimately, this Article argues that government actors should exploit the benefits of machine learning when they enjoy broad discretion in making decisions, while eschewing the tool for decision-making when government discretion is highly constrained.

This Article sounds a cautionary note. While machine learning and predictive algorithms have enormous beneficial potential in the appropriate circumstances, there exists a strong temptation to employ them broadly. We cannot allow the

[14] The field of machine learning is evolving rapidly. Perhaps computer scientists will find ways to modify these characteristics one day. Efforts to render opaque algorithms more intelligible to humans, for example, abound. *See* sources cited *infra* note 167.

[15] *See infra* Section II.B.2.a. Certain types of algorithms lend themselves more to transparency than others. Decision trees, for example, can be explained in plain language, whereas neural networks usually cannot.

[16] *See infra* Section II.B.2.b (discussing arbitrariness of errors in computer model predictions).

[17] *See infra* Section II.B.2.c (discussing manner in which human programmers shape output of machine-learning algorithms).

[18] *See infra* Section II.B.2 (discussing tension between machine learning and core values of legal system).

current enthusiasm and commercial push for the deployment of these tools to overwhelm concerns about their use in instances where they do not comport with our most basic values. Exactly which uses of machine-learning predictions fall into this category is a question on which reasonable minds will disagree, but the time is unquestionably past due for a serious conversation regarding when we, as a society, are comfortable authorizing our government to use these tools in its national security, counterterrorism, and criminal law enforcement efforts.

This Article proceeds in three Parts. Part I will describe the basic features of machine learning as well as the various ways in which these tools are being used in law enforcement and national security programs. Part II will first briefly examine in Section A some of the pragmatic, technological concerns about algorithmic predictions. Section B will then turn to the idea of the rule of law, setting out the basic values that the principle seeks to vindicate and explaining how machine-learning predictions come into tension with those values. Part III will then consider the implications of these tensions, arguing that they are at their nadir in contexts where government decision-makers have broad discretion, whereas the more constraints already applicable to government decision-making, the more intense those tensions become.

I. DATA MINING FOR SECURITY

One challenge in writing about the exploitation of data is the lack of precise, universal definitions of the central concepts. Machine learning means different things to different people. Section A will begin, therefore, by discussing what the term encompasses generally before going into more detail with respect to the specific form of machine learning at issue in this Article. Section B will then identify instances where machine learning is being employed in the national security state. This list will be necessarily incomplete, as there are sure to be many such programs shrouded from public view. It should, however, provide a sense of the ways in which the government uses the techniques described below.

A. *Defining Machine Learning*

Machine learning is a species of data mining. While the term "data mining" is often used to describe any means of extracting knowledge from a data set, computer scientists tend to define the term more narrowly, emphasizing the extraction of implicit knowledge by discovering patterns or relationships within a data set.[19] A broad conception of data mining includes two types of tasks:

[19] *See* Liane Colonna, *A Taxonomy and Classification of Data Mining*, 16 SMU SCI. & TECH. L. REV. 309, 311 (2013) (noting that some would limit use of term "data mining" to machine learning); Christoph Schommer, *An Unified Definition of Data Mining*, COMPUTING

hypothesis-driven tasks—such as use of a simple query-and-report tool to discover, for example, a list of all U.S. persons in the database who have traveled to Afghanistan since 2001, or to identify all emails to or from a particular email address; and data-driven tasks—for which the computer itself develops the hypothesis, and the analyst cannot predict in advance what knowledge will emerge.[20] Machine learning is an example of the latter form of data mining.[21]

Machine learning is itself an umbrella term that encompasses many different techniques. What unites all machine-learning data analysis techniques is that they generate their own computer models and (if well-constructed) improve automatically with experience—they learn.[22] Machine learning entails the deployment of an "algorithm"—"a sequence of instructions telling a computer what to do"[23]—that generates a "model" capturing the relationship among data

RESEARCH REPOSITORY 2 (2008), http://arxiv.org/pdf/0809.2696.pdf [https://perma.cc/Y32F-FGME] (defining data mining as "nontrivial extraction of implicit, previously unknown, and potentially useful information from data").

[20] *See* Colonna, *supra* note 19, at 340 (explaining for data-driven tasks, "analyst does not need to start with a hypothesis, but rather ask the system to create one"). Even within the U.S. government the definition varies. *See* Federal Agency Data Mining Reporting Act of 2007, 42 U.S.C. § 2000ee-3(b)(1)(A) (2012) ("The term 'data mining' means a program involving pattern-based queries, searches, or other analyses of 1 or more electronic databases . . . to discover or locate a predictive pattern or anomaly indicative of terrorist or criminal activity . . . ."); JEFFREY W. SEIFERT, CONG. RESEARCH SERV., RL31798, DATA MINING AND HOMELAND SECURITY: AN OVERVIEW 1 (2008) ("Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. . . . [and] consists of more than collecting and managing data, it also includes analysis and prediction."); DEP'T OF DEF., TECH. AND PRIVACY ADVISORY COMM., SAFEGUARDING PRIVACY IN THE FIGHT AGAINST TERRORISM 4 (2004) (defining data mining as "searches of one or more electronic databases . . . by or on behalf of an agency or employee of the government" but only when those searches "concern[] U.S. persons"); U.S. GENERAL ACCOUNTING OFFICE, GAO-04-548, DATA MINING, FEDERAL EFFORTS COVER A WIDE RANGE OF USES 1 (2004) (defining data mining as "application of database technology and techniques—such as statistical analysis and modeling—to uncover hidden patterns and subtle relationships in data and to infer rules that allow for the prediction of future results").

[21] Zarsky, *supra* note 7, at 1519-20 (explaining minimal role of analysts in certain applications of data mining). The use of data to extract information is a multi-step process, sometimes referred to as "knowledge discovery in databases" ("KDD") of which the "mining" is only one part. Colonna, *supra* note 19, at 315.

[22] *E.g.*, PETER FLACH, MACHINE LEARNING, THE ART AND SCIENCE OF ALGORITHMS THAT MAKE SENSE OF DATA 3 (2012) ("Machine learning is the systematic study of algorithms and systems that improve their knowledge or performance with experience.").

[23] PEDRO DOMINGOS, THE MASTER ALGORITHM: HOW THE QUEST FOR THE ULTIMATE LEARNING MACHINE WILL REMAKE OUR WORLD 1 (2015).

from a given data set.[24] A model is the mathematical depiction of the relevant relationships among the data that the computer extracts.[25] By continually analyzing data over time, machine-learning algorithms can refine their models to produce increasingly accurate results. Those results can be descriptive, meaning that they simply seek to identify properties of the available data set, or predictive, meaning the knowledge is extracted from the known data for the purposes of predicting properties of new data.[26] This Article focuses on predictive uses of machine learning because that type of use raises the thorniest legal and policy questions.

The most traditional forms of machine learning are supervised and unsupervised,[27] and within each of these categories, a variety of techniques can be used to extract information from the data set. Supervised machine learning, which is the most common means of generating predictive models, is a multi-step process for finding a model that captures the relationship between a set of "features" of the data—its descriptive characteristics—and the "target feature"—the information we want the computer to teach itself.[28] Ideally, the data set will be quite large; as a general rule, the more data available the more effective the algorithm.[29] The ability to make sense of data sets far too voluminous for humans to analyze unassisted is part of what makes machine learning so valuable. The process of generating a model begins with a "training set," or a data set of examples—or "instances"—for which we know both the features of the data and the target feature.[30] For example, we might want to train a computer to recognize whether a particular image is an image of a cat. The computer would be presented with a large database of images, some of which

[24] Surden, *supra* note 2, at 91 ("The goal of [a machine-learning] algorithm is to build an internal computer model of some complex phenomenon . . . that will ultimately allow the computer to make automated, accurate . . . decisions."). In essence, algorithms instruct computers to "figure it out on their own, by making inferences from data." DOMINGOS, *supra* note 23, at xi.

[25] Surden, *supra* note 2, at 91-92.

[26] DOMINGOS, *supra* note 23, at xv ("[A]t its core, machine learning is about prediction: predicting what we want, the results of our actions, how to achieve our goals, how the world will change."); FLACH, *supra* note 22, at 18-19 (explaining predictive models).

[27] In addition, there is "semi-supervised" machine learning, which combines aspects of supervised and unsupervised machine learning, as well as the somewhat newer fields of reinforcement learning and generative models. This Article does not address those techniques.

[28] KELLEHER, *supra* note 4, at 3 (describing supervised machine learning).

[29] STUART J. RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE, A MODERN APPROACH 807-08 (3d ed. 2010) (discussing difficulties inherent in using small data sets).

[30] KELLEHER, *supra* note 4, at 3-4 (explaining purpose of training set in machine learning).

depict cats, as well as the appropriate label for each image—cat or not-cat. The computer will apply an algorithm to the labeled data to develop a model that identifies which images are cats.

Once the computer generates one or more models based on the training data, those models must be verified. It is not enough that a model accurately identifies all the cat images in the training set.[31] For a model to be useful as a predictive tool, it must be able to analyze data accurately that was *not* in the training set—it must be able to "generalize."[32] To determine whether a model can do so, some of the data usually will be withheld from the training set to be used as a "test set," which will be given to the computer without the target variables.[33] The goal is to find a model that captures meaningful relationships between the data's features and the target variable such that when confronted with a set of features it has not seen before, the model will nevertheless produce the correct target variable. A successful model will therefore be able to consider new, unfamiliar images and accurately identify which ones are cats. There may be several models that fit this description, and the programmer will then have to choose among them.

Unsupervised learning takes place when there is no training set—no set of data that includes the target variable—on which to train an algorithm. Instead, the computer is simply asked to identify structures, relationships, patterns, or trends within a data set. For example, based on purchasing patterns, payments to veterinary clinics, and a history of retweeting links to cat videos on YouTube, an algorithm might cluster a large set of data into cat owners, dog owners, and households without pets.[34] Or it might detect a pattern indicating that when someone buys cat food at the grocery store, he is also likely to buy kitty litter and Cheetos. Whether these patterns are meaningful or useful, however, is left

[31] *Id.* at 8-9 (explaining that consistency with training dataset is not sufficient to prove model's effectiveness as predictive tool). Note that there may be several models that are consistent with the data set from which an analyst must choose one. *See id.*

[32] *Id.* at 9 ("[T]he goal of machine learning is to find the predictive model that generalizes best.").

[33] DAVID SKILLICORN, KNOWLEDGE DISCOVERY FOR COUNTERTERRORISM AND LAW ENFORCEMENT 76 (2009) (explaining use of test data). In the absence of sufficient data to have both a training set and a test set, a programmer can use "cross validation" to check a model's accuracy by slicing up the training data in various ways. *Id.* Some machine learning will also employ "validation sets" which are used to identify the best-performing model if there is more than one. *See* ETHEM ALPAYDIN, INTRODUCTION TO MACHINE LEARNING 40 (Thomas Dietterich ed., 3d ed. 2014).

[34] Clustering is a common unsupervised learning task. RUSSELL & NORVIG, *supra* note 29, at 694.

for humans to determine. It is unlikely, for example, that law enforcement could predict whether you are likely to engage in terrorist violence based on whether you have a dog, or a cat, or are pet-less. If you are marketing cat food, by contrast, the ability to identify individuals who belong in the cat cluster will be quite valuable. Unsupervised learning can also be predictive, as one might use the revealed patterns to predict relationships among new data, but there is no test set on which to verify the accuracy of such predictions.[35]

Regardless of what type of machine learning is being used, there are multiple types of algorithms. There is no one type of algorithm that will always outperform the others—an idea known as the No Free Lunch Theorem.[36] Rather, each has particular characteristics[37] and particular strengths.[38] Nevertheless, these algorithms all share at least three characteristics that are central to this Article. First is the tendency toward opacity. Algorithms vary in their comprehensibility—that is, to what extent they provide an explanation for their outputs and predictions that humans can understand.[39] Frequently, the more powerful the algorithm, the less comprehensible it will be.[40] This creates a black-box phenomenon where the inputs go in and the outputs emerge, but there is no means of tracking or describing what happens in between. Computer scientists have recently engaged in efforts to improve interpretability of algorithmic

[35] FLACH, *supra* note 22, at 27-29 (explaining predictive models based on known and unknown variables).

[36] KELLEHER, *supra* note 4, at 518.

[37] *See*, *e.g.*, SKILLICORN, *supra* note 33, at 84-107 (describing different types of algorithms).

[38] *See id.* at 75 (noting that different techniques have "different assumptions about the possible relationships and the complexity of the dependence between the ordinary attributes and the target attribute"). Support vector machines, for example, are said to excel at generalizing, even when the data contains a very large number of variables. RUSSELL & NORVIG, *supra* note 29, at 744.

[39] One form of machine-learning algorithm that is comprehensible is called a "decision tree." RUSSELL & NORVIG, *supra* note 29, at 707. Think of a decision tree as a flow chart for a game of twenty questions. Decision trees sort the data according to a series of (usually binary) choices. Each choice asks about one of the input factors from the data—does the image depict fur? Does the image have a tail? The decision tree continues to sort data into different buckets at each step until it eventually reaches a determination for each instance. SKILLICORN, *supra* note 33, at 84.

[40] *See* RUSSELL & NORVIG, *supra* note 29, at 728-32 (discussing comprehensibility in context of one of most popular and effective types of algorithm, "neural network," which usually includes one or more "hidden layers").

models.[41] These efforts may one day minimize or eliminate the opacity of machine-learning predictions, but as yet they have not done so.

Next, any model implemented to generate predictions will be one of several that might have been chosen. This means that the errors that emerge from machine-learning models will be, in a sense, arbitrary. Imagine that a programmer has used crime data to generate a risk assessment tool, and she has identified two equally effective models. Based on the same data, each will provide an accurate prediction in ninety-eight percent of cases. When it comes to the two percent of the outcomes that are erroneous, however, the algorithms will err in different ways. Whereas the first algorithm mistakenly identifies Individuals *A*, *B*, and *C* as not being flight risks, the second might correctly categorize Individuals *A*, *B*, and *C*, but inaccurately conclude that Individuals *X*, *Y*, and *Z* will not appear for trial. The costs of errors therefore fall in arbitrary places. Moreover, even positing an unrealistically accurate algorithm, the number of errors will be significant. A model that makes the correct prediction 99.999% of the time, for example, still will err in one of every one hundred thousand cases.[42] When applied to a sufficiently large population—such as everyone crossing the border into the United States, or every potential drone target—there will be a non-negligible number of mistakes. As one expert explained, current modeling capacity produces algorithms that are "statistically impressive, but individually unreliable."[43]

Finally, however automated machine learning becomes, humans will always play a critical role.[44] They will always determine what data is available, which type of algorithm(s) to employ, which features within the data are relevant,

[41] *See infra* note 167 and accompanying text (discussing efforts to explain artificial intelligence decisionmaking).

[42] *See* SKILLICORN, *supra* note 33, at 69.

[43] John Launchbury, *A DARPA Perspective on Artificial Intelligence*, DEF. ADVANCED RESEARCH PROJECTS AGENCY, at 11:06 (Feb. 15, 2017) https://www.darpa.mil/about-us/darpa-perspective-on-ai [https://perma.cc/SZ6V-EQ6S]; *see also* Brent Mittelstadt et al., *The Ethics of Algorithms: Mapping the Debate*, BIG DATA AND SOC'Y, July-Dec. 2016, at 5 (noting that algorithms infer correlations based on "populations while actions are directed towards individuals").

[44] *But see* Jagmeet Singh, *Google's AutoML Project Bears Fruits, Writes Better Machine Learning Code Than Humans*, NDTV: GADGETS360 (Oct. 18, 2017), http://gadgets.ndtv.com/science/news/google-automl-making-ai-smarter-1764553 [https://perma.cc/2367-DFYY] (reporting that machine-learning models designed by Google's automated machine-learning system outperformed human programmers) .

which model to select from among several accurate ones, and what the end result means.[45]

Note that this simple overview belies the complexity of the machine-learning process. As an initial matter, in practice, the line between supervised and unsupervised learning can blur. In "semi-supervised machine learning," for example, you may have target variables for only some of your data, so fully supervised learning is not viable. Instead, the program must examine relationships among the data in order to predict the missing target variables.[46] In this way, the computer uses a combination of supervised and unsupervised learning techniques. In addition, the process is not a linear one, moving methodically from training to verifying to implementing. Rather, machine learning is "a dynamic and iterative process" of "retrieving, excluding, comparing[,] reorganizing, digging and pulling, etc."[47] Finally, while machine learning can be either descriptive or predictive,[48] the two are very closely entwined. Descriptive models "explore the properties of the data examined," while predictive models create "a way to predict new properties."[49] Those predictions, however, are based on the properties of the data that have been identified. Thus, in many cases, the descriptive aspect of a project serves as a prelude to prediction. Once a sufficiently accurate descriptive model is generated, it can use the characteristics of the existing data set to assess new data and predict future outcomes.[50]

B. *Use of Machine Learning in the Security Context*

While discussions of predictive use of machine learning inevitably evoke the specter of *Minority Report*,[51] it is a topic of science fact, not science fiction. Data mining generally and the use of machine learning specifically are currently employed in a multitude of ways throughout society and across the

[45] Colonna, *supra* note 19, at 319 ("[E]ven if data mining is highly automated, a human still has a role in the interpretation of the end result."); *see also infra* Section II.B.2.c (discussing interactions between humans and machine-learning algorithms).

[46] RUSSELL & NORVIG, *supra* note 29, at 695 (discussing blurred distinction between supervised and unsupervised learning); Bellovin et al., *supra* note 7, at 596 (discussing semi-supervised machine learning in context of mobile communication and location data).

[47] Colonna, *supra* note 19, at 350-51.

[48] *Id.* at 341 ("Discovery-driven data takes two major forms: description and prediction.").

[49] *Id.*

[50] *See* Rich, *supra* note 7, at 881 ("The machine learning process then creates a model based on the labeled dataset that can be used to predict the proper classification of future objects.").

[51] *See* MINORITY REPORT (Twentieth Century Fox & DreamWorks 2002).

government.[52] But law enforcement and security-related uses of the tool raise multiple concerns unique to those contexts. First, its use is pervasive and poised to increase significantly. Not only is the government focused on expanding its own capabilities,[53] but the area also promises to be a huge moneymaker for suppliers of this technology.[54] "The market for . . . 'predictive analytics' technology [to predict crimes, terrorist acts, and social upheaval before they happen] is estimated to reach $9.2 billion by 2020 . . . ."[55] Since there will be no shortage of efforts to satisfy this demand, it is a critical time to consider where the use of these tools is appropriate and what safeguards are necessary. Second, the pervasiveness of government secrecy in this area might exacerbate concerns raised by lack of algorithmic transparency, especially because many of the legal safeguards implemented to preserve data privacy and government accountability do not apply to law enforcement agencies.[56] Third, the consequences for individuals on whom the government chooses to use its coercive powers can be particularly significant. Whereas a faulty Netflix prediction might result in two hours wasted watching a bad movie, the costs of false positives from government use of machine learning—adverse government action taken against an innocent

[52] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 20, at app. IV (listing federal government data-mining programs that use personal information, not including classified programs).

[53] JOHN PODESTA ET. AL., EXEC. OFFICE OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES 27-29 (2014) (noting that data storage programs that facilitate searches across databases and aggregation of databases allow Department of Homeland Security ("DHS") to "take on new kinds of predictive and anomaly analysis").

[54] The National Institute of Justice has been funding crime-prediction research, and IBM, Hitachi, and Lexis have all "begun to offer ways to predict crime through data." Maurice Chammah, *Policing the Future*, THE VERGE (Feb. 3. 2016), https://www.theverge.com/2016/2/3/10895804/st-louis-police-hunchlab-predictive-policing-marshall-project [https://perma.cc/V6ZX-Z7TM].

[55] Chris Strohm, *Predicting Terrorism from Big Data Challenges U.S. Intelligence*, BLOOMBERG: QUINT (Oct. 13, 2016, 5:00 AM), https://www.bloombergquint.com/technology/2016/10/13/predicting-terrorism-from-big-data-challenges-u-s-intelligence#gs.f9Ye2gY [https://perma.cc/Q3CT-BJT6]; *see also* Ione Wells, *Government Offers £2M for Scientific Research into Counter-Terrorism*, THE GUARDIAN (July 16, 2017, 7:01 PM), https://www.theguardian.com/uk-news/2017/jul/17/government-offers-2m-for-scientific-research-into-counter-terrorism [https://perma.cc/RT9H-EDBY] (noting British government will make up to £2 million available to fund research into technology that could identify possible terrorists in crowds).

[56] *See, e.g.*, Computer Matching and Privacy Protection Act of 1988, Pub. L. No. 100-503, 102 Stat. 2507, 2512 (codified at 5 USC § 552a) (exempting federal law enforcement agencies from regulatory requirements).

person—might be intrusions on fundamental rights or even deployment of the government's use of coercive force.[57]

Having said all of that, it is impossible to draw a complete picture of the use of machine learning for law enforcement and security. While information about the use of data mining analyses must be reported to Congress on an annual basis, those reports include a classified annex for law enforcement and security applications available only to congressional overseers.[58] And in any event, at least some experts view these reports—as well as other transparency-promoting mechanisms such as System of Records Notices ("SORNs")[59] and Privacy Impact Assessments ("PIAs")[60]—as too vague to be sufficiently informative regarding what the government is actually doing.[61]

---

[57] *See* David Cole, *We Kill People Based on Metadata*, N.Y. REV. OF BOOKS (May 10, 2014), http://www.nybooks.com/daily/2014/05/10/we-kill-people-based-metadata/ [https://perma.cc/GSW6-WXWA] (quoting General Michael Hayden, former director of NSA and CIA, as stating "we kill people based on metadata" when describing power of large amounts of metadata). There are other law-enforcement and security-related challenges to the use of machine learning that are less relevant to my argument: multiple scholars across disciplines have raised doubts with respect to whether the available data sets are sufficiently robust to support machine-learning techniques in some areas, particularly counterterrorism. *See infra* Section II.A.1 (discussing potential issues related to insufficient data to supply machine learning). And machine learning for law enforcement and security take place in an "adversarial" context, where groups or individuals will be highly motivated to subvert the algorithm, either by manipulating the data or its analysis to hide themselves or their actions. SKILLICORN, *supra* note 33, at 69.

[58] Federal Agency Data Mining Reporting Act of 2007, 42 U.S.C. § 2000ee-3(c) (2012) (requiring report from all federal agencies detailing specific use of data mining technology); *see also* Jacques Peretti, *Palantir: The 'Special Ops' Tech Giant that Wields as much Real-World Power as Google*, THE GUARDIAN (July 30, 2017, 9:59 AM), https://www.theguardian.com/world/2017/jul/30/palantir-peter-thiel-cia-data-crime-police [https://perma.cc/YF9F-ELXQ] (noting Air Force, CDC, CIA, FBI, IRS, Marine Corps, NSA, Special Operations Command, and West Point are all clients of Palantir—a company that has adapted algorithms and predictive tools for domestic use, which were originally developed to combat insurgency in Iraq).

[59] The Privacy Act of 1974, 5 U.S.C. § 552a (2012), regulates the collection, maintenance, use, and dissemination of records about individuals that are retrieved by personal identifier and collected, used, or disseminated by agencies and requires agencies to publish SORNs when they create or revise systems of records.

[60] 44 U.S.C. § 3501 note (2012) (E-Government Act of 2002 § 208) (providing for Privacy Impact Assessments).

[61] Citron & Pasquale, *supra* note 2, at 26 ("Kenneth Bamberger and Deidre Mulligan argue that Privacy Impact Assessments required by the E-Government Act are unsuccessful in part

Nevertheless, we know enough to determine ways in which machine learning is being employed. One area where algorithmic decision-making is becoming pervasive is in the criminal justice system. Police departments have used historical crime-related data to help allocate departmental resources and identify "high crime" areas or "hot spots" for years, but the application of predictive analytics to this data is now widespread. For example, as of October 2016, over sixty local police departments have begun using "a service sold by PredPol, which calls itself 'The Predictive Policing Company,' to forecast where crimes might occur based on past patterns."[62] In Chicago, the police department uses predictive analytics to identify not only places that are particularly vulnerable to crime, but also people more likely to be involved in gun violence.[63] Further, there have been several projects designed to integrate various law enforcement databases to support the use of sophisticated data mining, including machine learning. COPLINK, for example, is a long-established information-sharing and analysis system used by over six thousand law enforcement agencies; developed at the University of Arizona's Artificial Intelligence Lab, COPLINK is now

due to the public's inability to comment on the design of systems whose specifications and source codes remain obscured.").

[62] Strohm, *supra* note 55. PredPol was developed in cooperation with the Los Angeles Police Department and uses historical criminal activity reports to identify areas in which crime has been most prevalent during specific time periods. Cameron Albert-Deitch, *Predictive Policing Crime Prevention Software Successful for APD*, ATLANTA MAG. (Nov. 10, 2014), http://www.atlantamagazine.com/news-culture-articles/predictive-policing-crime-prevention-software-successful-for-apd/ [https://perma.cc/XQ7Y-3T6C]. PredPol does not, however, attempt to predict *who* will commit a crime. *See* Strohm, *supra* note 55. The LAPD has praised PredPol's effectiveness, while other police departments have been more lukewarm. Moreover, there have been some complaints about PredPol's aggressive sales tactics and provisions in their contracts that obligate PredPol users to engage in marketing and promotion on PredPol's behalf. Darwin Bond-Graham, *All Tomorrow's Crimes*: *The Future of Policing Looks a Lot Like Good Branding*, SF WEEKLY (Oct. 30, 2013, 4:00 AM), http://www.sfweekly.com/news/all-tomorrows-crimes-the-future-of-policing-looks-a-lot-like-good-branding/ [https://perma.cc/675E-KRR6]. The property crime rate in a Los Angeles suburb using PredPol reportedly dropped by thirteen percent over four months. Chammah, *supra* note 54; *see also* Simmons, *supra* note 7, at 954-58 (discussing PredPol and other predictive policing programs).

[63] Ferguson, *supra* note 5, at 384-85 ("In Chicago, analysts have identified young people at greater risk of being involved in gun violence."); Simmons, *supra* note 7, at 956 ("Chicago [Police Department] . . . us[es] predictive software to determine which individuals are most likely to be involved in a crime."); *see also* Zarsky, *supra* note 7, at 1516 ("Predictive modeling is applied to . . . decisions as to which individuals should be stopped for questioning.").

owned by a private firm.[64] It includes 1.25 billion shareable documents as well as facial recognition technology, and "sophisticated analytics . . . to allow investigators to discover hidden relationships and patterns that can be used to solve crimes."[65] COPLINK is just one of many such products being marketed to law enforcement agencies.[66]

[64] COPLINK was initially purchased by IBM. Press Release, IBM, IBM Brings One of the World's Largest Networks of More Than a Billion Law Enforcement Shareable Documents to the Cloud (June 22, 2015), *available at* https://www-03.ibm.com/press/us/en/press release/47156.wss [https://perma.cc/6SYZ-4YX6] (stating "COPLINK has transformed information sharing across more than 6,000 law enforcement agencies in North America"); Stephen Russo, *Creating a Safer Planet with Smarter Analytics Solutions*, IBM BIG DATA & ANALYTICS HUB BLOG (June 22, 2015), http://www.ibmbigdatahub.com/ blog/creating-safer-planet-smarter-analytics-solutions [https://perma.cc/9GSF-B9EA] (stating COPLINK "applies analytics to vast quantities of data to help officials piece together seemingly unrelated information and generate tactical leads"). It is now owned by Forensic Logic. Press Release, Forensic Logic, Forensic Logic Announces Acquisition of COPLINK Platform from IBM (Oct. 4, 2017), *available at* https://www.prnewswire.com/news-releases/forensic-logic-announces-acquisition-of-coplink-platform-from-ibm-300530930.html [https://perma.cc/W3 R4-DUYH]; *Data Warehousing - Coplink*/BorderSafe/RISC*, UNIVERSITY OF ARIZONA, ARTIFICIAL INTELLIGENCE LABORATORY, https://ai.arizona.edu/research/coplink [https:// perma.cc/8PYP-2SJG] (last visited Sept. 11, 2018) (detailing University of Arizona research into COPLINK).

[65] David Griffith, *IBM Introduces Cloud Version of Coplink*, POLICE MAG. (June 23, 2015), http://www.policemag.com/blog/technology/story/2015/06/ibm-introduces-cloud-version-of-coplink.aspx [https://perma.cc/4M6M-AXB9]. A similar tool, MATRIX, was developed by a consortium of state law enforcement agencies to combine multiple databases and analytic tools, but has since been discontinued. *Information Fusion Centers and Privacy*, ELECTRONIC PRIVACY INFORMATION CENTER, https://epic.org/privacy/fusion/ [https://per ma.cc/9EKN-ZC7N] (last visited Sept. 11, 2018).

[66] *See, e.g.*, Joh, *supra* note 7, at 24-25 ("*Social Media Monitor* claims to warn law enforcement clients of ongoing or potential threats of violence. . . . *Beware*[] promotes itself as a 'tool to help first responders understand the nature of the environment they may encounter' . . . . by assigning a 'threat rating' to a person based on an analysis of billions of commercial and public records[,] . . . . [including] registered cars and rap sheets . . . and . . . online comments, social media and recent purchases . . . ."); Simmons, *supra* note 7, at 955-56 (describing Beware); Justin Jouvenal, *The New Way Police Are Surveilling You: Calculating Your Threat 'Score,'* WASH. POST (Jan. 10, 2016), https://www.washingtonpost.com/local/public-safety/the-new-way-police-are-surveilling-you-calculating-your-threat-score/2016/01/10/e42bccac-8e15-11e5-baf4-bdf37355da0c_ story.html?utm_term=.b87d5a7130e5 (discussing Fresno Police Department's use of Beware).

It is not just at the crime-detection stage that predictive analytics are driving law enforcement decision-making. Algorithms are being employed throughout the criminal justice system in the form of "risk assessment" tools—sometimes labeled "evidence-based decision-making." These play a role in making decisions regarding which alleged criminals can be released pending trial, whether an offender should be incarcerated, what level of security an inmate should be placed in while incarcerated, as well as which offenders are likely to recidivate and therefore who should be denied probation or parole.[67] Proprietary programs, such as COMPAS, use variables related to criminal history, relationships, personality, family, and social exclusion to generate risk assessment scores.[68] COMPAS itself is used in multiple states, and other states have secured similar tools from other sources or developed their own risk-assessment algorithms.[69] Even the current revisions of the Model Penal Code "direct sentencing commissions to '[d]evelop actuarial instruments or processes . . . that will estimate the relative risk that individual offenders pose to public safety . . . and to incorporate them into the sentencing guidelines.'"[70]

Law enforcement methods are not, of course, entirely independent from national security efforts. After all, most threats to national security will also be crimes—crimes the government is particularly eager to prevent. The FBI has generated an "Indicators of Mobilization to Violence" that includes forty-eight

[67] NATHAN JAMES, CONG. RESEARCH SERV., RISK AND NEEDS ASSESSMENT IN THE CRIMINAL JUSTICE SYSTEM 4 (2015) (discussing risk and needs assessments as used in criminal justice system); Rebecca Wexler, Opinion, *When a Computer Program Keeps You in Jail*, N.Y. TIMES, June 13, 2017, at A27 ("At every stage—from policing and investigations to bail, evidence, sentencing and parole—computer systems play a role.").

[68] *Algorithms in the Criminal Justice System*, *supra* note 8 (stating risk assessment tools such as COMPAS use data on socioeconomic status, family background, neighborhood crime, and employment status to calculate individual's criminal risk); Angwin et al., *supra* note 8 (stating COMPAS assesses criminal personality, social isolation, substance abuse, and residence/stability).

[69] Simmons, *supra* note 7, at 965-67 (describing various risk assessment programs used to make decisions regarding bail, sentencing, level of supervision for incarcerated individuals, whether to grant parole, and what parole conditions apply); *Algorithms in the Criminal Justice System*, *supra* note 8 ("Jurisdictions have generally used one of three main systems, or adapted their own version of each: Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), Public Safety Assessment (PSA) and Level of Service Inventory Revised (LSI-R).").

[70] Petition for Writ of Certiorari at 12, State v. Loomis, 881 N.W.2d 749 (Wis. 2016), *cert. denied sub nom.* Loomis v. Wisconsin, 137 S. Ct. 2290 (2017) (No. 16-6387).

questions for agents to answer about terror suspects.[71] Based on those answers, individuals are assigned a score between zero and one hundred, which shows "'the subject's level of mobilization or likelihood of carrying out a violent act,' and 'the subject's likely level of radicalization or internal commitment to violent ideology.'"[72] Questions include, among others, whether a subject is a religious convert, whether the subject experienced a recent personal loss, and whether the subject has immediate access to weapons.[73] The score is generated through use of an undefined "statistical methodology" weighting—i.e., an algorithm—using data about historical terrorism suspects that have mobilized to violence as well as those that have not.[74]

Machine learning is also at work promoting national security beyond the law enforcement context. The Department of Homeland Security ("DHS") has many such programs. The individuals placed on the "No-Fly List" are selected using "predictive judgments,"[75] and Customs and Border Patrol ("CBP") is required "to use or investigate the use of advanced algorithms in support of its mission."[76] Thus, by exploiting historical data, CBP uses machine learning to generate rules "driven by algorithms to identify obvious and non-obvious relationships among data inputs"[77] to support its efforts to identify high-risk cargo[78] entering or leaving the country. Not unlike programs like COMPAS, the CBP's Automated Targeting System ("ATS") generates a risk-assessment score for incoming and

[71] Cora Currier & Murtaza Hussain, *48 Questions the FBI Uses to Determine if Someone Is a Likely Terrorist*, THE INTERCEPT (Feb. 13, 2017), https://theintercept.com/2017/02/13/48-questions-the-fbi-uses-to-determine-if-someone-is-a-likely-terrorist/ [https://perma.cc/A3WE-TU2N].

[72] *Id.*

[73] *Id.*

[74] *Id.* (explaining weaknesses of this methodology and noting that it likely leads to poor results).

[75] Defendants' Consolidated Memorandum in Support of Cross-Motion for Partial Summary Judgment and Opposition at 1, Latif v. Holder, 28 F. Supp. 3d 1134 (D. Or. 2014) (No. 10-cv-00750) (describing government attempts to combat terrorism by using sensitive intelligence reporting and investigative information to determine whether individuals should be allowed on aircraft).

[76] U.S. DEP'T OF HOMELAND SEC. PRIVACY OFFICE, 2016 DATA MINING REPORT TO CONGRESS 19 (2017) [hereinafter 2016 DATA REPORT].

[77] *Id.* at 25.

[78] High-risk cargo includes "weapons of mass effect, illegal narcotics, agents of bio-terrorism, threats to U.S. agriculture, or other contraband." *Id.* at 19.

outgoing cargo as well as vehicles and individuals crossing the land border.[79] Individuals entering the country by plane are screened by the Transportation Safety Administration's ("TSA") Secure Flight program, which generates "risk-based assessments" using both data the TSA receives and risk assessments generated by the airlines.[80] And DHS is constantly seeking to expand its use of machine learning. The agency's Analytical Framework for Intelligence system allows officials from several components of DHS (including the Customs and Immigration Services, the Coast Guard, and the TSA) to employ "enhanced search and analytical capabilities" in order to analyze information from various federal, state, and local law enforcement databases.[81] These analyses are used, in part, to study data regarding persons or cargo to understand whether there are "patterns or trends that could assist in the identification of potential law enforcement or security risks."[82] DHS has also pursued machine-learning technology to implement President Trump's "Extreme Vetting Initiative," seeking tools that can "determine and evaluate an applicant's probability of becoming a positively contributing member of society, as well as their ability to contribute to national interests."[83]

Other domestic agencies, such as the FBI and the DEA, are exploiting these technologies as well. But since classified and "law enforcement sensitive" information about government data mining programs "shall not be made available to the public,"[84] the details are harder to unearth. What we do know is that the FBI is using machine learning in its effort to identify individuals prone

[79] Complaint for Injunctive Relief at 8, Elec. Privacy Info. Ctr. v. U.S. Customs & Border Patrol, No. 17-cv-1438 (D.D.C. July 19, 2017), ECF No. 1 (alleging that photos of citizens and non-citizens can be retained in ATS); Zarsky, *supra* note 7, at 1515 (explaining that ATS uses government databases to generate predications assessing risk posed by those attempting to cross border).

[80] U.S. DEP'T. OF HOMELAND SEC., PRIVACY IMPACT ASSESSMENT UPDATE FOR SECURE FLIGHT 5 (2014) (explaining that "Secure Flight passenger prescreening computer system conducts risk-based analysis using, among other data," information received from aircraft operators, assessments, and pre-screening data).

[81] 2016 DATA REPORT, *supra* note 76, at 30.

[82] *Id.* at 31.

[83] Sam Biddle, *Trump's 'Extreme-Vetting' Software Will Discriminate Against Immigrants 'Under a Veneer of Objectivity,' Say Experts*, THE INTERCEPT (Nov. 16, 2017, 8:30 AM), https://theintercept.com/2017/11/16/trumps-extreme-vetting-software-will-discriminate-against-immigrants-under-a-veneer-of-objectivity-say-experts/ [https:perma.cc/X9DZ-WEPL].

[84] Federal Agency Data Mining Reporting Act of 2007, 42 U.S.C. § 2000ee-3(c)(3)(B)(ii) (2012).

to violent extremism,[85] and that machine learning is used to detect identity theft and fraud in a number of contexts.[86]

We also know that the FBI collects and stores massive amounts of information on a regular basis. All of that information—and much more—is available to the National Security Analysis Center ("NSAC"), a component of the Justice Department, that uses "advanced data mining software" to seek out national security threats.[87] NSAC has access to "over 130 databases and datasets of information comprising some two billion records," which it uses "to delve deeply into the activities and associations of foreigners and Americans alike."[88] In addition to state and local law enforcement databases, the U.S. government's intelligence, treasury, and commercial databases are also available to the NSAC. The NSAC, some of whose data mining capabilities were funded by the United States Department of Defense ("DoD") components, such as Defense Advanced Research Projects Agency ("DARPA"), partners with military and government contractors.[89] In addition to using link analysis to find individuals with connections to terrorist suspects, the NSAC also pursues "pattern analysis" using "predictive models."[90] Given its access to vast hordes of data and its data mining imperative, NSAC is surely employing machine-learning techniques in a number of ways. To what end and with what effect, one can only speculate.

Other parts of the intelligence community are also known to use machine learning. Thanks to Edward Snowden's leaks, we know that the NSA has

[85] *See supra* notes 71-74 and accompanying text (describing some questions asked by FBI when identifying potential terrorists).

[86] *See, e.g.*, FED. BUREAU OF INVESTIGATION, PRIVACY IMPACT ASSESSMENT FOR THE STAGED ACCIDENT DATA MINING INITIATIVE (2008), https://www.fbi.gov/services/records-management/foipa/privacy-impact-assessments/staged-accident [https://perma.cc/T3MC-UBBL] (creating data mining program to identify and analyze car insurance fraud).

[87] Homeland Security Presidential Directive-2: Combatting Terrorism Through Immigration Policies, 37 WEEKLY COMP. PRES. DOC. 1570 (Oct. 29, 2001). The directive established the Foreign Terrorist Tracking Task Force ("FTTTF"), which was subsequently subsumed by NSAC. William M. Arkin, *This Shadow Government Agency Is Scarier than the NSA*, PHASE ZERO (June 1, 2015, 2:10 PM), http://phasezero.gawker.com/this-shadow-government-agency-is-scarier-than-the-nsa-1707179377 [https://perma.cc/Z3P2-VREZ] (examining history of National Security Analysis Center).

[88] Arkin, *supra* note 87; *see* NAT'L SEC'Y AGENCY, DOMESTIC SURVEILLANCE DIRECTORATE, *Our Partners*, https://nsa.gov1.info/partners/ [https://perma.cc/NBW5-PAUC] (last visited Sept. 11, 2018).

[89] *See* Arkin, *supra* note 87.

[90] *Id*.

engaged in analysis of communications data.[91] The intelligence community as a whole has been working on ways to use machine learning more broadly. "Mercury" is a program at the Intelligence Advanced Researched Projects Activity ("IARPA"), whose mission is to develop means of automated analysis of signals intelligence ("SIGINT") "to anticipate and/or detect events such as terrorist activity, civil unrest, and disease outbreaks abroad."[92] According to a 2005 call for proposals, the Advanced Capabilities for Intelligence Analysis program seeks "'to construct and use plausible futures in order to provide additional, novel interpretations for today's collection' of intelligence information."[93]

DoD is also on the machine-learning bandwagon. Drone targeting already relies heavily on algorithmic calculation.[94] And there is an ongoing debate about the extent to which militaries should be permitted to use lethal autonomous weapons systems—weapons that would select and engage targets without

[91] *See* PRIVACY & CIVIL LIBERTIES OVERSIGHT BD., REPORT ON THE TELEPHONE RECORDS PROGRAM CONDUCTED UNDER SECTION 215 OF THE USA PATRIOT ACT AND ON THE OPERATIONS OF THE FOREIGN INTELLIGENCE SURVEILLANCE COURT 8 (2014) (noting NSA's telephone records program was intended to allow government to identify communications among known and unknown terrorism suspects). In Pakistan, the NSA reportedly uses algorithms to identify suspicious activity by monitoring cellular network traffic. Dave Gershgorn, *Can the NSA's Machines Recognize a Terrorist?*, POPULAR SCI. (Feb. 16, 2016), https://www.popsci.com/nsas-skynet-might-not-be-able-to-tell-what-makes-terrorist [https://perma.cc/3JW9-FJAW] ("The NSA project, disastrously named Skynet, uses cellular network traffic in Pakistan to identify and monitor potential threats, according to leaked documents on *The Intercept*.").

[92] Kristen Jordan, *Predictive Analytics: A New Tool for the Intelligence Community*, THE CIPHER BRIEF (Dec. 4, 2016), https://www.thecipherbrief.com/predictive-analytics-a-new-tool-for-the-intelligence-community.

[93] Shane Harris, *NSA Spy Program Hinges on State-of-the-art Technology*, GOV'T EXEC. (Jan. 20, 2006), https://www.govexec.com/defense/2006/01/nsa-spy-program-hinges-on-state-of-the-art-technology/20996/ [https://perma.cc/Y3KX-GUEQ].

[94] *See* Miranda Bogen, *Algorithms of War*, SLATE (Dec. 8, 2015, 3:05 PM), http://www.slate.com/articles/technology/future_tense/2015/12/the_dangers_of_enlisting_algorithms_in_statecraft.html [https://perma.cc/7VR2-RYEQ] ("Drone targeting is increasingly based on algorithmic calculations . . . ."); Martin Robbins, *Has a Rampaging AI Algorithm Really Killed Thousands in Pakistan?*, THE GUARDIAN (Feb. 18, 2016, 10:10 AM), https://www.theguardian.com/science/the-lay-scientist/2016/feb/18/has-a-rampaging-ai-algorithm-really-killed-thousands-in-pakistan [https://perma.cc/7THV-DLVJ] (describing use of machine learning to identify couriers for terrorist organizations).

human control over individual attacks—trained via machine learning.[95] DoD has also launched an Algorithmic Warfare Cross Functional Team, called "Project Maven," which is tasked with "using big data and machine learning to accelerate the process of discovering actionable intelligence" in aerial imagery, of which there is too much to be analyzed solely by humans.[96] If the program is successful, DoD foresees using similar tools across the armed services for "intelligence, for targeting, for collection management, [and] for sensor fusion."[97]

One issue raised by Project Maven is the capacity for machine learning to work with images. Video content analysis has applications across the law enforcement and national security terrain.[98] Algorithms have significantly improved in their ability to identify human faces, even in crowds. That capability has implications for surveillance feeds from drones in Afghanistan, but also for analysis of security camera footage in New York and Chicago. There are also efforts to develop predictive analytics that can identify threatening human behavior in public spaces, based on patterns from video footage of past events,

[95] INT'L HUMAN RIGHTS CLINIC, HARVARD LAW SCH., MAKING THE CASE: THE DANGERS OF KILLER ROBOTS AND THE NEED FOR A PREEMPTIVE BAN 4-21 (2016); Bogen, *supra* note 94 ("Discussions on autonomous weapons have reached the United Nations, and observers have reported that the U.K. and U.S. are attempting to water down international agreements banning the use of this technology."). For an analysis of issues surrounding use of machine learning for target selection, see Ashley Deeks, *Military Detention and Targeting in an Era of Predictive Algorithms* (on file with author).

[96] Andrew Tarantola, *The Pentagon Is Hunting ISIS Using Big Data and Machine Learning*, ENGADGET (May 15, 2017), https://www.engadget.com/2017/05/15/the-pentagon-is-hunting-isis-using-big-data-and-machine-learning/ [https://perma.cc/6BHH-AKCG]; Memorandum from Deputy Sec'y of Def. on Establishment of an Algorithmic Warfare Cross-functional Team (Project Maven) to Sec'ys of the Military Dep'ts, et al. (Apr. 26, 2017).

[97] *See* Tarantola, *supra* note 96. Sensor fusion is the combination of information from several sensors to generate a more accurate overall view.

[98] *See, e.g.*, Emerging Technology from the arXiv, *Machine-Learning Algorithm Aims to Identify Terrorists Using the V Signs They Make*, MIT TECH. REV. (Mar. 8, 2016), https://www.technologyreview.com/s/600971/machine-learning-algorithm-aims-to-identify-terrorists-using-the-v-signs-they-make/ (describing effort to identify individuals who cover their faces to hide their identity in photos posted online by examining hand gestures); Asha McLean, *How One Sheriff's Office Is Using Machine Learning to Uncover Persons of Interest*, ZDNET (Nov. 30, 2017, 11:31 PM), https://www.zdnet.com/article/how-one-sheriffs-office-is-using-machine-learning-to-uncover-persons-of-interest/ [https://perma.cc/4KJS-D49T] (describing Washington County Sheriff's Office's use of Amazon's machine-learning-based Rekognition image identification system to locate suspects).

and there is an ongoing effort to make police body-camera video more searchable and interoperable with other systems.[99]

Just as the private sector exploits predictive algorithms for uses from fraud detection to targeted marketing, the security state seeks to further its mission across a variety of contexts. In every context in which a threat is presented, there is an agency (or agencies) seeking machine-learning methods to predict, anticipate, and prevent those threats. Some of these projects are fully integrated into existing government policy, but as we see from ongoing FBI, DHS, and DoD research initiatives, the continued development of this tool remains a high priority for a range of law enforcement and security-focused agencies.

II. MACHINE LEARNING AND ITS DISCONTENTS

The use of predictive analytics has enormous potential. These techniques can generate novel insights about criminal behavior, identify with greater accuracy where to invest surveillance resources, reduce improper bias in policing, result in more equitable law enforcement by enabling detection of white collar crimes, and even perhaps replace disfavored tactics such as the use of informants or covert policing.[100] But, as they say, with great power comes great responsibility. This Part will highlight the challenges that arise when using predictive algorithms for government decision-making. Section A will briefly discuss what this Article refers to as "technological challenges"—concrete obstacles inherent in building and implementing accurate, effective machine-learning programs. In the course of this discussion, this Section calls attention to technological challenges that are particularly thorny in the law enforcement or national security context. These problems are significant and they deserve the attention they often get.

At the same time, however, many of these difficulties will likely diminish significantly in the coming years, as technology continues to advance. Therefore, Section B addresses what this Article refers to as "conceptual challenges"—challenges that arise not from the specific operations of machine-learning programs and their shortcomings, but rather from the very nature of machine learning itself. That Section asks the following question: if machine-learning methodology overcomes its technological challenges to achieve reliable accuracy rates of ninety-nine percent, will concerns about its use nonetheless remain? Answering in the affirmative, Section B explains the ways in which aspects of machine learning clash with fundamental rule-of-law principles.

---

[99] EXEC. OFFICE OF THE PRESIDENT, *supra* note 53, at 22.

[100] *See* Jane Bambauer, *Other People's Papers*, 94 TEX. L. REV. 205, 244 (2015) (stating that data mining "can lead to fairer enforcement of our criminal laws"); Joh, *supra* note 7, at 28-29 (explaining benefits of big-data tools); Zarsky, *supra* note 7, at 1516 (noting benefits of using big-data tools to detect white collar crimes like insider trading and money laundering).

A. *Technological Challenges*

Some scholars reject the "data mining" label, pointing out that the actual data analysis is just one step in a multi-step process of knowledge discovery in data ("KDD"). Technological challenges can come at any stage of this process. Thus, accuracy concerns can arise from the data, the methods of preparing the data for analysis, the analysis itself, or the way in which the results are interpreted and implemented. This Part will briefly mention some of these potential pitfalls and, in particular, how they are particularly likely to arise when using machine learning in law enforcement or national security programs.

1. Sufficiency of the Data

Machine learning is a "garbage in-garbage out" proposition.[101] An algorithmic analysis of incorrect or otherwise flawed data will generate incorrect or otherwise flawed predictions. And it is easy to end up with flawed data. Sometimes data will simply be inaccurate, outdated, or incomplete.[102] A common problem in the big data context is known as *incomplete matching*, which can arise when combining information from two different databases into one data set.[103] If one database enters names in the format "H. J. Potter" and the other as "Potter, Harry J.," will the computer realize those entries refer to the same person? Even if all data sets use the same format, some people have the same names; some change their names, phone numbers, and addresses; and a host of other inconsistencies are possible.[104] The federal government's long-running struggle to integrate twelve different watch lists maintained by nine federal agencies—a process described as "extremely difficult"—provides a vivid example.[105] The process of aggregating data into one source is partly

[101] This concept is a familiar one in the fields of computer science and information technology.

[102] DEP'T OF DEF., *supra* note 20, at 37 ("Transposed letters or numbers, transposed first and last names, missing address components (e.g., apartment number), and other errors can cause significant errors in records."); Ferguson, *supra* note 5, at 388-404 (reviewing positives and negatives of using data in traffic stops); Rich, *supra* note 7, at 871, 893-900 (explaining insufficiencies of "Automated Suspicion Algorithms," which apply machine-learning methods to government data "with the purpose of identifying individuals likely to be engaged in criminal activity"); Simmons, *supra* note 7, at 969-83 (discussing challenges inherent to predictive algorithms, including racial bias, use of forbidden factors, and pre-existing biases in underlying data).

[103] DEP'T OF DEF., *supra* note 20, at 37-38 (noting that many data records contain errors and listing factors contributing to difficulty of data integration).

[104] *Id.* at 37-38.

[105] *Id.* at 38.

regulated by the Computer Matching and Privacy Protection Act, but as is so often the case, there is a law enforcement exemption from its requirements.[106] Then, once the data is collected in one place, it must be "cleaned" or prepared for analysis. This might require various forms of data manipulation. If the data set is particularly voluminous or complex, for example, it may need to be simplified in order to expedite computation by eliminating or combining variables, or otherwise paring down the data. Such simplification sometimes will reduce accuracy.[107]

Concerns about data are particularly merited when using machine learning as a counterterrorism tool. First, to be successful, predictive algorithms might require more extensive amounts of historical data than is available.[108] Machine-learning models that ferret out credit card fraud, for example, can draw on data from the nine hundred million credit cards in the United States, about ten million of which are used fraudulently every year.[109] When it comes to terrorism, however, we (thankfully) have a much smaller data set.[110] According to many

[106] Computer Matching and Privacy Protection Act of 1988, Pub. L. No. 100-503, 102 Stat. 2507, 2512 (codified at 5 U.S.C. § 552a) (regulating use of computer matching agreements by federal agencies when system records match with other records).

[107] SKILLICORN, *supra* note 33, at 105. In fact, an entire field ancillary to machine learning called "data reduction" is devoted to finding ways of describing data with fewer numbers. But the method of reduction can affect the ultimate results. *See generally* Muhammad Habib ur Rehman et al., *Big Data Reduction Methods: A Survey*, 1 DATA SCI. & ENGINEERING 265 (2016) (providing review of data reduction methods).

[108] *See* Jeff Jonas & Jim Harper, *Effective Counterterrorism and the Limited Role of Predictive Data Mining*, POL'Y ANALYSIS (CATO Inst., D.C.), Dec. 2006, at 7-8 ("Without well-constructed algorithms based on extensive historical patterns, predictive data mining for terrorism will fail.").

[109] *See* Bruce Schneier, *Why Data Mining Won't Stop Terror*, WIRED (Mar. 9, 2006, 12:00 PM), https://www.wired.com/2006/03/why-data-mining-wont-stop-terror/ [https://perma.cc/59Z2-U67T] (describing credit card fraud as "one of data mining's success stories").

[110] *See* Jonas & Harper, *supra* note 108, at 7-8 (explaining that unlike consumer behavior, terrorism planning lacks meaningful patterns that show what behavior indicates planning or preparation for terrorist act); Letter from 18 Million Rising et al. to the Honorable Elaine C. Duke, Acting Sec'y of Homeland Sec. 2 (Nov. 16, 2017) ("Experts have also concluded that national security threats—in particular, acts of terrorism—are so rare that they are extraordinarily difficult, if not impossible, to predict, because the data are so scant that they do not provide a reliable basis for prediction."). This phenomenon is sometimes referred to as an "imbalanced data set," where the overwhelming majority of the data belongs to one class (here, not-terrorists). For example, in a database of one hundred people, one of whom is a terrorist, a model that identifies everyone as a not-terrorist will be ninety-nine percent accurate, but not particularly helpful. *See, e.g.*, Nitesh V. Chawla, Nathalie Japkowicz &

scholars, the number of attempted or successful terrorist attacks, when combined with distinctions regarding planning, personnel, and execution, is simply insufficient data from which one can extract meaningful patterns and build valid predictive models.[111] Second, data quality and reliability are likely to be an issue. Relevant data is often difficult to collect, and "the task of integrating data accurately is especially difficult in the counter-terrorism arena, which often involves matching data from disparate systems over which the intelligence community has no control."[112] Intelligence also comes from intercepts, unknown

Aleksander Kolcz, *Editorial: Special Issue on Learning from Imbalanced Data Sets*, 6 ACM:SIGKDD 1, 1 (2004) (explaining "class imbalance problem typically occurs when, in a classification problem, there are many more instances of some classes than others"); Jason Brownlee, *8 Tactics to Combat Imbalanced Classes in Your Machine Learning Dataset*, MACHINE LEARNING MASTERY (Aug. 19, 2015), https://machinelearningmastery.com/tactics-to-combat-imbalanced-classes-in-your-machine-learning-dataset/ [https://perma.cc/65MY-KL45].

[111] Bambauer, *supra* note 100, at 243 ("Predicting which people are terrorists is a futile task because virtually no one is."); Jonas & Harper, *supra* note 108, at 7-8; SCHNEIER, *supra* note 7; Kashmir Hill, *The Government Wants Silicon Valley to Build Terrorist-spotting Algorithms. But Is It Possible?*, SPLINTER (Jan. 14, 2016, 1:31 PM), https://splinternews.com/the-government-wants-silicon-valley-to-build-terrorist-1793854067 [https://perma.cc/PR65-T5FM].

[112] Newton N. Minnow & Fred H. Cate, *Government Data Mining*, *in* MCGRAW-HILL HANDBOOK OF HOMELAND SECURITY 19 (2005); *see* NAT'L RESEARCH COUNCIL, PROTECTING INDIVIDUAL PRIVACY IN THE STRUGGLE AGAINST TERRORISTS 38 (2008). This relative absence of reliable data has not prevented experts from attempting to find viable means to predict terrorist activity. Mercury—IARPA's program for automated analysis of SIGINT—aims to address the concern by developing data extraction techniques that focus on volume, rather than depth, and by identifying shallow features of data that correlate with events. *See* Jordan, *supra* note 92. Moreover, some studies have aimed to predict *where* terrorist attacks are most likely to occur. Press Release, Binghamton University, Researchers Can Predict Terrorist Behaviors with More Than 90% Accuracy (Mar. 2, 2017), *available at* https://www.binghamton.edu/news/story/440/researchers-can-predict-terrorist-behaviors-with-more-than-90-accuracy [https://perma.cc/P7WQ-CTBN]. Studies have also shown that *when* terrorist attacks are likely to occur can be predicted. *See, e.g.*, Michael D. Porter & Gentry White, *Self-exciting Hurdle Models for Terrorist Activity*, 6 ANNALS OF APPLIED STATS. 106, 106 (2012); Ana Swanson, *The Eerie Math that Could Predict Terrorist Attacks*, WASH. POST (Mar. 1, 2016), https://www.washingtonpost.com/news/wonk/wp/2016/03/01/the-eerie-math-that-could-predict-terrorist-attacks/?noredirect=on&utm_term=.2c59daa2f92a. Some studies have also attempted to identify the characteristics of future terrorist attacks (time of attack, weapon type, etc.). Salih Tutun, Mohammad T. Khasawneh & Jun Zhuang, *New Framework that Uses Patterns and Relations to Understand Terrorist Behaviors*, 78 EXPERT SYSTEMS WITH APPLICATIONS 358, 373 (2017) ("[W]e had more than

actors, and other sources where little or no identifying information is provided or in ways that prevent seeking or verifying additional identifying information.[113]

2. Selecting Features

There may be some data sets that come with ready-made features or variables that will constitute the inputs to the algorithm, but more often features "need to be constructed by the developer of the machine learning application."[114] Feature selection is a step in the KDD process that raises not only technological challenges but also contributes to concerns stemming from this Article refers to as "the human factor," discussed in more detail in Section II.B.2.c. To provide just two examples of the technological challenges, it will usually be impossible to include in a model all possible relevant factors, placing a thumb on the scale of features that are readily accessible and inexpensive, even if they are not the features best situated to capture the relevant relationships.[115] Features might interact with one another in ways that can distort the process. If, for example, two features are positively or negatively correlated with one another, using both

90% accuracy for most of the tactics."). Researchers have also sought to develop algorithms based on ISIS-supporters' online activity that might "eventually help predict attacks that are about to happen." Pam Belluck, *Fighting ISIS With an Algorithm, Physicists Try to Predict Attacks*, N.Y. TIMES, June 16, 2016, at A17. A December 2016 study that set out to identify factors that might predict radicalization found that relevant factors, such as psychological vulnerability, community or personal crisis, desire for status, and more are present in the lives of a host of people who will never turn to violence. MICHAEL JENSEN ET AL., FINAL REPORT: EMPIRICAL ASSESSMENT OF DOMESTIC RADICALIZATION (EADR) 69 (2016); Currier & Hussain, *supra* note 71. So, while these efforts are applauded as worthwhile, informative, and perhaps useful in some ways, when it comes to making reliable predictions, "they're [not] really there yet." Catherine Caruso, *Can a Social-Media Algorithm Predict a Terror Attack?*, MIT TECH. REV. (June 16, 2016), https://www.technologyreview.com/s/601700/can-a-social-media-algorithm-predict-a-terror-attack/; *see also* Nsikan Akpan, *This Computer Algorithm Might Be Able to Predict the Next ISIS Attack*, PBS (June 16, 2016, 5:16 PM), https://www.pbs.org/newshour/science/this-computer-algorithm-might-be-able-to-predict-the-next-isis-attack [https://perma.cc/9THM-6TYH]; Stew Magnuson, *Data Mining Not a Panacea for Catching Terrorists, Experts Warn*, NAT'L DEF. MAG. (Feb. 1, 2011), http://www.nationaldefensemagazine.org/articles/2011/1/31/2011february-data-mining-not-a-panacea-for-catching-terrorists-experts-warn [https://perma.cc/JE4X-79WJ] (noting that data mining might, at best, "improve the odds for the good guys").

[113] *See* SKILLICORN, *supra* note 33, at 13.

[114] FLACH, *supra* note 22, at 41.

[115] *See* Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 GA. L. REV. 109, 129 (2017).

of those features can result in over- or under-estimation of the amount of information conveyed by those variables.[116] Decisions featuring selection requirements can affect the accuracy of a computer model. More importantly for my purposes, their subjective nature makes feature selection one of the places where human impact is hugely significant.[117]

3. Choosing a Model

Constructing a model that not only accurately captures the data but also accurately generalizes to predict future events or behavior is far more art than science. In addition to determining which features of the data are relevant and how those should be represented, the analyst must decide which type of algorithm to employ and what the outcomes actually say about what policies to adopt.[118] Regardless of how the model is chosen, the choice will usually involve some sort of trade-off—the most accurate model, for example, might have a high false positive rate. Programmers must also resist the temptation of assuming that "results that seem plausible or interesting . . . must have found an underlying truth."[119] With a sufficient volume of data, "there are always patterns present by accident, and the danger is that an analysis algorithm will find these accidental patterns, rather than deeper, but perhaps less obvious, ones."[120] A good training and verification process can address this kind of issue, but it may come up if unaccounted for.

Again, law enforcement and security programs face particular challenges in choosing an appropriate model. First, whereas most of us make limited effort to mask our identity from Amazon—after all, if it does not know who we are, how can it recommend books we will like?—that is not the case when bad actors are seeking to evade government detection. Users in this area must therefore select models with an eye toward preventing manipulation by individuals who do not

[116] For example, a program that filters spam email by analyzing the words it contains might not use both "Viagra" and "blue pill" as relevant variables, because once the program recognizes the word "Viagra," considering whether "blue pill" is also included does not provide much additional knowledge. FLACH, *supra* note 22, at 44.

[117] *See infra* Section II.B.2.c (discussing impact humans have on machine learning).

[118] The model also must avoid common errors, such as overfitting and underfitting. *See* RUSSELL & NORVIG, *supra* note 29, at 705 (explaining that overfitting occurs when model incorporates features that are not actually relevant to relationship of data, while underfitting occurs when model is too simplistic to accurately capture relationships within data set).

[119] SKILLICORN, *supra* note 33, at 13.

[120] *Id.*

want to be identified.[121] A growing literature on adversary modeling has exposed how easily some models can be fooled.[122] This also makes ensuring continued accuracy over time more challenging, as bad actors will change their patterns if and when they discover that existing patterns have been identified. Thus, data accuracy and security—at both the collection and storage stage—becomes particularly important, and programmers must be sensitive to the presence of potentially misleading data.[123]

A further challenge in this regard differentiates national security from law enforcement. The cost of false negatives when it comes to national security programs is presumptively much greater than similar errors in the general law enforcement context. So whereas in general we might want to minimize false negatives to prevent the burdens of law enforcement from imposing too heavily on innocents, we might want programmers to strike a different balance in the national security context. Therefore model selection must be context specific, requiring a highly nuanced approach to government use of machine learning.

4. Verification

To ensure that a model is identifying useful relationships, it must be subjected to a rigorous verification process. A 2013 study of risk assessment algorithms like COMPAS found that "'in most cases, validity had only been examined in one or two studies' and that 'frequently, those investigations were completed by the same people who developed the instrument.'"[124] Like COMPAS, research on the PredPol's methodology has all been conducted by individuals who have financial stakes in the company.[125] Yet even objective efforts at verification can

[121] *Id.* at xvii (noting that one of most "surprising features of existing [data-mining technologies] is how fragile it is in face of actions by those who want to conceal themselves and their actions").

[122] *E.g.*, Richard Chirgwin, *Can You Get from 'Dog' to 'Car' with One Pixel? Japanese AI Boffins Can*, THE REGISTER (Oct. 30, 2017, 4:58 PM), https://www.theregister.co.uk/2017/10/30/fool_ai_image_classifier_by_changing_single_pixel/ [https://perma.cc/63VJ-KC68] (describing "pixel attack" that can trick one model into labeling image of car as dog by changing single pixel in image); Richard Chirgwin, *Another AI Attack, This Time Against 'Black Box' Machine Learning*, THE REGISTER (Dec. 18, 2017, 5:04 PM), https://www.theregister.co.uk/2017/12/18/black_box_ai_attack/ [https://perma.cc/G47U-WZ34] (describing same phenomenon but with images of celebrities).

[123] SKILLICORN, *supra* note 33, at 14.

[124] Angwin et al., *supra* note 8.

[125] Darwin Bond Graham, *Oakland Mayor Schaaf and Police Seek Unproven "Predictive Policing" Software*, EAST BAY EXPRESS (June 24, 2015), https://www.eastbayexpress.com/

overestimate the accuracy of the model in question, as their performance during verification does not always correspond to how models operate "in the wild."[126]

To be worthwhile, an algorithm's predictions must be at least as accurate as humans' predictions of the same phenomenon. There are times where this is likely to be the case. Beat cops trying to predict who is engaged in criminal activity or when someone is likely to have a weapon, for example, yield unimpressive results.[127] In determining whether an algorithm is accurate "enough," there must be a baseline against which to measure its performance. There is no predictive algorithm that will have a zero percent error rate. Even an algorithm that is correct ninety-nine percent of the time can still have both Type I and Type II errors—that is, it will both identify someone as a security threat who is not or fail to flag someone who does pose a security threat, or both. Determining what rates of each of these errors is acceptable is part of designing a predictive algorithm.[128] Any verification process must ensure that the computer model is a step forward, not back.

When it comes to verification, some law enforcement and national security programs will be particularly resistant to effective calibration. One benefit of machine learning models is that they can perpetually learn from new data. If Netflix predicts that you will enjoy a particular movie, for example, it will know it erred if you give it a one-star rating. Similarly, if you highly rate a show the algorithm would not have recommended for you, the error will be evident. Feeding this kind of information back into the data set allows the algorithm to learn from both its mistakes and successes, allowing it to become ever more accurate over time. Thus, once an algorithm is deployed in the field, ideally the relevant agency could track its performance and incorporate any new data into the model. A PredPol prediction that crime is likely to occur in a particular neighborhood can be assessed for its accuracy. Similarly, a determination that a particular individual should be subject to surveillance will either lead to evidence of a crime or threat, or it will not. Many security-related applications of machine learning, however, involve predictions that are not falsifiable.

oakland/oakland-mayor-schaaf-and-police-seek-unproven-predictive-policing-software/Content?oid=4362343 [https://perma.cc/XF48-QE72].

[126] Marco Tulio Ribeiro, Sameer Singh & Carlos Guestrin, *"Why Should I Trust You?" Explaining the Predictions of Any Classifier*, 2016 ACM SIGKDD 1135, 1136 ("[E]valuation on validation data may not correspond to performance 'in the wild.'").

[127] SKILLICORN, *supra* note 33, at 2-6 (noting that humans are not good at deceiving or detecting deception, with police only marginally better than random chance at identifying deception); Simmons, *supra* note 7, at 960-64 (discussing low "hit rates" for *Terry* stops and searches).

[128] Tal Z. Zarsky, *Automated Prediction: Perception, Law, and Policy*, 15 COMMS. OF THE ACM 33, 34 (Sept. 2012) ("[I]n some instances individuals would be wrongfully suspected and engaged due to a computer error.").

Consider a prediction that a foreign national poses a terrorism threat and is therefore denied entry to the United States. Or that a pre-trial detainee is a flight risk and is therefore denied bail. It is impossible to know whether the predictions were accurate or whether they were false positives. This means that the model will be able to learn from its false negatives, but not its false positives because those will be impossible to detect. Thus, in many of the machine-learning applications considered here, we will be at least partially blind to the actual success rate and therefore unable to exploit the tool's usual, inherent ability for self-improvement.

B. *Machine Learning and Rule-of-Law Values*

This Section will begin by identifying several fundamental values underlying the rule of law. It does not purport to provide an original account of what constitutes the rule of law. Rather, relying on leading scholars' existing accounts, this Section extracts from them shared principles. Moreover, this Article does not engage in an exhaustive discussion of these principles—each of them already is the subject of many volumes by philosophers, political scientists, and legal scholars. Instead, Section II.B.1 sketches out the rule-of-law principles with which machine learning for government decision-making might be in tension. Section II.B.2 will then explain the ways in which specific characteristics of machine learning engender that tension. Having set out these ideas, Part III turns to the implications of machine learning's potential inconsistencies with the rule of law for government decision-making.

1. Identifying Rule-of-Law Values

Chief Justice Marshall's assertion in *Marbury v. Madison*[129] that the United States is "a government of laws, and not of men"[130] is a manifestation of the universally accepted idea that, in America, individual conduct is governed by the rule of law, an idea that claims a historical pedigree dating back to Aristotle.[131] Appeals to the rule of law, as well as accusations that it has not been respected, are pervasive in legal and political argumentation, but contemporary authors do not always specify exactly what they mean by the concept and, indeed, it comes in many variations.[132] This is perhaps due to the fact that "rule

[129] 5 U.S. (1 Cranch) 137 (1803).

[130] *Id.* at 163.

[131] *See generally* ARISTOTLE, THE POLITICS: BOOK TWO (Benjamin Jowett, trans. 2000).

[132] *See* ALBERT VENN DICEY, INTRODUCTION TO THE STUDY OF THE LAW OF THE CONSTITUTION 179 (8th ed. 1915); JOHN LOCKE, TWO TREATISES OF GOVERNMENT §§ 131, 135-37 (P. Laslett ed., student ed. 1988) (1689) (emphasizing importance of governing by

of law" is not a single concept, but rather a web of interrelated principles. Moreover, it is not a binary inquiry. The question is usually not *whether* the rule of law exists, but to what extent it exists in any given context. Despite some variation at the margins, however, basic conceptions of rule of law aim to ensure that official power is exercised according to pre-defined legitimate authority rather than based on the whims of government officials.[133] More specifically, most modern accounts of the rule of law "generally emphasize five elements": 1) the law must be public and intelligible—people must be able to understand and comply with the law; 2) the law actually should guide people's actions; 3) the law must be stable, in the sense that rules are fixed and prospective; 4) the law must apply universally—to the powerful as well as the humble and to government officials as well as private citizens; and 5) there must be procedural mechanisms available to enforce the limits that the law imposes, including avenues to challenge government decision-making and demand reasoned justification.[134]

"established standing Laws, promulgated and known to the People," rather than rule by "extemporary Decrees"); MONTESQUIEU, THE SPIRIT OF THE LAWS Bk. 11, Ch. 6 (1748) (A. Cohler et al. eds., 1989) (insisting on separation of powers to preserve rule of law); JOHN RAWLS, A THEORY OF JUSTICE 207 (1999) (defining rule of law as "the regular and impartial, and in this sense fair" administration of public rules); JOSEPH RAZ, THE AUTHORITY OF LAW 214-18 (1979); Lon Fuller, *Positivism and Fidelity to Law: A Reply to Hart*, 71 HARV. L. REV. 630, 636–37 (1958).

[133] *See* Richard H. Fallon, Jr., *"The Rule of Law" as a Concept in Constitutional Discourse*, 97 COLUM. L. REV. 1, 7-9 (1997) (synthesizing "modern accounts" of rule of law into series of shared principles); Frank Lovett, *A Positivist Account of the Rule of Law*, 27 L. & SOC. INQUIRY 41, 42 (2002) (identifying "principles traditionally associated with the Rule of Law"); Colleen Murphy, *Lon Fuller and the Moral Value of the Rule of Law*, 24 L. & PHIL. 239, 239 (2005); Jeremy Waldron, *Is the Rule of Law an Essentially Contested Concept (in Florida)?*, 21 J. OF L. & PHIL. 137, 155 (2002) (noting that major theorists have "more or less the same conception" of rule of law). Lon Fuller's canonical discussion of the rule of law, for example, identifies eight principles: law should be general, publicly promulgated, prospective, intelligible, consistent, practicable, not too frequently changeable, and congruent with the behavior of government officials. *See* LON FULLER, THE MORALITY OF LAW 33-38 (2d ed. 1969).

[134] *See* Fallon, *supra* note 133, at 7-9 (identifying basic elements of most conceptions of rule of law). To be sure, like the idea of the rule of law itself, each of these principles can exist in greater or lesser degree along a continuum. There are, for example, regulatory regimes that are difficult for the lay person to understand, there are administrative rulings from which one cannot appeal, and there are judicial decisions issued by the Foreign Intelligence Surveillance Court ("FISC") that remain classified. But when government action moves too far toward the wrong end of the continuum, it can prompt objections, particularly when errors have high

Once the elements of the rule of law are identified, the next question focuses on what purpose those elements serve. There are at least three fundamental values that the rule of law vindicates: it ensures that individuals can plan their affairs effectively, imposes constraints on the arbitrary exercise of government power, and enhances government legitimacy.

a. *Ensuring Individuals Can Plan Their Affairs Effectively*

All standard rule-of-law accounts include the need for individuals to be able to order their affairs as one value served by the rule of law.[135] Some advocate this need on purely instrumental bases,[136] while others insist its value extends to non-instrumental goals like freedom, individual autonomy, and dignity.[137] The need for people to know how law will operate—and how to avoid running afoul of it—is evident in several rule-of-law elements. Insistence on a society governed by rules set out in advance in a way that is intelligible to everyone, that will remain relatively stable over time, and that will be enforced against private citizen and public official alike, reflects a commitment to advance notice of law's requirements. Prospectiveness and stability ensure that individuals can engage in planning for the future with "reasonable confidence that they can know in advance the legal consequences of various actions."[138] Public disclosure of intelligible rules recognizes both an individual's right and her ability to

costs. In other words, the degree to which each of these characteristics must exist for the system to conform to the rule of law will vary by context.

[135] *See, e.g.*, *id.* at 7-8; Jeremy Waldron, *Theoretical Foundations of Liberalism*, 37 PHIL. Q. 127, 134-35 (1987) (arguing that in liberal society, "principles must be amenable to explanation and understanding, and the rules and restraints that are necessary must be capable of being justified to people who are to live under them").

[136] *See, e.g.*, FULLER, *supra* note 133, at 31 ("Where penalties and deprivations are involved we are operating at the lower levels of human achievement where a defective performance can be recognized, if care is taken, with comparative certainty and formal standards for judging it can be established.").

[137] *See, e.g.*, JOHN FINNIS, NATURAL LAW AND NATURAL RIGHTS 273 (1980) (describing fairness and human dignity); F.A. HAYEK, THE CONSTITUTION OF LIBERTY 142-43 (1960) (describing freedom); RAZ, *supra* note 132, at 221 ("[O]bservance of the rule of law is necessary if the law is to respect human dignity.").

[138] Fallon, *supra* note 133, at 7-8; *see also* HAYEK, *supra* note 137, at 152 (noting that requirements, like generality, free individuals from dependence on will of others); RAZ, *supra* note 132, at 221 ("Respecting human dignity entails treating humans as persons capable of planning and plotting their future."); Jeremy Waldron, *The Rule of Law and the Importance of Procedures*, *in* JAMES FLEMING, GETTING TO THE RULE OF LAW 3, 21 (2001) (arguing that "freedom is possible" so long as "people know in advance how the law will operate and how they have to act if they are to avoid its application").

manage her own affairs according to her conscience, desires, and aspirations free from unjustified interference by the State.[139] When the law is certain and predictable, it generates an atmosphere within which individuals can control, at least in part, whether and when they will draw the State's attention.[140]

Procedural guarantees also serve to promote the necessary stability and predictability of the law. Such elements supply citizens with the confidence that the law will be applied as written and that when government officials deviate from it, they can be held accountable. And from the non-instrumental perspective, process rights "capture a deep and important sense . . . that law is a mode of governing people that treats them with respect, as though they had a view or perspective of their own to present on the application of the norm to their conduct and situation."[141] By enforcing legal norms through procedures that offer reasoned arguments and require justifications for government action, the State acknowledges citizens' capacity to reason by treating each as an entity "open to argument and persuasion, and deserving of reasoned explanations" rather than simply as objects to be coerced into compliance.[142] Any consequences of one's actions are the product of an individual with agency choosing that path, fully aware of the consequences. In short, the presence of these features confers on each individual the power to dictate the course of one's own life.

b. *Constraints on Arbitrary Exercise of Government Power*

Rule-of-law values constrain government power, thereby limiting government officials' ability to wield that power arbitrarily.[143] These constraints

[139] *See* Andrew E. Taslitz, *What Is Probable Cause and Why Should We Care?: The Costs, Benefits, and Meaning of Individualized Suspicion*, 73 L. & CONTEMP. PROBS. 145, 177 (2010) (explaining importance of transparency in Fourth Amendment context).

[140] *See id*. at 176 ("By refraining from criminal conduct or from actions that can be expected to raise suspicions of such conduct, the citizen gains some control over whether he will pay the added price for social safety and security of in fact being stopped by the police.").

[141] Waldron, *supra* note 138, at 15-16.

[142] Fallon, *supra* note 133, at 19; *see also* Murphy, *supra* note 133, at 250 ("Implicit in the idea of the rule of law is the view that an individual 'is or can become a responsible agent capable of understanding and following rules and answerable for his defaults.'"); Micah Schwartzman, *Judicial Sincerity*, 94 VA. L. REV. 987, 1012-18 (2008) (making this argument in context of judicial rulings).

[143] *See, e.g.*, DICEY, *supra* note 132, at 198 (arguing that rule of law must include "the absolute supremacy or predominance of regular law as opposed to the influence of arbitrary power"); HAYEK, *supra* note 137, at 205 (asserting that because "the rule of law means that government must never coerce an individual except in the enforcement of a known rule, it

flow from the public and universal nature of the rules, as well as citizens' right to contest their enforcement. When the government's exercise of power is effectively limited to action consistent with rules known to all citizens, those citizens are judged by well-known standards of behavior, rather than "the whims of officials."[144] The State is bound by fetters that reject unjustified exercises of power. It exerts coercive force because someone has taken specific action inconsistent with the recognized rules, not because that individual belongs to a disfavored group or offended a particular government official.

Insisting on procedural mechanisms to enforce the rules, such as the ability to seek redress for particular government actions from a neutral third-party or to appeal government decisions, renders this mission more effective.[145] The rule of law does not eliminate official discretion, but instead seeks to regulate and channel it, permitting the State's power to be brought to bear only when properly approved and authorized.[146] Determinations that are "backed with the collective and coercive force of political society" require justification and "must be defended in a way that those who are subject to it can, at least in principle, understand and accept."[147]

c. *Government Legitimacy*

While government legitimacy is not necessarily a substantive value underlying the rule of law, a certain level of consistency with rule-of-law principles is a necessary, though not sufficient, hallmark of legitimate government action.[148] Here this Article refers to legitimacy in the sense of

constitutes a limitation on the powers of all government"); Fallon, *supra* note 133, at 8 ("[T]he Rule of Law should guarantee against at least some types of official arbitrariness.").

[144] Murphy, *supra* note 133, at 250; *see also* Frank Lovett, *What Counts as Arbitrary Power?*, 5 J. POL. POWER 137, 139 (2012) ("[P]ower is arbitrary to the extent that its exercise is not reliably constrained by effective rules, procedures, or goals that are common knowledge to all persons or groups concerned.").

[145] Lovett, *supra* note 144, at 140 (explaining several procedural mechanisms to prevent arbitrariness).

[146] *See* SEAN COYLE, MODERN JURISPRUDENCE: A PHILOSOPHICAL GUIDE 122 (2014) (describing Professor Dworkin's belief that purpose of legality is "to guide and constrain the power of government" to employ coercive force of law in pursuit of its objective "except as licensed or required by individual rights and responsibilities flowing from past political decisions about when collective force is justified"); Jeremy Waldron, *Rule of Law*, *in* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY § 8.1 (Edward N. Zalta ed., 2016).

[147] *See* Schwartzman, *supra* note 142, at 990.

[148] *See, e.g.*, Jerry L. Mashaw, *Prodelegation: Why Administrators Should Make Political Decisions*, 1 J.L. ECON. & ORG. 81, 86 (1985) ("A consistent strain of our constitutional

"worthy . . . of both respect and public acceptance."[149] Multiple elements of the rule of law are building blocks for government legitimacy. Universality, the idea that law applies to everyone equally, regardless of their place in society, serves to legitimize legal rules.[150] When law is universal, the rules are less likely to serve as tools for one group to oppress another. When a burden is imposed or a benefit denied, it will be "for a reason other than the exercise of political power by the advantaged class."[151] Similarly, the predictability of the laws—their public, prospective, stable nature—is essential to their legitimacy. It ensures that the exercise of government power will be subject to limits and that citizens cannot be disadvantaged by government caprice. Publicly available rules are also more likely to reflect the deliberative consensus of the governed, rendering their enforcement a manifestation of the social contract. Procedures designed to ensure impartial justice are also a bulwark of legitimacy. When individuals can challenge government action and demand a reasoned justification, they are less likely to perceive that action as unjust or arbitrary.[152]

politics asserts that legitimacy flows from 'the rule of law.'"); Peter M. Shane, Chevron *Deference, The Rule of Law, and Presidential Influence in the Administrative State*, 83 FORDHAM L. REV. 679, 683 (2014) ("Because government officials may do only what the law permits, the rule of law also advances democratic legitimacy to the extent that the law is the product of democratic institutions."); *cf.* MAX WEBER, THE THEORY OF SOCIAL AND ECONOMIC ORGANIZATION 382 (1964) (arguing that political regime is legitimate when participants have faith in it); Tom R. Tyler, *Procedural Justice, Legitimacy, and the Rule of Law*, 30 CRIME & JUST. 283, 284 (2003) ("[P]eople's reactions to legal authorities are based to a striking degree on their assessments of the fairness of the processes by which legal authorities make decisions and treat members of the public."). In addition to conforming to the rule of law, various views of political legitimacy offer additional requirements—some procedural, such as democratic deliberation, and some substantive, such as protection of individual rights.

[149] Shane, *supra* note 148, at 681; *see* Waldron, *supra* note 135, at 134 (arguing that fundamental element of liberal thought is that authority must "answer at the tribunal of reason and convince us that it is entitled to respect").

[150] *See* Schwartzman, *supra* note 142, at 1004 ("[L]egal and political authorities act legitimately only if they have reasons that those subject to them can, in principle, understand and accept.").

[151] Cass R. Sunstein, *Beyond the Republican Revival*, 97 YALE L.J. 1539, 1579 (1988).

[152] Brennan-Marquez, *supra* note 7, at 1280 (emphasizing importance of providing explanation in context of Fourth Amendment probable cause because explanations promote rule-of-law values); Schwartzman, *supra* note 142, at 1004 (explaining value of having judges' opinions published for public).

2. Machine Learning's Tensions with Fundamental Values

The preceding discussion identified the ability to plan, rejection of arbitrary government action and government legitimacy as values that the rule of law protects. This Section turns to an explication of how some inherent characteristics of machine learning raise inevitable tensions with those values. Specifically, computer models' opacity, inevitable arbitrariness, and incorporation of subjective human decisions each conflict with the values identified above.

a. *Opacity*

Data mining generally, and machine learning specifically, often operate according to a black-box model—some inputs go in, an output emerges, but how the computer got from Point A to Point B is nearly always a mystery.[153] Sometimes this opacity is intentional—many algorithms are proprietary trade secrets[154]—while at other times secrecy flows from concerns that publicizing the process will allow bad actors to game the system.[155] But much of machine learning's opacity is a by-product of the technology. Some people simply do not understand what algorithms do. But even the most sophisticated computer scientists remain in the dark regarding the specifics of many algorithms.[156] Algorithms can find more complex relationships than a human ever would, which is part of what makes them so powerful. At the same time, algorithms frequently cannot explain in ways intelligible to humans how they reached their

[153] *See generally, e.g.*, FRANK PASQUALE, THE BLACK BOX SOCIETY (2015).

[154] *See* Jenna Burrell, *How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms*, BIG DATA & SOC'Y, Jan.-June 2016, at 1, 1-12 (identifying three sources of opacity: corporate or state secrecy, technical illiteracy, and inherent black-box nature of machine-learning algorithms); Citron & Pasquale, *supra* note 2, at 5 (noting that algorithms, such as financial scoring systems, are "shrouded in secrecy"); Wexler, *supra* note 67 (describing unsuccessful challenges by criminal defendants to use of proprietary algorithms for parole decisions and DNA matching); Letter from Elec. Privacy Info. Ctr. to Hon. John Thune, Chairman, and Hon. Bill Nelson, Ranking Member, U.S. Senate Comm. on Commerce, Sci., & Transp. (Mar. 22, 2017) (discussing heightened cybersecurity threat that algorithmic opacity poses to American consumers).

[155] Citron & Pasquale, *supra* note 2, at 26 (discussing "extent to which the public should have access to the data sets and logic of predictive credit-scoring systems" given "gaming concerns").

[156] *See* Rich, *supra* note 7, at 919 (noting that the most effective algorithms "are likely to operate in a way that is not comprehensible even to the people who programmed the algorithm").

conclusions.[157] In fact, the more complex and powerful an algorithm, the more opaque it is likely to be.[158]

This lack of transparency is a feature of complex computer models that cannot be wholly eliminated.[159] Numerous commentators have recognized the accountability concerns raised by the unintelligibility of algorithmic models.[160] Some legal scholars argue, for example, that we need judicial or public access to source code and the software applications it uses,[161] programmers' notes and

[157] *Id.*

[158] *Id.* at 928. Transparency concerns arise not just with respect to the algorithm itself, or with the collection, aggregation, and accuracy of data, but also with respect to whether the government discloses that it is *using* a predictive model. This is often true of information regarding any law enforcement or classified data mining programs. Zarsky, *supra* note 7, at 1523-24 (discussing legal requirements regarding transparency when employing predictive modeling processes).

[159] In some contexts, a statute requires the government to make public its use of data or data mining. *See, e.g.*, Privacy Act of 1974, 5 U.S.C. § 552a (2012) (requiring agencies to produce SORN for each group of records "from which information is retrieved by the name of the individual or by some other . . . identifying particular assigned to the individual"); Federal Data-Mining Reporting Act of 2007, 42 U.S.C. §2000ee-3 (2012) (requiring agencies to disclose to Congress all data mining programs, their goals, and what information they use); 44 U.S.C. § 3501 note (2012) (E-Government Act of 2002 § 208) (requiring agencies to conduct privacy impact assessments ("PIAs"); explain what personally identifying information they are collecting and why; and how it will be collected, used, accessed, shared, protected, and stored); *System of Records Notices*, U.S. DEP'T OF HOMELAND SECURITY, https://www.dhs.gov/system-records-notices-sorns (last updated August 30, 2018) [https://perma.cc/QTX8-SYYF]. But these requirements fail to result in true transparency because they frequently include provisions either exempting law enforcement entirely from their requirements or failing to require information about law enforcement's use of these tools to be made public. *See* 5 U.S.C. § 552a(k)(2) (law enforcement exemption from Privacy Act's requirements); Citron & Pasquale, *supra* note 2, at 26 (explaining under-inclusiveness of rules requiring SORNs and PIAs); Caleb Watney, *When It Comes to Criminal Justice AI, We Need Transparency and Accountability*, RSTREET (Dec. 1, 2017), https://www.rstreet.org/2017/12/01/when-it-comes-to-criminal-justice-ai-we-need-transparency-and-accountability/ [https://perma.cc/3LY8-TCJ7] (discussing whether transparency mandates should be imposed on government-employed algorithms in addition to private sector). In addition, these notices' descriptions of the programs to which they relate are couched in general language, failing to specify exactly how the data mining program at issue works or how the data are used. Zarsky, *supra* note 7, at 1526 (arguing these notice requirements often use language too broad to be helpful).

[160] *See* sources cited *supra* note 7.

[161] *See, e.g.*, Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1308 (2008) (arguing for new concept of technological due process and providing

the algorithms themselves,[162] information regarding how the models are used, including verification of their effectiveness,[163] and features that go into the algorithm and their respective weights.[164] Revealing an algorithm and all of its elements, however, does not necessarily render it transparent.[165] "Code can be complicated or obfuscated, and even expert analysis often misses" aspects of a program.[166] Nor is it feasible to expect that the government would divulge the features used to generate models designed to aid law enforcement or security at the border.

Others insist that algorithms should be interpretable. Indeed, machine-learning experts have begun to work on what is known as "XAI," or explainable artificial intelligence, but to date there is no such thing.[167] Moreover, because

framework of mechanisms to enhance transparency, accountability, and accuracy of rules embedded in automated decision-making systems); Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 117 (2014) (suggesting that due process guarantees should include "right to audit the data used to make the determination"); Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1376 (2018) (arguing that government should not be permitted to shield algorithmic details from public view on grounds that they are trade secrets); Zarsky, *supra* note 7, at 1524-26.

[162] Citron & Pasquale, *supra* note 2, at 14.

[163] Zarsky, *supra* note 7, at 1526-30 (discussing need for transparency with regard to algorithms).

[164] Citron & Pasquale, *supra* note 2, at 11 (criticizing opacity of credit scoring because, among other reasons, scorers do not specify relative weight of certain categories in scoring system); Simmons, *supra* note 7, at 997-99 (arguing to assess machine-learning algorithms, all that must be made transparent is factors that algorithm used and its historical accuracy); Letter from Elec. Privacy Info. Ctr., *supra* note 154 (arguing that absent access to details of models themselves, some alternative means to "recapture the purpose of transparency without simply relying on testing inputs and outputs" must be devised).

[165] *See* Lilian Edwards & Michael Veale, *Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For*, 16 DUKE L. & TECH. REV. 18, 18 (2018) (arguing that available means of explaining machine-learning outcomes are unlikely to provide meaningful information about their operation).

[166] Kroll et al., *supra* note 7, at 647; *id.* at 649-50 (stating that "transparency advocates often claim that by reviewing a program's disclosed source code, an analyst will be able to determine how a program behaves" but "this claim is belied by the extraordinary difficulty of identifying even genuinely malicious code ('malware'), a task which has spawned a multibillion-dollar industry based largely on the careful review of code").

[167] *See, e.g.*, Ribeiro, Singh & Gustrin, *supra* note 126 (proposing novel technique for understanding machine-learning predictions); Cliff Kuang, *Can A.I. Be Taught to Explain Itself?*, N.Y. TIMES MAG., Nov. 21, 2017 (describing some approaches to rendering AI more

the strongest predictive models are often opaque, limiting machine learning to interpretive predictive models threatens to sacrifice the very value of the tool, which is premised on the idea that algorithms can derive knowledge from data sets too large or complex for humans to analyze.[168] Still others seek technological means of ensuring procedural regularity, if not actual transparency.[169] But while those methods might increase public confidence in the accuracy and fairness of the resulting models, they ultimately do not represent actual transparency and cannot mitigate the rule-of-law-based concerns discussed below.

Before turning to that discussion, however, it is important to differentiate machine learning from other contexts in which opacity is widely accepted. As an initial matter, opacity of systems used by the government raises questions significantly different from those raised by reliance on opaque systems in other contexts. Indeed, there are many contexts outside the scope of this Article where we happily rely on "black-box" mechanisms. Most of us do not understand the details of how prescription drugs function, for example, or how our spam filter operates. Yet we continue to employ those tools. Each of these, however, is a decision an individual makes with full information about the risk of errors, rather than enabling an exercise of the government's use of coercive power.

transparent as well as some challenges subject area presents); David Gunning, *Explainable Artificial Intelligence (XAI)*, DEFENSE ADVANCED RESEARCH PROJECTS AGENCY, https://www.darpa.mil/program/explainable-artificial-intelligence [https://perma.cc/7X6D-J6R9] (last visited Sept. 11, 2018); Alfredo Vellido, José D. Martín-Guerrero & Paulo J.G. Lisboa, *Making Machine Learning Models Interpretable*, EUROPEAN SYMPOSIUM ON ARTIFICIAL NEURAL NETWORKS, COMPUTATIONAL INTELLIGENCE AND MACHINE LEARNING (Apr. 25-27, 2012), https://pdfs.semanticscholar.org/ce0b/8b6fca7dc089548cc2e9aaac3bae82bb19da.pdf [https://perma.cc/F9YS-YM6N]. Moreover, the explanation required in one context might not meet the needs of another—as one researcher put it, "the explanation a doctor needs from a machine isn't the same as the one a fighter pilot might need or the one an N.S.A. analyst sniffing out a financial fraud might need." Kuang, *supra*. In other words, there may be as many types of explanation as there are uses of machine learning itself.

[168] *See* SKILLICORN, *supra* note 33, at 78; *id.* at 313 ("[T]he strong predictors are opaque and the transparent predictors are weak. Confidence and explanatory power are important properties in practice, but we do not know how to get them both together."); Peter Margulies, *Surveillance By Algorithm: The NSA, Computerized Intelligence Collection, and Human Rights*, 68 FLA. L. REV. 1045, 1068-71 (2016) (explaining transparency paradox in algorithmic processes); Neil M. Richards & Jonathan H. King, *Three Paradoxes of Big Data*, 66 STAN. L. REV. ONLINE 41, 42 (2013) (describing transparency paradox).

[169] *See* Kroll et al., *supra* note 7, at 662-72 (proposing that technical tools can address some concerns raised by algorithms' opacity).

DNA testing, use of radar guns to measure speed, and breathalyzer tests to detect intoxication are examples of government use of opaque tools. When initially introduced, these tools also triggered opacity complaints. While they have come to be generally accepted, they differ from the government's use of machine learning in important ways. As an initial matter, there are standardized, nation-wide procedures, regulations, and oversight mechanisms in place to ensure that evidence is gleaned by these methods in a scientifically validated manner.[170] Moreover, we know that to avoid triggering a radar gun, we should not speed. When it comes to computer algorithms, there are no standardized requirements, and it is impossible to know what (sometimes entirely innocent) behavior the model might take into account.

There are also contexts where human decision-making is itself opaque—police officers might offer a false or pretextual explanation for their actions, and juries need not offer any justification at all. Our entire legal system, however, is designed with human decision-makers in mind. Human decision-makers can be asked to explain their actions, and procedures are in place to deter use of false evidence or improper motives. A law enforcement officer, for example, must have an objectively valid justification for her actions, even if the proffered reason was not her true motivation.[171] And when it comes to juries, we have elaborate rules regarding what evidence they may see, what kinds of questions they can be asked, and what factors they are allowed to consider. If a juror expresses racial prejudice, for example, a verdict might be overturned.[172] Majority or unanimity requirements also serve to check jurors' discretion and juries are not empaneled in the first place until after many other constraints are satisfied, such as probable cause, a grand jury indictment (in the federal system), and a decision to prosecute. No such constraints exist for predictive algorithms.

The most compelling parallel to machine learning is the use of dogs to detect illicit drugs.[173] There, the black box is actually an animal for which the only

[170] *See* Maryland v. King, 133 S. Ct. 1958, 1968 (2015) (describing uniform national quality and auditing standards laboratories must meet in order to participate in FBI's national Combined DNA Index System); STANDARDS FOR CRIMINAL JUSTICE, DNA EVIDENCE (AM. BAR ASS'N, 3d ed. 2007).

[171] *E.g.*, Whren v. United States, 517 U.S. 806, 813 (1996) (holding that under Fourth Amendment police motive does not matter so long as there is legitimate basis for stop).

[172] *E.g.*, Pena-Rodriguez v. Colorado, 137 S. Ct. 855, 856-57 (2017) (holding that when juror's statement indicates reliance on racial stereotypes to convict criminal defendant, trial court must consider whether juror's statement denied defendant's right to trial by jury).

[173] *See* Ferguson, *supra* note 7, at 267-69 (discussing various possible analogies for predictive policing tools); Michael L. Rich, *Machines as Crime Fighters, Are You Ready?*, 30

justification for its use is that historically it has successfully identified containers that held drugs.[174] And while the use of drug-sniffing dogs is by no means uncontroversial in itself, there are two reasons that it is less problematic from a transparency perspective than machine learning. First, dogs used for this purpose must successfully complete an accredited training regime.[175] No such certification exists for machine-learning algorithms. Second, like the radar gun or the breathalyzer, a drug dog is seeking a very specific type of wrongdoing: it is answering a binary question. So as an initial matter, we can avoid having our bags searched most of the time by declining to transport illegal drugs. More importantly, however, dog sniffs fall into the category of high-discretion government action where machine learning might be validly employed.[176] An erroneous determination by a drug-sniffing dog usually will have relatively insignificant consequences, it is easily falsifiable, and therefore inflicts no lasting harm on the innocent suspect. To be sure, such an experience is humiliating and disruptive. But its consequences pale in comparison to some non-falsifiable decisions currently informed by machine learning, such as not approving pre-trial bail or barring entry into the United States. Machine learning therefore represents a novel form of government opacity. Perhaps one day it will be as easily accepted as DNA evidence. Until that day comes, however, its use potentially conflicts with each of the identified values underlying the rule of law.

If the rule of law seeks to ensure that citizens can plan their lives based on public, stable, prospective rules, a framework of secret rules fails to provide a basis for understanding and following rules. To be sure, it is the means of administering the law that will be secret rather than the law itself. But not knowing what features an algorithm has deemed relevant to that assessment, or how heavily it weighs particular factors, will similarly undermine predictability.

CRIM. JUST. 10, 12-13 (Winter 2016) (discussing drug dogs as automated suspicion algorithms).

[174] Note that there is significant skepticism regarding the accuracy of drug-sniffing dogs. *See* Illinois v. Caballes, 543 U.S. 405, 410 (2005) (Souter, J., dissenting) (arguing that "assumption that trained sniffing dogs do not err" is "untenable"); United States v. $80,760.00 in U.S. Currency, 781 F. Supp. 462, 478 (N.D. Tex. 1991) (stating that drug detection dogs can be unreliable "when the dog receives poor training, has an inconsistent record, searches for narcotics in conditions without reliability controls or receives cues from its handler"). *See generally* Lisa Lit, Julie B. Schweitzer & Anita M. Oberbauer, *Handler Beliefs Affect Scent Detection Dog Outcomes*, 14 ANIMAL COGNITION 387 (2011) (describing study that revealed alerts by drug-sniffing dogs are affected by handler beliefs).

[175] *See* Rich, *supra* note 173, at 12 (noting that drug dogs must be trained and certified).

[176] *See infra* Section III.B (arguing more discretion allowed to government decision-makers, more appropriate use of machine learning predictions becomes).

Moreover, substantive choices embedded in an algorithm—such as the ones discussed in Section II.B.2.c—that government decision-makers rely on, whether consciously or unconsciously, operate much like a secret law or rule. Employing predictive algorithms prevents citizens from, as one scholar puts it, being "the authors of our own tales, and controlling how and when parts of it unfold."[177] We lose the ability to plan with any confidence. Instead of treating citizens as free individuals with the capacity to make reasoned decisions about their conduct, the government reduces individuals to a mosaic made up of a list of concrete data points that we are "unable to observe, understand, participate in, or respond to."[178]

The inability to convey the inner workings of computer models also undermines tools we rely upon to constrain the arbitrary exercise of government power. To lack arbitrariness, the government must have some justification for the exercise of its power, and that exercise of power must have some connection to the proffered explanation. Opacity thwarts both an individual's ability to challenge the initial decision—if the proffered justification is "the algorithm identified you," on what grounds could you appeal?—and the government's ability to legitimize its ultimate decision by providing reasoned justification. Basing government decisionmaking on factors not susceptible to reasoned explanation—and therefore not amenable to the usual protections against arbitrary action—runs contrary to the fundamental idea that government actors must justify the use of the State's coercive power.

Finally, the absence of transparency also threatens government legitimacy. To the extent that government action is considered legitimate when it adheres to universally known and understood norms, rules lacking these characteristics are viewed as less valid.[179] The unintelligibility of machine-learning algorithms means that there is no way to confirm whether government actors are following the rules. A slew of empirical studies show that individuals are more likely to accept as legitimate a government whose processes they perceive as fair.[180] Predictive analytics and their black-box processes undermine this perception due to the inability to explain their outcomes.

The opacity of algorithmic models therefore undermines each of the rule-of-law values identified. And if algorithms become less intelligible as they become more effective, an algorithm's effectiveness will vary inversely with its

---

[177] Taslitz, *supra* note 139, at 195.

[178] Ian Kerr & Jessica Earle, *Prediction, Preemption, Presumption: How Big Data Threatens Big Picture Privacy*, 66 STAN. L. REV. ONLINE 65, 71 (2013).

[179] *See* Waldron, *supra* note 135, at 135 (arguing that if social order cannot be justified to particular individual, its legitimacy with respect to that individual is undermined).

[180] *See, e.g.*, Tyler, *supra* note 148, at 283 (noting that "considerable evidence suggests that the key factor shaping public behavior is the fairness of the processes legal authorities use when dealing with members of the public").

compliance with rule-of-law norms. Moreover, if transparency in algorithmic decision-making generally is a problem, that problem is compounded in the law enforcement and security context. Secrecy is always at issue in law enforcement and security policy, but in addition to those persistent transparency concerns, agencies pursuing these goals are exempt from most of the data integrity and public reporting requirements that apply to the use of machine-learning tools elsewhere in the government. It is in this context that one justification for transparency—concerns over gaming the system—looms largest.

b. *Arbitrariness of Errors*

A computer model's prediction is based on a determination that certain traits, characteristics, events, or other known variables point toward a certain outcome. But errors, at least some of which are inevitable, reflect some level of arbitrariness for at least two reasons. First, recall that multiple models built from the same data set might produce equally accurate predictions. The incorrect predictions, however, will differ from one model to the next—the variation might be due to conscious decisions, such as whether to minimize false positives or false negatives, but they also may simply be the results of whatever black-box calculations that particular algorithm devised. In other words, two algorithms might each be ninety-nine percent accurate and yet produce inconsistent predictions in any particular individual instance.[181] These disparate outcomes result from how the computer model views the relationship between one person's data and the remaining data set. This relationship may or may not be meaningful, and no two algorithms are likely to define the relationship in exactly the same way. And while an error rate of one percent may seem negligible, if an algorithm that predicts who will become a terrorist with ninety-nine percent accuracy looked at the U.S. population of about three hundred million people, that algorithm would identify three million suspected terrorists.[182]

The second source of arbitrariness is the fact that no computer model can take into account all potentially relevant information. Machine-learning models have

[181] Imagine one hundred passengers are screened for placement on the No-Fly list. Two algorithms accurately identify safety threats eighty-five percent of the time. Despite their identical aggregate numbers, those fifteen individuals misidentified might be individuals one through fifteen in the first algorithm, and individuals thirty through forty-five in the other. Thus depending on the algorithm, a completely different set of people will be erroneously placed on, or excluded from, the No-Fly list.

[182] Jonas & Harper, *supra* note 108, at 7 ("[The] statistical likelihood of false positives is so high that predictive data mining will inevitably waste resources and threaten civil liberties.").

a finite number of inputs and can only consider the data supplied to it.[183] A data set may exclude potentially relevant information, for example, about unexpected or rare characteristics.[184] Other types of information absent from predictive models might also render something or someone an anomaly. Someone may exhibit all of the hallmarks of a flight risk, and yet actually would be a false positive because she feels an unusually significant amount of remorse, or because she has used this arrest as motivation to turn over a new leaf. If we are to acknowledge that individuals exercise free will, then we have to assume that in cases where a feature not part of the model is dispositive, the prediction will be inaccurate. To be sure, there are plenty of times that a human decision-maker will lack all of the relevant facts. But humans are capable of incorporating into their decision-making new or surprising information in a way that an algorithm is not.

Government decisions based on potentially arbitrary outcomes, rather than on individually justifiable considerations, directly conflict with the rule-of-law principle's demand that individuals know enough about relevant rules and their enforcement to plan their affairs. This includes making decisions regarding whether to transgress those rules and to accept the consequences. Machine learning divorces—at least in some instances—the expected cause-effect relationship between an individual's actions and what follows from those actions. It thereby removes from individual citizens the power to choose their path.

Reliance on an arbitrary decision-making process—even one that is nearly always accurate—is also inconsistent with our usual views of the permissible uses of the State's coercive power.[185] When the government relies on predictive

[183] Rich, *supra* note 7, at 897 (noting that algorithm "is limited in making its predictions to analysis of the data within its dataset, and it cannot consider other facts that might be relevant but that were not included"). This is undoubtedly true of human decision-making as well. Humans, however, can more easily incorporate unexpectedly relevant factors into decisions where they apply algorithms, which always consider only those data points determined to be relevant in most instances.

[184] This is known in the psychology literature as the "broken leg" problem. If Joe regularly attends a movie on Friday nights for a year, one might predict that he will do so this coming Friday as well. If, however, he breaks his leg before Friday and is home-bound for a month, that prediction will be incorrect. *See* DAVID FAUST, THE LIMITS OF SCIENTIFIC REASONING 53 (1984) (attributing idea to Paul Meehl's 1954 book *Clinical Versus Statistical Prediction*).

[185] Brennan-Marquez, *supra* note 7, at 1300-01 (noting that requiring justifications for individualized suspicion is likely to make policing less precise but that it will "vindicate a core promise of constitutional democracy: that governance is an outcome of popular sovereignty"); *see* Jane Bambauer, *Hassle*, 113 MICH. L. REV. 461, 507 (2015) (arguing that

analytics, some people will be subjected to seemingly arbitrary government intervention rather than paying the predictable consequences of engaging in certain behavior.[186] To be sure, human decision-making will produce false positives as well, but those false positives will be based on a human's rationale, rather than on the decision of which algorithm to employ. The affected individual will perceive no difference between law enforcement action informed by computer model and law enforcement action based on arbitrary whim. The rule-of-law insistence that government action be based on public, comprehensible, and predictable rules envisions government decisions based on meaningful justifications.

Legitimacy will also suffer from arbitrariness. When the process through which law is generated and enforced is viewed as fair, the law itself enjoys more legitimacy in the eyes of those that it governs.[187] The instinctive (negative) reaction that many people have to the idea of allowing machines to make consequential decisions indicates that the legitimacy of the method is on shaky ground from the start.[188] Indeed, studies show that seeing an algorithm err makes people less likely to trust it over a human forecaster, even when the algorithm outperforms the human overall.[189] Machine learning's inherent arbitrariness compounds this concern. A law-abiding citizen treated as a suspect or a danger to the community will expect an explanation for his treatment—an explanation

hassle imposed by law enforcement should be distributed randomly, rather than on discretionary basis that law enforcement currently uses).

[186] *See* Taslitz, *supra* note 139, at 176-77 (explaining this concept in context of probable cause).

[187] *See* Tom Tyler & Alan Lind, *Procedural Justice*, *in* HANDBOOK OF JUSTICE RESEARCH IN LAW 84 (2001) (exploring what makes authority legitimate to public); Tyler, *supra* note 148, at 284 (noting that "people's subjective judgments about the fairness of the procedures through which the police and the courts exercise their authority" strongly influence public's law-related behavior).

[188] *See* Tal Z. Zarsky, *"Mine Your Own Business!": Making the Case for the Implications of the Data Mining of Personal Information in the Forum of Public Opinion*, 5 YALE J.L. & TECH. 1, 18–50 (2002); *see also* Bambauer, *supra* note 185, at 466 (noting certain qualities not found in algorithmic processes still considered important, such as human intuition and case-by-case approach). *But see, e.g.*, Kate Goddard, Abdul Roudsari & Jeremy C. Wyatt, *Automation Bias: A Systemic Review of Frequency, Effect Mediators, and Mitigators*, 19 J. AM. MED. INFORM. ASS'N 121, 123 (2012) ("Automation bias [(the tendency to over-rely on automation)] appears to be a fairly robust and generic effect across research fields.").

[189] Berkeley J. Dietvorst, Joseph P. Simmons & Cade Massey, *Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err*, 144 J. EXPERIMENTAL PSYCH. 114, 114 (2014) (detailing research indicating that people often exhibit algorithm aversion and exploring why).

based on his actions—but no satisfactory explanation will be forthcoming. "The algorithm said I should do it" is unlikely to suffice as a satisfying reason that, for example, someone has been stopped and frisked. While innocents will be subject to state intervention regardless of how decisions are made, they are more likely to accept this treatment as part of the cost of living in a dangerous world when the State can explain what actions prompted the intervention. It is when the government "seem[s] to be acting without an adequate evidentiary basis[,] upon generalizations," or arbitrarily that its actions are likely to be perceived as unfair or punitive and therefore illegitimate.[190]

c. *The Human Factor*

The final aspect of machine-learning algorithms that creates rule-of-law problems is what this Article refers to as "the human factor." To state the obvious, computer models are built by people.[191] And while machine-learning algorithms do seem to develop a mind of their own, they do not burst forth like Athena fully formed. The role that humans play in the construction and deployment of the algorithm manifests in multiple ways, leaving human fingerprints on every stage of the process.[192] A programmer must make dozens of decisions that, consciously or unconsciously, impact the outcome. Questions such as which features to employ and calibrating their relative weight, how to address issues of incomplete or incorrect data, what types of models to use, which type of algorithm from among several plausible choices to employ, how to interpret the outputs, and how to measure the model's performance and

[190] Taslitz, *supra* note 139, at 190.

[191] Even when machine-learning models build other machine-learning models, humans must build the original models. *See* Cade Metz, *Building AI is Hard—So Facebook Is Building AI that Builds AI*, WIRED (May 6, 2016, 7:00 AM), https://www.wired.com/2016/05/facebook-trying-create-ai-can-create-ai/ ("Inside Facebook, engineers have designed what they like to call an 'automated machine learning engineer,' an artificially intelligent system that helps create artificially intelligent systems.").

[192] Citron & Pasquale, *supra* note 2, at 4 (explaining that because human beings program algorithms and algorithms mine data sets containing information reported and recorded by humans, humans' biases and values are embedded in algorithm's instructions and output); *see id.* at 14 ("Software engineers construct the datasets mined by scoring systems; they define the parameters of data-mining analyses; they create the clusters, links, and decision trees applied; they generate the predictive models applied. The biases and values of system developers and software programmers are embedded into each and every step of development."); Rich, *supra* note 7, at 885 (noting that "choices made by humans throughout the machine learning process can cause inaccuracies in the final predictions of a machine learning algorithm").

determine whether it is sufficiently reliable, all must be answered and all introduce a specific bias.[193] As a result, it is impossible to remove the human factor from algorithmic decision-making entirely. Machine learning is never going to be purely mathematical in its operation. The ways in which human "biases and values are embedded into the software's instructions," however, are not evident on the face of the algorithm.[194] Instead, those biases and values impact the result in multiple, frequently undetectable ways.

The form of human bias that has received the most attention is machine-learning models' demonstrated tendency to exhibit discriminatory bias, especially with respect to race.[195] This type of predisposition generates numerous concerns that must be addressed in any context where algorithmic models are employed. Biased models can reinforce discriminatory stereotypes, generate or perpetuate unjustifiable hierarchies, and disproportionately impose

[193] *See* KELLEHER, *supra* note 4, at 511. The issue of how each factor is weighted is one of the areas in which scholars champion increased transparency. *See* Citron & Pasquale, *supra* note 2, at 24-25; Mikella Hurley & Julius Adebayo, *Credit Scoring in the Era of Big Data*, 18 YALE J.L. & TECH. 148, 152 (2016) ("The use of non-transparent credit-assessment systems that judge consumers based on factors that they are not aware of and which may be beyond consumers' control, fundamentally conflicts with the American ideal of self-determination."); Kroll et al., *supra* note 7, at 646.

[194] Citron & Pasquale, *supra* note 2, at 4; *see* Citron, *supra* note 161, at 1258 (noting that "computer systems collapse individual adjudications into rulemaking, making it difficult, if not impossible, to determine whether a decision resulted from factual errors, distorted policy, or both").

[195] The risk of racial discrimination in machine learning is well documented. *See, e.g.*, Kelly Hannah-Moffat, *The Uncertainties of Risk Assessment: Partiality, Transparency, and Just Decisions*, 27 FED. SENT'G RPT. 244, 244 (2015); Selbst, *supra* note 115, at 5; Claire Cain Miller, *When Algorithms Discriminate*, N.Y. TIMES, July 9, 2015, at B1; ACLU et al., Predictive Policing Today: A Shared Statement of Civil Rights Concerns (Aug. 31, 2016), https://www.aclu.org/other/statement-concern-about-predictive-policing-aclu-and-16-civil-rights-privacy-racial-justice [https://perma.cc/2SJ4-YF2C] (expressing concerns of seventeen organizations regarding civil rights implications of predictive policing, including potential for racial bias). Researchers have worked to develop anti-bias measures to mitigate the risk. *See* Cynthia Dwork & Deirdre K. Mulligan, *It's Not Privacy and It's Not Fair*, 66 STAN. L. REV. ONLINE 35, 39 (2013); Kroll et al., *supra* note 7, at 689 (discussing such efforts); Moritz Hardt, Eric Price & Nathan Srebro, *Equality of Opportunity in Supervised Learning*, 30th Conf. on Neural Information Processing Systems (Oct. 7, 2016), https://arxiv.org/pdf/1610.02413.pdf [https://perma.cc/U6ZE-L6TC]. There is, however, no reliable means of entirely eliminating these concerns. *See* Kroll et al., *supra* note 7, at 678-79 (noting that ensuring "fair" classifications requires policy decision regarding what kind of fairness matters).

the costs of false positives on particular groups.[196] Bias can enter the system in several ways. First, using a database whose contents themselves are a product of bias will result in a model that reflects that bias.[197] If, for example, an algorithm that tells police whom they ought to stop and frisk is trained on a data set that over-represents the crime rate within certain groups—perhaps because those groups have historically been disproportionately targeted for enforcement—the resulting model can reflect the bias extant in the data.[198] Or government officials might overlook relevant data. In the national security context, one can imagine that disproportionate focus on threats emanating from the Muslim community might result in models that underestimate other potential sources of terrorist violence. Algorithms' risk of perpetuating discriminatory bias is a topic that deserves—and continues to receive—significant attention.

While systematization of discriminatory bias is one risk of using computer models, the news is not all bad. Indeed, proponents of algorithmic decision-making hail its potential to reduce some forms of bias, particularly improper discrimination, as one of its benefits. In this regard, computer algorithms present a means of actually reducing problematic government decision-making. If, for example, the bias is a known factor, computer models can take that into account in ways that the human mind cannot. Because humans exercise discretion based at least in part on conscious or unconscious bias, computerized decision-making can reduce bias by replacing human decision-making with data-driven

[196] *See, e.g.*, DEP'T OF DEF., *supra* note 20, at 39; Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 715-29 (2016); Ferguson, *supra* note 5, at 402 ("If data are collected only about certain classes of people [i.e., Muslims, people of color], then those people are more likely to become future targets of suspicion simply because of the initial selection bias."); Murphy, *supra* note 133, at 831 (noting that predictive models do not spread their burdens equitably if they are not fairly composed and adequately monitored). Even if features such as race are not included in a model, features highly correlated with race, such as socioeconomic status or home address, may nevertheless result in a model that will reflect that bias. *E.g.*, Simmons, *supra* note 7, at 969-80.

[197] *See, e.g.*, Simmons, *supra* note 7, at 980-83; Matt Cagle, *This Surveillance Software Is Probably Spying on #BlackLivesMatter*, ACLU NORTHERN CALIFORNIA BLOG (Dec. 15, 2015), https://www.aclunc.org/blog/surveillance-software-probably-spying-blacklivesmatter [https://perma.cc/4JWP-ZVBS] (exposing Fresno Police Department predictive algorithm's use of hashtag #BlackLivesMatter as increased risk factor for hate crimes against police). Moreover, due to the absence of transparency regarding the workings of algorithms, such models can simultaneously systematize discrimination and conceal its presence.

[198] *See* Kroll et al., *supra* note 7, at 680-81 (describing how model trained on NYPD's stop and frisk program's data—where eighty-three percent of those stopped were black or Hispanic and ten percent were white, resulting in more blacks and Hispanics found engaging in criminal behavior—might conclude that being black or Hispanic is predictor of criminal behavior).

decisions.[199] Algorithms' potential to mitigate the effects of conscious and unconscious bias in law enforcement and security-related decisions—where such bias is pervasive[200]—places a very strong thumb on the scale in favor of their use. So while it is true that bias cannot be eliminated entirely from algorithms, if they are able to act more objectively than humans exercising discretion—who are unable to shed bias entirely, particularly when it comes to protected categories like race, ethnicity, and religion—algorithmic-based decision-making will be an improvement.

Discriminatory bias is not, however, the only human judgment that might be incorporated into machine-learning models. Consider an algorithm seeking to implement a vague or ambiguous standard, such as "equality." That term means different things to different people. After *ProPublica* demonstrated that the COMPAS algorithm over-identified blacks as likely recidivists,[201] several teams of researchers sought to build a model that would "correct" that "error." As it turns out, it was not an error; it was a design choice. The algorithm was programed to avoid racial bias by producing equal *accuracy rates* for all races. That is a perfectly reasonable definition of "equality." Four sets of researchers independently concluded, however, that any model that guarantees equal *accuracy rates* will necessarily produce unequal *false-positive rates*.[202] Achieving equal false-positives—i.e., imposing the costs of machine-learning inaccuracy equally across races—is an equally plausible measure of fairness. Yet the necessary conclusion is that constructing an algorithm exhibiting racial "equality" still requires human actors to determine what "equality" means in that instance.[203]

---

[199] Bambauer, *supra* note 185, at 482.

[200] *E.g.*, Rich, *supra* note 7, at 897-900; Simmons, *supra* note 7, at 976 (noting that "racial biases of police officers and magistrates permeate every aspect" of their decision-making).

[201] Angwin et al, *supra* note 8.

[202] Julia Angwin & Jeff Larson, *Bias in Criminal Risk Scores Is Mathematically Inevitable, Researchers Say*, PROPUBLICA (Dec. 30, 2016, 4:44 PM), https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say [https://perma.cc/UBY4-AEJC].

[203] *See* Barocas & Selbst, *supra* note 196, at 715-29 (discussing challenges—both technical and legal—presented when trying to define what constitutes discrimination and what remedies are appropriate); Sam Corbett-Davies et al., *Algorithmic Decision Making and the Cost of Fairness*, 2017 ACM SIGKIDD 797, 799 (citing work of multiple researchers who have pointed out that "many notions of fairness are in conflict" with one another). Moreover, conforming to some definitions of fairness will require accepting lower accuracy rates. Corbett-Davies et al., *supra*.

To complicate things still further, using some definitions of fairness will yield less accurate results than others.[204] Would we rather maximize a model's accuracy or accept higher error rates because the algorithm is "fair," however that is defined? Similarly, the objections that multiple technology experts and civil libertarians raised to the use of algorithms in the new proposed "Extreme Vetting Initiative," found the vagueness and subjectivity of the proposed standard—determining the probability that an applicant would become a positively contributing member of society—problematic because "no computational methods can provide reliable or objective assessments" of those traits.[205] These and decisions like them can only be described as policymaking. For COMPAS, the policymakers were a team of programmers engaged in the for-profit venture that developed COMPAS. Presenting computer models as simply mathematical predictors conceals the policy decision behind the numbers.

Even very concrete policy goals that might have little room for interpretive argument are often difficult to translate into technical requirements. The expertise to build these models is rare in today's government officials. In the long run, society must endeavor to promote computer science literacy, but there is a long way to go before the relevant actors in government and the legal system will have the expertise to design or effectively evaluate predictive algorithms.[206] This means that for the foreseeable future, the translation process is likely flowing through translators that are less than completely fluent in one another's languages—a policy wonk and a computer geek (both terms of endearment).[207] Software code "both describes and causes the computer's behavior," whereas "public policies and laws are characteristically imprecise, often deliberately so."[208] As one programmer put it, "we can make stuff work—it's not our job to

[204] *See* Corbett-Davies et al., *supra* note 203 (explaining that models utilizing differing definitions of algorithmic fairness yield results of varying accuracy).

[205] Biddle, *supra* note 83 (warning also that resulting system would be "inaccurate and biased").

[206] *See* Crawford & Schultz, *supra* note 161, at 127; EXEC. OFFICE OF THE PRESIDENT, *supra* note 53, at 38 (noting that 2014 "assessment of the ability of the public and nonprofit sectors to attract and retain technical talent sounded a strong note of alarm," and encouraged government to "create a more attractive working culture for technologists").

[207] Citron, *supra* note 161, at 1261-62, 1268-69 (demonstrating risks of policy distortion using example of errors written into Colorado's public benefits system).

[208] Kroll et al., *supra* note 7, at 646 (explaining that in order to obtain a guaranteed property it must be built into specifications).

figure out if it's right or not. We often don't know."[209] Forcing the square policy peg into the round software-code hole requires substantive judgments to be made.

Human judgment is also a critical element of selecting both which features to use as inputs and which model among many to actually deploy. In addition to its technical challenges discussed above, feature construction is a highly subjective process whose effectiveness is "absolutely crucial for the success of a machine learning application."[210] If one input is an individual's annual income, for example, how should that information be represented? Should the algorithm use specific numbers? Should it view that feature as binary and divide the data into > x and < x? Or should it divide it into a set of ranges? If the latter, what should those ranges be? Ten thousand dollar increments? Fifty thousand dollar increments? In addition, there may be multiple models consistent with the data, from which the algorithm must know which to choose. There are a variety of bases on which this choice might be made—preference for the simplest model, or the model requiring the least computing power, or the most intelligible model.[211] Each of these determinations will affect the resulting model.

The very question of whether algorithms should seek to minimize false positives or false negatives (or strike a balance between them) is also a policy decision that must be built into the model. The rate of false positives and the rate of false negatives will be inversely proportional. Individual algorithms can therefore minimize either Type I or Type II error, but not both.[212] As the burdens of false positives in law enforcement and national security tend to fall on individuals and false negatives tend to fall on society at large, one might argue that the just policy decision is to prioritize algorithms with fewer false positives. But policymakers, especially those who are concerned with minimizing security risks, might argue that it is better to inconvenience ten men than allow one terrorist to board an airplane. On this view, false negatives are the lesser evil.

---

[209] Crawford & Schultz, *supra* note 161, at 105. *But see* SKILLICORN, *supra* note 33, at 38; KELLEHER, *supra* note 4, at 22 (opining that "analytics practitioner who is situationally fluent will have sufficient knowledge of the quirks of a particular domain to be able to competently build analytics solutions for that domain").

[210] FLACH, *supra* note 22, at 41.

[211] This instruction is called "inductive bias." KELLEHER, *supra* note 4, at 10.

[212] *See* SKILLICORN, *supra* note 33, at 78-79. This is another area where the technology is changing quickly. Using a technique called "ensemble learning," programmers can use training sets that model both types of errors and integrate the outputs of the resulting models later in the process. *See generally* ZHI-HUA ZHOU, ENSEMBLE METHODS, FOUNDATIONS AND ALGORITHMS (2012).

Who should decide what to prioritize or how many false positives are acceptable?

Professor Jeremy Waldron identifies an even more fundamental rule-of-law concern that emerges from the inevitable incorporation of human judgment into algorithmic decisionmaking. On his view, the identification of governing norms is merely one feature of the law. Equally important is the process by which we argue over those norms, and engage in "interpretive exercises about what it means to apply them faithfully."[213] Efforts to delegate this interpretive exercise to algorithms simply move that process of interpretive debate to the design of the algorithm itself.[214]

In each of these contexts, questions that deserve serious democratic, deliberative consideration are instead decided by law enforcement and national security officials or the vendors who supply them with analytic products. This is a troubling short-circuit of the democratic process around important policy debates. What is perhaps worse, however, is that embedding these normative and interpretive choices in the algorithms themselves obscures not only the policy outcomes that were chosen, but the very fact that a choice was made. It thus not only excludes robust public debate on significant policy questions, but also presents the process of resolving those questions as a mathematical calculation. As noted above, the possibility of controlling for improper bias makes the human factor a potentially valuable means of furthering the fundamental value of constraining arbitrary government action. And to the extent that bias reduction is evident, it will also yield legitimacy benefits. Nevertheless, the other implications of the hidden human factor are similar to those stemming from predictive models' opacity. It therefore undermines citizens' ability to plan and is less likely to be perceived as legitimate.

To summarize the preceding discussion, the rule of law requires limitations on arbitrary action by government officials in order to ensure our ability to plan our lives according to our own choices, to preserve the legitimacy of the government, and to match the burdens of government intervention to citizens' choices—to only incur consequences for things that we have actually chosen to do. Machine learning inevitably conflicts with these demands.

III. IMPLICATIONS

This Part will consider the implications of the foregoing discussion for the government's use of machine-learning predictions. This Part concludes—perhaps counterintuitively—that when a government official enjoys a high level of discretion in making decisions, machine learning can be a valuable tool.

---

[213] Waldron, *supra* note 138, at 22.

[214] *See* Mireille Hildebrandt, *Law as Computation in the Era of Artificial Legal Intelligence*, 68 U. TORONTO L.J. 12, 21-23 (2018).

Where the law already imposes significant constraints on discretion through other means, however, the potential value of machine learning is outweighed by its costs.

Part III begins with the assertion that the more significant the consequences of a government action, the greater the harm inflicted if that action is lawless, and thus the more important rule-of-law constraints become. An arbitrary decision to stop and frisk someone on the street, for example, is problematic from a rule-of-law perspective, but not as problematic as intercepting their electronic communications. The necessary robustness of the rule-of-law values—predictability, non-arbitrariness, legitimacy—thus rises and falls with the severity of the consequences.

As it turns out, this principle is often reflected in existing law through the level of discretion with which government officials are permitted to act. As a rule, the higher the costs of false positives to individuals, the less discretion is afforded to the government. To stop and frisk someone, a single police officer must merely have reasonable suspicion of criminal activity, but the rigorous process of securing a surveillance order from the Foreign Intelligence Surveillance Court ("FISC")[215] must be satisfied in order to collect an American's electronic communications for foreign intelligence purposes.[216] As a result, the importance of the rule of law correlates inversely with the level of discretion afforded to government decision-makers—the more important the rule of law, the less discretion government officials enjoy.

Section A draws examples from existing legal doctrine to illustrate that, in the context of law enforcement and national security, the wide array of legal rules—constitutional, statutory, and regulatory—that govern decisionmaking fall along a spectrum, where one end represents highly discretionary decisions and the other end represents highly constrained decisions. This Article refers to this

[215] Contrary to recent reports questioning legitimacy of certain Foreign Intelligence Surveillance Act ("FISA") Court orders, see MAJORITY STAFF OF H. PERMANENT SELECT COMM. ON INTELLIGENCE, 115TH CONG., MEMO ON FOREIGN INTELLIGENCE SURVEILLANCE ACT ABUSES AT THE DEPARTMENT OF JUSTICE AND THE FEDERAL BUREAU OF INVESTIGATION 3-4 (Jan. 18, 2018), the FISA application process is a demanding one that DOJ officials take very seriously. *See* Asha Rangappa, *It Ain't Easy Getting a FISA Warrant: I Was an FBI Agent and Should Know*, JUST SEC. (Mar. 6, 2017), https://www.justsecurity.org/38422/aint-easy-fisa-warrant-fbi-agent/ [https://perma.cc/W5KF-XLZ6] (discussing steps necessary for FBI to get warrant from FISC).

[216] *See* Rangappa, *supra* note 215 (noting that FISA permits FBI to obtain warrant from FISC by showing only "that the target *might be* spying for a foreign government or organization").

spectrum as "the discretion continuum.".[217] Where decisions fall along the discretion continuum reflects, in most instances, the severity of the decision's consequences.[218] And while reasonable people can disagree whether the law currently places each type of government decision in exactly the right place on the spectrum, this Article argues that the general framework it establishes makes sense. Section B will explain why it is appropriate to use predictive analytics in high-discretion contexts. Section C will then argue that when it comes to decisions already constrained by objective factors, such as evidentiary burdens or due process requirements, reliance on machine-learning predictions is inappropriate.

A. *The Discretion Continuum*

There are several types of national security and law enforcement decisions where government officials exercise broad discretion. First, there are decisions based on geography, rather than on individual behavior. This includes determinations such as whether to increase police presence or investigative attention on a certain block or in a certain neighborhood. A computer model might indicate that criminal activity is likely to take place at a particular location,[219] or that when it is raining on Fridays after dark, the intersection of State and Main is likely to be the site of traffic accidents. In tests run in conjunction with the LAPD, predictive models outperformed humans when it came to predicting locations of future crimes.[220] Other cities have had more mixed results.[221]

---

[217] *See infra* Figure 1 (depicting discretion continuum).

[218] Note that whether any particular decision qualifies as a high- or low-discretion decision accounts for all constraints on discretion. In contexts where government actors have significant amounts of discretion—prosecutorial decisions and military targeting decisions, for example—but that discretion is cabined by the overarching legal regime, such as the totality of criminal procedural protections and the international laws of war, it qualifies as low-discretion in my taxonomy.

[219] *See supra* notes 62-66 and accompanying text (discussing police use of predictive analytics and data mining in solving crime).

[220] Stuart Wolpert, *Predictive Policing Substantially Reduces Crime in Los Angeles During Months Long Test*, UCLA NEWSROOM (Oct. 7, 2015), http://newsroom.ucla.edu/releases/predictive-policing-substantially-reduces-crime-in-los-angeles-during-months-long-test [https://perma.cc/4E7F-UHH2] (noting that model predicted twice as much crime as analysts were able to predict).

[221] Chammah, *supra* note 54 (noting that in some cities like Lincoln, Nebraska, police officers "have found [predictive analytics software] mostly tells them what they already know").

Investigative activity that does not qualify as a search or seizure under *Katz v. United States*,[222] and is therefore not regulated by the Fourth Amendment,[223] represents another high-discretion context. This is the case, for example, when the police decide to place an individual under physical surveillance in public spaces.[224] A third potentially high-discretion scenario is when the Fourth Amendment requires only that the government act "reasonably," such as when the government executes a search or seizure that qualifies for an exception to the warrant requirement.[225] While at times, reasonableness limits discretion by imposing some level of individualized suspicion, this is not always the case. When CBP officials decide to search one suitcase and not another at the border, for example, there is no individualized suspicion requirement.[226] Also in this category are some searches authorized under the so-called "special needs" doctrine. When a search or seizure is undertaken for reasons beyond routine law enforcement—for a "special need"—those searches and seizures do not require a warrant and, in some cases, also lack individualized cause requirements. The NYPD's subway passenger searches and (arguably) the electronic surveillance of non-U.S. persons located overseas are examples.[227]

[222] 389 U.S. 347 (1967).

[223] *See id.* at 352-53 (finding that use of recording device in private telephone booth constituted search and seizure under Fourth Amendment despite lack of any physical trespass).

[224] United States v. Knotts, 460 U.S. 276, 277-78, 281 (1983) (finding no search or seizure when police tracked defendant's travels using beeper transmitter placed in chloroform container purchased by co-defendant, because "governmental surveillance conducted by means of the beeper . . . amounted principally to the following of an automobile on public streets and highways"); Joh, *supra* note 7, at 21 (noting that "suspects can emerge from the data for purposes of investigation" and that they "can appear even if police do not seek a particular person for a particular crime").

[225] *See, e.g.*, Mich. Dept. of State Police v. Sitz, 496 U.S. 444, 450-53 (1990) (requiring that government act reasonably in "special needs" beyond law enforcement); United States v. Martinez-Fuerte, 428 U.S. 543, 565-67 (1976) (requiring that government act reasonably in border searches).

[226] *E.g.*, United States v. Cotterman, 709 F.3d 952, 960-61 (9th Cir. 2013) (en banc) (noting that routine searches of containers and electronics crossing border do not require individualized suspicion).

[227] *See, e.g.*, MacWade v. Kelley, 460 F.3d 260, 271 (2d Cir. 2006) ("Where . . . a search program is designed and implemented to seek out concealed explosives in order to safeguard a means of mass transportation from terrorist attack, it serves a special need."); In re Directives, 551 F.3d 1004, 1011-12 (FISA Ct. Rev. 2008) (finding exception to "warrant requirement for surveillance undertaken for national security purposes and directed at a foreign power" by "[a]pplying principles derived from the special needs cases"). Some special

In each of these situations, law enforcement and national security personnel enjoy broad discretion as a matter of law. In the first two situations, the government enjoys essentially complete discretion. It may make decisions based on any criteria (or on no criteria at all), save those that are barred by constitutional provisions such as the Equal Protection Clause.[228] In the third, the reasonableness requirement imposes some limits on the government, but they are not individualized, cause-based limits. Statutes or regulations might impose additional requirements, but the standards under such rules tend to be extremely forgiving—that the information the government seeks is "relevant" to an investigation, for example[229]—and therefore reduce discretion only minimally.

The reason the government enjoys so much discretion here is that rule-of-law constraints do relatively little work in contexts where false positives do not impose heavy costs on individuals. Recall that the presence or absence of what is considered "the rule of law"—and hence the stringency with which its principles must be respected—is not a binary state, but rather can exist in more or less robust forms.[230] We are not entitled to be able to predict, for example, the location of police patrols at any given moment. A determination where they will go does not interfere with anyone's ability to anticipate the operation of the law and carry out their (lawful) plans. Thus, the seemingly arbitrary presence of law enforcement officials does not usually pose a threat to the legitimacy of the legal regime. To be sure, over-policing of particular communities does, in fact,

needs programs require individualized suspicion, while some do not. *See* Eve Brensike Primus, *Disentangling Administrative Searches*, 111 COLUM. L. REV. 254, 278-79 (2011) (noting that because Court has established standard for administrative searches that does not require individualized suspicion, it "has begun the process of removing the individualized suspicion requirement entirely, even in the context of searches within special subpopulations with reduced expectations of privacy").

[228] Government action is always limited by constitutional constraints, such as equal protection and the First Amendment. *See* Hassan v. City of New York, 804 F.3d 277, 295 (3d Cir. 2015) ("A 'claim of selective investigation' by the police draws on 'ordinary equal protection standards." (quoting Flowers v. City of Minneapolis, 558 F.3d 794, 798 (8th Cir. 2009))); United States v. Scopo, 19 F.3d 777, 786 (2d Cir. 1994) (Newman, C.J., concurring) ("Though the Fourth Amendment permits a pretext arrest, if otherwise supported by probable cause, the Equal Protection Clause still imposes restraint on impermissibly class-based discriminations."); Joh, *supra* note 7, at 34 ("[S]urveillance discretion cannot be 'exercised in a discriminatory fashion.'").

[229] *See, e.g.*, 50 U.S.C. § 1861 (2012) (requiring any application for document production in intelligence investigation to include "statement of facts showing that there are reasonable grounds to believe the tangible things sought are relevant to an authorized investigation").

[230] *See* discussion *supra* Section II.B (noting that question of whether rule of law exists is not binary, but rather exists on spectrum).

undermine legitimacy.[231] But, as discussed below, this provides an argument in favor of predictive analytics, not against them. So, the legal framework affording significant discretion to certain government decisions reflects an assessment that government legitimacy is not necessarily undermined by imposing relatively lax versions of the rule-of-law principles.

One reason that rule-of-law concerns are lowest in high-discretion decisions is that the consequences imposed on citizens tend to be relatively minimal as well—i.e., the costs of false positives are low. To be sure, erroneous decisions may generate inconvenience but no catastrophic results. Individuals' bags might be searched, or one might find oneself in a neighborhood with a large police presence, but the government cannot read your emails or search your home.

As the costs of false positives increase and government action becomes more targeted at specific individuals, decisions move toward the other end of the discretion continuum. Similarly, the necessary enforcement level for rule-of-law principles rises. In the criminal investigation context, the mechanism for limiting government discretion is often the Fourth Amendment, which dictates the degree of individualized evidence the desired government action requires.[232] When executing a search or making an arrest, for example, the government must have probable cause indicating that the search or arrest is justified.[233] The less consequential the government action, the less justification is demanded—so-called *Terry* stops, for example, require only reasonable suspicion of criminal activity.[234] In other areas, these stricter limits might come from statutes or regulations. Such rules indicate that the issue is sufficiently sensitive that we as a society have determined we need limitations on the government's ability to intervene. In other words, to be considered legitimate, government decisions must reflect more transparency and increased levels of constraint on arbitrariness as the stakes of those decisions rise. Electronic surveillance of Americans' phone

[231] *See supra* Section II.B.1.c (discussing role of government legitimacy in rule-of-law context).

[232] U.S. CONST. amend. IV (barring "unreasonable searches" of people's "persons, houses, papers, and effects"). There are, of course, many other variables at work as well. Outside the criminal procedure context, for example, decisions are subject to Fifth Amendment due process limits. *Id.* amend. V ("[N]o person shall . . . be deprived of life, liberty, or property, without due process of law . . . .").

[233] WAYNE R. LAFAVE, SEARCH AND SEIZURE: A TREATISE ON THE FOURTH AMENDMENT § 3.1 (5th ed. 2012) (explaining that under the Fourth Amendment the police may not make an arrest or search unless they have probable cause to do so).

[234] Terry v. Ohio, 392 U.S. 1, 30 (1968) (holding that Fourth Amendment permits police officer to stop and frisk suspect on street when officer has "reasonable suspicion that criminal activity is afoot").

calls, medical records, and emails are examples of statutorily protected categories.[235]

Outside the investigative context, the Due Process Clause of the Fourteenth Amendment,[236] the guarantees embodied in the Fifth and Sixth Amendments,[237] and the Administrative Procedure Act[238] play the same role. And again, as the intensity of the individual interest grows, more process is due.[239] After an individual is added to the No-Fly List, restricting her constitutional right to travel, due process guarantees the listee is provided "notice regarding their status on the No–Fly List and the reasons for placement on that List . . . reasonably calculated to permit" the listee to contest the designation with relevant evidence.[240] When the constitutional right at stake is an individual's liberty, however—in the pre-trial detention context, for example—due process requires the government to prove "by clear and convincing evidence that an arrestee presents an identified and articulable threat to an individual or the community" before making a bail determination.[241] And the ultimate consequence—long-term deprivation of liberty or even deprivation of life—of course requires procedures in which the government proves all elements of a crime beyond a reasonable doubt.[242]

---

[235] Stored Wire and Electronic Communications and Transactional Records Access Act, 18 U.S.C. §§ 2701-2712 (2012) (limiting ability to access and disclose communications); Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936, 1991-2037 (1996) (imposing regulations designed to prevent health care fraud and abuse); Foreign Intelligence Surveillance Act of 1978, Pub. L. No. 95-511, 92 Stat. 1783 (1978) (providing regulation and limitation to government's ability to electronically surveil foreign powers).

[236] *See* U.S. CONST. amend. XIV ("[N]or shall any State deprive any person of life, liberty, or property, without due process of law . . . .").

[237] *See id.* amends. V, VI.

[238] *See* Administrative Procedure Act § 1, 5 U.S.C. § 551 (2012).

[239] *See* Matthews v. Eldridge, 424 U.S. 319, 348 (1976) (determining amount of process due by weighing interests of individual, risk of error based on procedures used, costs of additional process, and interests of government).

[240] Latif v. Holder, 28 F. Supp. 3d 1134, 1162 (D. Or. 2014).

[241] United States v. Salerno, 481 U.S. 739, 751 (1987) (upholding Federal Bail Reform Act of 1984 against due process challenge by emphasizing Act's requirement of procedures specifically designed to further accuracy of judicial officer's determination of dangerousness).

[242] *See, e.g.*, Hurst v. Florida, 136 S. Ct. 616, 621 (2016) (explaining that Sixth Amendment, "in conjunction with the Due Process Clause, requires that each element of a crime be proved to a jury beyond a reasonable doubt").

Figure 1. The Discretion Continuum

B. *High-Discretion Decisions and Machine Learning*

Perhaps counterintuitively, this Section argues that the more discretion current law allows government decision-makers, the more appropriate use of machine-learning predictions becomes. One might assume that if rule of law is designed to constrain government discretion, and machine learning is inconsistent with values underlying our commitment to the rule of law, then use of machine learning would be *least* justified in contexts where government discretion is at its height.

To the contrary, the reason the government enjoys so much discretion in these contexts is that its decisions pose less of a threat to the rule of law.[243] The mere presence of law enforcement agents in a public place, for example, neither renders the operation of the law a secret nor subjects citizens to the State's coercive power.[244] Moreover, the appearance of law enforcement officials does not pose a threat to the legitimacy of the legal regime.[245] Once that police force significantly interferes with a citizen's liberty or engages in unpredictable or arbitrary action, however, legitimacy concerns intensify.[246] Thus, in high-discretion decisions, where the stakes are low and unpredictability and arbitrariness pose less of a threat to government legitimacy, the rule-of-law shortcomings of machine learning are less problematic.

[243] *Cf.* Maryland v. King, 133 S. Ct. 1958, 1969 (2013) (noting that need for warrant is lowest when search involves no discretion and therefore no facts for neutral magistrate to evaluate).

[244] To be sure, some will see the mere presence of government agents as a deployment of coercive force that chills individual rights. By coercion here I refer to physical coercion.

[245] *See supra* Section II.B.2 (discussing rule-of-law framework and government legitimacy).

[246] *See supra* Section II.B.2.c (noting that predictability of operation of law is essential to rule-of-law legitimacy).

At the same time, predictive models can improve the quality of high-discretion decisions in at least two ways. Predictive models promise to reduce the impact of improper bias by channeling that discretion.[247] Predictive models encourage data-driven decisions, rather than decisions based on hunches at best and discriminatory motives at worst. It is when decisions are broadly discretionary that concerns about improper motives, conscious and unconscious bias, or unjustified privacy intrusions are at their height.[248] Indeed, one of the Fourth Amendment's primary purposes "is to impose a standard of 'reasonableness' upon the exercise of discretion by government officials . . . in order 'to safeguard the privacy and security of individuals against arbitrary invasions.'"[249] These concerns exist not only in the context of Fourth Amendment regulated action but in all official decisionmaking. When neither the Constitution nor any statute imposes objective constraints on government action, however, there is significant risk (and a history) of discriminatory

[247] Bambauer, *supra* note 100, at 244 (arguing that "[p]attern-driven data mining of third-party records can lead to fairer enforcement of our criminal laws" in part by "reducing the opportunities for human bias to infect decision-making"); Simmons, *supra* note 7, at 961 (discussing pervasiveness of inaccurate assumptions police use in trying to identify or predict criminal activity); Zarsky, *supra* note 7, at 1516 (pointing out that data mining reduces opportunities for human bias to infect decision-making and can lead to more equitable law enforcement by ferreting out different sorts of crimes, like white collar crimes of insider trading and money laundering).

[248] *See* Barry Friedman & Maria Ponomarenko, *Democratic Policing*, 90 N.Y.U. L. REV. 1827, 1858-65 (2015) (discussing history of rulemaking surrounding police operations and discretion, and noting studies in 1950s and 1960s found that "real 'rules' of policing were made 'bottom-up' in an ad hoc manner through hundreds of individual decisions made by cops on the beat").

[249] Delaware v. Prouse, 440 U.S. 648, 653-54 (1979) (footnote omitted) (quoting Marshall v. Barlow's Inc., 436 U.S. 307, 312 (1978)); *see* Skinner v. Railway Labor Execs.' Ass'n, 489 U.S. 602, 613-14 (1989) (explaining Fourth Amendment "guarantees the privacy, dignity, and security of persons against certain arbitrary and invasive acts by officers of the Government or those acting at their direction"); Camara v. Mun. Court of S.F., 387 U.S. 523, 528 (1967) ("The basic purpose of [Fourth Amendment], as recognized in countless decisions of this Court, is to safeguard the privacy and security of individuals against arbitrary invasions by governmental officials."); Tracey Maclin, *The* Pringle *Case's New Notion of Probable Cause: An Assault on Di Re and the Fourth Amendment*, 2003-2004 CATO SUP. CT. REV. 395, 411 (2003-2004) (summarizing history of Fourth Amendment as reflection of "Framers' desire to control the discretion of ordinary law enforcement officers and to eliminate governmental intrusions lacking particularized suspicion").

decision-making.[250] Deploying computer models in these circumstances can lead to fairer, more balanced policing.

The risk of improper discrimination in the absence of individualized suspicion requirements explains why courts have insisted that special-needs programs lacking such a requirement include policy-based limits to officials' discretion. Police officers at a drunk-driving checkpoint, for example, might be obligated to stop every tenth car, rather than allowing individual officers to select the targets of their searches. And in approving New York City's suspicionless searches of subway passengers as a valid special-needs program, the court pointed to the fixed policy by which the NYPD determined where it would set up check points on a day-to-day basis, rather than leaving those decisions to individual officers carrying out the searches themselves.[251] In the absence of either a warrant or a cause requirement, this limitation on discretion has emerged as the primary safeguard against abuse.[252] Reliance on (accurate, unbiased) computer models would similarly channel government discretion. Note that to the extent one believes that existing constitutional doctrine or statutory regimes are under-protective of individual rights, one can argue that the relevant decisions should fall further to the left along the spectrum than exiting law places them.[253]

One could argue that in the absence of constitutional or statutory limits on discretion, political forces might operate to impose extra-legal constraints. As a result, these "high-discretion" decisions might not be so high-discretion after all. The use of suspicionless subway searches or the location of police patrols, for example, might seem sufficiently public and politically salient to generate political checks. Unfortunately, such checks are susceptible to the same

[250] *See* Friedman & Ponomarenko, *supra* note 248, at 1858-65 (discussing history of rulemaking and lack of legislation regarding police discretion).

[251] *See* MacWade v. Kelley, 460 F.3d 260, 273 (2d Cir. 2006) (finding program minimally invasive in part on grounds that "police exercise no discretion in selecting whom to search, but rather employ a formula that ensures they do not arbitrarily exercise their authority").

[252] *See* Friedman & Ponomarenko, *supra* note 248, at 1878 (noting that while policing necessarily involves discretion, "much of modern policing—from the use of checkpoints and administrative inspections, to reliance on new technologies like drones . . . –is entirely amenable to . . . regulation"); Primus, *supra* note 227, at 263-70 (using dragnet search example of "checkpoints where government officials stop every car (or every third car) driving on a particular roadway" and discussing various limitations on exercise of discretion in administrative searches).

[253] *See supra* Figure 1 (depicting discretion continuum). More examples of high-discretion decisions include No-Fly List nominations, suspicionless border searches, and the aggregation of location information about one person over time.

(conscious and unconscious) biases that render official decision-making potentially problematic. Thus, the potential benefits of machine-learning tools can serve as a corrective for bias driving the political winds as well.

Effective algorithmic predictions could also yield significant efficiency benefits. Because any algorithm actually employed—even algorithms with accuracy rates far below our hypothetical ninety-nine percent threshold—should more accurately identify likely crimes and threats than humans, following its recommendations will result in more effective use of limited investigative resources. Currently, because high-discretion decisions need not be motivated by specific facts,[254] it is difficult to discern whether they represent efficient resource allocation. Data-driven decisions promise less time spent following fruitless investigative paths, thereby reducing the opportunity costs of such activity. So, if a predictive model identifies particular behaviors as predictors of criminal activity or security threat—certain patterns of driving, purchasing, or movement, for example—police might allow that model to guide their discretionary decision-making, leading to increased attention towards those places, individuals, or groups who fit the pattern.[255] As an example, if the rules governing NYPD's subway search program came from a predictive model's suggestion, it might suggest establishing checkpoints at the West 4th Street subway station on the second Friday of every month between March and October, rather than on some more arbitrary, human-derived rule. This might render the program more effective while still avoiding the evils of executive discretion.[256]

While the benefits of channeling executive discretion and more efficiently allocating resources are significant, the costs of predictive models in the context of high-discretion decision-making are relatively low. The limited applicability of the Fourth Amendment suggests that privacy concerns are arguably less acute than when the Constitution demands a warrant or individualized cause. Moreover, the costs to individuals of false positives—increased police presence in a particular area, more scrutiny of certain individuals while they are in public, additional screening at the airport or on the subway—will be relatively low. This is especially true if predictive models are crafted to result in equitable

[254] *See supra* Section III.A (noting that certain high-discretion decisions do not require any individualized suspicion).

[255] *See* Joh, *supra* note 7, at 21 (arguing that if correlation between behaviors and crime or threat is high enough, law enforcement need not know why correlation exists before relying on it).

[256] The argument assumes that the data and the algorithms employed are themselves free from improper influence, such as racial bias.

distribution of the burden of false positives, rather than letting them fall disproportionately on certain populations.[257]

Machine-learning programs applied to discretionary decision-making retain their opacity, potentially arbitrary outcomes, and inherent human bias. So long as care is taken to prevent improper or discriminatory biases from being built into the algorithm, however, operating based on predictive analytics promises to be less arbitrary than government officials exercising discretion.[258] It is therefore here, where rule-of-law demands are least stringent, that machine-learning tools can be most effective without running afoul of fundamental values.

C. *Low-Discretion Decisions and Machine Learning*

The calculus changes when addressing low-discretion decision-making. As the consequences of government action become more severe, government officials' discretion tends to fall. Often, limits on discretion exist to protect a constitutional right, such as the privacy right protected by the Fourth Amendment[259] or due process rights.[260] At other times, the requirement is imposed by statute or regulation.[261] Whatever their source, limits on discretion indicate that protected individual rights or entitlements are on the line. Rather

[257] *See supra* text accompanying note 196 (discussing potential high-cost ramifications of false positives and noting need for fair predictive models). Of course, errors impact not only the individual, but also society as a whole. Every false positive drains resources that might otherwise have borne fruit elsewhere, and every false negative allows someone with malicious intent to go undetected. But society already suffers from these errors, and if computer models can reduce the number of errors and deliver on the promise of curbing discriminatory decision-making while simultaneously making it more efficient, the net cost to society will be less than it is without algorithms.

[258] In some instances, these choices will be relatively easy, such as refraining from using constitutionally protected categories—race, gender, ethnicity, religion—and their proxies for input features is an example. Others, as the COMPAS program demonstrates, will be much more difficult. *See supra* notes 201-05 and accompanying text (discussing programmers' decision to reduce racial bias in COMPAS program by achieving equal accuracy rates, which yields unequal false positives rates). We must recognize that some of these choices are policy choices that merit public debate and discussion—they are not mathematical calculations. No algorithm is going to be perfect or eliminate all forms of bias. But so long as it is designed to act more objectively than humans—who are also unable to shed bias entirely—it is an improvement.

[259] *See* Brennan-Marquez, *supra* note 7, at 1255 (arguing that forcing police to articulate theories of wrongdoing and suspicion are means by which courts enforce Fourth Amendment).

[260] *See* U.S. CONST. amends. V, XIV (providing each U.S. citizen with right to due process).

[261] *See, e.g.*, 5 U.S.C. § 706 (2012) (instructing reviewing courts to hold unlawful agency action that is arbitrary and capricious or abuse of discretion).

than simply directing government surveillance to a particular place, for example, it might be an invasive search of your home, arrest, denial of parole, the right to board an airplane, or collection of your electronic communications. Discretion-channeling is even more important in this context in order to minimize the instances in which someone is wrongly singled out. Rather than relying on machine-learning predictions to channel discretion here, however, we can rely on existing law, which already reflects the need for more accurate decision-making as the stakes of those decisions increase. This is why the higher the consequences of false positives, the higher the standard the government must meet to justify its action. So, in low-discretion decisions, government action is already subject to constraints using mechanisms that do not conflict with rule-of-law values. These rules also increase efficiency, because the government cannot take action absent objective evidence that it is justified. As a result, any discretion-channeling or efficiency gains that machine learning can offer in this context are less significant, while at the same time, the costs of the conflict with rule-of-law values are higher.

1. Limited Added Value of Machine Learning in Low-Discretion Decisions

Machine learning has less value to offer in low-discretion contexts than it does in high-discretion contexts because those decisions already are governed by mechanisms designed to channel official discretion. One such mechanism is the requirement of individualized suspicion. Under the Fourth Amendment, the government must usually have individualized suspicion before invading individual privacy rights. Whether the necessary standard is reasonable

suspicion[262] or probable cause,[263] robust correlations are insufficient; the Fourth Amendment requires the government to have some individualized justification for its actions. To qualify as "individualized," courts consistently have rejected mere statistical likelihood.[264] Instead, "the state should judge each citizen based upon his own unique character, thoughts, and situation" rather than on

[262] Reasonable suspicion exists so long as there is a "moderate chance" that criminal conduct has occurred, is occurring, or will occur. Safford Unified Sch. Dist. #1 v. Redding, 557 U.S. 364, 371 (2009) (noting standard for law enforcement officer's evidence search at schools could "readily be described as a moderate chance of finding evidence of wrongdoing"); *see* United States v. Sokolow, 490 U.S. 1, 7 (1989) (finding that whether DEA agents were justified in making stop depended on "whether the agents had a reasonable suspicion that respondent was engaged in wrongdoing when they encountered him on the sidewalk"). This chance can be established by "a series of acts, each of them perhaps innocent" if seen in isolation, but that "warrant[] further investigation" when viewed in combination. *Sokolow*, 490 U.S. at 9-10 (quoting Terry v. Ohio, 392 U.S. 1, 22 (1968)); *see* Ferguson, *supra* note 7, at 298 (noting in context of using predictive methods to establish reasonable suspicion, "[l]ower courts have upheld arrest warrants on DNA matches and other forensic science matches based on probabilities, but there has never been a Supreme Court case in which the probability of crime explicitly has been used as the sole justification for a stop").

[263] Probable cause exists when there exists "a 'fair probability' or a 'substantial chance' of discovering evidence of criminal activity." *Safford*, 557 U.S. at 371 (quoting Illinois v. Gates, 462 U.S. 213, 238, 244 n.13 (1983) (citations omitted)). It is "a fluid concept—turning on the assessment of probabilities in particular factual contexts—not readily, or even usefully, reduced to a neat set of legal rules." *Gates*, 462 U.S. at 232. While based on the idea of "probability," that term is not used in a statistical sense and courts have been reluctant to quantify the requisite probability. *See id.* at 235 ("[A]n effort to fix some general, numerically precise degree of certainty corresponding to 'probable cause' may not be helpful. . . ."); Brinegar v. United States, 338 U.S. 160, 175 (1949) (noting probabilities dealt with in determining probable cause "are not technical; they are the factual and practical considerations of everyday life on which reasonable and prudent men, not legal technicians, act"). Some scholars, however, argue that the courts should define both reasonable suspicion and probable cause using a specific numerical probability. *See* Sherry F. Colb, *Probabilities in Probable Cause and Beyond: Statistical Versus Concrete Harms*, 73 L. & CONTEMP. PROBS. 69, 100, 105 (2010) (discussing human tendencies to put more weight on actual rather than statistical harm in context of legal decisionmaking); Simmons, *supra* note 7, at 999-1009 (discussing judicial adoption of predictive algorithms and noting difficulties associated with establishing exact percentage to assign to reasonable suspicion or probable cause).

[264] *See* Rich, *supra* note 7, at 896-901 (discussing automated suspicion algorithms and noting that "in most circumstances the Fourth Amendment entitles *each* suspect to an assessment of whether individualized suspicion exists based on *all* available facts relating to her potential guilt," rather than whether "*on average* probable cause existed").

"stereotypes, assumptions, guilt-by-association, or other generalities."[265] To establish individualized suspicion, "[t]here must be something specific to the defendant to create the probability as to him," rather than basing individualized suspicion on probabilities.[266] In each case, there must be "facts that, when combined with the supportable generalizations, establish[] reasonable suspicion to believe that *this suspect*, *at this time*," has misbehaved "*in the specific location*" law enforcement identified.[267] The presence or absence of individualized suspicion thus requires highly fact-specific, case-by-case assessments of the totality of the circumstances in order to determine whether the facts of the situation suffice to meet the applicable standard.[268]

Not only are generalities insufficient, but mere probabilities—even if the probability is fifty-one percent (i.e., it is more likely than not that there is

[265] Taslitz, *supra* note 139, at 146; *see* Bambauer, *supra* note 185, at 507 (arguing that "[r]andomly distributed hassle is preferable to the nonrandom distribution brought about by common police practices"); Simmons, *supra* note 7, at 984-85 (noting that individualized suspicion "cannot be based only on who the person is; it must also be based on what the person does"). One school of thought argues that relying on probabilities and algorithms to generate individualized suspicion is perfectly acceptable, so long as the inconveniences of such a program are distributed relatively randomly. *See* Christopher Slobogin, *Policing and the Cloud*, NAT'L CONSTITUTION CTR. 10 (2017), https://constitutioncenter.org/media/files/sloboginfinal5.pdf [https://perma.cc/D5WM-XYKQ] (discussing issues with using risk factors in algorithms in law enforcement's access of private database information and noting that "[t]ransparent algorithms that can produce the relevant hit rates and that avoid obviously illegitimate variables are very likely to be an improvement" on traditional law enforcement's reliance on "static factors" that are "more subject to invidious manipulation"). On this view, false positives are not problematic so long as the intrusion they represent is proportional to the harm being investigated. Professor Bambauer gives the example that, if we know that sixty percent of Harvard dorm residents use drugs, law enforcement should be able to search sixty percent of Harvard's dorm rooms, selecting those rooms at random. Bambauer, *supra* note 185, at 462-63, 483 (characterizing individualized suspicion requirement as consisting of two parts: suspicion, or "chance that evidence will be discovered in the course of a stop and search" and individualization, or "chance that an innocent person will have to undergo a stop and search").

[266] Arnold H. Loewy, *Rethinking Search and Seizure in a Post-9/11 World*, 80 MISS. L.J. 1507, 1518 (2011).

[267] *See* Taslitz, *supra* note 139, at 149 (making assertion in context of establishing reasonable suspicion to carry out search of high school student).

[268] Dunaway v. New York, 442 U.S. 200, 213-14 (1979) (rejecting "multifactor balancing" approach to probable cause); *see* Taslitz, *supra* note 139, at 169 (noting that because "it is rare that police have nonspurious generalizations upon which to rely" when making decisions, individualized suspicion is especially important step for "*the state to justify* invading individuals' privacy, property, and locomotive rights").

evidence of a crime (probable cause) or that crime is afoot (reasonable suspicion))—also fail to establish individualized suspicion. In *Wong Sun v. United States*,[269] for example, an informant told law enforcement officials that "Blackie Toy," the proprietor of a laundry run on Leavenworth Street, was selling heroin.[270] However, Leavenworth Street was apparently home to more than one Chinese laundry run by someone named "Toy," and the record held no evidence that the agents had any information indicating that the arrested laundry proprietor was, in fact, "Blackie Toy."[271] As a result, there was no probable cause to arrest any one of them, because there was no showing that the officers "had some information of some kind which had narrowed the scope of their search to this particular Toy."[272]

Ybarra v. Illinois[273] provides an example from a situation requiring only reasonable suspicion.[274] There, police with a valid warrant to search a tavern for drugs also stopped and frisked every individual inside the tavern, assuming their actions were justified by these individuals' presence in an establishment suspected of selling drugs.[275] The Supreme Court rejected this argument, holding that officers must have reasonable suspicion "directed at the person to be frisked."[276] Thus, individualized cause does not exist until a single person is identified based on a combination of factors unique to them.[277] What the

[269] 371 U.S. 471 (1963).

[270] *Id.* at 473-75 (noting that federal narcotics agents pursued petitioner following lead that owner, known as "Blackie Toy," of laundromat on certain street was dealing heroin, despite fact that there was "nothing in the record which identifie[d] [petitioner] and 'Blackie Toy' as the same person").

[271] *Id.* at 480-81 (finding that informant's "accusation merely invited the officers to roam the length of Leavenworth Street . . . in search of one 'Blackie Toy's' laundry").

[272] *Id.* at 481; *see* Mallory v. United States, 354 U.S. 449, 456 (1957) (stating police must have probable cause at time of arrest, rather than "arrest, as it were, at large and to use an interrogating process at police headquarters in order to determine whom they should charge before a committing magistrate on 'probable cause'").

[273] 444 U.S. 85 (1979).

[274] *Id.* at 92-93 (finding that officers' first frisk of appellant "was simply not supported by a reasonable belief that he was armed and presently dangerous, a belief which [the] Court has invariably held must form the predicate to a patdown of a person for weapons").

[275] *See id.* at 88.

[276] *Id.* at 94.

[277] *See* Colb, *supra* note 263, at 105 (discussing scenario in which "an officer confronts [a] two-people-one-innocent situation" and "select[s] one of the two people" to arrest and noting that this would yield "typical probabilities case, acceptable to everyone, even though

government must show is thus articulated in the language of case-by-case decision-making, rather than pattern-based predictions.

To be sure, minimal standards of individualized suspicion, such as reasonable suspicion, are relatively easy for the government to satisfy and therefore might not impose a particularly stringent constraint on discretion. Even assuming agreement with the premise that machine-learning predictions are inappropriate for low-discretion decision-making, one could argue whether an action requiring only reasonable suspicion is a sufficient constraint to treat it as a low-discretion decision. In other words, low-discretion does not mean no-discretion. And reasonable minds can disagree with respect to where on the continuum the costs of machine learning begin to outweigh their benefits. Some may argue that actions constrained only by a reasonable-suspicion requirement should be treated the same as actions not subject to any constitutional constraints at all, while others would treat such actions the same as actions governed by probable cause.[278]

The Fifth and Fourteenth Amendments' due process clauses serve a similar discretion-reduction function outside the criminal investigation context whenever a liberty or property interest is at stake.[279] What process is due will vary depending on the relative strength of the government's and the individuals' interests,[280] but the *purpose* of due process is to ensure that the government reaches accurate—i.e., individualized, non-arbitrary—decisions to the extent possible.[281] Indeed, inquiry into what process is due in any given situation includes an assessment of the extent to which additional procedures would

there would be no reason to distinguish between the person [the officer] is arresting and the person [the officer] is not arresting").

[278] This Article does not attempt to determine definitively when machine learning should be used and when it should not. Rather, it aims to establish the level of government discretion as the variable most relevant to making that determination.

[279] *See* Matthews v. Eldridge, 424 U.S. 319, 332 (1976) ("Procedural due process imposes constraints on governmental decisions which deprive individuals of 'liberty' or 'property' interests within the meaning of the Due Process Clause of the Fifth or Fourteenth Amendment.").

[280] *See id.* at 334 ("[R]esolution of the issue whether the administrative procedures provided here are constitutionally sufficient requires analysis of the governmental and private interests that are affected.").

[281] *Id.* at 348-49 (finding that to meet due process requirements "[a]ll that is necessary is that the procedures be tailored, in light of the decision to be made, to 'the capacities and circumstances of those who are to be heard,' to insure that they are given a meaningful opportunity to present their case" (quoting Goldberg v. Kelly, 397 U.S. 254, 268-69 (1970))).

reduce the risk of error.[282] Process requirements usually include notice of and justification for the government's actions as well as a right to challenge the determination, either before or after that determination is final.[283] Even when constitutional rights are not at stake, the Administrative Procedure Act bars government agencies from making "arbitrary and capricious" decisions.[284] Here too arbitrary deprivations of individual rights are impermissible.

Throughout these rules, whether they establish a substantive standard like probable cause or a procedural limit like the need for a hearing, one of the primary enforcement mechanisms—i.e., the method with which we ensure that the government adheres to the rules—is the reason-giving requirement. Individuals who, for example, are subject to arrest are entitled to know the basis of their detention—perhaps not at the moment of arrest, but shortly thereafter at the latest.[285] State officials must be able to come before a judge and provide reasons, demonstrating that there were facts that added up to probable cause to arrest.[286]

In contexts where decision-making discretion is limited, the benefits of machine learning—increased efficiency and the ability to control for at least some aspects of the human factor—are less pronounced. When government officials are constrained by the need to point to objective evidence and make their case before neutral decision-makers, those mechanisms already reduce discretion by forcing the government to present an evidence-based justification for its actions. In other words, we rely on things like individualized suspicion

[282] *Id.* at 343 ("An additional factor to be considered . . . is the fairness and reliability of the existing . . . procedures, and the probable value, if any, of additional procedural safeguards.").

[283] *See id.* at 346 (discussing procedural safeguards against risk of mistaken decision in context of disability benefits appeals and finding that potential disability recipient must have opportunity "to challenge directly the accuracy of information in his file as well as the correctness of the agency's tentative conclusions").

[284] *See, e.g.*, 5 U.S.C. § 706 (2012) ("The reviewing court shall . . . hold unlawful and set aside agency action, findings, and conclusions found to be arbitrary, capricious, or otherwise not in accordance with law."). Litigants will often challenge government action under both the Administrative Procedure Act and the Due Process Clause. *See, e.g.*, Villarreal v. Horn, 203 F. Supp. 3d 765, 773 (S.D. Tex. 2016) ("A procedural Due Process challenge to a federal regulation may be brought under the APA.").

[285] *See* Cty. of Riverside v. McLaughlin, 500 U.S. 44, 55-57 (1991) (requiring prompt judicial determination of probable cause after arrest).

[286] Papachristou v. City of Jacksonville, 405 U.S. 156, 169 (1972) ("We allow our police to make arrests only on 'probable cause,' a Fourth and Fourteenth Amendment standard applicable to the States as well as to the Federal Government.").

and administrative or judicial process to channel official decision-making rather than predictive tools. And with respect to efficiency, these existing limits on discretion are designed to ensure that government officials expend limited resources only when the facts call for it—when there has been some concrete indication that a particular individual, organization, or circumstance merits further attention or investigation. We have thus already found ways to avail ourselves of the benefits that machine learning otherwise might provide.

Of course, as noted above, there will be instances in which this system fails to play out according to its ideal. As noted above, police officers are not always sincere in the explanations they provide for *Terry* stops.[287] Legal opinions authorizing electronic surveillance pursuant to the FISA usually remain classified, and the subject of the surveillance is rarely offered the opportunity to challenge the validity of the surveillance order.[288] Nevertheless, police officers must offer a valid reason if the constitutionality of a *Terry* stop is challenged.[289] Even if that reason was not the true motivation, it places concrete constraints on what qualifies as lawful activity. And even FISC orders can be challenged if their fruits are used in a criminal prosecution.[290] Moreover even imperfect instantiations of these human-reason-based limits are less costly from a rule-of-law perspective than much more accurate predictive analytics would be.

2. Undermining the Rule of Law

Low-discretion decisions pose more of a potential threat to the rule of law because of the important interests the limits on discretion are designed to protect. In circumstances with such high stakes, it is critical that government action is more predictable, less arbitrary, and more likely to be accepted as legitimate. Discretionary constraints like individualized suspicion and due process are the means by which the legal system serves to preserve these rule-of-law values. In other words, conducting searches without probable cause would conflict with

[287] *See supra* Section III.A (noting that *Terry* stops require only law enforcement's reasonable suspicion of unlawful activity).

[288] William C. Banks, *The Death of FISA*, 91 MINN. L. REV. 1209, 1230-31 (2007) (describing FISC procedures and noting that "target may eventually learn of the FISA targeting only if the FISA surveillance is used by the government in a criminal or other proceeding against him before its use against the target").

[289] *See* Terry v. Ohio, 392 U.S. 1, 19 (1968) ("The scope of the search must be 'strictly tied to and justified by' the circumstances which rendered its initiation permissible." (quoting Warden v. Hayden, 387 U.S. 294, 310 (1967))).

[290] *See* 50 U.S.C. § 1806(e) (2012) ("Any person against whom evidence obtained or derived from an electronic surveillance to which he is an aggrieved person is to be, or has been, introduced or otherwise used or disclosed in any trial, hearing, or other proceeding . . . may move to suppress the evidence obtained or derived from such electronic surveillance . . . .").

those values, so the probable cause standard is imposed to prevent that kind of government behavior.

Given the argument in Section II.B that machine learning raises tensions with rule-of-law values, it should come as no surprise that its opacity, arbitrariness, and reflection of human judgments raise similar conflicts with the ways those values have been embedded into existing law. Due to this conflict, use of machine learning in low-discretion decisions would undermine the basic principles put in place to ensure that the rule of law prevails when the government imposes significant burdens on citizens. The problem is not that machine learning will be less accurate—indeed, it is possible that predictive models would result in even more accurate results than our current mechanisms. But even assuming ninety-nine accuracy, the characteristics of opacity, arbitrariness, and the fact that human judgments are embedded in all computer models are inconsistent with the way our legal framework protects individual rights by minimizing discretionary decision-making.

The most obvious way that machine-learning models clash with discretion-narrowing rules are in their frequent inability to provide intelligible reasons for their determinations. When it comes to reasonable suspicion, for example, courts determine whether the standard is met by looking at the totality of the circumstances, and considering whether, taken together, all factors that the government relied upon add up to reasonable suspicion.[291] When the government relies on an unintelligible algorithm, however, there is no way to ascertain what factors led to the decision. When we rely upon a dog sniff to provide probable cause to search a bag, we know why the "black box" (i.e., the dog) believes probable cause exist. The dog is telling us, "I smell drugs." An algorithm that suggests there is probable cause for a search, or that a bail applicant is a flight risk, however, tells us nothing more than that conclusion itself. If a judge decided that Capricorns between the ages of 18 and 25 with no stable family relationships are more likely to jump bail, we would reject the astrological element of that description as irrelevant—Capricorns are no more likely to represent flight risks than Scorpios or Leos. But if an algorithm interpreted a data set to indicate astrological signs as a factor in recidivism, we may never know the role that (irrelevant) factor played in the decision-making. Even if we know what features were used, we do not know how they are weighted. And because algorithms will err arbitrarily, not only is an algorithm incapable of telling you why you were erroneously singled out, but it is also

[291] *See, e.g.*, *Terry*, 392 U.S. at 27 (noting that law enforcement's "reasonable suspicion" in searching for weapons on person relies on "whether a reasonably prudent man in the circumstances would be warranted in the belief that his safety or that of others was in danger").

incapable of telling you why, in another state, using another algorithm, a completely different outcome might prevail.[292]

Moreover, recall that a computer model will have access to only a limited universe of facts.[293] Only the data provided to the algorithm will factor into its decision-making, leaving out potentially relevant—indeed, potentially dispositive—data points that humans can take into account.[294] This means that, almost by definition, algorithms cannot take into account the totality of the circumstances in any given instance. As one machine-learning expert put it, "[p]rediction technologies are estimates of likely future regularities based on observed regularities in the past," but "humans are deeply unpredictable" and "someone may show all of the signs of setting out to perform some act, but draw back at the last minute" for reasons not accounted for in the algorithm.[295] Moreover, to the extent people know that such decisions are being made by a program that operates on the basis of determining what "normal" behavior consists of, it risks either unfairly singling out the eccentric or creating pressure to conform.[296]

Nor do we know what additional choices individual programmers have embedded within the algorithmic code. Whether the question is how an algorithm has defined critical terms such as "equal" or "likely," how it has translated complex policy goals into ones and zeros, whether it is instructed to minimize false positives or false negatives, or simply what unconscious reflections of the programmers' worldview affect the model, the human element of machine-learning predictions obscure the basis on which decisions are

[292] *See supra* Section II.B.2.b (discussing predictive model's arbitrariness and its impact on government legitimacy and rule of law).

[293] *See supra* Section II.B.2.b (noting that limitation on predictive models are constrained by access to data).

[294] *See supra* notes 183-186 and accompanying text. In "active learning"—a particular form of machine learning—an algorithm chooses its own training data, but it is still limited to choosing from among the data it is given. *See* Burr Settles, *Active Learning Literature Survey*, 1648 COMPUT. SCI. TECH. REPORT 27 (2010), http://burrsettles.com/pub/settles.activelearning.pdf [https://perma.cc/MQ5H-8YMA] (discussing empirical analysis done on active learning and noting that "training set built in cooperation with an active learner is inherently tied to the model that was used to generate it").

[295] SKILLICORN, *supra* note 33, at 68.

[296] *See id.* at 74 (discussing opposing ideas regarding anomalies in data).

made.[297] If a judge is asked to verify that the outcome of a computer model establishes probable cause, on what grounds can she make such a determination?

These same problems become even more pronounced when government decisions are challenged. The rule of law requires fora for challenging government action to ensure that incorrect, insufficiently supported, or arbitrary decisions can be corrected.[298] When an algorithm is not comprehensible, a trade secret, or a reflection of subjective human judgment, however, it is not clear how an adversely impacted individual can challenge its conclusions. Citizens cannot effectively contest the output of computer models that rely on features they do not know about, use data whose accuracy they cannot verify, or perform analyses that they cannot understand.[299] Traditionally, a criminal defendant can argue that the factors that underlie a probable cause determination were insufficient or that the government relied upon an incorrect piece of data—i.e., the defendant was not actually carrying a weapon.[300] But how is a defendant to make such an argument when the government's justification is a black box? Moreover, even if the government wanted to do so, it could not provide a satisfactory explanation.[301] As one scholar put it, "reasons are what we typically give to support what we conclude precisely when the mere fact that we have concluded is not enough."[302] And whether the context is a challenge to the validity of a search warrant, or the decision to place someone on the No-Fly list, or to deny bail, all an algorithm can provide is a conclusion. It cannot provide reasons.

---

[297] *See supra* notes 201-05 and accompanying text (discussing issues surrounding definitions in predictive models and impact that varied definitions for same term can have on outcomes across different models).

[298] *See* Fallon, *supra* note 133, at 9 (identifying impartial court as essential element of rule of law, as "[c]ourts should be available to enforce the law and should employ fair procedures").

[299] *See* Citron & Pasquale, *supra* note 2, at 5 (explaining that public is unable to test scoring systems "because the algorithms are zealously guarded trade secrets"); Crawford & Schultz, *supra* note 161, at 123 (discussing procedural opportunities that are available for persons harmed by automated systems and Big Data adjudication); Zarsky, *supra* note 7, at 1531 (discussing idea of fundamental right to government transparency and noting broadly that "'default' state for governmental actions should indeed be that of transparency").

[300] *See, e.g.*, United States v. Abernathy, 843 F.3d 243, 251 (6th Cir. 2016) (agreeing with defendant's argument that "Warrant was not supported by probable cause, because the marijuana roaches and T2-laced plastic bags the police recovered from Defendant's garbage were insufficient to create a fair probability that drugs would be found in Defendant's home").

[301] *See* Crawford & Schultz, *supra* note 161, at 123 (noting that government agencies "are unlikely to share the evidence and reasoning for the predictions that were made").

[302] Frederick Schauer, *Giving Reasons*, 47 STAN. L. REV. 633, 637 (1995).

Several scholars have documented the ways in which use of machine learning undermines due process protections.[303] Citizens are explicitly entitled to sufficient notice of the reasons for the government's action to allow an appeal.[304] With non-transparent algorithms, it is not clear how the government can provide such notice. And while there may be great debate regarding what process is due in any given instance,[305] notice that a benefit has been denied or a burden imposed will have little benefit if neither the affected party nor the government actor can understand exactly how that decision was made. The same is true for efforts to meaningfully challenge a computer model's output. In effect this could mean, for example, denial of bail with no means of contesting the result.[306]

With low-discretion decisions, we have a long-established framework of overlapping rules designed to enforce the rule-of-law values of predictability and constraint on arbitrary exercises of power to bolster government

[303] *See* Citron & Pasquale, *supra* note 2, at 8-16 (illustrating how automated credit scoring systems can be inaccurate and unfair to minorities); Crawford & Schultz, *supra* note 161, at 124-28 (discussing how Big Data can impact procedural due process and suggesting that notice, opportunity for hearing, and access to impartial adjudicator can alleviate some of negative impact Big Data can have on due process); Daniel J. Steinbock, *Data Matching, Data Mining and Due Process*, 40 GA. L. REV. 1, 78 (2005) ("Challenges to the algorithms used in data matching or data mining . . . may not fit well with the kind of individualized hearings that are the due process paradigm.").

[304] The constitutional notice requirement might be fulfilled by permitting individuals to see the data about them and confirm its accuracy or by maintaining audit trails. Citron, *supra* note 161, at 1308 (suggesting that software's failure to generate audit trails may be grounds for due process violation under balancing test); Crawford & Schultz, *supra* note 161, at 125 (noting that option for giving notice includes permitting customers "to petition Big Data providers to check and see if their data was being included or used in any predictive adjudications, and whether that data was accurate"). But even if such access were to become routine, it does little to explain the government's justifications for its action.

[305] National security measures such as the No-Fly List and seizure of assets of alleged terrorism supporters have led to years of litigation regarding what type of notice and redress procedures are constitutionally required. *See, e.g.*, Al Haramain Islamic Found., Inc. v. Dep't of Treasury (*AHIF II*), 686 F.3d 965, 965-66 (9th Cir. 2012) (discussing asset seizure); Ibrahim v. Dep't of Homeland Sec'y, 669 F.3d 983, 986 (9th Cir. 2012) (discussing No-Fly List); Latif v. Holder, 28 F. Supp. 3d 1134, 1162 (D. Or. 2014) (discussing No-Fly List); KindHearts for Charitable Humanitarian Dev., Inc. v. Geithner, 647 F. Supp. 2d 857, 858 (N.D. Ohio 2009) (discussing asset seizure).

[306] *See* Letter from Elec. Privacy Info. Ctr., *supra* note 154, at 4-5 (noting that while many legal systems rely on algorithms to "assess forensic evidence, determine sentences, [and] to even decide guilt or innocence," algorithmic decisions are often "entirely opaque, leaving individuals to wonder whether the decisions were accurate, fair, or even about them").

legitimacy.[307] This discussion illustrates how certain characteristics of machine-learning predictions come into conflict with these fundamentals. As a result, use of machine-learning predictions in low-discretion decisions will undermine the rule of law.

To be sure, the accuracy of human decisions in the aggregate almost certainly falls far short of the "statistically impressive" outcomes of predictive analytics, but these decisions lack the element of arbitrariness inherent in computer modeling.[308] At the same time, the benefits to be gained from using machine learning in low-discretion contexts are arguably less substantial than the benefits in high-discretion decisions. By definition, the government already is operating under at least some constraints.[309] Moreover, the more impactful the government decision, the more important it is both to be viewed as legitimate and to be justified by non-arbitrary factors. In addition, to the extent that a preference for data-driven analyses stems from the desire to limit government discretion to prevent state officials' conscious or unconscious bias from driving low-discretion decisions, there is less to be gained here than there is in high-discretion decision-making contexts. In high-discretion decisions, the risk of discrimination is at its height and the burdens imposed by Type I errors are at their nadir,[310] while the opposite holds true for low-discretion decisions, where

[307] *See, e.g.*, Kerr & Earle, *supra* note 178, at 70-71 ("[P]rivacy and due process values seek to limit what the government (and, to some extent, the private sector) is permitted to presume about individuals absent evidence that is tested . . . ."); Taslitz, *supra* note 139, at 176 ("Probable cause and reasonable suspicion thus help to protect citizen autonomy.").

[308] Even if human decisions are less accurate than algorithms, those decisions are individualized in a way that can be explained and, if necessary, challenged and corrected.

[309] *See supra* note 224 and accompanying text (discussing constitutional constraints on government).

[310] Some may argue that these concerns can be alleviated by ensuring that a human is always the final decision maker, and that algorithmic predictions are simply one factor to take into account. *See* Ferguson, *supra* note 7, at 263 (arguing that while ideally predictive policing will "become an important factor in a court's Fourth Amendment calculus," it is "never enough alone"); Simmons, *supra* note 7, at 1009, 1013 (suggesting that legal system ought to set numerical probability thresholds corresponding with reasonable suspicion and probable cause and noting that "there will be some situations in which the predictive algorithm is merely one of the factors that the judge considers"). In such circumstances, however, how will judges considering warrant requirements or beat officers deciding whether a car search is permissible, know how heavily to rely on the model's results and how much corroborating information to insist upon? Moreover, the human tendency to over-rely on the results of computerized processes means that even decisions ultimately made by humans will tend to give the algorithmic results great weight. *See* EXEC. OFFICE OF THE PRESIDENT, *supra* note 53, at 10 (discussing implications of reliance on big-data technologies on privacy and noting that

the government is expected to offer a non-biased explanation that affected individuals have the right to contest.

CONCLUSION

Machine learning has the capacity to work wonders. Concerns that others have raised regarding transparency, accountability, and accuracy provide ample justification for treading carefully when integrating machine learning into the public sector. But even highly accurate algorithms can threaten more fundamental values when they remain opaque, at times arbitrary, and shaped by subjective human decisions. These characteristics in turn undermine the legitimacy of government action relying upon the algorithms. Due to these characteristics, in government decision-making contexts in which eschewing the rule of law would do the most harm, the value that machine learning can add will be outweighed by the costs of undermining predictability, non-arbitrariness, and legitimacy. In circumstances where the cost of erroneous government action is lower, however, machine learning can provide a valuable tool to minimize discriminatory bias and maximize efficient allocation of resources. When determining whether and how to employ predictive analytics, policymakers should thus consider the different circumstances we find at the different ends of the discretion continuum.

"[t]echnology can be used for the public good, but so too can it be used for individual harm"); Ferguson, *supra* note 7, at 324-25 (arguing that machine-learning outputs will have outsize weight given their appearance of objectivity and scientific value); Goddard, Roudsari & Wyatt, *supra* note 188, at 123 (documenting "automation bias" generally and discussing relationship between automation bias, user accuracy, and erroneous decisionmaking); Kroll et al., *supra* note 7, at 680 ("[D]ecisions made by computers may enjoy an undeserved assumption of fairness or objectivity."); Murphy, *supra* note 133, at 261 (concluding that value of rule of law lies in its ability to "impose[] specific restrictions on the content of the law"); Frank Pasquale, *Secret Algorithms Threaten the Rule of Law*, MIT TECH. REV. (June 1, 2017), https://www.technologyreview.com/s/608011/secret-algorithms-threaten-the-rule-of-law/ (noting that while algorithms "may seem scientific, an injection of computational rationality into a criminal justice system riddled with discrimination and inefficiency," they are problematic because of issues of due process).