VIRGINIA JOURNAL OF LAW & TECHNOLOGY

FALL 2022 UNIVERSITY OF VIRGINIA VOL. 26, NO. 4

Law's Computational Paradox

FRANK FAGAN†

† Associate Professor of Law, South Texas College of Law (Houston); Research Associate, EDHEC Augmented Law Institute (France). Email: frank.fagan@edhec.edu. For comments, I thank Ramsi Woodcock, Jerold Soh, Joe Leahy, and the participants of the Global Faculty Seminar of the Singapore Management University School of Law.

ABSTRACT

Artificial intelligence (AI) and machine learning will bring about many changes to how law is practiced, made, and enforced. However, machines cannot do everything that humans can do, and law must face the limitations of computational learning just as much as any other human endeavor. For predictive learning, these limitations are permanent and can be used to ascertain the future of law. The basic tasks of lawyering, such as brief writing, oral argument, and witness coaching, will become increasingly precise, but that precision will eventually plateau, and the essential character of lawyering will remain largely unchanged. Similarly, where machines can be used to clarify application of law, they simply will limit judicial discretion consistent with moves from standards to rules or from rules to personalized law.

In each of these scenarios—lawyering and case clarification—enhanced precision is made possible through systemic closure of the machine's domain and AI will ascend easily. In scenarios where law's architecture is open, and systemic closure is not possible or worth it, machines will be frustrated by an inability to discern patterns, or by a powerlessness to comprehend the predictive power of previously discerned patterns in newly changed contexts. Lawmakers may add new variables to compensate and encourage attempts to model future environments, but open innovation and social change will undermine even a determined empiricism. In response to these limitations, lawmakers may attempt to actively impose closure of dynamic legal domains in an effort to enhance law's precision. By limiting admissibility of evidence, black-listing variables, requiring specific thresholds of white-listed variables, and pursuing other formalist strategies of closure, law can elevate its predictive precision for a given environment, but this elevation comes at the expense of openness and innovation. This is law's computational paradox.

This Article introduces the paradox across machine learning applications in lawmaking, enforcement, rights allocation, and lawyering, and shows that innovation serves as a self-corrective to the excessive mechanization of law. Because innovation, change, and open legal domains are necessary ingredients for continual technological

ascendance in law and elsewhere, fears of AI-based law as an existential threat to human-centered law are exaggerated. It should be emphasized, however, that there is ample room for quantification and counting in both closed and open settings; the products of innovation will always undergo measurement and machine learning algorithms will always require updating and refinement. This is the process of technological becoming. The goal for law is to never fully arrive.

The uncertainty of dynamic legal environments, even if diminishing with growing predictive power in law, forms the basis of an interpersonal constitutional authority. Understanding that some disruptions will always be unplanned prevents the construction of blind pathways for longer-term legal error, and relatedly, prevents empirical and technical rationales from overrunning a human-centered public square. A growing awareness of paradoxical error generated by precise, but closed, computational environments will generate a societal response that seeks to balance the benefits of precision and innovation. This balancing—what might be termed a "computational legal ethics"—implies that tomorrow's lawyers, more so than their counterparts of the past, will be called upon to discern what should be considered versus ignored.

TABLE OF CONTENTS

I. INTRODUCTION

Machines can now predict, with increasing precision, the outcomes of legal decision-making. The choices of the litigator, transactional attorney, legislator, regulator, judge, parole commissioner, and police officer have each been exposed to algorithmic learning.[1] Examples are numerous. Choices over the persuasive expression of lawyers—as subtle as word selection, writing style, and voice intonation—are being measured and then used for the prediction of judicial outcomes.[2] Today's lawmakers and regulators, perhaps incredibly for earlier programmers and computer scientists, rely on data and predictive models in order to choose among regulatory alternatives.[3] Chosen rules are then calibrated with the help of

[1] Each, with relevant references, are discussed presently *infra* §§ I.A-I.C.

[2] On word choice *see* Hannah Laqueur & Anna Venancio, *A Computational Analysis of California Parole Suitability*, *in* LAW AS DATA: COMPUTATION, TEXT, AND THE FUTURE OF LEGAL ANALYSIS 193, 221-30 (Michael A. Livermore & Daniel N. Rockmore eds., 2019) (finding a relationship between specific words spoken during parole hearings and the probability of being released). On writing style *see* Elizabeth C. Tippett, Charlotte S. Alexander, L. Karl Branting, Paul Morawski, Carlos Balhana, Craig Pfeifer & Sam Bayer, *Does Lawyering Matter? Predicting Judicial Decisions from Legal Briefs, and What That Means for Access to Justice*, 100 TEX. L. REV. 1157, 1173 (2022) (finding a relationship between a given writing style manifest in legal briefs of litigants to employment disputes, such as a repetitive or wordy style, and the probability of receiving a favorable judicial outcome). On voice intonation *see* Daniel E. Chen, Yosh Halberstam, Manoj Kumar & Alan C. L. Yu, *Attorney Voice and the U.S. Supreme Court*, *in* Livermore & Rockmore, *supra* at 367 (finding a relationship between voice intonation of the utterance "may it please the court" and a favorable judicial outcome).

[3] Examples abound. *See, e.g.*, Ryan Calo, *Modeling Through*, 72 DUKE L.J. 1390, 1390-91 (2022) (noting that lawmakers can leverage advances in artificial intelligence to model effects of new law and then select among alternatives); Fabiana Di Porto, *Algorithmic Disclosure Rules*, ARTIFICIAL INTELLIGENCE & L. at 4-5 (2021) (suggesting that machine learning and natural language processing can help lawmakers create, and then update, optimal disclosure rules); Jonathan H. Choi, *An Empirical Study of Statutory Interpretation in Tax Law*, 95 N.Y.U. L. REV. 363, 363 (2020) (demonstrating how natural language processing can help judges select among various interpretations of statutes); Michael A. Livermore, Vladimir Eidelman, & Brian Grom, *Computationally Assisted Regulatory*

machines.[4] Judges and other decision-makers are deploying risk assessment models to guide their decisions over sentencing and bail,[5] while police departments are using algorithms to allocate limited enforcement resources and gain other efficiencies.[6] The overall trend is toward computational law.

Law professors and legal scholars have responded with suggestions, often normative, in each of these areas.[7] These suggestions have most often been confined to a single legal

---

Participation, 93 NOTRE DAME L. REV. 977, 977 (2018) (demonstrating how algorithms can assess public comments in regulatory rule-making and use those comments as inputs to select among rules).

[4] *See* Di Porto, *supra* note 3, at 1, which describes the interplay of periodic application of natural language processing and the periodic updating of regulated mandatory disclosures.

[5] For a high-profile example of using risk assessments in sentencing, *see* State v. Loomis, 881 N.W.2d 749, 769-70 (Wis. 2016) (use of algorithmic risk assessment tool permitted so long as judge considers other non-algorithmic factors). On the use of algorithms for making bail-or-jail determinations, see, for example, Jon Kleinberg, et al., *Human Decisions and Machine Predictions*, 133 Q. J. ECON. 237, 237-38 (2018). The use of algorithms by the state for determining is not limited to liberty interests. Property interests are implicated, too. *See, e.g.*, Houston Fed'n of Teachers, Local 2415 v. Houston Indep. Sch. Dist., 251 F. Supp. 3d 1168, 1180 (S.D. Tex. 2017) (state's use of algorithmic assessment of teacher performance to justify termination of teachers prohibited); *accord* Trout v. Knox Cnty. Bd. of Educ., 163 F. Supp. 3d 492, 497-99 (E.D. Tenn. 2016).

[6] *See* Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. 1115, 1128-30 (documenting the development of algorithmic tools to predict places of property crime, places of violent crime, and persons involved in criminal activity). On how errors can be unevenly distributed across protected classes, see Rashida Richardson, Jason Schultz & Kate Crawford, *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*, 94 N.Y.U. L. REV. ONLINE, 192, 192 (2019). For recent work that describes additional conflict with constitutional rules, see Céline Castets-Renard, *Human Rights and Algorithmic Impact Assessment for Predictive Policing*, *in* CONSTITUTIONAL CHALLENGES IN THE ALGORITHMIC SOCIETY *passim* (Hans-W. Micklitz, et al., eds., 2021).

[7] Consider that today over 2,800 articles have been archived by the SSRN eJournal "Artificial Intelligence – Law, Policy, & Ethics." *See* SSRN, https://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalBrowse&journal_id=2874401 (last visited June 30, 2022).

domain.[8] After all, law professors and lawyers build careers around specializations and knowledge tends to organize itself into categories until overlap becomes worthwhile. The upshot is that systematic and comprehensive treatments of computational law remain under-represented in the literature. To borrow an analogy from economics, there is plenty of "micro," but there is a bit less "macro." This Article is macro. It broadly examines the deep incursions artificial intelligence has made into lawmaking, enforcement, rights allocation, and lawyering.[9] When taking a stepping back, a fundamental trade-off comes into sharp relief. Computational law gives up innovation in exchange for precision at its deepest level. This inherent paradox is built into its fabric: On the one hand, enhanced predictive capability with artificial intelligence furthers a fundamental goal of law upon which its most revered procedures stand, *viz.* the minimization of error.[10] At the same time, the predictive perfection of the legal process, which can be imagined with machines, demands systemic closure and suppression of innovation.

Consider that sophisticated algorithms of chess and the game of *go* triumphed over the world's best humans precisely

[8] *Id.*; *see also* AI LAW BLAWG, http://www.ailawblawg.com (last visited June 30, 2022) (sorting and archiving "law and technology" scholarship by largely by standard doctrinal divisions).

[9] Throughout the Article, enforcement is used as an all-encompassing term that includes the application of criminal and other rules that implicate a person's bodily and property interests. In addition, enforcement as used here includes the allocation of scarce societal resources where that allocation is made by judges or other agents of the state on the basis of legal rights defined by law.

[10] Apart from the magnitude of private and government interests, determination of whether a process is constitutionally sound rests on its tendency to generate errors. *See* Mathews v. Eldridge, 424 U.S. 319, 335 (1976) (noting the importance of minimizing error in civil contexts); Medina v. California, 505 U.S. 437, 445 (1992) (giving deference to state criminal procedures on basis of state expertise and ability to minimize error).

because chess and *go* can be perfectly modeled.[11] Even if magnificently complex, they are closed systems whose environments remain fixed over time. If humans add new variables, or occasionally change the rules of the game, then of course the algorithms can be easily frustrated. If Gary Kasparov or Ke Jie were to declare a new method for a game piece to move, such as allowing a king to move two squares instead of one in chess, or in *go*, by allowing new governance of the standard rules of liberty and empty intersections, then these expert human players could easily defeat an untrained machine. An unexpected innovation can befuddle the deepest of learning algorithms and reduce the most magnificent computational edifices to noise.

It is tempting to assume that innovative behavior relevant to law can be modeled, and if not, then humans will eventually model the totality of life. There are at least two objections to these assumptions. Data may be insufficient and computational power may be limited. Consider that many times in law sample sizes are too small.[12] And when data may be big, simply accounting for one additional feature or nuance of an observation will consistently lead to exponential increases in the number of possibilities that the machine must evaluate.[13] Apart from small data and limits to computational power, innovation

[11] *See* David Silver et al., *Mastering the Game of Go with Deep Neural Networks and Tree Search*, 529 NATURE, 484, 484 (2016) (documenting the computational architecture of *go*).

[12] *See* Saul Levmore, *The Eventual Decline of Empirical Law and Economics*, 38 YALE J. ON REGUL. 612, 612 (2021) (noting the persistence of small sample sizes in law); *accord* Frank Fagan, *The Un-Modeled World: Law and the Limits of Machine Learning*, 4 MIT COMPUTATIONAL L. REPORT (2022).

[13] *See* LESLIE VALIANT, PROBABLY APPROXIMATELY CORRECT 74 (2013) (explaining that the addition of one Boolean variable doubles the number of candidate algorithms that accurately describe the data); *see also* PEDRO DOMINGOS, THE MASTER ALGORITHM: HOW THE QUEST FOR THE ULTIMATE LEARNING MACHINE WILL REMAKE OUR WORLD 73-74 (2015) (noting the problem of exponential growth of candidate algorithms when adding new variables and the resultant inability to create general learners).

presents yet another categorical objection to the eventual arrival of perfectly predictive models. Innovation generates novelties and surprises and uncovers unimaginable unknowns that are not available as computational inputs earlier in time. Even if data were big and computational power limitless, an innovation may be spontaneous and hence unpredictable with even the best of models.

There is a danger that in its effort to elevate precision, law may attempt to impose closure and forbid consideration of certain variables—much the same way a well-trained data scientist will attempt to limit the number of potentially correct algorithms that she puts forward as suggestions of the true state of the world by means of ignoring certain features of her data.[14] Law, in the same vein, may attempt to exclude certain types of evidence, forbid judges from considering various features of a case, or insist that personalized rules must apply on the basis of white-listed variables only.[15] These strategies of closure elevate precision at the expense of innovation, and paradoxically generate errors in their immoderate pursuit of computational accuracy. In some scenarios, closure may be worthwhile for a time, as when the benefits of precision outweigh those of innovation in temporally stale and static domains of human activity and law. But in other scenarios, as when law and its practice are situated within dynamic, fast-moving, and uncertain domains, closure can prove socially costly.

It may be conventional to classify a player of games who innovates around the rules as a cheater, but in life, innovation is often socially valuable. Of course, not all innovation is desirable. The inventive use of machines may help one privately benefit at another's expense. In law, machines can help people evade ideal

[14] *See* DOMINGOS, *SUPRA* note 13 at 73-75 (noting that data scientists subject candidate algorithms to hypothesis testing in order to reduce the total number of algorithms that they evaluate).

[15] *See infra* § I.A.2.

application of legal rules by revealing exploitable ambiguities and enforcement gaps. Predictive tools can help litigants mislead the judge toward favoring their cases, deceptively present circumstantial facts, and advocate for privately favorable but socially wasteful rules.[16] The optimal design of computational law depends to some extent on what type of innovation the lawmaker faces.

This distinction between good and bad innovation is used throughout the Article. Good innovation leads to genuine improvements in life and law that satisfy some broadly accepted criterion such as welfare maximization. Bad innovation is novelty directed at gaining a strategic advantage over one's adversary and produces few, if any, broader social benefits. As we will see, bad computational innovation between adversarial parties collapses to an arms race and provides no long-term advantage (public or private). Closure lawmaking strategies that suppress it provide fewer benefits than might be expected. A better role for law under bad innovation is to accelerate the computational arms race by equipping adversaries with open access data and cheap data processing tools.[17]

---

[16] While these problems are already present and familiar in our system of law, AI will enhance the capacities of bad actors and magnify the effects of their corrosive efforts. In general, my idea of bad innovation tracks the public choice concept of rent-seeking. Bad innovation simply represents effort expended toward gaining a larger share of a pie. Good innovation increases the size of the pie. *See* NICHOLAS MERCURO & STEVEN G. MADEMA, ECONOMICS AND THE LAW: FROM POSNER TO POST-MODERNISM 96-97 (1997) (explaining that public choice theorists view rent-seeking activity as wasteful because effort directed at gaining private advantage could be channeled toward societal wealth creation).

[17] *See* Tippet et al., *supra* note 2 at 1194-95 (noting that paywalls such as PACER may create uneven playing fields and that law should favor open access to court data). *See also* Charlotte S. Alexander & Mohammad Javad Feizollahi, *On Dragons, Caves, Teeth, and Claws: Legal Analytics and the Problem of Court Data Access*, in COMPUTATIONAL LEGAL STUDIES: THE PROMISE AND CHALLENGE OF DATA-DRIVEN LEGAL RESEARCH 97 (Ryan Whalen ed., 2019) (2019) (AK: not sure if same is referring to the previous note that paywall such as PACER may create uneven playing fields and that law should favor open access to court data. I think it might be referring to

With respect to good innovation, closure strategies such as black-listing variables for consideration by judges and other lawmakers, requiring threshold magnitudes of certain white-listed variables, limiting judicial creativity and hunches, deploying social credit systems and surveillance with few safety valves for novel deviations, and so on, implicate the computational paradox. There is a tradeoff, under good innovation, between short-term gains in computational precision and the longer-term costs of dampening socially beneficial innovation.

***

A blind tendency toward the computational closure of law shares much in common with several themes that have been explored elsewhere. There are intersections with modern philosophy's conception of *tekhnē*,[18] especially as elaborated by

"same" date as 2019?). However, inasmuch as PACER represents a natural monopoly, elimination of fees may reduce data collection effort and the total data available to all parties. *See* David Freeman Engstrom & Jonah B. Gelbach, *Legal Tech, Civil Procedure, and the Future of Adversarialism*, 169 U. PENN. L. REV. 1001, n. 248 (2021).

[18] Aristotle considers *tekhnē* as the art, or the means, of bringing something into existence: "Every art [*tekhnē*] is concerned with bringing something into being, and looks for technical and theoretical means of producing a thing which belongs to the category of possibility and the cause of which lies in the producer and not in what is produced." Aristotle, *Nichomachean Ethics, Bk. 6*, *in* THE COMPLETE WORKS OF ARISTOTLE, JONATHAN BARNES (ED.) 4 (1984). For Heidegger, *tekhnē* is a process of disclosure: "[*Tekhnē*] reveals whatever does not bring itself forth and does not yet lie here before us…what is decisive in *tekhnē* does not at all lie in making and manipulating, nor in the using of means, but rather in the revealing…" Martin Heidegger, *The Question Concerning Technology*, in MARTIN HEIDEGGER, BASIC WRITINGS, DAVID FARRELL KRELL (ED.) 318 (2008). He then equates this process of disclosure to technology: "Technology is the mode of revealing. [It] comes to presence in the realm where revealing and unconcealment take place." *Id.*

Bernard Stiegler,[19] as well as Jürgen Habermas'[20] related idea of "system" or technical "colonization" of the world of human life and interactions.[21] Stiegler and Habermas describe a process of systemization, or continual technological becoming, which is always switched on, and with the advent of industrialization, is always accelerating.[22] Apart from new styles of bureaucratically managing large pools of lawyers, law has remained relatively immune to the dramatic accelerations seen elsewhere. But law's continued inoculation is not guaranteed, and perhaps even unlikely. Current tools that can precisely measure large bodies of non-numerical data[23] pose a serious challenge to the historically unhurried pace of development of legal practices and rules.[24]

Processes of technological becoming are generally presented as inevitable, and both Stiegler and Habermas offer some guidance on the ethics of systemization and active empiricism. They describe how human societies should respond.

---

[19] (Former) Professor, University of Technology at Compiègne; Director, Institute of Research and Innovation (Centre Georges-Pompidou).

[20] Professor, University of Frankfurt; Director, Institute for Social Research; Permanent Visiting Professor, Northwestern University; Theodor Heuss Professor, The New School, New York.

[21] *See infra* § II.A.

[22] *See id.*

[23] The availability of textual data and tools to process it has been growing. *See* Frank Fagan, *Natural Language Processing for Lawyers and Judges*, 119 MICH. L. REV. 1399, 1400 (2021) (noting the growth of textual data and cataloging advances for processing).

[24] Earlier challenges, such as automated decision support systems for lawyers and judges, were wildly optimistic and relied too heavily on the arrival of general learners that could replicate the processes of the human mind. *See id.* (noting that the slow development of legal technology generally tracked the slow development of artificial intelligence). General learners, sometimes referred to as "strong AI," have yet to materialize; they may not ever. MAX TEGMARK, LIFE 3.0 129-30 (2017) (noting that the world's leading AI scientists are divided on the possibility). What has changed in the past decade is the dramatic increase in our ability to process large bodies of data, including textual data. Basic advances in machine learning and the arrival of "weak AI" have the capability to rapidly accelerate technological advance in law. *See* Fagan, *supra* note 23 at 1400.

In short, Habermas represents a strain of thought that legitimizes technology in law on the basis of intersubjectivity and human mediation.[25] Stiegler represents a softer form of empiricism that is partly tempered with intersubjectivity.[26] These intellectual undercurrents support different paces of technological incursion into law. Thus, many today believe law is faced with a choice. It can accelerate or decelerate the enlargement of machines in lawmaking, lawyering, judging, policing, and rights allocation by setting small or large limits to machine learning.[27] The point of this Article is that the imperative to innovate, which itself has given us AI, operates as a self-corrective—a built-in feature that limits the ability of technology to fully dominate human life and law. Nonetheless, awareness of the computational paradox is helpful inasmuch as it guides lawmakers away from poor strategies of imposed closure and the social losses that come from suppressed innovation.

There are antecedents in history of technological closure followed by innovation and openness that can serve as models. Gilmore, channeling Llewellyn, speaks of the "Ages" of American law.[28] The Age of Discovery, running from the Revolution to the Civil War, embodied the transformation of an agricultural to industrialized society and milieu. Its lawyers, having relied on Blackstone's commentaries and wholesale

[25] *See infra* § II.B.

[26] *See id.*

[27] There is a common fear that AI-based rules threaten the rule of law and require actively imposing limitations. *See, e.g.*, Robert F. Weber, *Will the "Legal Singularity" Hollow Out Law's Normative Core?*, 27 MICH. TECH. L. REV. 97, 149 (2020) (suggesting that the legal order of National Socialist Germany "resonate[s] with the description of … predictability presented here"); John Danaher, *The Threat of Algocracy: Reality, Resistance and Accommodation*, 29 PHIL. & TECH. 245, 245 (2016) (suggesting that algorithmic governance devoid of human input "pose[s] a significant threat to the legitimacy of such processes").

[28] Gilmore drew from Llewellyn's tripartite division of American legal history. The three periods run from 1800 to the Civil War, from the Civil War to World War I, and from World War I to today. *See* GRANT GILMORE, THE AGES OF AMERICAN LAW 11 (1977) (noting the general outline of these periods is contained within Llewelyn's book, *The Common Law Tradition*).

importation of English rules, were furnished with new treatises and commentaries built from an indigenous base of American law—a base that had not existed yet a generation earlier—by the efforts and genius of Story and Kent.[29] The treatises proved popular among American attorneys. Additional treatises on obscure and growing areas of law in response to industrialization proliferated and boomed.[30] West Publishing was established. It quickly built up a National Reporter System shortly after the Civil War, which became "the essential stock-in-trade of the working lawyer."[31]

As cases multiplied, the profession sought systemization. It found its guide in Langdell, who Gilmore wryly notes, "seems to have been an essentially stupid man who, early in his life, hit on one great idea," that is, that law is a science.[32] There arose under his tutelage Gilmore's Age of Faith, whose mechanical application of laboratories (law schools) to materials (printed case reports) reduced law's "unruly diversity to a manageable unity"[33] much the same way that today's computational legal scholars seek to shrink the dimensionality

---

[29] Early American lawyers had little of their own law to draw upon. The most readily available sources were found in "the crabbed and incomprehensible pages of Coke on Littleton" and the only somewhat superior "elegant superficialities of Blackstone." *Id.* at 19. By the 1820s, American legal materials, especially the decisions of American state and federal courts, had sufficiently accumulated for a decidedly American law to emerge. *Id.* at 22. Gilmore explains that "[t]he Story treatises, like Kent's *Commentaries*, were quite consciously designed to lay the foundations for an American law derived from but in no sensed confined by the principles of English law." *Id.* at 28.

[30] Treatises also appeared and multiplied as enterprising writers sought financial success:

> There was evidently money to be made… For the most part these writers were not academics nor were they distinguished practitioners or judges. They were hacks who would run up a book on negotiable instruments this year, a book on corporations next year, and a book on insurance the year after that. […] The best of them, however, were astonishingly good.

[31] *Id.* The Reporter System emerged in the 1880s.

[32] *Id.* at 42.

[33] *Id.* at 43.

of legal texts and other sources.[34] Where Langdell provided organization, Holmes provided intellectual imprimatur, distilling, for instance, all theories of liability to a "philosophically continuous series."[35] The Age of Faith, loosely discerned from the Civil War until World War I, treated law as a closed system whose principles could be scientifically deduced and applied.[36] The primary tool of the faithful was the citation string, and it worked extremely well for a society whose technological progress was slowing from 1860 to 1910.[37] With fewer environmental changes, law became more closed and accordingly more easily made and applied. Perhaps Langdell was a genius who sensed a period of technological stagnation and responded in kind.

Much is made of Legal Realism and its intersection with ideology,[38] but Gilmore notes that it easily arose as a response to the limitations of the exactitude imposed on law by Langdell and his followers.[39] Limitations were laid bare by the growing

[34] *See, e.g.*, Eliot Ash & Daniel L. Chen, *Case Vectors: Spatial Representations of Law Using Document Embeddings*, in LIVERMORE & ROCKMORE, *SUPRA* note 2 at 337 (applying embedding models to legal documents in order to reduce the high dimensionality of legal texts); Frank Fagan, *Big Data Legal Scholarship: Toward a Research Program and Practitioner's Guide*, 20 VA. J. L. & TECH. 1, 62-74 (2016) (documenting the application of classifiers and topic models to judicial decisions in order to categorize those decisions and distill their core features).

[35] GILMORE, *SUPRA* note 28 at 53.

[36] *Id.* at 62.

[37] On the success of the citation tool, see *id.* at 41. On slowing progress, see *id.* at 65.

[38] *See, e.g.*, Michael A. Livermore, *Realist Lawyers and Realistic Judges: A Brief Rebuttal to Judge Posner*, 59 DUKE L. J. 1186, 1187-88 (2010) (describing a tension between legalist and realist models of judging, where the former eschews judicial ideology as an input into decision-making while the latter acknowledges and even embraces it).

[39] GILMORE, *SUPRA* note 28 at 87 ("The 'conceptualism' of the Langdellian period was … held up to scorn. The great treatises and Restatements … were pilloried as nonsensical attempts to portray the life of the law as having been logic rather than experience… The idea that the process of judicial decision was much more irrational than it was rational had a fashionable currency.").

volume and diversity of case law and the expansive and novel technological advances following industrialization. Law's indeterminacy was once again made apparent, but instead of discovering principles and unraveling logic, the task of jurists from World War I to today became creation. Gilmore notes that in this "Age of Anxiety" legal scholars became social scientists; legal activists drafted statutes and administered agencies; and judges, following them both, surged toward activism.[40] So long as law resists a conservative urge to foreclose innovation, much of that creative energy will continue to flow downward in the "Age of Artificial Intelligence," even as—we will see—its overstated promises of precision by those who ignore the technical limitations generated by innovation and life remain startlingly familiar.[41]

II. JURIDICAL TECHNICS

Traditional statistical approaches measure characteristics and search for correlations and causal relationships between variables. In law, empirical work often focuses on using statistics to search for good policy. It asks questions such as how mental health affects the way in which a criminal perceives the risks and "costs" of committing a crime,[42] whether rules like the one articulated in *Alice Corp. Pty. Ltd. v.*

[40] *Id.* at 91-92.

[41] There is in addition the strategic response of humans, the so-called Lucas critique, who resist revealing or even carrying out patterned behavior since they know that machines will govern any discernible pattern that they manifest. Think of a criminal who can perfectly randomize her actions so as to not be observed. This would be an instance of bad innovation. *See supra* note 16 and accompanying text. On the Lucas critique, named for the economist who first noted the endogeneity of strategic behavior within a closed system, see Robert E. Lucas, Jr., *Econometric Policy Evaluation: A Critique*, 1 CARNEGIE-ROCHESTER CONF. SERIES ON PUB. POL'Y 19, 41-42 (1976).]

[42] *See* Jeffrey Fagan & Alex R. Piquero, *Rational Choice and Developmental Influences on Recidivism Among Adolescent Felony Offenders*, 4 J. EMPIRICAL LEGAL STUD. 715, 716 (2007).

CLS Bank International[43] targets patent trolls,[44] and whether non-lawyers make good panelists for adjudicating WTO disputes.[45] These types of studies support good policy. We might increase public spending on mental health to lower crime, adjust *Alice* to provide protections for inventors, and encourage non-lawyers to become WTO panelists. The general pattern of research, normative policy conclusion, and practical influence is a familiar one to law professors, though in most cases, the path from research to policy change takes time because political processes must endorse research, however convincing it may be.[46] In matters of policy, there is a delay between empirical rationale and human action.

This limitation has partly been the result of available measurement tools. Statistical instruments are good at measuring and manipulating categorical variables that represent one of two possible states of an observed phenomenon. They are proficient in handling continuous variables that represent an infinite gradation of states. But in either circumstance, these instruments most often rely on numbers. The empirical legal scholar routinely collects numerical data (or converts continuous observations to real numbers and categorical observations to zeros and ones) prior to analyzing the data in STATA or similar statistical software.[47] Empirical legal scholarship's focus on

[43] Alice Corp. Pty. Ltd. v. CLS Bank Int'l, 573 U.S. 208 (2014).

[44] *See* Mark A. Lemley & Samantha Zyontz, *Does* Alice *Target Patent Trolls?*, 18 J. EMPIRICAL LEGAL STUDS. 47, 47 (2021).

[45] *See* Julian Nyarko & Jerome Hsiang, *Conforming against Expectations: The Formalism of Nonlawyers at the World Trade Organization*, 48 J. LEGAL STUDS. 341, 341 (2019).

[46] Thus, Gilmore's concern with formalist approaches to law is overstated, or at least, he understates the importance of political buffers. *See* GILMORE, *supra* note 28, at 108-109 (stating the concern). [it seems like the beginning of this paragraph, which starts on 108 is also addressing his concern about formalism, so might expand pin cite to include it?]

[47] *See* Jens Frankenreiter & Michael A. Livermore, *Computational Methods in Legal Analysis*, 16 ANN. REV. L. & SOC. SCI. 39, 41 (2020) (noting the distinction between earlier data deployed for computational legal study and new textual data).

slow-moving policy outcomes reflects the character of its popular tools.

Lawyers and judges, however, rely heavily on non-numerical data such as textual cases and briefs as well as aural and visual cues from clients, witnesses, and parties to litigation. This means that an overwhelming proportion of empirical legal research has had little immediate impact on the day-to-day tasks of lawyers and judges. While this may be obvious, it remains unclear to what extent new advances in legal empirical research portend changes to law.[48] To be sure, the new tools will partly shift focus away from policy studies and toward immediate applications for practitioners inasmuch as research output remains at prevailing levels and enterprising law professors look for new opportunities.[49]

Adoption delays in some domains should prove comparatively shorter as well. As mentioned, political process stands between policy research and implementation for policy selection. Nothing comparable stands between the research and implementation of new techniques for legal practice. In fact, the contrary is true; that is, attorneys are pressured by competitive forces to quickly adopt superior techniques in order to successfully represent clients while researchers and developers seeking profit face competitive pressure to lower the economic costs of adopting the new techniques in order to make them even more attractive for practitioners. For lawyers and judges, interest group pressures are comparatively smaller and competitive forces more strongly drive technological advance.

[48] *See* Fagan, *supra* note 23, at 1401 (noting that entrepreneurial activity and investment in legal technology is growing, but the scope of transformation remains unclear for practitioners).

[49] *See, e.g.*, Richard A. Posner, *Legal Scholarship Today*, 115 HARV. L. REV. 1314, 1324-26 (2002) (noting the broad rise in interdisciplinary legal scholarship for this reason).

In contrast to law practice, the use of artificial intelligence in lawmaking and enforcement will require broader consensus, and changes will likely come about through the traditional channels of political process as has already been seen, for instance, in the legislative mandates to use predictive risk assessment tools in sentencing.[50] The variation in these political-economic barriers to AI proliferation in law may suggest variation in its proliferation, but the inherent paradox presented by computational law (and the resultant shape of a computational legal ethics as it responds to that paradox) will play a more fundamental role. The following subsections consider recent advances across lawmaking, enforcement, rights allocation, and lawyering, and theorize each of their longer-term

[50] Some states require judges to be provided with predictive risk assessments at sentencing, while others permit, but do not require, their use. For examples of required provision, see KY. REV. STAT. ANN. § 532.007(3)(a) (2016) (requiring sentencing judges shall consider the results of risk assessment carried out during presentence investigation); OHIO REV. CODE § 5120.114(A)(1)-(3) (2015-16) (stating that the Ohio Department of cCrrection "shall select a single validated risk assessment tool for adult offenders" that shall be used for sentencing); 42 PA. CONS. STAT. § 2154.7(a) (2016) (adopting risk assessments to help judges determine appropriate sentences); ARIZ. CODE OF JUD. ADMIN. § 6-201.01(J)(3) (2016) ("For all probation eligible cases, presentence reports shall [] contain case information related to criminogenic risk and needs as documented by the standardized risk assessment and other file and collateral information"); OKLA. STAT. tit. 22, § 988.18(B) (2016) (requiring the use of predictive risk assessments to determine appropriateness of any community punishment).

For examples of permissible use, see IDAHO CODE § 19-2517 (2016) (if judges deploy risk assessments for imprisonment or probation decisions, the report must include current Idahoan recidivism rates differentiated on risk levels of low, moderate, and high); Malenchik v. State, 928 N.E.2d 564, 566, 571-73, 575 (Ind. 2010) (encouraging the use of predictive risk assessments at sentencing) ; LA. REV. STAT. § 15:326(A) (2016) (presentence investigation validated risk assessment tool permissible for sentencing adult offenders); WASH. REV. CODE § 9.94A.500(1) (2016) (requiring judges to consider predictive risk assessments at sentencing if available); State v. Rogers, No. 14-0373, 2015 WL 869323, at *4 (W. Va. Jan. 9, 2015) (Loughry, J., concurring) (probation officers required to use risk assessments pursuant to W. VA. CODE § 62-12-6(a)(2) (2014), but judges maintain the discretion to use them at sentencing).

trajectories in terms of the paradox. Section II will consider the character of a computational legal ethics driven by the paradox.

A. Lawmaking

1. From Causality to Prediction

For the past several decades, empirical legal scholarship has been growing in popularity and importance.[51] This growth is evidenced by the founding of new academic associations, such as the Society of Empirical Legal Scholars, the creation of legal research chairs dedicated to empirical legal work, and increased levels of empirical scholarship published in academic law journals.[52] As mentioned in the beginning of this Section, the overwhelming majority of this scholarship to date has focused on discovering causal relationships and then normatively changing law in response to these new discoveries.[53] For empirical legal scholars, identification of causality begets good lawmaking. But this pattern of causal research followed by policy recommendations has paid only modest dividends to law.[54] Research in law is unlike experimental medicine; legal rules are not randomly applied.[55] Random application of a drug over a large enough test population can ensure that unseen or

[51] Description of a course on empirical legal studies at Stanford Law School, for example, notes that "[e]mpirical legal studies have become trendy in the U.S. and are now spreading to law faculties in other countries as well." *Empirical Legal Studies: Research Design*, STANFORD LAW SCHOOL: COURSE CATALOG, https://law.stanford.edu/courses/research-design-for-empirical-legal-studies/ (last visited June 30, 2022).

[52] *The Journal of Empirical Legal Studies* is the leading outlet, though empirical legal scholarship is often published in leading law and economics journals as well as law reviews.

[53] *See, e.g.,* notes 42-45 and accompanying text *see also* Fagan, *supra* note 23, at 1401-04 (discussing the early development of predictive, versus causal, research agendas by empirical legal scholars).

[54] *See* Levmore, *supra* note 12, at 614 (noting that "[g]ood empirical work is simply hard to do" because of inherent limitations in data collection and sample sizes).

[55] *See* RONALD M. DWORKIN, LAW'S EMPIRE 178-84 (1986) (discussing random law's uneasy relationship to equal protection and problems with arbitrary application).

omitted variables are uncorrelated with the pharmaceutical researcher's (random) choice to treat or not treat a person.[56] Legal scholars do not have that empirical luxury. Random application of a legal rule runs afoul of the constitutional requirements of equal protection and non-arbitrary application.[57]

This barrier leaves empirical legal scholars with limited options. While experimental settings such as laboratories of graduate students permit random application of rules, these are of course imperfect because the students are not representative of the true population. Graduate students hardly represent a cross-section of citizens where an actual legal rule would be applied. Even if experimental (and random) application of a rule were tested over a more representative sample, such as people who respond to a survey or receive payment for interacting with Amazon's Mechanical Turk, or if representativeness were addressed with a creative experimental design or corrected with additional variables, experimental subjects (people) in those instances will, nonetheless, continue to be observed while responding to an application of a hypothetical rule in an experimental setting. Even if researchers adjust the experiment by providing cash payments or other rewards in exchange for various behaviors, additional assumptions must be made in order to infer the expected behavior of an experimental subject who is faced with the same dilemma in real life.[58] These assumptions can be attacked. In short, there is a difference between a real-life persona and an experimental persona and identifying a causal relationship by simulating randomness is therefore challenging.

---

[56] *See* JEFFREY M. WOOLDRIDGE, INTRODUCTORY ECONOMETRICS: A MODERN APPROACH 88-92 (5th ed. 2013).

[57] *See* DWORKIN, *supra* note 55, at 178-84.

[58] *Cf.* Peter A. Diamond & J.A. Hausmann, *Contingent Valuation: Is Some Number Better Than No Number?*, 8 J. ECON. PERSPECTIVES 45, 49 (1994). (noting that people respond strategically to hypothetical survey questions and rarely reveal true preferences in the process).

To avoid this problem, the researcher can, on occasion, stumble into a natural experiment. Law's random application might occur by chance when nature, instead of the researcher, "selects" to whom a law is applied and thereby side-steps constitutional rules of equal protection and non-arbitrariness. Bad weather, for instance, may cause intermittent electricity outages at a port so that cargo is only randomly scanned by electronic sensors even though law may require the sensors to scan all incoming cargo. Because of the weather, a causal relationship between the mandated scanning rule and the level of illegal contraband may be able to be inferred.[59] Instead, if the researcher decided when to turn the electronic scanners off and on, even at random, constitutional rules could be violated. In life, natural experiments are rare. While useful for identifying causal relationships, they are difficult to find and many important empirical questions in law remain unanswered precisely because their proper identification is problematic.

It should be easy to see the growing stores of data cannot side-step the "identification" problem. More data alone cannot answer causal questions that demand random application of a rule. And observing every variable, so that the omitted variable problem disappears, is a fanciful enterprise within the fast-changing environments often found in law.[60] New variables can appear every day. For this reason, researchers "structurally" model the behavior of people on the basis of data combined with

[59] Maybe not: The careful researcher might suspect a relationship, for instance, between bad weather and the ability of small ships to port. If small ships are systematically different from other ships, say because smaller ships are better (or worse) for smuggling, then the natural experiment will fail.

[60] VALIANT, *supra* note 13, at 74 (explaining that the number of hypothesized states of the world captured in a model specification most often *exceeds* the number of available observations because additional observations lead to the introduction of additional variables that explode the hypothesis set); *see also* DOMINGOS, *supra* note 13, at 73-74 (noting this problem and explaining that researchers often submit their hypotheses (think candidate algorithms or model specifications) to hypothesis testing in order to reduce their astronomical number).

various assumptions.[61] It is true that greater stores of data can sharpen those models, but no algorithm situated within a changing environment can be purely data-driven.[62]

Recognition of this limitation, combined with growing stores of data, has begun to shift some of the research agendas of empirical legal scholars away from causal inference and toward predictive inference.[63] The goals of predictive inference are less ambitious and its methods are consequently less demanding. For instance, prediction may help a judge assess the probability that a person accused of a crime may flee if granted

[61] See, for example, the work of Susan Athey, which uses machine learning to model the characteristics of people as well as how those characteristics come about. Susan Athey, *Machine Learning and Causal Inference for Policy Evaluation*, 21 A.C.M. INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING, at 5 (2015), https://doi.org/10.1145/2783258.2785466.

[62] *Id.*; *see also* Levmore, *supra* note 12, at 615; Saul Levmore & Frank Fagan, *Competing Algorithms for Law: Sentencing, Admissions, and Employment*, 88 U. CHI. L. REV. 367, 404 (2021) (discussing the development of synthetic algorithms that leverage large data in combination with theoretical assumptions in order to infer counterfactuals). Researchers in computer science, notably Judea Pearl, are actively working toward supplying algorithms with the tools of causal reasoning, but this work is still in its early stage and its feasibility remains unclear. *See* Judea Pearl, Theoretical Impediments to Machine Learning With Seven Sparks from the Causal Revolution 1 (Jan. 15, 2018) (unpublished manuscript), https://arxiv.org/pdf/1801.04016.pdf (noting if counterfactuals can be made computationally tractable, then machines can be equipped with the tools of causal reasoning). On feasibility, *see* DOMINGOS, *supra* note 13, at 73-74 (discussing the difficulty of developing a powerful algorithm that can reason like a human, and highlighting the fact that simply by adding one variable to a dataset explodes the number of candidate hypotheses for explaining that data exponentially). Even equipped with causal reasoning, machine decision accuracy would still be plagued by unforeseen changed circumstances. *See infra* note 66 for an example.

[63] Much of this work can today be traced to computational legal scholars who use natural language processing to manipulate textual data found in judicial decisions; legislation, and its supporting textual documentation, such as congressional debate; regulation and public commenting; and other sources of legal textual information. For an excellent overview of this research, *see* LIVERMORE & ROCKMORE, *supra* note 2. For a review of the book, Fagan, *supra* note 23.

bail.[64] A typical bail-or-jail algorithm is built from a number of variables that describe a person's social and criminal histories.[65] While an algorithm may predict an increase in the probability of flight by several percentage points if the person has, for example, been convicted of a felony in the past, it cannot tell the judge that the presence of a felony conviction will *cause* flight. Nonetheless, prediction can obviously be useful to lawmakers inasmuch as it provides a basis for justifying legal rules even if that basis is not as robust as one grounded in causal inference. What might be expected then, as the popularity of predictive algorithms grows, is that law will become increasingly justified on the basis of probabilistic outcomes. But the probable effects of a legal change represent nothing new to law inasmuch as causal inference itself is susceptible to changing circumstances and innovation.[66]

2. Prediction's Domain

A model's capacity to predict is a direct result of its ability to discern a pattern and the stability of the environment in which that pattern occurs.[67] Only when these two conditions

---

[64] *See, e.g.*, Kleinberg et al., *supra* note 5, at 237-38 (developing algorithmic simulations that reduce crime rates holding jailing rates constant, or reduce jailing rates by holding crime rates constant); *see also* Himabindu Lakkaraju & Cynthia Rudin, *Learning Cost-Effective and Interpretable Treatment Regimes*, 54 PROCS. MACH. LEARNING RSCH. 166, 166 (2017); Jongbin Jung et al., *Simple Rules for Complex Decisions* 1 (Stanford Univ., Working Paper, 2017).

[65] The COMPAS algorithm, built by Northpointe, Inc., for instance, observes 137 variables related to social and criminal history. Alec MacMillen, *Can an Algorithm Identify Repeat Offenders?*, CHI. POL'Y REV. (Mar. 12, 2019), https://chicagopolicyreview.org/2019/03/12/can-an-algorithm-identify-repeat-offenders/.

[66] For instance, a causal inference of the impact of prison square footage and inmate violent crime may suggest the construction of large prisons in order to reduce inmate violent crime if supporting data were collected from 2005-2018, prior to a pandemic. Following a pandemic, if prisoners practice higher levels of social distancing regardless of prison size, the strength of the causal inference will be reduced. *See* Fagan, *supra* note 23, at 1402-03.

[67] Valiant refers to these, respectively, as the Learnable Regulatory Assumption and the Invariance Assumption. VALIANT, *supra* note 13, at 62.

are sufficiently met does the size of a data sample and the precision of models matter. Just as three or four observations of a stable game of chess will not predict much, one billion observations of an opening chess move are of no predictive value if the rules of chess are constantly changing. While a pattern may be discerned from the past moves, it is not predictive if placed within a different context. And while a different context may be captured by an additional variable, new stores of data must be accumulated when the context variable itself is new. This is the essential challenge of developing general AI. Adding one additional (even discrete) variable exponentially increases the number of candidate algorithms for explaining the state of the future world.[68] A data scientist who predicts a legal outcome on the basis of 20 Boolean (yes/no) case features must check 2^{20} possible configurations of the algorithm. That amounts to roughly 1 million configurations, or hypotheses, of the future. But the introduction of a new variable increases the number of possibilities to 2^{21}, or about 2 million. Adding one new Boolean feature *doubles* the number of possible algorithmic specifications. The challenge, as widely recognized by data scientists, is that oftentimes adding an additional observation introduces a new variable, which means that bigger data exponentially increases the number of candidate algorithms to evaluate.[69] For this reason, data scientists routinely impose closure on the possibilities that they need to check by means of hypothesis testing.

Given these limitations, lawmakers may seek to likewise close systems. In some legal domains, imposed (or guided) closure will make for good law inasmuch as it maximizes welfare or satisfies some other widely agreed upon normative criterion. For instance, the use of self-driving cars may be limited, by law, to a closed network in order to minimize

[68] See *supra* note 60.

[69] New variables arise inasmuch as observations are unique. People, as units of study for instance, are different. *See id.*

accidents and other problems associated with tort liability.[70] Once closed, patterns can be easily evaluated. More importantly, those evaluated patterns cannot be placed in different contexts absent systemic change. Their earlier predictive capability can therefore remain stable.

In other domains, adequate closure may already have taken place, but law may have been unable to take full advantage of data because of collection or processing limitations. For instance, various laws related to tort liability, such as an applicable standard of care, may be settled, but still applied with comparatively high judicial discretion or with flexible legal standards as opposed to rigid rules. As more data is collected and processed, patterns may be revealed. A discernable pattern, even if contingent on a dozen or even hundreds of variables, can serve as a rule. This is the promise of micro-directives[71] and personalized law.[72] Both are built from contingent features that tell lawmakers something about human behavior. These features may be environmental, such as whether it is raining or sunny, or personal, such as whether a driver is slightly drunk or has a headache.[73] Indeed, the driver's optimal standard of care in that

---

[70] *See* Mark Geistfeld, *A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation*, 105 CA. L. REV. 1611, 1612 (2017) (noting that a network operating system could guide the entire fleet of autonomous vehicles in order to resolve difficult tort questions). Similarly, Ben-Shahar and Porat note that coordination of autonomous driving at the "programing" level has the potential to improve coordination. *See* OMRI BEN-SHAHAR & ARIEL PORAT, PERSONALIZED LAW: DIFFERENT RULES FOR DIFFERENT PEOPLE 172-73 (2021). They cite the additional example of class actions in which the class membership is closed by a "pattern or practice" (established by a legal finding) even if individuals vary in injury and remedy. *Id.* at 173. *See* Int'l Bhd. of Teamsters v. United States, 431 U.S. 324, 336 (1977); United States v. City of New York, 717 F.3d 72, 82-85 (2d Cir. 2013).

[71] *See* Anthony J. Casey & Anthony Niblett, *The Death of Rules and Standards*, 92 IND. L. J. 1401, 1401 (2017) (theorizing that law will select from a catalog of context-specific micro-rules crafted on the basis of big data and artificial intelligence).

[72] *See* BEN-SHAHAR & PORAT, *supra* note 70 at 23-25.

[73] *Id*. at 24.

moment can be calculated on the basis of billions of observations of past behavior. In a closed system, law can fully benefit from this data. More data about drivers and their environments will yield more optimal standards of care precisely because they are, in that moment, comparable to that data.

Adequately closed legal domains provide an opportunity for creating better law.[74] But lawmakers must distinguish closed from open systems. Big data is not an adequate marker for assuming that a legal domain is closed and ready for detailed and contingent rules.[75] Once an environment

[74] Of course, patterns must be discernible. State-of-the-art machine learning requires 5,000 observations per category to achieve adequate performance, and roughly 5,000,000 observations to match or exceed human performance. IAN GOODFELLOW, YOSHUA BENGIO & AARON COURVILLE, DEEP LEARNING 20 (2016) ("As of 2016, a rough rule of thumb is that a supervised deep learning algorithm will generally achieve acceptable performance with around 5,000 labeled examples per category and will match or exceed human performance when trained with a dataset containing at least 10 million labeled examples.") In law, sample sizes are often much smaller. Consider that a recent study of recidivists in Wisconsin tabulated 2,379 over a three-year period. Only 840 were labeled as violent reoffenders. Only 111 were incarcerated for more than five years before paroled. This is far less data than what is needed. For this reason, state-of-the-art sentencing and bail-or-jail algorithms are synthetic, that is, they deploy assumptions and theory in order to hypothetically label data, which is then used for large-scale algorithm building. Of course these algorithms are only as good as their underlying theory. To gather more data and rely less on theory, we might wait several years, or perhaps a decade, and collect more observations; or we might rely on data from jurisdictions other than Wisconsin. However, the environment must be stable enough across time and space in order for this new data to be helpful. Someone who is paroled in 2015 must act in a sufficiently similar manner to someone a decade later if her data will help predict behavior in 2025. To the extent that future is unlike the past, the data will not be helpful. The same is true for space. If Texan parolees behave in systemically different ways than Wisconsin employees, say because Texas is hotter, or it has a better tracking system for those who have been released, then Texan data will be of less use for predicting recidivism in Wisconsin. On the persistence of limitations of sample sizes in law, see Levmore, *supra* note 12 at 615-16, 623-24.

[75] *See* Frank Fagan & Saul Levmore, *The Impact of Artificial Intelligence on Rules, Standards, and Judicial Discretion*, 93 S. CAL. L. REV. 1, 6 (2019) (noting that legal standards will prevail over rules when past is unlike the future and new variables appear over time).

introduces new and previously un-modeled variables, predictive rules can no longer be calculated against the backdrop of past environmental data unless the new variables can be safely ignored. Thus, legal domains of personalized law, micro-directives, and contextualized rules must satisfy the two conditions for machine learning to take place and for predictive law to be accurately applied: discernible regularity and contextual stability.[76] When the legal domain offers up new variables, including those that result from environmental dynamism and human innovation, personalized standards of care, mandatory contractual terms and disclosures, evolving community-based definitions of contractual counterparty good faith and unconscionability, contingent caps on pollution, and so on, cannot be calculated and expected to continue to yield accurate results unless the new variables have no bearing on the situation.

It bears reemphasis that simply adding to a model an additional variable, or set of variables, in order to perfectly predict a future environment may not be possible—not only because of the exponential increase in hypothesized states of the world—but also because a change may be entirely unexpected.[77] It is easy to see that even the most sophisticated hedge fund cannot perfectly predict a future asset price; the economy is an open system. Likewise, a legal or social change may only fully reveal itself to legal model builders in the future. In many scenarios, personalized law and micro-rules will do well for a time, though data and models will require periodic updating. The trick is to time the updates so as to maximize the benefits of AI-

[76] *See* VALIANT, *supra* note 13 at 61-62 (identifying these two conditions).

[77] *See* Fagan, *supra* note 12 (explaining the impossibility of modeling unknown unknowns).]*See also* Mirielle Hildebrandt, *Code-Driven Law: Freezing the Future and Scaling the Past, in* IS LAW COMPUTABLE: CRITICAL PERSPECTIVES ON LAW AND ARTIFICIAL INTELLIGENCE 67, 74-75 (Simon Deakin & Christopher Markou eds., 2020) (positing that unknown unknowns arise from cascading strategic interactions between people and make much of contingent law intractable).

enabled rules-based architectures, but humans will need to discern when AI requires overruling. Otherwise, errors will accrue when earlier models and new environments are mismatched.

3. Judicial Precedent

There are parallels to judging and judicial philosophy. Advocates of judicial precision and the stability of rules might suggest that judges should strictly adhere to precedent. Again consider Gilmore's 1974 Storrs Lecture, in which he describes that wish from the lawyer's perspective:

> Within the legal profession most practicing lawyers (who are interested in winning cases or in advising their clients in such a way that they don't have cases) prefer a formalistic approach to law. That approach holds out the promise of stability, certainty, and predictability—qualities which practitioners value highly. Judges, on the other hand, are paid to decide cases. Apart from such practices as bribery and corruption (which at times become institutionalized), judges want to decide the cases which come before them sensibly, wisely, even justly. Sense, wisdom, and justice are community values, which change as the community changes.[78]

This is an expression of a judicial form of the computational paradox. Judicial intuitions, hunches, and creativity around ideas of sensibleness and justice are innovative. While tethered to changes in community values and the dynamism of public sentiment, judges work within existing law in order to frame its novel extensions. Lawyers in search of certainty and precision may prefer static communities and less innovative judges. Good law balances this trade-off.

It was Cardozo who, having recognized the illusiveness of stability, certainty, and predictability (at least for law "in its

[78] GILMORE, *supra* note 28 at 17.

highest reaches"), grew to see the judicial process not as discovery and application of "correct" rules, but as innovation and creation.[79] But it should be noted that Cardozo reserved creation for the rare cases. Ample room for certainty and precision is available to the routine. He explains:

> Of the cases that come before the court in which I sit, a majority, I think, could not, with semblance of reason, be decided in any way but one. The law and its application alike are plain. Such cases are predestined, so to speak, to affirmance without opinion. In another and considerable percentage, the rule of law is certain, and the application alone doubtful. A complicated record must be dissected, the narratives of witnesses, more or less incoherent and unintelligible, must be analyzed, to determine whether a given situation comes within one district or another upon the chart of rights and wrongs. […] Finally, there remains a percentage, not large indeed, […] where a decision one way or the other will count for the future…[80]

Cardozo helpfully divides cases into three groups: predestined affirmances, doubtful application, and creation. Machines will provide judges with little advantage for the predestined affirmances. Where law and fact are certain, a judge of standard ability, however well-equipped with the latest legal tools, can easily decide the case because the merits of the case are clear. For these easy cases the state may even economize by deploying machines to assist the judge with more rapid decision-making or develop and incentivize the use of predictive tools to

---

[79] BENJAMIN N. CARDOZO, THE NATURE OF THE JUDICIAL PROCESS 166-67 (1967). Gilmore documents the scandal that Cardozo's recognition caused: "The thing that is hardest to understand about *The Nature of the Judicial Process* is the furor which its publication caused […] Cardozo's hesitant confession that judges were, on rare occasions, more than simple automata, that they made law instead of merely declaring it, was widely regarded as a legal version of hard-core pornography." GILMORE,*supra* note 28 at 77.

[80] *Id.* at 164-65.

encourage adversaries to settle.[81] It is when law or fact is uncertain that machines can help. Machines can help the judge ascertain a circumstantial record with predictive models. In other instances, machines might help judges think about future behavior in response to their rulings and suggest optimal precedent.

In common law jurisdictions, the parties themselves will likely drive the process more than the judge. Consider an algorithm that evaluates a set of facts surrounding an employment dispute and predicts that the defendant will prevail with a probability of 85% in front of a particular judge.[82] If damages, say in lost wages, are $100,000, then the defendant will offer 15% of those wages, or $15,000, as a settlement, especially if she believes that the judge will deploy the same algorithm.[83] Supposing that the plaintiff is also equipped with the algorithm and shares the same confidence in its capabilities, settlement will be reached. Nonetheless, as litigants rely more and more on algorithms to predict judicial decisions, there will be fewer decided cases for future litigants (and smarter algorithms) to review; and prior beliefs will remain

[81] On the benefits of developing tools to assist the judge and the broader merits of judge-machine partnerships, see Fagan & Levmore, *supra* note 75 at 7. On the development of tools for encouraging settlement, see Charlotte S. Alexander, Khalifeh al Jadda, Mohammad Javad Feizollahi & Anne M. Tucker, *Using Text Analytics to Predict Case Outcomes*, in LIVERMORE & ROCKMORE, *supra* note 2 at 305-07.

[82] *See, e.g.*, Alexander, al Jadda, Feizollahi & Tucker, *supra* note 81 at 306-08 (creating an algorithm to predict case outcomes of employment disputes).

[83] These figures set aside expected legal fees for representation and court filings. Of course in the American context this example assumes that both parties are able, financially, to conduct unconstrained discovery in order to supply the prediction machine with the facts of the case. It should be noted, however, that discovery costs will very likely fall as evidence becomes more susceptible to algorithmic processing and as the costs of that processing fall. *See* Engstrom & Gelbach, *supra* note 17 at 1050-51 ("as [AI tools] continue to proliferate and improve, … the discovery cost curve is likely to bend down more quickly than the digitization curve bends up).

undisturbed.[84] This is not a problem if cases are alike. However, precedent stocks will degrade in value to the extent that new disputes present undetected nuances or completely new sets of facts. Only in these novel scenarios—and the attendant appearance of new variables—will disputants and their lawyers be discouraged from settling and will precedent stocks buoy upwards until there is enough data to mine and encourage settlement anew.[85]

Any attempt by judges and other lawmakers to make the algorithm more precise by ignoring nuances or novel features of the case paradoxically generates error. It is true that discernible patterns that remain in stable legal environments will become increasingly mechanized, but novelty, or older patterns relocated to novel environments, will remain in hands of the human judge. Changing environments and human innovation make prior machine calculations obsolete. This is a built-in feature of technology that serves as a brake on machines overrunning human life and law. Creative judging works the same way. It serves as a brake on formalism.

B. Enforcement and Rights

1. Behavioral Predictions and Enforcement

In addition to lawmaking, the computational paradox is present in enforcement and rights allocation. Consider the basic decision of a parole commissioner to grant an inmate's petition

[84] Note that the above example considers the use of prediction algorithms for disputes decided by a given judge. It is true that the problem can increase in complexity once forum shopping is introduced, but if parties select fora for victory, then precedent stocks for building accurate algorithms in unselected fora will diminish. *Id.* at 1066-67.

[85] For this reason, artificial intelligence will do better in legal domains with less change than others with more change. *See* Fagan & Levmore, *supra* note 75 at 1; Dana Remus & Frank Levy, *Can Robots Be Lawyers?: Computers, Lawyers, and the Practice of Law*, 30 GEO. J. LEGAL ETHICS 501, 538, 541 (2017) (noting machine learning's limits to make sense of "unanticipated contingencies").

for early release. In California, parole commissioners base their early release decisions on whether they believe an inmate presents a public safety risk. This belief is predicated on a state psychiatric evaluation and an interview.[86] With recent advances in machine learning, analysts can now identify phrases uttered by inmates during parole hearings that predict whether commissioners choose to grant or deny parole.[87] Inmates and their attorneys, armed with the predictive algorithm, might choose language to mislead the commissioners into thinking that they will not risk harm to the public.[88] This possibility will be discussed below, but suppose for the moment that only the commissioners can use the algorithm. Perhaps public use of the algorithm is prohibited by law or its cost is prohibitive to inmates. Suppose further that the algorithm predicts public safety risk on the basis of the inmate's interview language as well as other factors related to criminal and social history.

Armed with the tool, it is easy to see that the commissioners would want to hold language patterns constant in order to preserve its predictive power, perhaps by limiting interviews to a handful of pre-determined questions like a COMPAS questionnaire. If the interview is wide-ranging and inmates provide answers to new questions with new vocabularies, then commissioners will have greater difficulty predicting public safety risk with current models. Interview questions and language, however, usefully evolve for predicting flight risk. Few would argue that the benefits of algorithmic precision outweigh the benefits of increased learning from asking new interview questions and allowing for expansive linguistic variation in discussions with inmates. This is a

[86] *See* In re Shapitus, 190 P.3d 573 (Cal. 2008) ("[T]he paramount consideration for both the Board of the Governor under the governing statutes is whether the inmate currently poses a threat to public safety.")

[87] Laqueur & Venancio, *supra* note 2 at 193. The predictive model of Laqueur and Venancio achieves an accuracy rate of 66%. *Id.* at 223.

[88] This example of bad innovation is discussed below at notes 92-96 and accompanying text.

relatively straightforward example that demonstrates an absurd area for computational closure.[89] On the other hand, closure may be worthwhile if inmate behavior captured by questions and interview language evolve slowly enough to allow for updating the algorithm at a sufficiently low cost. If so, then precision and innovation benefits can both be efficiently captured. The point is to see that the tradeoff persists and that the only way to capture the innovation benefits is to create an open architecture that leaves some amount of room for change.[90]

When considering the trade-off, California may believe that interview language in parole hearings evolves too quickly

[89] One only needs to think of the evolution of slang for numerous examples of rapid evolution in meaning. *See* MAX DÉCHARNÉ, VULGAR TONGUES: AN ALTERNATIVE HISTORY OF ENGLISH SLANG (2016). See also HERBERT MARCUSE, ONE DIMENSIONAL MAN: STUDIES IN THE IDEOLOGY OF ADVANCED INDUSTRIAL SOCIETY 91-92 (2d ed. 1991), which provides an example with the nouns "freedom," "equality," "democracy," "and "peace." Writing in 1964 before the dissolution of the Soviet Union, Marcuse observes that in the West, these words are associated with "free enterprise, initiative, elections, [and the] individual." *ID.* at 91. In the East, they are associated with "workers and peasants, building communism or socialism, [and] abolition of hostile classes." *ID.* Even though the words are the same, their meaning is contingent upon space and time.

[90] Suppose interview language matters little to the parole decision and closing off behavioral and linguistic innovation presents few costs. In other words, the interview would receive little weight in a model that predicts public safety risk. Today in California, the psychiatric test is used as the primary screening tool. Each inmate receives a risk assessment, valid for five years, expressed in one of three categories: low, moderate, and high. Laqueur & Venancio, *supra* note 2 at 202. As might be expected, the inmate's psychiatric risk score is highly statistically significant for predicting release. These are the easy cases like Cardozo's pre-determined affirmances. For example, Hannah Laqueur and Anna Venancio find that 56% of 1,673 inmates who are assigned a low risk score are granted parole. This proportion stands in contrast to a grant rate of less than 3% of 1,193 inmates with high or moderate risk scores. *Id.* at 207. These statistics show that for the easy cases, words uttered during the interview do not matter. If the inmate is of high or moderate risk, it is very unlikely that he or she will be granted parole no matter what is said to the commissioner. Thus, it is for the low risk cases, in the California parole context, where words can be decisive. Laqueur and Venancio accordingly set the high and moderate risk inmates aside, and then examine the conversations between low risk inmates and commissioners. *Id.* at 222.

and the costs of periodically updating the predictive algorithm are too great. In response, commissioners might retreat to something broader, like language style. In addition to identifying specific phrases predictive of release, analysts have found that comparatively high usage of verbs and nouns by prisoners, as opposed to adjectives and adverbs, also lead to grants of parole.[91] But now commissioners, armed with the looser and underfitted predictive model, would want inmates to at least maintain consistent style and grammar over time in order to maximize its predictive capabilities. No matter how indirectly the predictive model relies on inmate language, the algorithm demands that the relationship between inmate language and future behavior remain sufficiently consistent over time in order to maintain predictive accuracy. Greater dynamism requires more frequent updates. In terms of economic costs and benefits, the computational paradox presents a clear trade-off. If a predictive domain in law evolves slowly, as might be expected with inmate language and recidivism, then the benefit of using a predictive model is likely great. Errors generated by dynamic inaccuracies will be few and the periodic cost of updating will be low. As dynamism increases, errors and updating costs increase. At some threshold, the costs are prohibitive, and modeling is not possible or worth it.

An Aside on Cheating

Consider now the possibility that inmates are armed with the algorithm. Even if today's tools are imperfect, it is easy to imagine technological advances that arm the inmate with a substantial capability to persuade commissioners in hearings, especially in situations where the interview matters.[92] As the

[91] Verbs and nouns are perhaps more indicative of a clearer statement of facts, which may make an inmate's attempt at expressing remorse appear more objective and sincere. Adjectives and adverbs in contrast may suggest greater subjectivity and signal an inability to empathize. *Id.* at 228.

[92] One can imagine the unprincipled use of successful phrases in other areas of legal practice. Litigants in oral argument might possess finely

predictive technology continues to sharpen, parole boards might grow concerned that clever attorneys unrestricted by law or professional responsibility rules will coach their clients to use successful vocabularies regardless of the inmate's expected behavior.[93] In the extreme case, we can imagine that the inmate's selection of language bears no relationship to future inmate behavior, but even the more likely intermediate cases could deceive the undiscerning commissioner.[94]

Deception, however, is only a winning strategy for the short- and medium-term, since subsequent rounds of machine learning will show that the use of particular words are no longer predictive of good behavior. This is an important point and worth dwelling on. Suppose that by 2030, machine learning tools are able to predict with 90% accuracy that low-risk inmates who use a certain vocabulary (and avoid another one) increase their chances of parole by 20%. Suppose further that in 2030, 1,000 inmates are interviewed by the parole board. Without machine learning tools, 560 would be released (if release rates remain consistent with the 2011-14 rates observed in earlier

---

grained data on what pleases a particular judge. Parties to a trial may attempt to successfully coach a witness with specific words that cue credibility. And of course brief writing may attempt ideal word selection much the same way the appellate advocates of today attempt ideal case citation. On writing style and citations *see* Tippet et al., *supra* note 2 at 1157.

[93] Model rules of professional conduct could be implicated. Lawyers, for instance, may refuse to offer evidence that they reasonably believe to be false. MODEL RULES OF PROF'L CONDUCT R.3.3 (2020). For an opposite view, where lawyers are ethically required to exercise technological competence with respect to artificial intelligence, see Mark L. Shope, *Lawyer and Judicial Competency in the Era of Artificial Intelligence: Ethical Requirements for Documenting Datasets and Machine Learning Models*, 34 GEO J. LEGAL ETHICS 191, 191-92 (2021).

[94] Contemporary linguists have explained that the meaning of language is socially constructed, which implies that the meaning of signifiers like words or voice intonation can depend on human relationships. *See, e.g.*, FERDINAND DE SAUSSURE, COURSE IN GENERAL LINGUISTICS 15 (Perry Meisel & Haun Saussy eds., Wade Baskin trans., Colum. Univ. Press 2011) (1959). *See* Fagan, *supra* note 23, at 1406-07 for additional discussion.

studies).[95] Let us assume that of those 560, 460 do not recidivate, but 100 are released in error. Now with machine learning, inmates are able to increase their chances of parole by 20%. However, the number of errors is not evenly distributed across recidivists and non-recidivists. Non-recidivists use the predictive vocabulary without coaching. Their words better reflect their future behavior. After all, the predictive model was built with their language. It is the recidivists who are now deceptively using the winning vocabulary. Errors fall more heavily on them. In subsequent rounds of model building, predictive words like "thank" and "care" will now be less predictive of non-recidivism and commissioners will rely on that language less. Future recidivists and their attorneys, aware that they can no longer fool commissioners, will forgo the use of machine learning if only to save costs and time.

Today, California "looks most favorably on clients who are respectful, penitent and truly appear to be remorseful."[96] It should be expected that commissioners will need to become more discerning as tools for eluding their conventional judgments grow in strength. California's valuation of precision will drive its decision to equip the commissioners with the same machine (and skill to use it).[97] Once equipped, they will more

---

[95] *See* Laqueur & Venancio, *supra* note 2 at 207.

[96] POST-CONVICTION JUSTICE PROJECT, USC GOULD SCH. L., SUMMARY OF THE CALIFORNIA PAROLE PROCESS: PAROLE ATTORNEY VERSION 9 (2016), https://pcjp.usc.edu/wp-content/uploads/2019/09/364d6-parole-manual-and-case-law-chart.pdf (last visited June 8, 2021). The Summary states: "It is crucial that [inmates] not seem hostile, disrespectful, or defensive. […] If your client maintains their innocence, it is possible for them to be remorseful about what happened to the victim even if they are not responsible for it." *Id.* In addition, while the parole board cannot require an inmate to admit guilt as a condition of parole, the Summary notes that it is more difficult to express remorse for something an inmate claims that he did not do. *See also* CAL. PEN. CODE § 5011(b) (West 2019).

[97] California might also respond with a ban on the use of machine learning for determining the words and phrases used during interviews that predict release—a strategy of closure. Perhaps it does not wish to purchase the software used by defendants, which would uncover their strategic word choices, or is unable to train its commissioners to use the software in a cost-

easily observe that the inmate and his attorney are strategically deploying empty words, and both sides will more readily search for other signals. Thus, a prohibition on the use of the word-choice algorithm by inmates and their attorneys provides few, if any, long-term social benefits. A better policy may be to arm the public and commissioners with open access to data and data-processing tools.

2. Automated Rights Allocation

Like their human counterparts, machines suggest early release on the basis of data. If data is protected, then people have greater control over what the machine can and cannot see. The European approach to data protection is often seen as furthering a social preference for privacy,[98] but it is partly grounded in a suspicion towards machine-based reasoning and decision-making. Early legislation, such as the French 1978 law on "computing, files, and liberties" was less concerned with privacy and more concerned with abuses of automated profiling and personality screening used for rights determinations.[99] Article

effective way. With a ban in place, California may believe that it can preserve the commissioners' ability to precisely identify public safety risks during parole hearings and encourage honesty. In the long-run, this seems extreme and contrary to notions of adversarial justice. After all, if dangerous inmates begin to cheat the hearing process by fooling commissioners with empty words, then future rounds of machine learning will show that those words are bad predictors of public safety risk as shown above.

[98] *See generally* Charles Whitman, *The Two Western Cultures of Privacy*, 113 YALE L. J. 1151 (2004).

[99] *See* Loi 78-17 du 6 janvier 1978 relative a l'informatique, aux fichiers et aux libertés [Law 78-17 of January 6, 1978 on Information Technologies, Datafiles and Civil Liberties], Journal Officiel De La République Française [Official Gazette of France], Jan. 7, 1978, p. 227. "Aucune décision de justice impliquant une appréciation sur un comportement humain ne peut avoir pour fondement un traitement automatisé d'informations donnant une définition du profil ou de la personnalité de l'intéresse." ("No judicial decision involving an appraisal of human conduct may be based on any automatic processing of data which describes the profile or personality of the person concerned."). *Id.* art. 2. The rule applied to administrative decisions as well. *Id.* Other European jurisdictions have codified similar rules, which led to a European-wide directive that obliged Member States to

22 of the General Data Protection Regulation (GDPR) preserves the purpose of that law today by providing European residents with recourse to human review of machine decisions.[100] The California Consumer Privacy Act of 2018, passed two years after the GDPR, also demonstrates a concern with profiling and automated decision-making in commercial transactions, even if it stops short of guaranteed human review and instead provides for an opt-out.[101] These rules can be understood as safety valves. Human review and opt-outs express open architectures that suppress errors potentially generated by models built from past behavior—prior to an innovation that takes place in the future.

---

create a right so that European residents would not to be subject to a decision which produces legal effects concerning [them] or significantly affects [them] and which is based solely on automated processing of data intended to evaluate certain personal aspects relating to [them], such as [their] performance at work, creditworthiness, reliability, conduct, etc. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, art. 15, 1995 O.J. (L 281) 43. The Directive has since been repealed.

[100] *See* Regulation 2016/679, of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), art. 22(1), 2016 O.J. (L 119) 46 [hereinafter GDPR].

[101] CAL. CIV. CODE § 1798.185 (a)(16) [hereinafter CCPA]. While the relevant policy instrument is consent, a concern with automated decision-making is manifest in several sections of the CCPA. Profiling is defined as "any form of automated processing of personal information, as further defined by regulations pursuant to paragraph (16) of subdivision (a) of Section 1798.185, to evaluate certain personal aspects relating to a natural person and in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location, or movements." *Id.* at § 1798.140 (z). Californian regulators are then tasked with: Issuing regulations governing access and opt-out rights with respect to businesses' use of automated decision making technology, including profiling and requiring businesses' response to access requests to include meaningful information about the logic involved in those decision making processes, as well as a description of the likely outcome of the process with respect to the consumer. *Id.* at § 1798.185 (a)(16).

Prediction examines the past in order to make sense of the future. When a person is algorithmically profiled in the present, their profile is generated on the basis of the past behavior of others. Even if the people of today share features and behaviors with the people of the past, the algorithm will be accurate only to the extent that the present resembles the past. Any innovative behavior found in contemporary life or new features of present-day environments may be enough to generate algorithmic error. Statutory provisions for human review and opt-outs implicitly recognize that things can change. They provide for automated rights determinations within a closed algorithmic environment but provide a secondary layer of open decision-making for additional protection against error. These are straightforward examples of how law attempts to address the computational paradox.

Nevertheless, automated profiling (and the attendant, sometimes intrusive, surveillance that often makes it possible) strongly implicates long-standing constitutional norms.[102] These norms are hardly seen as technocratic tools for error minimization and mechanisms for quality control for out-of-date models and stale data. Indeed, normative recalibrations that tilt more towards administrative law principles are often predicated on precision and efficiency.[103] But as already seen with the

[102] Surveillance implicates privacy norms and profiling implicates due process and equal protection norms. Aziz Z. Huq, *Constitutional Rights in the Machine-Learning State*, 105 CORNELL L. REV. 1875, 1905, 1919 (2020) (noting that privacy, due process, and equality norms are consistently implicated in the state's use of predictive machine learning tools). *See* CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION (2016) and FRANK PASQUALE, THE BLACK BOX SOCIETY (2016) for arguments against the use of algorithmic surveillance tools.

[103] *See* Huq, *supra* note 102 at 1875 (suggesting a recalibration and retooling of fundamental norms by means of a new mix of regulation to avoid algorithmic harm coupled with aggregate litigation to redress wrongs at the group (as opposed to individual) level). Similarly, David Engstrom and Daniel Ho have suggested that administrative law, as opposed to constitutional law, will increasingly be used to hold the state accountable for its enforcement activities. *See* David Freeman Engstrom & Daniel E.

parole example, precise and low-cost rights allocation with machine learning is susceptible to the computational paradox. One implication is that greater frequency and severity of the paradox induces greater demand for straightforward constitutional protections.

Consider, for instance, machine-enabled strategies for reducing the investigation costs of enforcing tax fraud.[104] By combing through millions of tax returns, the Internal Revenue Service can identify and triage filers who are more or less likely to have committed fraud. Identification and triaging is efficient because it allows the IRS to smartly allocate investigation resources to those filers who are more likely to have fraudulently reported their tax liabilities. Perhaps instead of employing 100 people in order to monitor 100,000 returns, the IRS can employ 10 people and deploy a small algorithm development team.

But suppose instead that the IRS deploys an algorithm to identify tax fraud without providing a safety valve for human review, or with a very limited one based upon new (and stronger) administrative law rules, after observing that the algorithm has been extremely successful in identifying fraudulent activity over the past five years. Not only does the state economize on investigation resources, but it additionally economizes on enforcement resources by relying more heavily on the algorithm. Accuracy may only last for a time, however, as people and their environments change. On the one hand, bad innovation that helps dishonest people circumvent tax rules will lead to activity being diverted away from surveilled locations in response to the algorithm. On the other, benign innovation will lead to new patterns of behavior, or identical patterns in new environments, previously undiscerned by machines. The state's precision

---

Ho, *Artificially Intelligent Government* in RESEARCH HANDBOOK ON BIG DATA LAW 57 (Roland Vogl ed., 2021).

[104] *See, e.g.*, JANE MARTIN & RICK STEPHENSON, INTERNAL REVENUE SERV., RISK-BASED COLLECTION MODEL DEVELOPMENT AND TESTING 141, 142 (2005), https://www.irs.gov/pub/irs-soi/05stephenson.pdf.

benefits will erode. Meanwhile, errors generated by stale and incomplete computational profiling (and related tools such as social credit scoring) will encourage greater demand for basic constitutional protections of privacy, due process, and equal protection,[105] which serve as brakes on the formalism of predictive models.[106]

3. Between Constitutional and Administrative Norms

Insofar as constitutional norms are seen as safeguards against formalist machine error, their resultant balance with administrative norms can be understood as largely shaped by a "computational legal ethics" that considers (i) the efficiency gains from lower cost machine-based implementations of

[105] This demand partly explains the emergence of Article 22 of the GDPR, which provides for a conditional right to human review of machine decisions. *See* GDPR, *supra* note 100, and accompanying text.

[106] This is true even if the profiling and social credit scoring takes place within the context of an exchange, as it does when the state requires a background investigation prior to granting employment, benefits, or other rights. For a different view, see Michael C. Munger, *Evolutionary or Not, Exchange Is Just*, 28 SOC. PHIL. & POL'Y 192, 192 (2011) (arguing that exchange, even if lopsided, "may be the only means by which people in desperate circumstances can improve their position"). To see the tension, imagine that a potential government employee owns all of her data. She possesses, thanks to agreements with her private telecommunication provider, the entire history of her network messaging and internet browsing actions. In addition, she owns her complete fitness and sleep information by means of an agreement with her favorite health-wristband operator, as well as her complete medical history through mandated access to centralized health records. For the sake of the argument, assume, too, that she owns her complete purchasing history via agreements with banks and other payment processors. If the state demands access to all of this data in exchange for a screening interview, she may very likely characterize this demand as an imposition, especially so if she has few or no other options for employment or income. Imagine further that all sources of income, including from private employers, public assistance, or even contractual access to a family trust, require her to disclose her data. The only option left in this hypothetical scenario is the foregoing of income, which will appear to her as an imposition as opposed to a choice. As voluntary exchange appears to be more characterized as an imposition, constitutional protections and norms become more important to her, especially so if she believes the algorithm does not capture her true characteristics.

enforcement and rights allocation, and (ii) how much weight society wants to give to the novelty and innovation that outpaces aging machine models. In general, this ethical frame will be the outcome of a conflict between hardened empiricism and resultant technocratic rules on the one hand, and community-based rationalization anchored in intersubjective dialog and exchange on the other.[107]

Consider again the GDPR. On the one hand, it is clearly based upon a fundamental right to privacy. As a fundamental right, privacy no longer depends on legitimization through continual dialog and norm mediation among people.[108] Privacy and data protection in Europe are empirically sanctioned on the basis of shared history and experience. At the same time, the GDPR legitimizes machine action by providing for its human review.[109] This important provision, which creates an intersubjective forum for human and machine dialog, can be understood as an enlargement of the public sphere consistent with familiar institutions that legitimize and rationalize action by expanding democratic participation even if the work now includes greater levels of machine input.[110] Again, this form of

[107] Examination of the modes of societal rationalization has largely been the domain of professional sociologists, and the ensuing section will draw on their work. *See* JÜRGEN HABERMAS, THE THEORY OF COMMUNICATIVE ACTION, VOLUME ONE: REASON AND THE RATIONALIZATION OF SOCIETY 7 (Thomas McCarthy trans. 1984) [hereinafter COMMUNICATIVE ACTION] (noting that his theory of communicative action is meant to "take up, once again…the problematic of societal rationalization, which was largely ousted from the professional sociological discussion after Weber"). American theorists such as Richard Posner have unfairly criticized Habermas for generating polarizing views. *See* Richard A. Posner, *Legal Scholarship Today*, 45 STAN. L. REV. 1647, 1651 (1993). Habermas is at his best when read descriptively.

[108] The GDPR, by its own terms, is based upon Article 8(1) of the Charter of Fundamental Rights of the European Union and Article 16(1) of the Treaty on the Functioning of the European Union. *See* GDPR, *supra* note 100 art. 1. Both enshrine a right "to the protection of personal data concerning him or her." *Id*.

[109] *See* GDPR, *id.* art. 22(1).

[110] *See* Fagan & Levmore, *supra* note 75 at 1 (exploring the benefits of human-machine partnerships in legal decision-making). A similar, relatively

intersubjectivity (here, the human review of machine decisions) can be understood as an example of an open legal architecture that provides a safety valve for unanticipated innovations in behavior and the legal environment.[111] Precision benefits are captured with settled models and past data, while errors are minimized with extended dialog and human input. Multiple requests for human review of a particular prediction system, will, in all likelihood, serve as an early warning for a need to update the existing model and data.

Broadly speaking, one might wonder whether the jurist is firmly in control of the empiricism, or will predictive science overwhelm the bar.[112] Empiricism is clearly on the move, but it

---

recent example can be seen in Europe's approach to the discipline of "law and economics." The European Doctorate in Law and Economics, which was funded by the European Union in its inaugural year in 2010, implemented the explicit scientific profile of balancing efficiency with other societal goals. Manifestations of efficiency, often empirically grounded in the law and economics literature, are balanced against other societal aims such as equity and justice, which are often inter-subjectively established.

[111] Lawmakers may, in the future, attempt to hold automated decision-making and other legal technologies in check by other means. *See* Aziz Z. Huq, *A Right to a Human Decision*, 106 VA. L. REV. 611, 611-12 (2020) (asserting that in contrast to human review, the "limits to machine decision making are appropriately found in the technical constraints on predictive instruments"). Whether those means consist of fundamental rights or rigid forms of evaluation such as "[un]flawed training data and common standards of industry performance," they represent empirical rationalizations and stand in contrast to alternatives grounded in intersubjective deliberation between humans and their machines. *Id.* at 687. Huq's grounds for empirical legitimization may appear less authoritative than those of fundamental rights—given that we have accumulated comparatively less experience with training and evaluating the performance of our machines—but they are empirical in character nonetheless. One meaningful difference is that a fundamental right provides a person with a protection that is less contingent on machine capabilities and tilts more towards a hard rule.

[112] My use of the term bar here is expansive and inclusive. It is a social unit that determines the outcome of legal practice and consists of lawyers, judges, law professors, other scholars, police, lawmakers, and yet others. Comparativists might recognize parallels to what Rudolfo Sacco has referred to as "legal formants". *See* Rudolfo Sacco, *Legal Formants: A Dynamic Approach to Comparative Law* 39 AM. J. COMPAR. L. 1, 22 (1991) (positing that legal rules are formed from many elements).

is important to recognize that there is an emerging form of rationalization and legitimization that is more flexible than hardened empiricism, albeit less flexible than limitless and fully open intersubjectivity. It acknowledges that computational precision and exactitude are constrained by possibility and change;[113] that we have entered a permanent state of innovation;[114] and perhaps more controversially, that we live today in a world where technical processes increasingly appear to precede social ones and impose themselves thereupon.[115] These are the core assumptions of a computational legal ethics, and they will be discussed further in Part II. For now, it is sufficient to see that these assumptions promise to expand the bases for legitimizing the use of predictive learning in law, precisely because they temper predictive technology's assertiveness and sharp edges. This is especially true in fast-changing and dynamic legal domains. When law recognizes that things change, and that a strong undercurrent of empirical imposition can therefore generate error, then it is more likely to sidestep strategies of closure and more confidently expand its use of computational learning by focusing on sufficiently static domains and deploying open architectures.

C. Lawyering

1. Writing Style

As already noted, the computational paradox presents less of a problem when sufficiently identical prediction algorithms are deployed by adversaries.[116] Symmetrical competition erodes one-sided computational advantage because the judge can more easily observe when a party's use of machine prediction is erroneous. The parties do the uncovering work. Competitive revelation limits the judge's reliance on

[113] *See generally supra* Part I.
[114] *See infra* Part II.C.
[115] *See infra* Part II.C.2.
[116] *See supra* notes 92-95 and accompanying text.

computational learning in turn and can weaken the real effects of the paradox. By contrast, when parties are asymmetrically financed, and competitive pressures with respect to algorithmic precision are low—as they may be when an inmate without an algorithm is pitted against an algorithmically equipped state parole board—then the paradox presents more of a problem because the weaker party is less able to challenge the algorithm's validity and demonstrate that its predictions are stale.

By way of illustration, consider a recent study carried out by Elizabeth Tippet et al. that examines writing style in federal employment litigation.[117] Cardozo would likely say that this type of dispute falls within his second category; the law is mostly settled and easily applied, but often the larger challenge for the judge is to discern the facts.[118] Ideally, the judge would simply grasp the true facts and apply the law. In practice, the judge must establish the facts from a circumstantial record. The good lawyer responds with skillful portrayal and contextualization, and that presentation can be enhanced with a machine.

In the aforementioned study, the research team examined 864 disputes and found evidence that a particular writing style is correlated with greater summary judgment success.[119] The successful writing style was characterized by frequent use of hedging words like "regardless" and "however." It deployed positive intensifiers such as "unmistakable" and "hastily";

[117] *See* Tippett, Alexander, Branting, Morawski, Balhana, Pfeifer & Bayer, *supra* note 2 at 1159.

[118] Of course, some areas of employment discrimination law present uncertainty. *See*, *e.g.*, Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671, 704 (2016) (noting circuit court differences in application of the business necessity defense). However, discernment of facts remains the primary challenge for case disposal. *Id.* at 696 (noting that factual proof of discrimination is difficult to discern because of masking and pretext, especially when it is reflected in years of historical data).

[119] *See* Tippett, Alexander, Branting, Morawski, Balhana, Pfeifer & Bayer, *supra* note 2 at 1166.

negative intensifiers, including provocations like "woefully" and "frivolous"; as well as repetition words such as "again" and "also."[120] The study also found—perhaps unsurprisingly for a skilled litigator—that longer briefs on average performed better than shorter ones.[121]

While the researchers clarify a number of limitations to their results and explain how a careful empiricist might poke holes and invalidate their conclusions,[122] let us assume that a sustained research effort is able to identify the writing style that maximizes the chances of adversarial success. Perhaps the attorney furnishes the machine with the known facts of the case including all circumstances, as well as "meta-data" that helps predict how the facts will be interpreted, including, for instance, the name and track record of the presiding judge, and even various obscurities—apparently decisive on occasion!—such as what the judge ate for breakfast.[123] In other words, assume the machine is able to soak up every variable in the environment and predict the winning writing style.[124] It is the ideal writing style machine.

---

[120] *Id.* at 1186-88.

[121] *Id.* at 1191.

[122] *See id.* at 1192 (noting that the 864 cases were imbalanced, that is, more were decided in favor of defendants as is usually the case for employment litigation).

[123] *See* Alex Kozinski, *What I Ate for Breakfast and Other Mysteries of Judicial Decision Making*, 26 LOY. L.A. L. REV. 993, 993 (1993) (noting that "[i]f the judge has a good breakfast and a good night's sleep, he might feel lenient and jolly, and sympathize with the downtrodden," but also noting the frivolousness of the theory). However, there may be some truth to the idea. *See* Daniel L. Chen, *This Morning's Breakfast, Last Night's Game: Detecting Extraneous Factors in Judging*, (Inst. for Advanced Stud. in Toulouse (IAST) Working Paper No. 16-49, 2016) (finding a 1.5% increase in the probability of asylum grants if the NFL team of the court's city won the night before).

[124] Assume, too, that the prediction algorithm was built from complete data, that is, past writing styles were observed alongside every environmental variable that affected the judge's decision.

When law is certain, but facts are not, the machine's usefulness depends upon how difficult it is for the judge to discern the facts. The plainer the circumstances, inferences, and cues, the less impact the lawyer's brief will make. Even the best writing cannot reveal (or obscure) what the judge can easily see. Given the results of the Tippet et al. study, this seems to be a reasonable assumption to make.[125] However, when facts are less clear, a lawyer's writing ability can contextualize the facts to her client's favor and steer the judge toward a winning result.[126] The machine's usefulness, however, must be compared to what the profession already has, i.e. legal research and writing professors who, among other things, teach lawyers how to use positive intensifiers, repetition, and hedging words.

In order to accurately test the hypothesis that the machine truly offers an advantage, we would need to test it with two lawyers of identical writing ability—maybe proxied by comparing half of a 1L writing class of a particular law school against the other half—allowing just one side to use the machine. If, all other things equal, the machine-equipped side consistently triumphs over the other, then we might say the machine is useful. This seems unlikely. After all, legal research and writing professors teach their students to do the same things as the machine: hedge, positively intensify, and use repetition. Put differently, the machine is doing nothing more than what humans already do. Research on legal writing style so far has not uncovered any hidden connections between successful outcomes and a peculiar, lesser known writing style.[127] This means that, at the moment, the best use of the ideal writing-style machine is to assist poor writers and place them on more equal

[125] *See* Tippett et al., *supra* note 2, at 1191 (noting that some of the results could be traced to the underlying merits of the dispute).

[126] *Id.* at 1191.

[127] Note, however, that a machine which could reveal an optimal writing style given things the lawyer cannot observe directly, such as what the judge had for breakfast, could be useful precisely because it could reveal the impact of hidden variables.

footing with superior writers.[128] Careless writers who inconsistently deploy hedging words, positive intensifiers, and repetitive phrases benefit from a machine that alerts them of those deficiencies just as poor spellers benefit from the red lines of a spellcheck tool.

Suppose further research uncovers a quirky writing style that comparatively persuades a judge more strongly than good use of intensifiers, repetition, and word counts. Perhaps the algorithm relies heavily on meta-data and demonstrates that a particular judge favors a given rhetorical style. Say the judge was a fan of Lincoln throughout high school and systemically favors briefs that remind her of Lincoln's debates with Douglas. Use of the Lincoln writing style will matter little to the outcome of the dispute if both sides deploy it (insofar as the style is neutral to status as a plaintiff or defendant and other asymmetrical features of the case). The judge will simply be

[128] Indeed, the researchers of the employment litigation study advocate wider use of machines in lawyering in order to level the playing field and provide wider access to justice. *See id.* at 1192. On the other hand, patterns of competition in computational analytics and constraints presented by data access—and resultant asymmetrical benefits from legal data economies of scale—may continue to generate uneven benefits across litigants according to their wealth in the short to medium term. *See* Frank Fagan, *Standardized Data Collection: Legal Requirements, Guidelines, or Competition?*, 2 GUJARAT NAT'L L. U. L. & ECON. REV. 69, 70 (2019). Nonetheless, widespread availability of predictive tools is likely in the long-term as costs continue to fall. While electronic access to public case-law is still a commercially viable product, and archived court filings are even more costly to access, there is growing support for low-cost access to data and the construction of low-cost natural language processing tools. *See* Brief for 36 Computational Law Scholars as Amici Curiae Supporting Respondent, at 5-6, Georgia v. Public.Resource.Org, Inc., 590 U.S. __ (2020) (No. 18-1150). Restrictions to legal data seem to be on the wane, and it should be expected that more data will become available over time. Open access to case law is readily available. *See Data Coverage — What's in CourtListener?*, COURTLISTENER.COM, https://www.courtlistener.com/coverage/#opinions (last visited June 2, 2021). Open access to federal court documents has been proposed. *See* Open Courts Act of 2020, H.R. 8235, 116th Cong. (2020) (a proposed bill that would eliminate fees for accessing federal court documents). *See infra* note 138 and accompanying text.

affected equally (in terms of style) by the identically styled briefs of plaintiff and defendant.

Returning to the chess example, when identical algorithms are pitted against each other, victory will be determined by the rules of the game.[129] In chess, first-mover white has an advantage over second-mover black based upon the rules of chess, not the independent capabilities of the players. Clearly, if black uses a superior algorithm, then black retains an advantage. But it should be expected that the economic motivations of algorithm developers will drive costs toward zero, and both sides will easily equip themselves with the best tools. Advantages only accrue to the extent that adversaries use algorithms of variable quality and worth, but in the long-run, it is more likely that both sides will be able to equip themselves with tools of comparable quality.[130] Judges, meanwhile, will be less susceptible to the algorithm's influence because adversaries are able to expose each other's doubtful predictions on the basis of symmetrical capabilities. As a consequence, the effects of the computational paradox are less pronounced in a world of competing algorithms, especially when both sides are comparably equipped.

The paradox is more of a problem when algorithms are deployed in asymmetrical competitions, as is presently the case when a criminal defendant faces the state and its algorithm. Today, no one meaningfully competes with the state's algorithms. Adversarial parties to criminal sentencing, parole determinations, bail-or-jail decisions, and so on, have mounted

[129] *See* Ernst Zermelo, *On the Application of Set Theory to the Theory of the Game of Chess* [1913] *in* READINGS IN GAMES AND INFORMATION 79 (Eric Rasmusen ed., 2001) (observing this eventuality first and giving birth to modern game theory).

[130] In the event that the newly discovered style systemically favors plaintiffs or defendants, then the judge would learn of the bias as awareness proliferates—perhaps with the help of further academic study and alerts from the bar. For these reasons, any one-sided rhetorical advantage is severely reduced over time.

attacks on the basis of constitutional rules and have not (yet) presented competing algorithms of their own.[131] Personal features and the nuances of a case continue to predict behavior with no real challenge to the algorithm on its own terms. If the state's algorithm ignores meaningful changes to patterns of behavior or environments, then the algorithm will generate unexpected errors in decision-making. And if a competing algorithm of comparable capability is needed for detecting those errors, then the judge will be none the wiser.

2. Citations

As a final example, consider a machine that instructs a lawyer on ideal citation selection. Suppose that an expensive and sophisticated algorithm finds a hidden connection between the citation of an obscure case in contract and the outcome of a remotely related antitrust dispute. Each time a lawyer cites to the lucky contracts case, the presiding judge or her clerk receives a signal that varies in intensity on the basis of the judge or clerk's past experiences with the case. Perhaps some Bayesian updating occurs when the court receives a new brief, so that each time the signal is sent, its predictive power changes. Suppose the citation algorithm is regularly updated. It will register a successful or failed event in terms of the ability of the lucky contracts case to generate a desired ruling in conjunction with other variables and combinations that it measures. Over time, as the algorithm undergoes updating and lawyers learn of the citation's ability to impact a case, its usage will settle into an equilibrium. If the citation is unhelpful and its lack of predictive power becomes widely acknowledged by the bar, then the citation's use will decrease. Taking this process to its limit, the algorithm will simply distill the universe of applicable law to the antitrust

[131] *See* Levmore & Fagan, *supra* note 62, at 367 (noting that criminal defendants and other disappointed parties have challenged the state's use of algorithms on the basis of due process and other constitutional rules, and that it should be expected that they will eventually challenge the state's machine-based decision-making with algorithms of their own).

dispute by removing the contracts case from the set of helpful citations.[132]

When the citation is instead helpful, the citation's use will increase, but its persuasive powers will only last for a time. Taking the increase to its limit, everyone with access to the algorithm or its predictions will use the citation. The result of this process is similar to how every lawyer eventually learns, perhaps first at law school, that *Celotex Corp. v. Catrett*[133] should be cited for an assertion that affidavits are not required for a summary judgment motion. Citations to precise rules become boilerplate, and their absence often signals low quality lawyering if anything at all.

When rules are less settled and crafting legal argument matters—as it does in fast-changing and innovative areas of human activity and law—ideally judges and their clerks would evaluate the use of a popular and ever-present citation on the basis of how well the lawyer uses it to analogize or distinguish the novel case at hand. Perhaps somewhat troublingly, the Tippet et al. study suggests that courts can be persuaded by boilerplate alone.[134] While their research shows that constructing a careful relationship between a cited case and a present dispute is helpful for persuading a court, so is the mere insertion of widely-used citations contained within the winning

[132] It bears repeated emphasis that this distillation process is contingent on the stability of a legal domain. *See* VALIANT, *supra* note 13, at 61-62 (noting that machine learning cannot occur when the context of a generalization is changing); Fagan & Levmore, *supra* note 75, at 19 (applying this concept to law). The pattern of distillation has been theorized within the context of machine learning's ability to reduce legal standards to rules. *Id.* at 31; *see also* Casey & Niblett, *supra* note 71, at 1433 (theorizing further distillation from rules to micro-rules); BEN-SHAHAR & PORAT, *supra* note 70, at 24-25 (theorizing further distillation from societal rules to personalized rules).

[133] 477 U.S. 317 (1986).

[134] Tippett et al., *supra* note 2, at 1192.

briefs of the past.[135] One implication of this empirical result is that judges and their clerks may, to some extent, interpret the inclusion of citations alone as a signal for a correct argument.[136] By relying too heavily on a signal, they are engaging in a closed form of mechanical decision-making.

However, the persuasive power of this signal should cancel out, or at least diminish, when algorithms instruct adversaries to use identical citations for making and acknowledging contrary arguments. Identical usage can reduce the citation's signaling power because the judge observes that adversarial parties rely on the same signal to make contrary assertions. Moreover, if a particular citation only favors one side of an argument, then the citation tool will alert adversaries that they need to provide counterargument, which will also reduce the citation's power as a signal. Finally, as with the writing style tool, judges will learn over time if the citation systemically biases plaintiffs or defendants for meritless reasons or if adversaries are otherwise strategically using those citations detached of all substantive connection to the underlying facts and law of a dispute. In short, a citation's ability to independently persuade the judge on its own should diminish as both parties become equipped with the best tools; any asymmetric advantage conferred by citation algorithms should wane as their use proliferates.[137] While it is true that excessive

[135] *Id.* at 1175 (noting that inclusion of particular citations increases the performance of a model used to predict the outcome of employment disputes).

[136] This possibility may help explain Langdell's enthusiasm for the citation string. *See supra* note 37 and accompanying text. Judges and clerks can rely on citation inclusion as a proxy for a legal argument.

[137] This point is well-considered by Tippett et al., and it leads them to normatively conclude that citation recommendation tools should be provided as "open access" in order to level the playing field between unevenly matched adversaries. Tippett et al., *supra* note 2, at 1192. A tendency toward convergence of applicable case law to represent a legal rule, and that case law's broad revelation, will further level the playing field between adversaries. *See* Michael A. Livermore, Peter Beling, Keith Carlson, Faraz Dadgostari, Mauricio Guim & Daniel N. Rockmore{, *Law*

economic costs of developing widely-accessible tools may confer a temporary advantage to well-funded lawyers, so long as everyone else eventually learns to use a citation, perhaps by following the cues of its successful advocates, then signaling capability will degrade.[138] All of these pressures reduce judicial over-reliance on a boilerplate citation, despite its pedigree and initial association with an expensive prediction algorithm.

Citation tools are even less helpful in faster changing areas of life and law, where existing case law is, of itself, less specific, and where no case has ever ascended to the realm of boilerplate for statement or suggestion of a legal rule. This is true no matter how accurately constructed is the algorithm, or how better equipped one party is than the other (and able to wield their algorithm strategically). Clearly no advantage can be conferred when there is no "on point" citation, and this remains true when there is no citation sufficiently suggestive of a correct

Search in the Age of the Algorithm, 2020 MICH. ST. L. REV. 1183, 1209–11 (2020). The authors consider that parties may optimally stop their search for good case law in order to save costs. *Id.* at 1210. However, over time search costs will tend toward zero in settled legal domains inasmuch as the popular and winning citations are revealed by others, including the developers of citation algorithms in a competitive market whose development costs will decrease.

[138] Similarly, advances in discovery, notably technology-assisted review, are already reducing discovery costs for plaintiffs and defendants alike. *See* Engstrom & Gelbach, *supra* note 17, at 1053 (noting cost reductions and caveats). Predictive tools used for brief writing and other forms of argument presentation will follow the same course. Thus, in the short- to medium-term, asymmetries may persist, but in the long-term, a technological arms race will reduce, if not eliminate, any adversarial advantage as innovation, competition, and perhaps a public option, continue to drive the economic costs of data access and processing downward. For this reason, some scholars have suggested that PACER fees should be reduced or eliminated. *See* Charlotte S. Alexander & Mohammad Javad Feizollahi, *On Dragons, Caves, Teeth, and Claws: Legal Analytics and the Problem of Court Data Access*, *in* COMPUTATIONAL LEGAL STUDIES: THE PROMISE AND CHALLENGE OF DATA-DRIVEN LEGAL RESEARCH 95, 97 (Ryan Whalen ed., 2019). However, inasmuch as PACER represents a natural monopoly, elimination of fees may reduce data collection effort. *See* Engstrom & Gelbach, *supra* note 17, at 1064 n.248. This unfortunate caveat may persist in the development of low-cost citation and writing style tools.

judicial decision. Citation-tool indeterminacy is yet another instance of legal dynamism and novelty that renders closed decision tools which rely on the past less valuable. The more pronounced the dynamism, the less valuable the decision tool. Insistence on its use, based upon an erroneous belief in its continued precision, generates mistakes. In contrast, open decision-making architectures that leave space for flexible legal rules may appear less precise, but they paradoxically reduce error if Valiant's two conditions for machine learning remain unsatisfied.[139]

III. COMPUTATIONAL LEGAL ETHICS

A. Preliminaries from Weber to Marcuse

Lawyers do things in certain ways. They research case law with books and databases, compare and contrast cases with analogic reasoning, plead complaints and present arguments with persuasive rhetoric, interview clients and examine witnesses with empathy, and interpret rules with grammar and other contextual cues. All of these tasks are relatively standardized means for obtaining loosely pre-determined results; that is, they are *techniques*, but each is partly determined by the others.[140] When taken together as whole, individual techniques and their relationships with each other encapsulate the entirety of legal practice: law as a technical system is comprised of its parts. In the past, it may have been sufficient to think of individual legal techniques as constituting legal practice[141] in the same way that carpentry constitutes woodworking. There is independence between the techniques of

[139] *See* note 67 and accompanying text.

[140] For an early elaboration of a theory of technical systems, *see* 1 BERTRAND GILLE, HISTORY OF TECHNIQUES: TECHNIQUES AND CIVILIZATIONS 19 (1986) (describing basic technical acts as standard techniques, which, when grouped together, form a technical system).

[141] By legal practice, I mean lawyering activity as distinct from lawmaking and enforcement activities.

hammering, measuring, cutting, and so forth on the one hand, and the technical system of woodwork on the other. If woodworking as a system were to change the way carpenters carry out individual techniques, then woodworking would impact the various tasks of the carpenter much like macroeconomic policy can influence the microeconomic decisions of people. When tasks and systems become dependent upon each other, elaboration of the system as a concept can be useful since change and transformation can happen from the top as well as the bottom.[142]

Whether law's technical system is currently transforming the day-to-day tasks of lawyers is largely irrelevant because the practices themselves are changing from the bottom up. By contrast, sociologists, in particular Weber, Marcuse, and Habermas, identify instances of rationalization where system-level thinking and organization justify (and shape) human-level adoption and elaboration of individual tasks.[143] They provide examples of top-down transformations. For instance, Weber describes how the stock-market system, born of small-scale investment and the limited liability company, ended up "rationalizing speculation" even as its architects sought to provide for mobility of capital.[144] This example, and others like it, is often presented as evidence of the imposition of system-wide logic on daily human life in which the consequences of technical development lead to unexpected and sometimes unintended interferences into human activity.[145] To take another

[142] GILLE, *supra* note 140; *see also* 1 BERNARD STIEGLER, TECHNICS AND TIME: THE FAULT OF EPIMETHEUS 31 (Richard Beardsworth & George Collins trans., 1998) (noting that description of a technical system becomes useful when technical structures, ensembles, and channels propagate new and interdependent technical structures).

[143] Examples are given immediately below.

[144] MAX WEBER, THE PROTESTANT ETHIC AND THE "SPIRIT" OF CAPITALISM AND OTHER WRITINGS 22 (Peter R. Baehr & Gordon C. Wells trans., 2002).

[145] *See, e.g.*, 1 JÜRGEN HABERMAS, THE THEORY OF COMMUNICATIVE ACTION: REASON AND THE RATIONALIZATION OF SOCIETY 340–41 (Thomas

example closer to legal practice, consider the doctor-patient relationship. This relationship, from the perspective of many sociologists, has been disrupted by the technical system of medicine.[146] Medicine, comprised of individual techniques such as research, testing, and drug delivery, "imposes" itself on patients who are powerless to influence their own treatment even as they describe (and embody) their symptoms.[147] The more complicated an illness, the less likely the doctor will rely on the patient for discerning the best course of action. In the abstract, Habermas speaks of rationalized technical systems that are colonizing the lifeworld.[148]

But none of this top-to-bottom influence is necessary to see that advances in artificial intelligence and machine learning are reshaping the various techniques of law practice. There is no imposition of research conclusions on the ideal word-choice for

---

McCarthy trans., 1984) [hereinafter COMMUNICATIVE ACTION] (noting that the institutionalization of knowledge has led to internal systems of validity—immune from critique—across the activities of science, art, ethics, political theory, jurisprudence, and public lawmaking).

[146] *See* Graham Scrambler & Nicky Britten, *System, Lifeworld and Doctor-Patient Interaction: Issues of Trust in a Changing World*, *in* HABERMAS, CRITICAL THEORY AND HEALTH 45, 56 (Graham Scrambler ed., 2001) ("[T]he voice of medicine relies exclusively on the biomedical model . . . . [which reflects] the technical-instrumental framework of the biosciences, strips away social contexts of meaning on which a full and adequate understanding of patients and their illnesses depend." (quoting E. Mishler, SOCIAL CONTEXTS OF HEALTH, ILLNESS AND PATIENT CARE 192 (Elliot Mishler et al. eds., 1984))).

[147] *Id.*

[148] Habermas uses the concept of a lifeworld to refer to "a complex of interpenetrating cultural traditions, social orders, and personal identities." JÜRGEN HABERMAS, BETWEEN FACTS AND NORMS 23 (William Rehg trans., 1996) [hereinafter FACTS AND NORMS]. Sometimes colonization is interpreted as a "late-stage" capitalist phenomenon. *See, e.g.*, Joseph Heath, *System and Lifeworld*, *in* JÜRGEN HABERMAS: KEY CONCEPTS 74, 74 (Barbara Fultner ed., 2011). This is perhaps an overstatement. Habermas himself has suggested that modern society satisfies its need for social integration and goal-setting with three resources: "money, administration, and solidarity." FACTS AND NORMS, *supra*, at 299. What is important for Habermas, is that solidarity "should be able to hold its own against the two other mechanisms of social integration, money and administrative power." *Id.*

appellate briefs, voice intonation for oral arguments, or citations for judicial opinions. Lawyers freely choose to adopt new techniques merely to prevail on behalf of their clients; judges do so to avoid reversal and articulate clear law. Law's social processes still precede its technical ones. In other words, there is no colonization of law by the very logic of law's own technical system. In contrast, when the system moves downward, technical processes can crowd out social ones even if humans set the empirical structures and processes in motion and sanction them. This pattern of top-down system logic is expressed in lawmaking when machines strongly influence goal-setting and law's purpose,[149] as well as enforcement in scenarios where humans are unable to adapt and compete with machine surveillance, profiling, and other administrative techniques or when there are no safety valves for error such as human review of machine decisions, opt-outs from computational profiling, and other methods for checking the decision-making power of programmed systems.[150] By contrast, day-to-day lawyering and basic judging remains a bottom-up endeavor.

This comparison is worth enlarging and helps clarify an emerging soft empirical legal ethics and a tentative movement from constitutional to administrative norms. For Weber, techniques such as electronic surveillance, enforcement quotas, social credit scoring, and hand-tying statutory purpose and constitutional norm creation so as to allow machines to set goals, can be justified according to a broadly accepted criterion of rational decision-making.[151] Their deployment is irresistible

[149] *See infra* § II.C.2.

[150] *See supra* § I.B.2.

[151] See, e.g., WEBER, *supra* note 144 at 26, who describes the modern process of system rationalization, which determines societal ideals and action:

> This rationalization process in the field of technology and economics undoubtedly also determines a significant proportion of the "ideals" of modern civil society: in the minds of the representatives of the "spirit of capitalism," labor in the

when benefits are stacked against costs, for instance, or when some other method of validation is used such as wealth maximization, environmental damage minimization, or measures of equity. This is especially true when top-down deployment of computational law fails to acknowledge its longer-term suppression of innovation benefits.

Where private life may have been off-limits in the past, system-level rationalization provides the means for in-roads to be made. Surveillance and quotas can enhance safety for all. Social credit scoring can be used to promote system-wide equity. Pre-commitments to granting machines permission to fashion statutory goals can be justified to avoid rent-seeking.[152] These are instances of the technical system reaching down to rationalize and reshape the day-to-day techniques of enforcement and lawmaking. Weber ties rational decision-making to the industrialization of work and "late" capitalism,[153] but it is easy to see that any mode of organization can leverage top-down rationality. Socialists rationalize surveillance and social credit just as well.[154]

> service a rational structuring of the provision of the material needs of humanity has always been on the guiding purposes…

Other modes of organization can be just as easily rationalized as well.

[152] For a discussion of legislative pre-commitment devices across a number of examples, including budget law, environmental law, health law, banking law, and criminal sentencing, *see* Frank Fagan, *Legal Cycles and Stabilization Rules*, *in* THE TIMING OF LAWMAKING 11, 11–32 (Frank Fagan & Saul Levmore eds., 2017).

[153] *See* WEBER, *supra* note 144 at 26-27 (noting that industrialism is characterized by rational thinking applied to traditional forms of work and that the leaders of capitalism suppose "the structuring of provision of the material needs of humanity" as a profession [*beruf*]).

[154] *Cf.* JAMES BURNHAM, THE MANAGERIAL REVOLUTION: WHAT IS HAPPENING IN THE WORLD 79-80 (2021) [1941] (noting that capitalist and socialist systems alike tilt toward "managerialism," which is defined as a system-wide logic embodied in the actions of the day-to-day managers of productive assets).

Thus, Marcuse, perhaps ingenuously, depicts rationalization as a hidden form of political domination.[155] Science and technology, once used to moderate nature and liberate humanity, are now used to legitimize authority.[156] Marcuse is particularly interested in how the human want for comfort and productivity justifies (technologically enabled) limitations on human freedom:

> In this universe, technology also provides the great rationalization of the unfreedom of man and demonstrates the "technical" impossibility of being autonomous, of determining one's own life. For this unfreedom appears neither as irrational nor as political, but rather as submission to the technical apparatus which enlarges the comforts of life and increases the productivity of labor. Technological rationality thus protects rather than cancels the legitimacy of domination…[157]

In response to Marcuse, we might insist that society needs to develop new modes of computational surveillance, quotas, and social credit scoring free from political domination. For instance, people may choose to use social media platforms that scrape personal data, but rely on legal rules like the GDPR to take their data with them or delete their data entirely.[158] These

---

[155] *See* MARCUSE, *supra* note 89 at 162 (2d ed. 1991). In particular, Marcuse notes that:

> The scientific method which led to ever-more-effective domination of nature thus came to provide the pure concepts as well as instrumentalities for the ever-more-effective domination of man by man through the domination of nature. Theoretical reason, remaining pure and neutral, entered into the service of practical reason. The merger proved beneficial to both. Today, domination perpetuates and extends itself not only through technology but *as* technology, and the latter provides the great legitimization of expanding political power, which absorbs all spheres of culture.

[156] *Id*.

[157] *Id.*

[158] The relevant legislation for data portability is found in Article 20 of the GDPR. *See* GDPR, *supra* note 100, art. 20(1) ("The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and

rules appear motivated by Marcuse inasmuch as they weaken the rent-seeking capabilities and attendant political authority of the platform while retaining the comfort and productivity increases for the consumer.[159]

This example represents a type of separation strategy, which permits law to capture the benefits of both forms of rationalization—empirical and intersubjective. According to Habermas, however, pursuit of separation schemes is a mistake because strategies for decoupling the benefits of a technical system (like surveillance) from its domination costs (with separation methods like data portability, for instance) are hopelessly utopian in the long-term. Habermas suggests that the historical evolution of technical systems is embodied by an incremental and inevitable objectification of rational action.[160] In terms of personal data protection, even if consumers are given

machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided…"). Deletion of data, also known as the "right to be forgotten," is found in Article 17. *See* GDPR, *id.*, art. 17(1) ("The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay…").

[159] Ambitiously, Marcuse proposes an entirely new science free from the use of technical systems as a means of political domination. *See* MARCUSE, *supra* note 89 at 235 ("Under such conditions, the scientific project itself would be free for trans-utilitarian ends, and free for the "art of living" beyond the necessities and luxuries of domination.").

[160] *See* JÜRGEN HABERMAS, TOWARD A RATIONAL SOCIETY: STUDENT PROTEST SCIENCE AND POLITICS 87 (Jeremy J. Shapiro trans. 1971) [hereinafter RATIONAL SOCIETY] ("If we comprehend the behavioral system of action regulated by its own results as the conjunction of rational decision and instrumental action, then we can reconstruct the history of technology from the point of view of the step-by-step objectification of that very system."); STIEGLER, *supra* note 142 at 11 ("Habermas finds [Marcuse's] project utopian. The history of technics represents that of a progressive but ineluctable objectification of [empirically sanctioned] action in technical systems.") The idea that technology leads to the objectification of humans is first elaborated by Heidegger. *See* Heidegger, *supra* note 18 at 332 (noting that adherence to an empirical perspective leads to humanity itself "to be taken as 'standing-reserve'", which is Heidegger's conception of a stockpile in service to technological purpose).

the ability to transfer or delete their data, the platform, its competitor, or a substitute product will eventually rationalize a form of sovereignty over consumers one way or another. Political domination itself is legitimized by the very progress of technical systems.[161] It is the sine qua non of empirical-technical rationality.

B. Between Intersubjectivity and Empiricism

Contemporary social theorists sometimes apply this critique to human processes that work from the bottom-up, but it bears emphasis that intersubjective rationality is a form of social authority based upon communication among people and emergent social norms, habits, and customs.[162] For Habermas, it is a grassroots endeavor characterized by apposition to empirically rationalized technical rules.[163] It shares its logic with the wisdom of the crowd and other forms of practical reason, but its predominant purpose is to assert the primacy of consensual legitimacy. Empirically sanctioned authority, on the other hand, is based upon observation, testing, and empirical refinement. This type of authority produces technical rules that are largely dissociated from communicative life.[164]

Lawyers and judges in association with a bar engage in "communicative action."[165] That is, their practices and

---

[161] *See, e.g.*, HABERMAS, *supra* note 145 at 340-42 (noting that political consensus is met more and more frequently by agreement because participants now customarily accept empirical and expert machine-based knowledge).

[162] *See, e.g.*, STIEGLER, *supra* note 142 at 11 (noting that as a theoretical construct social norms are "grounded upon intersubjectivity alone" and "cannot be put on the same level as technical rules").

[163] *See infra* notes [189]-[190] and accompanying text.

[164] *See* STIEGLER, *supra* note 142, at 11.

[165] Communicative action is Habermas' term. It forms the basis of a social authority derived from intersubjective dialog among humans, and represents an alternative to the "purposive-rational action" of technical systems derived from empiricism. *Id*. Stiegler helpfully notes that:

> All human history can be analyzed, accordingly, as the history of the varying set of relations between communicative action on the

techniques are more or less shaped by their interactions with one another. The norms and customs that emerge from these interactions characterize the basis of the bar's social authority. Legal research of case law and statutes, the comparing and contrasting of cases, the various rhetorical modes of pleading and argument presentation, interactions with clients and witnesses, and the reliance on context for interpreting legal texts are professional techniques developed from the bottom up by members of the profession.

In contrast, the enforcement of law is fractured. On the one hand, there is community-oriented policing, where focus is placed on symbolically (or otherwise human) mediated interaction between citizens and police.[166] On the other is professional policing, where decisions and communication flow downward from the police to the community.[167] Community policing, however flawed, is more strongly characterized by communicative action and intersubjective social authority derived from people. The professional model adheres to

---

one hand and purposive-rational action on the other. The difference between traditional societies and modern societies is characterized by the fact that, in the former, communicative action forms the basis of social authority (whether it be mythical, religious, or metaphysico-political), whereas in the latter, legitimation is dominated by technical and scientific rationality, which progressively spreads across all areas of life, including those so-called "communicative" aspects whose specificity is thereby denied. *Id*. at 11-12.

[166] *See* Adam Dorbin, *Professional and Community Oriented Policing: The Mayberry Model*, 13 J. CRIM. JUSTICE & POPULAR CULTURE 19, 21 (2006) (noting that police outcomes in the community model are reached on the basis of joint assessment of problems, co-production of solutions, and joint-responsibility for results).

[167] Lawmaking also exhibits joint tendencies of communicative and purposive-rational action. While judge-made law is more socially sanctioned and administrative law more empirically sanctioned, legislation sits somewhere in between. However, to the extent that lawmaking pre-commits to allowing machines to select the goals and purposes of statutes and set constitutional norms, it becomes more empirically hardened.

technical rules that derive their social authority on the basis of system-wide rationalization.

It is important to restate Habermas' identification of the problem: political domination itself is legitimized by the progress of technical systems. People who develop technical rules on the basis of scientific rationalization find themselves developing rules in order preserve and maintain the integrity of the technical system.[168] These rules often escape public discussion and engender dissatisfaction with the administrative power centers involved in maintenance of the system. The situation ultimately leads to a "systemic closing [in which] [individual human] interests define the social system so much as a whole that they coincide with the interest in maintaining the system."[169] As the values of the system progressively continue their ascendance, people become increasing depoliticized: "the industrially most advanced societies seem to approximate the model of behavioral control steered by external stimuli [and resultant empirically sanctioned technical rules] rather than guided by [intersubjective] norms."[170] Thus, use of force may be technically justified, but intersubjectively abhorrent. This is a good description of the relatively recent incident where an Ohioan police officer shot and killed a young assailant who was attempting to stab an acquaintance.[171] Systemic rules required use of force; intersubjective valuations forbade it.

Where Marcuse might identify an instance of political domination and naively search for a way to decouple the benefits and costs of technical rules,[172] Habermas insists that the system

---

[168] *See* HABERMAS, RATIONAL SOCIETY, *supra* note 160, at 101.

[169] *Id.* at 105.

[170] *Id.* at 107.

[171] *Teenage Girl Is Fatally Shot by Police in Columbus, Officials Say*, N.Y. TIMES, Apr. 20, 2021. https://www.nytimes.com/2021/04/20/us/columbus-ohio-shooting.html.

[172] *See* MARCUSE, *supra* note 89, at 235.

will always triumph unless we decouple the intersubjective and empirical rationalizations themselves:

> The process of development of the productive forces can be a potential for liberation if, and only if, it does not replace rationalization on another level. *Rationalization at the level of the institutional framework* can occur only in the medium of symbolic interaction itself, that is, through *removing restrictions on communication*.[173]

The challenge for Habermas is to find a reasonable way to reject system-wide justifications for restricting the intersubjective authority of human beings. Put differently, "[t]he question is one of liberating communication from its technicization."[174] So long as humans remain firmly in control of their persuasive language and rhetorical arts, then for Habermas, they will remain firmly in control of their machines (and laws and police officers).[175]

This idea has been expressed in terms of having a debate about the use of technology with a liberated form of communication.[176] For instance, we might consider the merits and demerits of professional policing in terms of the personal values of community members. There is a localist flavor to this approach that can analogize to principles of federalism and subsidiarity: we see that the best level for finding solutions for a community is at the community level and we shun the admonitions of technical rules based upon a system-wide

---

[173] HABERMAS, RATIONAL SOCIETY, *supra* note 160, at 118.

[174] STIEGLER, *supra* note 142, at 12.

[175] This is similar to Heidegger's response. He suggests that humans can avoid becoming objectified (as "standing-reserve"), and then controlled, by their machines by remaining firmly in control of their aesthetic arts. *See* HEIDEGGER, *supra* note 18, at 339-41.

[176] STIEGLER, *supra* note 142, at 13 (noting a need to create a new relationship to technology that "rethinks the bond originally formed by, and between, humanity, technics, and language").

rationale.[177] But in contrast, an empirical rationality provides that the solution simply amounts to adding another variable that captures local conditions and appropriately measures the values of the local community. For scholars like Habermas who favor intersubjectivity over empiricism, power should tilt fully toward the community in order to avoid human objectification.[178] It remains to be seen if a Marcusean middle-ground will broadly hold in other areas beyond data protection, but law's computational paradox all but guarantees that empiricism cannot completely efface intersubjectivity.

C. Permanent Innovation and the Triumph of Intersubjectivity

1. Legal Domains Opened and Closed

Of course, artificial intelligence may never bring crisis to law if it cannot do better than the very lawyers, lawmakers, and enforcers that it could potentially displace. Small data, or the inability to recognize patterns; environmental change, or the instability of machine learning contexts; the Lucas critique,[179] or the endogeneity of behavior within a closed system, are each common to open and innovative legal domains that presuppose change. Recognition of the limits of machine learning in these

[177] *See, e.g.*, Aurielen Portuese, *The Principle of Subsidiarity as a Principle of Economic Efficiency*, 17 COLUM. J. EUR. L. 231, 231 (2010); Roger Van den Bergh, *Subsidiarity as an Economic Demarcation Principle and the Emergence of European Private Law*, 5 MAASTRICHT J. EUR. & COMP. L. 129, 130 (1998).

[178] The necessity of elevating intersubjectivity has also been cast in terms of a broader communal political strategy and model of citizenship: individuals need to ensure that they retain a sufficient amount of intersubjectivity required for preventing the rationales of technical systems from overrunning the public square. *See* STIEGLER, *supra* note 142, at 13 (noting that one must "ensure for oneself a 'minimum of subjectivity {or "will and mastery"} … required for a democratic thought to fix limits' to technological expansion 'through public decisions based themselves on public discussion and argumentation between subjects'") (quoting LUC FERRY & ALAIN RENAULT, HEIDEGGER AND MODERNITY 42 (1991) (Franklin Philip trans.)).

[179] *See supra* note 41 and accompanying text.

contexts suggests that while law will become decidedly more empirical in closed domains, blanket empiricism is not possible without imposed closure that generates error.

That innovation will recede is not an adequate answer. Consider that today we have entered a state of permanent innovation and technological development. In his history of technology, Bertrand Gille documents the ever-shortening delays between the discovery of a basic physical phenomena and its industrial application:

> One hundred and two years elapsed between the discovery of the physical phenomenon applied to the photograph and photography itself (1727-1829) … [but only] fifty-six years [elapsed] for the telephone, thirty-five for the radio, … six for the uranium bomb, five for the transistor.[180]

The reduction in delay is the result of a new relationship between technology, politics, and the economy. Prior to the modern era, the development of applications from basic inventions required waiting for the ideal economic and social conditions to materialize. Today, that process is inverted. We create favorable economic and social conditions in order to incite innovation.[181] Today's world is constantly under development as a matter of policy, in a state of "perpetual modernization or constant innovation."[182]

Consider again the two necessary conditions for machine learning to occur. A pattern must be observable (usually because there is sufficiently big data and a well-specified model), and the

---

[180] GILLE, *supra* note 140, at 39 (*quoted in* STIEGLER, *supra* note 142, at 40).

[181] *See* STIEGLER, *supra* note 142, at 41. *See also* GILBERT SIMONDON, ON THE MODE OF EXISTENCE OF TECHNICAL OBJECTS 93 (Ninian Mellamphy trans. 1980) (noting that in the modern era, humans guide innovation by organizing relationships between technical stages, "instead of being, as artisan, of those technical stages himself").

[182] STIEGLER, *supra* note 142, at 39.

environmental context in which that patterned regularity occurs must be sufficiently stable.[183] Obviously, if the context is unstable, as it is when new variables appear over time, then confidence in the earlier pattern will weaken or dissolve. An algorithm may confidently predict flight risk and recommend that the accused should be jailed, but if a new ankle bracelet is able to track those who are released, then the algorithm's predictions may fail. It will take time to gather sufficient data in the ankle bracelet context. Some people may flee even wearing the bracelet, but it can be safely assumed that less will flee with it than without. Of course, the present algorithm could be adjusted on the basis of a theory (or structural model) of how tracked people will behave, but accuracy will suffer to the extent that the theory is wrong. The point is that any innovation changes the context of a legal prediction. If an innovation is small, then the algorithm will easily adapt. The opposite is true if an innovation is large. Suggesting normatively that law should be excessively precise is to suggest that law should foreclose innovation.

If law were concerned with only immediate precision, then it would prohibit the use of ankle bracelets by the state. Without the bracelets, the machine learning algorithm would continue to predict flight risk accurately. It is easy to see that a prohibition would be bad policy. The bracelet innovation reduces flight risk, which in turn, can reduce the economic costs of sitting in jail to the accused and the incarceration costs to the state. Refusing to embrace the innovation in order to maintain the earlier algorithm's precision can only be justified if the bracelet has a small effect.[184] If the bracelet's effect is large, so that nobody flees at all, then the algorithm becomes obsolete, or

[183] *See* VALIANT, *supra* note 13, at 62.

[184] Specifically, any beneficial effects must outweigh the cost of updating the algorithm after accumulating sufficient data with the bracelet.

can be adjusted to one variable. Wearing the bracelet perfectly predicts no flight.

It may be tempting to assert that the use of the bracelet represents a closed system, where law is now sufficiently precise, and innovation is no longer required. But that assertion could only be justified in a static world where the environment is fixed. Perhaps the bracelet is later found to interfere with new radio signals or an innovative technique is developed to remove it. In legal domains characterized by higher levels of stasis, machine learning will perform better,[185] but caution should be exercised if law overtly seeks to close a legal domain for precision's sake. Permitting law's technical system to overrun innovation is likely a sign that the innovation is not understood.

Similarly, law can discourage an empirical rationale from overrunning the public square by, counter-intuitively, encouraging and investing in technical innovation.[186] Greater levels of innovation intensify the dynamism of the legal environment and render existing data and machine learning obsolete more quickly. This should be welcomed rather than feared inasmuch as the benefits of innovation outweigh those of the heightened computational precision available in closed domains.

2. Pre-committing to Machines

In other work, Saul Levmore and I have suggested that humans will retain the ability to override machines as a prerequisite for allowing machines into law.[187] Suppose instead

[185] *See* Fagan & Levmore, *supra* note 75, at 1.

[186] *See, e.g.*, Robert D. Atkinson, *Why the United States Needs a National Advanced Industry and Technology Agency*, INFORMATION TECHNOLOGY & INNOVATION FOUNDATION 1 (Jun. 2021), available at https://itif.org/publications/2021/06/17/why-united-states-needs-national-advanced-industry-and-technology-agency (suggesting that the United States should create an administrative agency for the development of technology).

[187] *See* Fagan & Levmore, *supra* note 75, at 7.

that a powerful and decisive interest group, perhaps from an earlier generation, creates a rule that imposes a pre-commitment to the machine's selection of social goals and assigned weights. A machine may determine that taxes should be levied so as to achieve, for instance—75% of possible equity maximization and 25% efficiency—on the basis of a pre-commitment to maximal social cohesion.

It is hard to imagine humans ceding control to machines by means of random pre-commitments, but it is possible. A group of technology enthusiasts in power may direct the machine to form its own larger objectives and then encourage it to assign weights to efficiency, equity, and other variables accordingly. It should be immediately apparent that the ability of the pre-commitment to bind is at least a partial function of the quality of its outcome. If people are content with the machine's work, then they are more likely to be satisfied with its selection of goals and adhere to its directions. If the machine does a poor job, then the pre-commitment is likely to be overturned or amended much the same way lawmakers repeal a statute or a court overturns constitutional precedent. The same enthusiasts may cede the power of repeal and amendment to the machine, but then of course basic patterns of revolution remain in place for change.[188] Demand for change is a product of the machine's inability to adequately adapt and keep up with human innovation. While the machine's robustness and defenses against change can be increased by foregoing narrow precision and tailoring of rules, it must sacrifice precision by broadening its mandate. The computational paradox persists no matter how much control is ceded to the machine.

By way of analogy, consider an earlier generation of political leaders who set the trajectory of an organization through formally binding documents like a corporate charter or

[188] *See* HAROLD J. BERMAN, THE INTERACTION OF LAW AND RELIGION 14 (1974) (noting that "[l]aw has to be believed in, or it will not work").

constitution, or informally through supporting but influential writings such as works like the Bible or the Federalist Papers. They set the goals of a future generation and generate a dialog between the present and past. Much the same way, humans that program machines earlier in time to offer goals or provide instruction on goal-setting are simply setting the stage for a dialog between machines and future generations of humans. When conflict arises, the trajectory of lawmaking will be shaped by the various outcomes of these exchanges much the same way constitutional law is developed today. Courts, legislatures, and other polities on occasion overrule broader, constitutional, pre-commitments set by earlier humans. The analogy easily carries to machines that have been programmed to offer lawmaking goals. They will be overridden like earlier generations of leaders. Innovation and change erodes the applicability and precision of earlier rules, and law efficiently responds with updates.[189]

A form of this conflict has already arrived via unexplained AI that discriminates.[190] This conflict is not between two goals, but between a human goal and the outcome of a machine decision that results from earlier human goal-setting and data selection. This is an easy conflict to resolve. Law can either override the machine, restate the goal, or feed the machine different data. If machines begin to set goals, however, the conflict is more challenging because the desire to refine the goal will not be based upon what is clearly an error, but rather upon the need for updating a model of goal-setting in response to innovation and social change. Obvious errors like discriminatory machine profiling incite broad consensus for updating, but an innovation that degrades machine precision may or may not be so total. Consider, too, that if machines are

[189] *See* Fagan, *supra* note 152, at 11.

[190] *See, e.g.*, O'NEIL, *supra* note 102, at 116, *passim*; *see also* Gianclaudio Malgieri & Jędrzej Niklasb, *Vulnerable Data Subjects* 37 COMP. L. & SEC. REV. 105415, 105415 (2020) (noting the variable susceptibility of people to machine bias).

programmed to select their own data in addition, then the option to provide the machine with new data sources will require override as well. In dynamic environments the computational paradox will be amplified insofar as humans cede more control to their machines and more strongly tie their hands with an inability to quickly and flexibly update rules.

IV. CONCLUSION

This Article has suggested that the benefits of machine learning will accrue in legal domains that are sufficiently closed. In these domains, familiar patterns of distillation from standards to rules, or perhaps rules to micro-directives or personalized commands, will proceed apace. The closed character of the legal environment will enable machine learning to do what it does best. Because machine learning is conditioned upon the discernment of observable regularities and the sufficient stability of legal environments, innovation and change serve as a self-corrective to human overreliance on machines. Nevertheless, lawmakers may attempt to impose closure of legal domains by prohibiting improvements or erecting barriers to innovation in order to enhance law's precision. This would be a mistake insofar as innovation yields greater benefits than imposed precision. When imposed precision grants a deficit of benefits, lawmakers inefficiently fall prey to the computational paradox.

Heightened innovation in law prevents its total empirical rationalization. Innovation leaves data periodically stale and machine learning tools obsolete, but it also prevents a purely empirical ordering from overrunning the public square. Even as humans increasingly draw upon empirical and technical rationales for their laws, openness and change necessarily places an intersubjective social authority based on human dialog and exchange at a deeper layer. This is true across lawmaking, enforcement, rights allocation, and lawyering. In response, a computational legal ethics will emerge to capture the benefits of machine precision while simultaneously balancing the benefits of innovation and human dynamism. Tomorrow's lawyers will be called upon, more so than their counterparts of the past, to

decide what should be considered versus ignored. A human-centered constitutional authority will remain, however far technology may advance.

* * *