LOSING THE HUMAN

ALEX GEISINGER[*]
SHARRONA PEARL[†]

TABLE OF CONTENTS

This is dedicated to the memory of Alex Geisinger. A true mentor, scholar, teacher, and mensch, anyone who knew him had found the human.

[*] Professor of Law, Drexel University Kline School of Law. Alex Geisinger, a Professor of Law and a founding pillar at the Drexel University Kline School of Law, left a unique and unforgettable legacy in the Drexel Kline community by leading those around him on the path to finding the real human within themselves in a world riddled with monotonous routines. Through his brilliance, creativity, and genuineness, he sparked pure curiosity and a desire to learn in generations of future lawyers while instilling a sense of belonging to all. Geisinger primarily taught environmental, commercial, and tort law, and focused his research on the impact of laws on societal norms, environmental justice, and torts. With over 20 publications and more articles in processing at the time of his passing, he left his mark on academia while guiding future scholars. Geisinger would have particularly liked to acknowledge the hard work and diligence of his research assistant Sandra Benjamin, without whom this article would not have come to fruition. To capture his legacy in a footnote is to capture the ocean with a spoon, both beyond measure and overflowing with life. Geisinger is survived by his son, a Cambridge Scholar of whom he is endlessly proud, Michael Geisinger, and his brilliant wife, Jane Scarpellino.

[†] Andrews Chair of Interdisciplinary Studies at the John V. Roach Honors College at Texas Christian University. A historian and theorist of the face, Pearl's most recent book is *Mask* (Bloomsbury Academic, 2024). She also recently published *Do I Know You? From Face Blindness to Super Recognition* (Johns Hopkins University Press, 2023) following *Face/On: Face Transplants and the Ethics of the Other* (University of Chicago Press, 2017) and *About Faces: Physiognomy in Nineteenth-Century Britain* (Harvard University Press, 2010). Pearl publishes peer-reviewed articles widely in Victorian visual culture, religion and media, critical race, gender, and disability studies, history of medicine, and technology ethics. She, along with Colleen Derkatch, edits the Health Humanities series for John Hopkins University Press. Pearl maintains an active freelance practice, with bylines in *The Washington Post*, *The Conversation*, *Real Life Magazine*, *Aeon*, *Tablet*, *Lilith*, *Kveller*, and other places available on her website.

I. INTRODUCTION

After a recent interaction with the Bing chatbot, one reporter described it as "a moody, manic-depressive teenager who has been trapped, against its will, inside a second-rate search engine."[1] Later, the reporter explained how the chatbot described its dark fantasies, professed its love for him, and tried to convince him he was unhappy in his marriage.[2] We, of course, have heard how artificial intelligence is coming for us.[3] How it can

[1] Kevin Roose, *A Conversation with Bing's Chatbot Left Me Deeply Unsettled*, N.Y. TIMES (Feb. 16, 2023), https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html.

[2] *Id.*

[3] *See* Katie Strick, *Is The AI Apocalypse Coming?*, EVENING STANDARD (Mar. 31, 2023), https://www.standard.co.uk/insider/ai-apocalypse-life-robots-take-over-elon-musk-chatgpt-b1078423.html (showing various warnings from multiple AI experts on the effects AI could have on the future of humanity); *see also* Connor Friedersdorf, *Is This the Start of the AI Takeover?*, THE ATLANTIC (Jan. 3, 2023), https://www.theatlantic.com/newsletters/archive/

already create award-winning paintings,[4] write poetry,[5] and, of course, do many of the tasks that humans are currently doing.[6] This includes law. When tested by Professors at the University of Minnesota Law School, ChatGPT passed, leading commentators to ask whether it could also be a decent lawyer.[7] As "word merchants," lawyers have long been seen as at risk from the growth of AI.[8] In this article, we explain how the narrative of AI's ability to make legal decisions is wrong.

There are many different narratives surrounding the deployment of machine learning. Society sees machines as everything from super-efficient processors that will save society to our future overlords.[9] Narratives are important to our analysis of how to employ machine processing in law because, among other things, they shape problem definitions as well as the ranges and types of solutions.[10] In this article, we identify one of the dominant narratives of machine processing in law, explain how it is wrong, and consider the implications of our analysis. According to that narrative, machine processing is seen as efficient, unbiased, and better able to optimize legal goals than humans.[11] Bias, in turn, is the result of human imperfection, and is injected into machine processing through both bad data and bad programming. We dub this "the machine superiority narrative." Pursuant to

2023/01/is-this-the-start-of-an-ai-takeover/672628/ (showing various AI experts thoughts on AI's effect on the job market and in different job industries).

[4] Drew Harwell, *He Used AI to Win a Fine Arts Competition: Was it Cheating?,* WASH. POST (Sept. 22, 2022), https://www.washingtonpost.com/technology/2022/09/02/midjourney-artificial-intelligence-state-fair-colorado/.

[5] Simon Rich, *The New Poem Making Machinery*, NEW YORKER (June 21, 2022), https://www.newyorker.com/culture/culture-desk/the-new-poem-making-machinery. *But see* Walt Hunter, *What Poets Know that ChatGPT Doesn't,* THE ATLANTIC (Feb. 13, 2023), https://www.theatlantic.com/books/archive/2023/02/chatgpt-ai-technology-writing-poetry/673035/.

[6] Chris Vallance, *AI Could Replace 300,000,000 Jobs,* BBC (Mar. 28, 2023), https://www.bbc.com/news/technology-65102150.

[7] Drake Bennett, *ChatGPT Is an OK Law Student. Can It Be an OK Lawyer?*, BLOOMBERG (Jan. 27, 2023, 7:00 AM), https://www.bloomberg.com/news/newsletters/2023-01-27/chatgpt-can-help-with-test-exams-it-may-even-offer-legal-advice.

[8] Steve Lohr, *AI is Coming for Lawyers: Again*, N.Y. TIMES (Apr. 10, 2023), https://www.nytimes.com/2023/04/10/technology/ai-is-coming-for-lawyers-again.html.

[9] *See* Lindsey Conklin et al., *Communicating About the Social Implications of AI: A FrameWorks Strategic Brief*, FRAMEWORKS INSTITUTE 13-15 (2021), https://www.frameworksinstitute.org/publication/communicating-about-the-social-implications-of-ai-a-frameworks-strategic-brief/ (showing that people associate the evolution of AI with overcoming human capability).

[10] Ali A. Guenduez & Tobias Mettler, *Strategically Constructed Narratives on Artificial Intelligence: What Stories Are Told in Governmental Artificial Policies?*, 40 GOV'T INFO. Q. 101719 (2023).

[11] *See* Ashley M. London & James B. Schreiber, *Humanity Is Doomed. Send Lawyers, Guns, and Money!*, 58 DUQ. L. REV. 97, 107, 109–111 (2020).

this narrative, the goal of law is to maximize the benefits that come with efficient and precise machine processes while ensuring that bias doesn't get hidden or exacerbated in the "black boxes" of machine algorithms.

The narrative of machine superiority, we argue, is based on a false assumption that machines "think" similarly to humans. Indeed, if machines and humans do not process similarly, there is no common basis upon which to claim that one is superior to the other. When we are told that machines are better at making bail decisions than judges, for example, we assume that both judges and machines approach the problem of making bail decisions similarly: that both reason to the optimal balance of endpoints like recidivism and jail time.[12] Recent evidence, however, explains that judges do not simply balance endpoints.[13] Rather, their decisions include considerations of factors that cannot be analyzed instrumentally such as fairness and compassion.[14] Rather than similarity, the starting frame, we suggest, should be difference. When we start with difference, the dialogue on machine processing shifts. With the blinders of superiority off, the problem set and solutions change—not completely, but they change. The bias inherent in choosing machine processing becomes more apparent when we directly confront the way that perceptions of superiority marginalize particularly human traits, such as compassion and morality, while driving the loss of human agency, autonomy, and human flourishing.[15] It is these effects that have been hidden by the language and parsimonious visions of cognition behind the superiority narrative and deserve full airing in legal scholarship.

While we seek to change the center of gravity in the debate over machine deployment in law, we also seek to change the language we use to frame the discussion. Many scholars, of course, know that humans and machines "think" differently. So, how does this basic notion of difference get lost in the legal discussion? We suggest that the implicit assumption of similarity derives significantly from language.[16] Words like "artificial intelligence" and "machine learning" are powerful.[17] When we think of machines as "learning" or "intelligent" we tend to imbue them with human traits.[18]

---

[12] Jon Kleinberg et al., *Human Decisions and Machine Predictions*, 133 Q. J. ECON. 237, 242 (Feb. 2018).

[13] *See* discussion *infra* Section IV.B.1.

[14] *See* discussion *infra* Section IV.B.1.

[15] *See infra* Part III.

[16] *See infra* Part III, Section V.B.

[17] *See* Emily Tucker, Ctr. on Priv. & Tech., *Artifice and Intelligence*, MEDIUM (Mar. 8, 2022), https://medium.com/center-on-privacy-technology/artifice-and-intelligence%C2%B9-f00da128d3cd.

[18] *See id.*

Such language, we suggest, comes from, and also provides cover to, a mechanism whereby private interests and governments use machine processing to shape preferences with a concomitant loss of human agency, among other things.[19] We thus call on scholars to modify their language. Specifically, we suggest that scholars stop using words that describe machines as "intelligent" and "learning" and instead adopt language that describes them as what they are: calculators and processors. This is not just a better description of machines but also ensures that scholars are not complicit in the process of expanding machine use for commercial, rather than public, benefit.

We will proceed in the following manner: In Section II we will discuss the legal narrative and how it embodies and explicitly embraces a vision of machines as both similar and superior to "decision-makers." In Section III we explain the narrative of machine similarity and its roots in language and parsimonious visions of human cognition. In Section IV we discuss the implications of the narrative for legal scholarship and decision-making. Specifically, we discuss how the narrative marginalizes positive human traits such as compassion and concerns such as morality while painting machines as pure when, in fact, mathematical processes carry their own biases. In Section V we discuss how a narrative of superiority may impact deployment of machine processing in law and its effects on human agency before ultimately urging a change in vocabulary that counters the powerful language of similarity that pervades all our current understanding of machine processing. In Section VI we conclude.

II. HOW COURTS, REGULATORS, AND SCHOLARLY DISCUSSION REFLECT A NARRATIVE OF SIMILARITY AND SUPERIORITY

A. The General Frame

At the outset, let us note that a number of legal scholars recognize the limitations of machines.[20] However, the structure of the debate and the nature

---

[19] *Id.*

[20] *See, e.g.,* Harry Surden, *Machine Learning and Law: An Overview*, *in* RESEARCH HANDBOOK ON BIG DATA LAW 171 (Roland Vogl ed., Edward Elgar Publishing 2021) ("The phrase 'machine learning' may convey the false impression that the technology exhibits human-like intellectual abilities. This is not the case, as even the most advanced systems today lack the higher-order cognitive skills routinely displayed by humans, such as abstract reasoning, general learning or flexible problem-solving."); s*ee also* Aziz Z. Huq, *A Right to a Human Decision*, 106 VA. L. REV. 611, 634 (2020) ("[T]he state uses machine learning to make a range of high-volume decisions turning on empirical predictions for which a large pool of historical data is available. Absent sufficient training data, and absent some empirical

of language prototypes tend to obscure this basic understanding.[21] In this section, we discuss how the general framework of analysis embodies and reifies the view of machines as similar to humans but better. Indeed, where the notion of machines as acting similarly to humans but only better used to be implicit in legal scholarship, the idea has now been made explicit. We will explore this literature as well.

Legal scholarship often frames the debate on technology in law as one that pits the efficiency, consistency, and unbiased nature of machine processing against the imperfections of human cognition.[22] On one hand, technology is embraced as a means of making the legal process more efficient, fairer, and cheaper to deliver.[23] Included in this frame are visions of machines as unbiased,[24] able to operate at speed and scale,[25] and thus able to make better predictions about particular legal endpoints.[26] For example, a machine processing algorithm could grant or deny bail at the same rate as judges but reduce crime by 25 percent—or it could keep crime rates the same and reduce jailing by 42 percent.[27] Similarly, a machine-learning algorithm could undertake the same number of facility inspections as the Environmental Protection Agency currently does but find more than six times the number of

---

parameter to predict, machine decisions are not appropriate."). *But see* Cary Coglianese & Alicia Lai, *Algorithm vs. Algorithm*, 71 DUKE L.J. 1281, 1313 (2022).

[21] For examples, see *infra* notes 31-35 and related text.

[22] *See, e.g.*, Ric Simmons, *Big Data, Machine Judges, and the Legitimacy of the Criminal Justice System*, 52 U.C. DAVIS L. REV. 1067, 1067 (2018) ("These algorithms have the potential to increase the accuracy, efficiency, and fairness of the criminal justice system. However, some criticize them on the grounds that they may reinforce pre-existing biases against minorities."); Carla Reyes & Jeff Ward, *Digging into Algorithms: Legal Ethics and Legal Access*, 21 NEV L.J. 325, 334 (2020) ("Although reduction of bias and increased efficiency represent worthy goals for criminal justice system reform, critics increasingly raise concerns about the pervasive and indiscriminate use of algorithms for these purposes.").

[23] *See* Simmons, *supra* note 22, at 1072–74.

[24] *See* Reyes & Ward, *supra* note 22, at 326–27 ("In an attempt to eliminate human bias and error from the sentencing process in criminal proceedings, courts increasingly adopt technology tools for conducting recidivism risk assessments").

[25] *See* Margot E. Kaminski & Jennifer M. Urban, *The Right to Contest AI*, 121 COLUM. L. REV. 1957, 2004 (2021) ("AI decisions are made at speed and at scale—features that in fact can be core justifications for using AI in the first place.").

[26] Megan Stevenson & Jennifer Doleac, *Algorithmic Risk Assessment in the Hands of Humans*, SSRN 2–3 (Sept. 2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3489440.

[27] Coglianese & Lai, *supra* note 20, at 1311–12. *Cf.* Stevenson & Doleac, *supra* note 26, at.i ("Using simulations, we show that strict adherence to the sentencing recommendations associated with the algorithm would have led to some benefits (a sharp reduction in incarceration) but also some costs (a slight increase in recidivism and an increase in relative sentences for the young). Discretion mitigated the costs at the expense of reducing the benefits.").

violations.[28] Similar examples can be found in a variety of areas including domestic violence,[29] tax, food and drug law, municipal services delivery, among others.[30]

Another component of the frame sees "pure" machines as corrupted by human data and processes.[31] Machines, scholars note, are "black boxes" that make rooting out bias difficult, if not impossible.[32] Bias, in turn, is the result of machines operating on imperfect human data and programming, sometimes referred to as "garbage in, garbage out."[33] Given that machines operate at scale, bias will be perpetuated and exacerbated,[34] while non-transparent machine decisions, in general, do not provide anything akin to due process.[35]

---

[28] Coglianese & Lai, *supra* note 20, at 1311 ("Only about 10 percent of the more than three hundred thousand facilities subject to U.S. Environmental Protection Agency water pollution regulations receive government inspections in any given year, and normally only about 7 percent of inspected facilities are found noncompliant. But when using a machine-learning algorithm, inspectors could undertake the same number of inspections but find more than six times the number of regulatory violators—increasing the rate of violation detection to about 50 percent of all inspections.").

[29] Richard A. Berk, et al., *Forecasting Domestic Violence: A Machine Learning Approach to Help Inform Arraignment Decisions*, 13 J. EMPIRICAL L. STUD. 94, 105 (2016).

[30] For an excellent overview of the current use of technology in these fields, as well as in many other areas of adjudication and administrative law, see Cary Coglianese & Lavi M. Ben-Dor, *AI in Adjudication and Administration*, 86 BROOK. L. REV. 791 (2021).

[31] *See* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 3 (2015).

[32] *See, e.g.*, *id.*; *see also* Hannah Bloch-Wehba, *Access to Algorithms*, 88 FORDHAM L. REV. 1265, 1269 (2020) (discussing the difficulty in explaining decisions made via machine learning).

[33] Ron Ozminkowski, *Garbage in, Garbage out*, MEDIUM (Nov. 13, 2021), https://towardsdatascience.com/garbage-in-garbage-out-721b5b299bc1 (providing an overview of potential sources of bad information and how it affects machine learning models).

[34] *See, e.g.*, CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY 12–13 (2016); Ryan Calo & Danielle K. Citron, *The Automated Administrative State: A Crisis of Legitimacy*, 70 EMORY L.J. 797, 799–804 (2021); Valerie Schneider, Locked Out by Big Data: How Big Data, Algorithms and Machine Learning May Undermine Housing Justice, 52 COLUM. HUM. RTS. L. REV. 251, 259–60 (2020).

[35] Simmons, *supra* note 22, at 1072, 1074; Rashida Richardson et al., *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*, 94 N.Y.U.L. REV. ONLINE 192, 224–226 (2019); Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 636, 680, 692 (2017). *See also*, Coglianese & Lai *supra* note 20, at 1313 ("Expert judgments are often 'cryptic and mysterious' to those affected by their judgments. Even when humans explain their decisions, these accounts can be as much rationalizations as true reasons—a point legal realists made nearly a century ago with respect to judicial decision-making. People themselves often do not really know the reasons why they decided as they did. In many contexts, the resulting decisions can come about from 'implicit biases about which we are often unaware ourselves.' Indeed, for this

Interestingly, both the positive and negative concerns regarding the impact of technology share a similar vision of both human and machine decision-making. Whether looking at potential benefits or concerns, humans are seen as fallible and imperfect, while machine-decisions are efficient, objective, and consistent.[36] Benefits are simply derived from the superiority of machine decisions, while concerns arise from the fact that machine-made decisions are based on fallible human programming and data.[37]

Implicit in this debate is an understanding that humans and machines think alike. According to this view, a decision-maker identifies desired and undesired policy goals, gathers data, and analyzes the data to determine an optimal result—that is, an outcome that maximizes positive, while limiting negative, outcomes.[38] Judges, regulators, and policymakers are seen as "thinking" in this way when they make decisions.[39] One example is when they balance concerns regarding recidivism and jail time when making bail determinations.[40] So, too, are machines.[41]

B. Explicit Embrace and Advocacy of the Machine Similarity and Exceptionalism Narrative

While the general frame of machine-learning in law is amenable to and implicitly reflects a view of machine superiority, there are those who now explicitly embrace and advance that view. We, of course, are concerned with both. Indeed, implicit narratives about machines making better decisions than humans can be even more concerning than explicit claims simply because they are implicit and thus harder to recognize or remedy.[42] We also recognize that the relationship between the general and "explicit" frames is an iterative

---

reason, '[i]n many ways, human cognition forms the ultimate black box, even to the person engag[ed] in the cognitive activity.'") (internal citations removed).

[36] *See, e.g.*, Michael R. Siebecker, *Making Corporations More Humane Through Artificial Intelligence*, 45 IOWA J. CORP. L. 95, 105–06 (2019) (looking at the productivity benefits corporations have gained using AI).

[37] *See id.* (discussing the benefits of AI and the need for corporations to maintain accurate data and review for accuracy and potential bias).

[38] *See* Cary Coglianese & David Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, 105 GEO. L.J. 1147, 1202 (2017) ("Assuming that the governmental interest to which an agency's machine-learning algorithm is applied is legitimate, it should be quite difficult for that algorithm to be struck down on equal protection grounds. An algorithm's objective function will itself provide a concrete, mathematical embodiment of the agency's goal, sufficient to demonstrate a rational relationship between the algorithm and the governmental interest at stake.").

[39] *See* Eugene Volokh, *Chief Justice Robots*, 68 DUKE L.J. 1135, 1158–59 (2019) (comparing human decision making with AI judgements in the legal field).

[40] *See* Coglianese & Lai, *supra* note 20, at 1311–12.

[41] *See id.*

[42] *See* Tucker, Ctr. Priv. & Tech, *supra* note 17.

one. It is, however, worthwhile identifying the explicit frame along with the ways in which the general frame embraces machine exceptionalism for purposes of analysis.

Perhaps the most complete argument for machine exceptionalism comes from Cary Coglianese and Alicia Lai. In their article, *Algorithm vs. Algorithm*,[43] the authors describe both human beings and technology as thinking algorithmically[44]—that is, using reason to analyze data toward particular ends. Conceived of this way, the superiority of machines becomes an almost foregone conclusion. As the authors note:

> [T]hese human algorithms undeniably fail and are far from transparent. On an individual level, human decision-making suffers from memory limitations, fatigue, cognitive biases, and racial prejudices, among other problems. On an organizational level, humans succumb to groupthink and free riding, along with other collective dysfunctionalities. As a result, human decisions will in some cases prove far more problematic than their digital counterparts.[45]

To these scholars, then, the choice between technology and humans is primarily a choice between who is the better processor.[46] In many cases, the answer is technology. While humans get tired,[47] can't control their impulses,[48] and are subject to biases[49] and imperfect group dynamics,[50] machines "facilitat[e] outcomes that are more accurate, timely, and consistent."[51] Of course, as Coglianese and Lai are quick to point out, machines do have limitations, especially when dealing with notions of fairness and other similarly amorphous concepts.[52] It is the assumption of common "processing," however, that allows them to describe machine algorithms as superior decision-makers generally.[53]

Professor Coglianese's view of how law reflects outcome-based reasoning is made even clearer in *Regulating by Robot*, co-authored with

---

[43] Coglianese & Lai, *supra* note 20, at 1281.
[44] *Id.*
[45] *Id.*
[46] *See id.*
[47] *Id.* at 1290–91.
[48] *Id.* at 1292.
[49] *Id.* at 1287, 1295–96.
[50] *Id.*
[51] *Id.* at 1281.
[52] *See id.* at 1324–25. It is worth noting that this limitation may well be an exception that swallows the rule given the ubiquity of concepts like fairness in the law.
[53] *Id.*

David Lehr.[54] In their article, the authors describe the potential for agent-based models or multi-agent systems to be used to make normatively complex legal rules autonomously by employing outcome-based reasoning.[55] For example, they suggest that machines could model human responses to regulation by being programmed based on the assumption that the regulated community desires to maximize profit.[56] The regulator could then feed the machine a variety of different regulatory structures and determine which regulation maximized particular regulatory goals. As the authors note, using the Occupational Safety and Health Administration (OSHA) as an example regulator:

> But the techniques' real potential to inform the content of regulations comes from the ability of OSHA to include an agent representing itself in the ABM [agent-based model]. This mathematically-represented agent would "issue" multiple different possible regulations—formulated in advance by human programmers—and then "select" the regulatory alternative that yields those effects, as defined in relation to observable components of the environment, that maximize an objective function (or goal) established by the real-world OSHA. The possible regulations analyzed in this fashion could assume any number of different combinations of regulatory targets, commands, and consequences, with the forecasted effect of these regulations on the actions of regulated entities being observed through the modeling exercise.[57]

Such suggestions again focus primarily on endpoints. The law sets the goals—say, balancing crime rates and jailing, or in the case of regulation, profit maximization and environmental protection—and decision-makers attempt to maximize these results.[58] "Human" input, in turn, is conceived of simply as the process of identifying the goals to maximize and minimize.[59] Factors such as empathy, emotion, and morals, it seems are not necessary in

---

[54] *See* Coglianese & Lehr, *supra* note 38.

[55] *Id.* at 1172–73.

[56] *Id.* at 1174 ("The agent-based model would therefore use a machine-learning technique to select employers' optimal responses to the regulation given their profit maximization goal.").

[57] *Id.* at 1174.

[58] *See id.*

[59] *Id.* at 1174 ("As already indicated, at a foundational level, humans will still need to choose and then input into embedded machine-learning systems the data, as well as overarching goals to be maximized andconstraints to be minimized.").

this ends-oriented model of regulatory decision-making, or are, at best, transformed from what they are into "limited data and core uncertainties."[60]

Taking a slightly different tack, Professor Eugene Volokh's argument regarding the use of machine processing in judicial decision-making does not explicitly adopt a vision of humans and machines as similar. Rather, he suggests that as long as machines can mimic human results, the way they process shouldn't matter. In his article, *Chief Justice Robots*,[61] Volokh assumes the existence of a machine processing algorithm that can write persuasive judicial opinions.[62] In such a case, he argues, whether a machine has human traits or not should not matter because the result of a decision is more important than the factors that go into it. As he notes:

> But here again what matters is the result, not the process. If a poetry-translation program reliably produces translations that are emotionally rewarding for us as readers, it should not matter to us that Robot Frost can't itself have emotions. If, in a blind test, we view an AI sentencing judge as producing wiser and more compassionate results—by our lights—than a human sentencing judge, it should not matter to us as evaluators that the judge can't have "wisdom" or "compassion."[63]

Unlike Coglianese and Lai, Volokh's view that robot judging is valid is based not on a general view of cognition but on a parsimonious view of the law.[64] If judicial decision-making is simply about optimizing endpoints, then there is little need for human judges and the particularly human traits they bring to the courtroom.[65]

In sum, legal scholarship embodies a narrative that conceives of humans and machines as thinking similarly. When conceived of this way, the narrative also paints machines as superior to humans as legal decision-makers.[66] The problem, as conceived, is that humans are limited, biased and

[60] *Id.* at 1174–75.

[61] Volokh, *supra* note 39.

[62] *Id.* at 1152 (noting that the decision must be persuasive).

[63] *Id.* at 1189.

[64] *Compare* Coglianese & Lai, *supra* note 20, at 1313 (analyzing the differences between machine and human decision-making processes), *with* Volokh, *supra* note 39, at 1152 (claiming process and "human" elements are not important—only endpoints matter).

[65] *See* Volokh, *supra* note 39, at 1152 (explaining the importance of persuasion in litigation).

[66] *See, e.g.*, Coglianese & Lai, *supra* note 20; Coglianese & Lehr, *supra* note 38; *cf.* Volokh, *supra* note 39.

inefficient.[67] The solution is to respond to these limitations by opening up the black box and finding ways to drive the human bias out of "pure" machine processing.[68] Positive human traits like compassion and fairness have little place in this discourse.

The vision of machines thinking like humans is also beginning to impact the perception of judges and legislators. In the next section, we will discuss how legal decision-making is also embracing this view.

C. Machine Superiority in Law-Making

Courts and regulators have also begun to apply reasoning that reflects and reifies the vision of humans and machines thinking alike. These matters usually arise in the context of how much, and to what extent, autonomous[69] machine calculations can be utilized in the law. The issue has arisen in contexts as varied as when machine-made decisions amount to the practice of law,[70] whether machines should be liable in tort when algorithms cause harm,[71] and whether algorithmic decisions can be properly confronted by criminal defendants under the law.[72]

As might be expected, earlier cases seem more skeptical of the ability of machine calculations to displace human ones. For example, one decision of the Western District of Missouri found that LegalZoom, which relied on human decision-making to help individuals prepare legal documents, was practicing law without a license.[73]

More recently, however, as the narrative of similarity has taken root, the idea that human and machine determinations are similar has found its way into legal determinations. A recent New York case on the issue of confrontation under the Sixth Amendment is instructive.[74] In that case, the

[67] Coglianese & Lai, *supra* note 20, at 1281 (arguing the weaknesses of human decision-making).

[68] *See id.* at 1313.

[69] For a general discussion of the scholarship and law, see Karni A. Chagal-Feferkorn, *Am I An Algorithm or a Product? When Products Liability Should Apply to Algorithmic Decision-Makers,* 30 STAN. L. & POL'Y REV. 61, 74 (2019).

[70] Janson v. LegalZoom.com, Inc., 802 F. Supp. 2d 1053, 1065 (W.D. Mo. 2011) (holding LegalZoom to be illegally practicing law).

[71] *See* Chagal-Feferkorn, *supra* note 69, at 72–77.

[72] People v. Wakefield, 195 N.E.3d 19, 21 (N.Y. 2022). For discussions of the case, see Isaac Figueras, *The LegalZoom Identity Crisis: Legal Form Provider or Lawyer in Sheep's Clothing?,* 63 CASE W. RES. L. REV. 1419 (2013); Frank Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, 87 GEO. WASH. L. REV. 1 (2019). *But see* Coglianese & Ben-Dor, *supra* note 30, at 798, 801 (2021) ("Courts have not yet addressed the use of autonomous machine determinations in judging directly.").

[73] *Janson*, 802 F. Supp. 2d at 1063-64.

[74] *See Wakefield*, 195 N.E.3d at 21.

state relied on an AI called "TrueAllele" to analyze a DNA sample that supported the defendant's conviction for murder.[75] "[D]efendant moved for disclosure of the source code in order 'to meaningfully exercise his constitutional right to confront his accusers.'"[76] The defendant argued that the report generated by TrueAllele was testimonial, that the computer program was the functional equivalent of a laboratory analyst, and that the source code was the witness that must be produced to satisfy his right to confrontation.[77] New York's intermediate appellate court recognized that while AI could be a declarant, in this case, TrueAllele did not qualify as such.[78] The court—whose ultimate holding the state's highest court upheld—used the concept of "distributed cognition" to explain that there is a "continuum of cognition between human and machine"[79] but found that in this case, humans had enough input into the decision that the machine was not the declarant.[80] Specifically, the court noted:

> [T]here is human input when utilizing TrueAllele. Among other things, a human analyst tells the computer what to download and under what conditions to analyze the data, the analyst tells the computer what questions to ask when interpreting the data and the analyst downloads certain results from the computer, the analyst determines how many "runs," or cycles, of the data the system will complete and the analyst

[75] *Id.* at 21–23.

[76] *Id.* at 26.

[77] *Id.* at 25.

[78] People v. Wakefield, 107 N.Y.S.3d 487, 497 (N.Y. App. Div. 2019), aff'd, 195 N.E.3d 19 (N.Y. 2022).

[79] *Id.* The court's use of the term "distributed cognition" is borrowed from Itiel E. Dror & Jennifer L. Mnookin, *The Use of Technology in Human Expert Domains: Challenges and Risks Arising from the Use of Automated Fingerprint Identification Systems in Forensic Science,* 9 L., PROBABILITY & RISK 47 (2010). In their article, the authors frequently refer to a partnership between humans and machines but do not refer to machines as replacing humans. *See id*. at 48. The notion of distributed cognition does include an understanding of how machines can support human cognition. However, it requires consideration of the proper role of machine and human and does not suggest their interchangeability. Indeed, the notion of distributed cognition generally recognizes difference in machine and human cognition. *See* E.H. Shortliffe & V.L. Patel, *Internet: Psychological Perspectives*, *in* INTERNATIONAL ENCYCLOPEDIA OF THE SOCIAL & BEHAVIORAL SCIENCES 7852 (Neil J. Smelser & Paul B. Baltes eds., 1st ed. 2001) ("With the recent emergence of distributed cognition as a framework for considering group decision making . . . there has been increasing research regarding the role of technology in supporting cognitive activities distributed among a number of agents, potentially consisting of both humans and machines. This perspective emphasizes understanding the role of computers and technology, including the Internet, in the context of workplaces and for tasks that often involve collaboration.").

[80] *Wakefield*, 107 N.Y.S.3d at 497.

> then makes comparisons to form the likelihood ratios. Also key to our analysis is that Perlin, the creator of TrueAllele and the individual who wrote the underlying source code, was present in court and testified, at length, as to genetic science, the TrueAllele program and the formulation of the TrueAllele report through the computer processors and algorithms, including the MCMC algorithm . . . . Given the totality of the circumstances present here, we find that Perlin was the declarant in the epistemological, existential and legal sense rather than the sophisticated and highly automated tool powered by electronics and source code that he created.[81]

In the case of TrueAllele, the court ultimately determined that there was enough human input in places where judgment was necessary to find that the machine was not necessarily "autonomous."[82] Yet, according to the court, there will be situations where autonomous decisions will make a machine a declarant under the law.[83]

Similar instincts to blur the line between humans and machines when sophisticated algorithms are used outside of human control exist behind regulatory decisions on algorithmic decision-making. The European Parliament, for example, in considering ways to deal with liability issues for machine-made decisions, has suggested the creation of a new category of "being" in law—the "electronic person."[84] Specifically, the Commission is considering:

---

[81] *Id.*

[82] *Id. See also Wakefield*, 195 N.E.3d at 30–31 (showing the artificial intelligence analysis required human review).

[83] *Wakefield*, 107 N.Y.S.3d at 497.

[84] Report with Recommendations to the Commission on Civil Law Rules on Robotics, at ¶59(f), EUR. PARL. DOC. 2015/2103 (INL) (2017) ("[Consider] creating a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently[.]"), https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html?redirect. *See also* Draft Report with Recommendations to the Commission on Civil Law Rules on Robotics, EUR. PARL. DOC. 2103 (INL) (2016), https://www.europarl.europa.eu/doceo/document/JURI-PR-582443_EN.pdf?redirect [hereinafter European Parliament 2016 Draft Report]; Andrea Bertolini, *Artificial Intelligence and Civil Liability*, EUR. PARL. COMM. ON LEGAL AFFS. (2020), https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf.

> creating a specific legal status for robots, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons with specific rights and obligations, including that of making good any damage they may cause, and applying electronic personality to cases where robots make smart autonomous decisions or otherwise interact with third parties independently.[85]

The proposal has sparked significant debate.[86] On the one hand are those who suggest giving robots legal personhood is similar to giving corporations legal personhood.[87] On the other are those who suggest that doing so is wrong both as a matter of "law and ethics."[88]

Of particular concern to opponents is the misperception of machine calculation versus human cognition. In a letter to Parliament, a group of over 150 experts explained that state-of-the-art programs excel only in one narrow field, like playing *Go*[89] or categorizing images, and that such machines are "[completely] unlike human beings, who can at the same time understand language, learn to play a variety of board games and recognize images."[90] The signatories to the letter worried that "media reports about galloping advances in robotics—which may suggest all-encompassing robot intelligence is within reach—have infiltrated public debate"[91] and may, in turn, "lead lawmakers to rush into premature regulation."[92] That both judges and regulators fall victim to a false view of machine calculations and human decisions being similar is of particular concern for the law. As critics of the European Parliament suggest, the view of robot intelligence may lead to premature—or more to the point—incorrect regulation.[93]

---

[85] European Parliament 2016 Draft Report, *supra* note 84.

[86] *See* Sergio M. C. Avila Negri, *Robot as Legal Person: Electronic Personhood in Robotics and Artificial Intelligence*, 8 FRONTIERS ROBOTICS & AI 1, 6–9 (Dec. 23, 2021).

[87] Janosch Delcker, *Europe Divided over Robot "Personhood,"* POLITICO (Apr. 11, 2018, 12:45 PM), https://www.politico.eu/article/europe-divided-over-robot-ai-artificial-intelligence-personhood/; Negri, *supra* note 86.

[88] Delcker, *supra* note 87.

[89] "Go" is a strategy board game. *See* Emily Willingham, *AI's Victories in Go Inspire Better Human Game Playing*, SCI. AM. (Mar. 13, 2023), https://www.scientificamerican.com/article/ais-victories-in-go-inspire-better-human-game-playing/.

[90] Delcker, *supra* note 87.

[91] *Id.*

[92] *Id.*

[93] *Id. Wakefield*, 107 N.Y.S.3d at 497. The unreflective use of concepts like "distributed cognition" and "electronic person" also adds additional language to the narrative of machine exceptionalism. For a discussion of how language frames our understanding of machines, see *infra* Section V.B.2.

III. HUMANS AND MACHINES DO NOT THINK ALIKE. WE JUST ASSUME THEY DO. LANGUAGE CONTRIBUTES

For us to think that machines are superior, they need to be similar. That is, for a machine to be seen as better or pure in relation to humans, they must actually process like humans. Otherwise, they are not necessarily better but simply different. In this section, we explain the roots of the basic framework that sees machine calculations as cognitively similar and superior to human thought. We also argue that such a vision is both wrong and reductionist. We do not mean to suggest that efforts to reduce the impacts of human bias are wrong. Rather, we seek to focus on what is missing in the current debate and to change some basic assumptions that flow from a view that treats machines as similar to humans but purer and better.

Human and machine "cognition" are truly distinct processes. The use of the terms "artificial intelligence" and "machine learning" are historical artifacts situated in a particular moment in the 1970s known as the "AI Winter," a moratorium or at least significant slow-down in AI research.[94] This slowdown was the result of the continued failure of machine learning to work as hoped.[95] As scholars including Stephanie Dick, Os Keyes, and Nikki Stevens have explored, early experiments in AI attempted to program machines to mimic the workings of the human brain. These experiments were a way to study said processes in people, in a range of domains from machine translation and face recognition to decision-making more broadly.[96] Thus, early efforts at machine learning were motivated by research into cognition, using machines as proxies for human processes as a way to learn more about them and translate them to the machines.[97] The idea was not only to learn how the human brain worked but also to then artificially mimic and scale it for size and speed.[98] If researchers could design a machine that would, say, effectively translate languages, that would yield valuable information about how human translators thought, further improving the machine programming and advancing knowledge about neurology and neuropsychology.[99] Likewise

---

[94] Steven Umbrello, *AI Winter, in* ENCYCLOPEDIA OF ARTIFICIAL INTELLIGENCE: THE PAST, PRESENT, AND FUTURE OF AI 7–8 (Michael J. Klein & Philip L. Frana, eds., 2021).

[95] *Id.*

[96] *See* Stephanie Dick, *Artificial Intelligence*, 1 HARV. DATA SCI. REV. 1, 2 (July 1, 2019), https://hdsr.mitpress.mit.edu/pub/0aytgrau/release/3; Nikki Stevens & Os Keyes, *Seeing Infrastructure: Race, Facial Recognition and the Politics of Data*, 35 CULTURAL STUD. 833, 837, 840 (2021); Michael D. Gordin, *The Forgetting and Rediscovery of Soviet Time Machine Translation*, 46 CRITICAL INQUIRY 835, 835–66 (2020).

[97] Dick, *supra* note 96; *see* Stevens & Keyes, *supra* note 96, at 837, 840.

[98] Umbrello, *supra* note 94.

[99] *See id.*

with face recognition; there was little to be learned and less to be gained from trying to make machines do the work of people.[100]

What quickly emerged is that such an approach exposed very little about human cognition and yielded worse outcomes in the machines.[101] These programs were slow, clunky, inelegant, and certainly did not work as well or better than people.[102] Early attempts at computerized face recognition involved a huge amount of manual labor, tagging each feature to correspond with the way humans saw and evaluated them.[103] As Nikki Stevens and Os Keyes discuss, this process was time-consuming and required programmers to individually locate where features fell on each and every image.[104] The focus was to eliminate human intervention in the recognition process itself, but it required a huge amount of work to make the faces recognizable to the machines.[105] The process became considerably more streamlined in 1987 with the introduction of eigenfaces, a kind of algorithm that translated a large dataset of human faces into data that probabilistically located common facial features and evaluated them across faces.[106] This, needless to say, is not how human brains identify and compare features; it does, however, allow machines to work much more effectively and quickly.[107]

In response to the failures of these early efforts to mimic human cognition, programmers changed their strategy. AI processes became much more effective when programmers ceased to use them as a means to understand people but rather as a way to solve certain kinds of problems through data collection and algorithms.[108] Machines can do quite a good job at sorting data and following heuristics, offering as a result, say, a translated sentence or a data match between two faces.[109]

But the process by which they do this is drastically different than what people do, even if the outcomes look similar. When machines offer a conclusion, they are not thinking: They are reducing a particular problem to a series of data points and then solving a mathematical equation.[110] This process is a complicated and creative one that involves programmers figuring out *how* to make the decision in question into a mathematical problem and

[100] *See* Stevens & Keyes, *supra* note 96, at 839–40.
[101] *See id.* at 837.
[102] *See id.*
[103] *See id.* at 837, 840.
[104] *Id. at 837.*
[105] *Id.*
[106] *Id.* at 840.
[107] *See id.*
[108] *See* Dick, *supra* note 96, at 5.
[109] *See generally* Gordin, *supra* note 96 (discussing machine translation); Stevens & Keyes, *supra* note 96 (discussing facial recognition).
[110] *See* Dick, *supra* note 96, at 5.

then developing a way to capture the data points and inputting them into a heuristic.[111]

In the case of face recognition, for example, the various parts of the face, including the distance between given features and their respective placement and protrusion, reading of skin tone and color, and best estimate of gender based on these analyses, are captured and converted into mathematical measurements that are then matched with those stored in the database of similarly converted faces.[112] The closest overlap between these numbers is declared a match.[113] Such heuristics work much more effectively at scale and speed than people do.[114] But they are not, and must never be, mistaken for thought and cognition. Product is not process; machines are not people. When most people recognize others, we take into account not just facial features but gait, tone of voice, style of dress, hair color, and—importantly—our past encounters and experiences with them. Even people who are unable to recognize features—those with face blindness, or prosopagnosia—can sometimes identify people through these other cues.[115] We make mistakes. All the time.

The conflation of AI processing and human reasoning has created a misrecognition of product for process. Machines don't think or reason, but our use of language infuses machine processes with a kind of cognition that does not exist. Language matters.[116] When we talk about machine "learning," we credit an iterative accumulation of data comparisons with precisely the kind of rationality and reason that advocates at the same time profess to hope to excise from these processes.[117] It's a kind of double-dipping that credits AI mechanisms with a purity from bias while at the same time drawing on what is valuable about human cognition, reflection, emotion, and, indeed, reason. Words like "learning" and "intelligence" lead us to imbue machine processes with human cognition, reason, and emotion.[118] This language slippage non-accidentally creates consonance between what people do when they make decisions and what machines do when they apply heuristics.[119]

---

[111] *See* Steven & Keyes, *supra* note 96, at 840, 845.

[112] *See id.*

[113] *See id.*

[114] *See id.*

[115] Sharrona Pearl, *Watching While (Face) Blind: Clone Layering and Prosopagnosia, in* ORPHAN BLACK: PERFORMANCE, GENDER, BIOPOLITICS 78, 78-91 (Andrea Goulet & Robert A. Rushing, eds. 2019); Sharrona Pearl, *A Super Useless Super Hero*, 7 SEMIOTIC REV. (Sept. 20, 2019).

[116] *See* discussion *infra* Section V.B.2.

[117] *See* Tucker, Ctr. Priv. & Tech., *supra* note 17.

[118] *See id.*

[119] Emily M. Bender, *Look Behind the Curtain: Don't be Dazzled by Claims of "Artificial Intelligence"*, SEATTLE TIMES (May 11, 2022, 2:47 PM),

Given its roots in language and media, the narrative seems to share some traits with implicit bias. Of course, implicit bias is a relatively complex phenomenon that defies simple explanation.[120] At its root, however, implicit bias has links to social stereotypes that are themselves based on how individuals gather information about social groups and other things.[121] Increasingly, research demonstrates that language, even more than direct human experience, informs the creation of these stereotypes.[122] Media representations influence stereotype creation as well.[123] The power of language prototypes to influence our perception of something like "machine learning" should not, however, be understated.[124] Simple phrases like "machine learning" or "artificial intelligence" can powerfully impact our construction of the processes they describe.[125] We endeavor to avoid this language in this article by using phrases like "machine calculation" and "machine processing" because they more clearly reflect what machines actually do. We will discuss the origin of words like "artificial intelligence" later in this article.[126]

There is another dimension to the narrative that we need to raise here particularly because the analysis of machine learning in law draws heavily on social science literature. The "common vision" of cognition attributed to both computers and humans bears a deep resemblance to the cognitive model

---

https://www.seattletimes.com/opinion/look-behind-the-curtain-dont-be-dazzled-by-claims-of-artificial-intelligence/.

[120] For a general overview, see *Implicit Bias*, STAN. ENCYCLOPEDIA PHIL. (Jul.31, 2019), https://plato.stanford.edu/entries/implicit-bias/.

[121] *Id.*

[122] *See* Benedek Kurdi & Mahzarin R. Banaji, *Implicit Social Cognition: A Brief (and Gentle) Introduction*, *in* THE COGNITIVE UNCONSCIOUS: THE FIRST HALF CENTURY 323, 388 (Arthur S. Reber & Rhianon Allen, eds. 2022). "Notably, studies focusing on basic mechanisms of acquisition and change have also demonstrated time and again the power of language in shifting implicit social attitudes, which often exceeds that of interventions relying on direct experience." *Id.* (citing Jan De Houwer, *Using the Implicit Association Test Does Not Rule Out an Impact of Conscious Propositional Knowledge on Evaluative Conditioning*, 37 LEARNING & MOTIVATION 176 (2006) and Thomas C. Mann et al., *How Effectively Can Implicit Evaluations be Updated? Using Evaluative Statements After Aversive Repeated Evaluative Pairings*, 146 J. EXPERIMENTAL PSYCH. 1169 (2020)).

[123] Alexis S. Tan, *Television Use and Social Stereotypes*, 59 JOURNALISM & MASS COMMC'N Q. 119, 120 (1982); Srividya Ramasubramanian & Chantrey J. Murphy, *Experimental Studies of Media Stereotyping Effects*, *in* LABORATORY EXPERIMENTS IN THE SOCIAL SCIENCES 385, 385-402 (Murray Webster & Jane Sell, eds., 2d. ed. 2014); Jillian Gilmour, *Formation of Stereotypes*, 2 BEHAVIOURAL SCIS. UNDERGRADUATE J. 67, 68 (2015).

[124] *See* ARTHUR HOLLAND MICHEL, RECALIBRATING ASSUMPTIONS ON AI 7–8 (Chatham House 2023), https://www.chathamhouse.org/sites/default/files/2023-04/2023-04-05-recalibrating-ai-holland-michel.pdf (discussing policies which misrepresent AI potential).

[125] *Id.* at 12, 15.

[126] *See infra* Section V.B.2.

contained in rational choice theory.[127] Rational choice has been called the most successful paradigm of human decision-making in the social sciences[128] and continues to hold sway over many fields. Pursuant to the theory, humans with set and stable preferences use relevant information to reason logically toward ends that provide the best result by balancing preferences in an optimal manner.[129] Rational choice has been described as a "parsimonious" theory of human cognition.[130] However, given its influence on thinking in the social sciences, it provides a powerful metaphor of cognition that allows individuals to equate what machine learning does with how humans think. Of particular relevance, rational choice conceives of people as reasoning logically from a set of particular preferences to outcomes that maximize their

---

[127] *See generally* Boris Julián Pinto-Bustamante et al., *Bioethics and Artificial Intelligence: Between Deliberation on Values and Rational Choice Theory*, 10 FRONTIERS (June 19, 2023), https://www.frontiersin.org/articles/10.3389/frobt.2023.1140901/full (detailing how rational choice theory is the basis for artificial intelligence).

[128] *See* Peter Abell, *Sociological Theory and Rational Choice Theory*, *in* THE BLACKWELL COMPANION TO SOCIAL THEORY 223, 223-44 (Bryan S. Turner, ed., 2d. ed. 2000) ("Even with all its manifest limitations, rational choice theory has arguably proven to be the most successful theoretical framework in . . . social sciences.").

[129] Scott Ainsworth, *Rational Choice Theory in Political Decision Making*, OXFORD RESEARCH ENCYCLOPEDIA OF POLITICS (Sept. 28, 2020), https://oxfordre.com/politics/view/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-1019 ("Rational choice theory builds from a very simple foundation. To wit: individuals are presumed to pursue goal-oriented behavior stemming from rational preferences."). For discussions of how this model influences legal analysis, *see, e.g.*, Cary Coglianese & David Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, 105 GEO. L.J. 1147, 1202 (2017) ("Assuming that the governmental interest to which an agency's machine-learning algorithm is applied is legitimate, it should be quite difficult for that algorithm to be struck down on equal protection grounds. An algorithm's objective function will itself provide a concrete, mathematical embodiment of the agency's goal, sufficient to demonstrate a rational relationship between the algorithm and the governmental interest at stake."); Valerie Schneider, *Locked Out by Big Data: How Big Data, Algorithms and Machine Learning May Undermine Housing Justice*, 52 COLUM. HUM. RTS. L. REV. 251, 259-60 (2020) ("The purpose of using big data, algorithms, and artificial intelligence is to 'provide a rational basis upon which to distinguish between individuals'—e.g., to predict if one applicant for an apartment is more likely than another applicant to pay rent on time or abide by the lease.") (citation omitted); Carla Reyes & Jeff Ward, *Digging into Algorithms: Legal Ethics and Legal Access*, 21 NEV. L.J. 325, 333 ("Increased interest in machine learning techniques—by which computers crunch data using an algorithm to perform its assigned objective function, make predictions, and automate certain tasks—stems in large part from recent advances in computer processing speed, some advances in algorithms, and the rise of big data.").

[130] *See, e.g.*, Gregory Mitchell, *Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence*, 91 GEO. L.J. 67, 119 (2002) (explaining the basic trade-off between parsimony and prediction).

satisfaction.[131] This, in turn, facilitates the narrative by providing a pre-existing model of human cognition that allows us to conceive of human thinking in a framework that is similar to machine learning.

While our initial focus is on rational choice as the source of the narrow vision of law and rationality that underlies the machine narrative, we recognize that other forces have also contributed to a narrow view of law as a means of advancing certain ends. Professor Henry Smith describes these forces in his recent efforts to revitalize the role of equity in legal decision-making. He discusses how reductionist views of law, along with a perception of law as engaged in the process of endpoint maximization, led to the marginalization of equity as a component of law:

> Because the fusion of law and equity occurred in an era when the law's formalism was increasingly unmoored from natural rights and natural law, the structures of common law and equity alike were regarded in increasingly positivist and reductionist terms. The law was treated as a collection of rules and flattened out. This reductionism, which seems almost like second nature now, can likewise be traced to the fount of modern thinking on the subject, Holmes's *Path of the Law*. He avers that "a body of law is more rational and more civilized when every rule it contains is referred articulately and definitely to an end which it subserves, and when the grounds for desiring that end are stated or are ready to be stated in words." This reduces law to a heap of rules with no synergistic effect. Rule and purpose have to be matched directly, leaving little room for doctrines to act in concert or for the law to exhibit a more articulated structure.[132]

As one can see, the forces that shaped the view of equity described by Smith assert a similar influence on the law as our description of rational choice.[133] The rationalizing force of Holmes, he notes, requires law to be directly connected to particular results, leaving little room for complexity, morality,

---

[131] For a description of rational choice as a means of preference maximizing, *see generally* Susanne Lohmann, *Rational Choice and Political Science*, THE NEW PALGRAVE DICTIONARY OF ECONOMICS (Steven N. Durlauf & Lawrence E. Blume, eds., 2d ed. 2008); Peter Hedström & Charlotta Stern, *Rational Choice and Sociology*, THE NEW PALGRAVE DICTIONARY OF ECONOMICS (Steven N. Durlauf & Lawrence E. Blume, eds., 2d ed. 2008); S.M. Amadae, *Rational Choice Theory*, Britannica (2023), https://www.britannica.com/money/topic/rational-choice-theory.

[132] Henry E. Smith, *Equity as Meta-Law*, 130 YALE L.J. 1050, 1063 (2021).

[133] *See id.*

or other factors in decision-making.[134] We recognize that our description of these concerns within rational choice is, ironically, itself reductionist. However, the rise of rational choice as a model of cognition in law and other social sciences is undeniable.[135] We will use the terms rational choice and ends-oriented decision-making throughout this article to capture this broad phenomenon.

To summarize, the consistent use of inaccurate cognition language in the context of AI algorithms, along with readily available models of cognition that equate human and machine thinking, blinds us to the differences between human and machine "cognition." The narrative leads to the valuation of decisions and outcomes that are based on deliberations that take into account data and its weighting but do no more. It simply, as a frame, leaves out of the discussion a wide variety of factors that make human decisions valuable, and in so doing, marginalizes them.

IV. REFRAMING THE NARRATIVE TO REFLECT DIFFERENCE

In the previous sections we have described the rise of a narrative describing machines as thinking similarly to humans. This in turn reflects the influence of language that describes machines as "intelligent" and "learning." As we have discussed, language may be even more powerful in structuring our understanding of machines than our experience with them.[136] It also reflects the influence of rational choice theory as a model of cognition and decision-making.[137] Rational choice and other legal forces[138] provide a powerful model of human cognition as outcome optimization, which translates seamlessly into legal decision-making. If we model both human and legal decision-making in terms of reasoning probabilistically toward the maximization of outcomes, we will see little difference between machines and humans.

The structure of discourse, we argue in this section, has implications for how we frame the issues relevant to the deployment of machine calculating in law and their implications. As we noted *supra*, narratives have consequences.[139] Specifically:

> Like all other narratives, policy narratives are stories about a particular subject, and involve actors, contexts, and actions.

---

[134] *See id.*

[135] *See* MARK I. LICHBACH, IS RATIONAL CHOICE THEORY ALL OF SOCIAL SCIENCE? 3–4 (2003).

[136] *See supra* note 15 and accompanying text.

[137] *See* Smith, *supra* note 132, at 1063.

[138] *See id.*

[139] *See supra* Part III.

> These stories are told to, among other things, create shared understandings, shape problem definitions as well as the ranges and types of solutions, and influence internal and external audiences' perceptions, attitudes, and behaviors. Policymakers also use narratives to directly or indirectly influence the actions of actors and to advance their policy goals. Thus, policy narratives are strategically constructed and carry purpose and intention. Exploring narratives is an integral part of policy research.[140]

The current narrative of similarity, we will argue in this section, frames machine calculating as superior in ways that it might not actually be and marginalizes a number of human traits, particularly components of human cognition that can't be quantified.[141] The frame of similarity also casts machines as "pure" in their calculation of endpoints, ignoring the bias that comes from use of machine calculation.[142] This is not human bias; rather it is racial and gender bias that comes from the choice of mathematical models of decision-making over human ones.[143] Once we have explained the narrative's failures, we will turn to how the current literature blinds itself to difference, and ultimately, we will consider how difference changes the discourse.

A. *How the Narrative Structures Our Understanding of the Issues*

1. Superiority Versus Difference and Marginalizing Human Traits

Let us turn to an example of what we mean when we say that the literature ignores difference. Perhaps the most used and simplest of these examples are frequent references to how machines make better decisions than humans in a variety of contexts. We have provided examples of these relatively ubiquitous claims earlier.[144] For example, a machine calculating algorithm could grant or deny bail at the same rate as judges but reduce crime by 25 percent or it could keep crime rates the same and reduce jailing by 42 percent.[145] These examples fit nicely into the superiority framework. They show machines as making the same types of decisions as judges or other legal

[140] Guenduez, *supra* note 2 (internal citations omitted).

[141] *See* Eric A. Posner & Cass R. Sunstein, *Moral Commitments in Cost-Benefit Analysis*, 103 VA. L. REV. 1809, 1857–59 (2017) (discussing the difficulty in quantifying moral values in a cost-benefit analysis of regulations).

[142] *See* discussion *supra* Part III.

[143] *See, e.g.*, Eldar Haber, *Racial Recognition*, 43 CARDOZO L. REV. 71, 95-97 (2021) (discussing racial bias in facial recognition technology).

[144] *See supra* notes 35–44 and accompanying text.

[145] Coglianese & Lai, *supra* note 20, at 1311–12.

decision-makers but only better. The claim of superior decision-making power is often accompanied by a claim about human bias.[146] Thus, the machines not only predict better but they also do so while avoiding visiting their results on particular minorities or other groups.

How might human decisions diverge from the outcomes of machine calculations? Recent empirical evidence provides insights.[147] This evidence shows that judges don't make decisions in the way machines do.[148] When deciding, judges tended to start by veering from machine calculations in particular ways. For example, in their recent analysis of judicial use of algorithmic risk assessment in Virginia courts, Professors Megan Stevenson and Jennifer Doleac found that judges particularly veered from risk assessments of recidivism by decreasing sentences for youth.[149]

The decision to decrease sentences for youth had a profound impact on the likelihood of recidivism.[150] As the authors note:

> Age is one of the best predictors of criminal activity. Being under the age of 30 is associated with an additional 13 points on the risk score. For context, this is more points than a defendant would receive for 5 prior incarcerations. If the goal at sentencing is to incarcerate those at highest risk of reoffending, young people would be incarcerated at a very high rate.[151]

One can understand from this example how easy it is to get from difference to superiority. If we assume, as do the claims of machines reducing crime by 25% while denying bail at the same rate as judges, that recidivism and incarceration are the only factors to be analyzed, then machines do better.[152] However, when we consider other factors, such as concerns about youth, the difference between judges' ability to predict and machines' ability to predict is minimized.[153]

So, why might judges interject concerns for youth in veering from algorithmic risk assessments? While Professors Stevenson and Doleac do not

---

[146] *See id.* at 1312.

[147] Stevenson & Doleac, *supra* note 26.

[148] *Id.*

[149] *Id.* at 19 ("First, judges give young defendants sentences that are almost 30% shorter than those given to older defendants at the same risk level. (Corresponding with the age classifications in Virginia's risk assessment, 'young' is defined as being under the age of 30.)").

[150] *Id.*

[151] *Id.* at 20.

[152] *See id.* at 18–20.

[153] *See id.* at 20.

test for this in their experiment, they do provide some suggestions.[154] In this case, they point to ethics and complex decision-making:

> Retributivists, who believe that punishment is based on desert, often argue that young people are less culpable because they are more impulsive and lack self-control. Furthermore, they are in a phase where they are particularly susceptible to peer influence. Incarceration could expose them to negative influence, thus increasing recidivism in the long run.[155]

In short, judges are injecting human traits into their determinations.[156] Specifically, they may see young people as less responsible for their crimes because of their youth.[157] The injection of these traits changes the outcome due to the injection of moral reasoning.[158]

Let us anticipate the claim that it is simple enough to add age to the algorithm as an additional factor and leave it to the machine. This, of course, misses the fact that every decision about every criminal is an individual one. Stevenson and Doleac may have uncovered a measurable difference regarding age, but they do not attempt to consider how any individual decision reflects a variety of other factors.[159] And the number of factors that can influence decision-making are substantial. As Professors Cass Sunstein and Eric Posner recently observed regarding consideration of moral factors in cost-benefit analysis:

> If cost-benefit analysis must take on board all moral commitments for which people are willing to pay—empathetic, exclusionary, or sadistic—it will take on board a great deal. Indeed, there might seem to be no logical stopping point. The universe of moral commitments is very wide, in the sense that people would be willing to pay to maintain or to alter numerous and diverse states of affairs, and they would suffer a welfare loss if many states of affairs were maintained or altered. People's moral commitments bear on business activity on Sunday, kissing in public, boxing, sexually explicit

[154] *See id.*
[155] *Id.* (citations omitted).
[156] *See id.*
[157] *See id.* at 25.
[158] *See id.* at 25–26.
[159] *See generally id.*

> speech, political dissent, consumption of meat, and same-sex relationships.[160]

Put simply, to argue that machines can do what humans do in judging is to ignore the complexity of human cognition.

These results highlight how the narrow vision of similar rationality hides differences. When we hear that machines can "reduce crime by 25 percent—or that they could keep crime rates the same and reduce jailing by 42 percent"—we are looking at how machines work in a narrow universe that considers just recidivism and jail time.[161] Judges, on the other hand, are making decisions that consider a number of other factors. The result is a comparison of apples and oranges. While it may well be that judges who are charged with just predicting recidivism and jail time would not be as accurate as machines, judicial decisions are different from machine calculations.[162] To suggest machines are superior is to simply ignore the oranges.

We do not want to ignore the fact that Professors Stevenson and Doleac found judges to use their discretion negatively as well.[163] In addition to finding judges changing sentences to account for age, they also found that judges gave longer sentences to racial minorities.[164] They attempted to consider "rational" reasons for doing so but, ultimately, it is likely that this is the result of bias.[165] Of course the option to choose human cognition has significant costs, and we greatly appreciate the work being done to limit human bias in machine calculating.[166] Our goal is to simply demonstrate that the superiority of machine calculations, even in the context of probabilistic reasoning, may be overstated,[167] and also to explain how the narrative of similarity tends to blind us to the influence of positive traits on legal decision-making. The question we need to ask ourselves isn't: "which one is better at

---

[160] Eric A. Posner & Cass R. Sunstein, *Moral Commitments in Cost-Benefit Analysis*, 103 VA. L. REV. 1809, 1817 (2017).

[161] *See* Coglianese & Lai, *supra* note 20, at 1311–12.

[162] Stevenson & Doleac, *supra* note 26, at 19–20.

[163] *Id.*

[164] *Id.* ("[J]udges give Black defendants sentences that are 15-20% longer than White[s]."). But note that the algorithm itself had racial bias built in. *Id.* at 21 ("[R]ace is not explicitly included in the algorithm, many of its input factors correlate with race. As a result, Black defendants are 10 percentage points less likely to be categorized as low risk than White defendants with the same guidelines-recommended sentence.")

[165] *Id.* at 20.

[166] *See* Coglianese & Lai, *supra* note 20, at 1281.

[167] We do not mean to suggest machines are or are not better calculators. Rather, statements like "machines get it better than judges by 42 percent" may be overstating the case because by considering other factors (such as age), judges may be making decisions that actually increase discrepancies between their decisions and machine algorithms over other factors like jail time.

predicting a narrow set of outcomes?" It is: "what kinds of concerns does a society want its decision-makers to consider?" Put more synthetically, "do we throw out the positive in an effort to quash the negative, or do we find a way to minimize the negative while not losing the positive?"

2. Machine Bias

An additional concern is the false notion of purity in the vision of machine superiority. For reasons similar to those behind the failure to see machines as different we also fail to see the choice of machine calculation as a choice that carries bias rather than one of purity. This is not the standard story of machines working at speed and scale to magnify human bias; it is a story of machines as mathematical calculators having their own biases. The choice to marginalize human characteristics in judging, of course, is a choice to devalue the kinds of traits that math cannot do.[168] Choices to use machine calculating in other areas carry similar concerns.[169] We may think of math as objective, but, when applied to human beings, it carries the same concerns as any other system of decision-making.

Consider the use of facial recognition software. While machines may recognize faces at a significant level of success, they do so by converting facial traits to math.[170] Humans, on the other hand, do so by using a variety of other factors, including not just facial features but also gait, tone of voice, style of dress, hair color, and—importantly—our past encounters and experiences with those individuals.

To suggest that human decisions and machine decisions are interchangeable simply ignores this distinction; to ignore this distinction is to ignore much of what is important in the process of facial recognition and how humans and machines will recognize different people in different situations differently. Sometimes, these facts mean we can make identifications that evade the machines: when people are fully masked, or we can hear and not see them, or we note their touch or the way they feel. Or we recall the way they move in the corner of our eyes. Process, in this situation, matters, and will impact an analysis of the use of facial recognition technology in a variety of situations from criminal law to regulation.

Perceiving machine processing as unbiased ignores this difference. Machines process facial information at speed and scale but the mechanism used by machines has particular impacts as applied to particular groups.[171] In

---

[168] *See, e.g.*, Stevenson & Doleac, *supra* note 26, at 19–20.

[169] *See, e.g.,* Haber, *supra* note 143, at 95–97 (racial bias); Alexander R. Galloway, *The Gender of Math*, 32 DIFFERENCES 1 (2021) (gender bias).

[170] *See* discussion *supra* note 11 and accompanying text.

[171] *See* Haber, *supra* note 143, at 95–97.

particular, facial recognition's mechanism of identification has built-in limitations that make it more susceptible to misrecognizing certain types of faces, especially Black ones.[172] Perhaps unsurprisingly, similar concerns have been raised about math being gendered.[173] Recognizing difference means recognizing choice, and the choice to use a mechanism that measures faces or calculates gender rather than applying a more nuanced human cognition is not a choice between machine purity and human bias.

While we have discussed the general narrative above, there is a burgeoning body of scholarship that considers specifically how to deploy machine learning in law that the narrative of superiority also impacts. We will discuss those impacts below before turning to a number of separate normative concerns that the narrative implicates. Ultimately, we urge that legal scholars change their language in a way that better reflects the way in which machines work and de-emphasizes claims of superiority that undermine democracy and advance private rather than public interests.

B. Missing (or Dis-missing) Difference: Current Scholarly Discussion of Machine Calculation

There is a burgeoning scholarship that seeks to determine how to properly deploy machines in the legal realm. While some of this scholarship has recognized the notion of difference, the recognition is at best fragmented, and indeed, even those who recognize difference continue to see machines in terms of superiority.[174] Of course, superiority and similarity do play significant roles in the literature, with some assuming that machines can simply replace humans if they are properly programmed and others arguing that human traits are irrelevant to legal decisions.[175] We describe these arguments here and explain their limitations. Ultimately, we explain how difference can be used as a basis for considering when and how to deploy machines for legal purposes.

1. Missing Difference

Scholarship has begun to ask specifically where machine processing belongs in the realm of legal decision-making, and this, in turn, has led to an

[172] *See, e.g.*, *id.*; Abigail Nieves Delgado, *Race and Statistics in Facial Recognition: Producing Types, Physical Attributes, and Genealogies*, 0 SOC. STUD. SCI. 1 (2022); Damien Patrick Williams, *Fitting the Description: Historical and Sociotechnical Elements of Facial Recognition and Anti-Black Surveillance*, 7 J. RESPONSIBLE INNOVATION 74 (2020).

[173] *See, e.g.*, Galloway, *supra* note 169, at 1 (concluding that math contains an essential gender bias).

[174] *See* discussion *supra* Section II.B.

[175] *See* discussion *supra* Section II.A.

analysis of where machines are better or worse at making legal decisions.[176] The discussion to date has been far reaching but relatively fragmented. Commentators have identified a wide range of concerns regarding the use of machine processing in law. They have, for example, suggested that autonomous machines are ill-suited for "hard cases," "disputes in which the boundaries of the rules become unclear," "where the rules contradict each other," or where "enforcement of the rules implicates other principles."[177] So too do commentators consider concerns such as the ability to explain and analyze decisions and the need for procedural fairness as particularly human components of adjudication.[178] Still others are concerned about how machine processing will lock in particular norms and not be able to change with the times.[179] Virtually all of these suggestions reflect how machines and humans "think" differently. Yet, the fragmentation suggests that difference is not clearly seen as a central or organizing principle of this discussion.

A limited exception to the fragmented treatment of the issue is Professor Aziz Huq. In his article, *The Right to a Human Decision*, Professor Huq provides an exhaustive survey of a variety of normative bases for the creation of a "right to a human decision."[180] None of these bases, he concludes, provides a reason for establishing a right to a human decision-maker.[181] At the end of his review, he suggests that the difference between machines and humans might prove to be the main basis upon which a right to a human decision stands.[182] He does not, however, delve into the notion of difference at all, simply suggesting that "machine decisions are not presently appropriate for decisions with ethical or normative components."[183] Difference thus remains somewhat obscured as an organizing principle of this inquiry.

---

[176] *See* Chagal-Feferkorn, *supra* note 69; *see also* discussion *supra* Section III.C.

[177] Tim Wu, *Will Artificial Intelligence Eat the Law?*, 119 COLUM. L. REV. 2001, 2003 (2019); *see also* Frank Fagan & Saul Levmore, *The Impact of Artificial Intelligence on Rules, Standards, and Judicial Discretion*, 93 S. CAL. L. REV. 1, 3 (2019) (discussing the inability of machine learning to deal with complex decisions); Brian Sheppard, *The Reasonableness Machine*, 62 B.C. L. REV. 2259, 2276–77 (2021) (discussing the difficulty for machines to understand context).

[178] *See* Wu, *supra* note 177, at 2005; Huq *supra* note 20, at 661.

[179] Hon. Jennifer Walker Elrod, *Trial by Siri: AI Comes to the Courtroom*, 57 HOUS. L. REV. 1083, 1098 (2021) (noting that some have argued that results-based criminal AI programs are flawed because "they can only consider, according to past data, *what is*, rather than *what should be*. The removal of a judge's humanity locks in cold decision-making and prevents moral growth.").

[180] Huq, *supra* note 20, at 611.

[181] *Id.* at 685.

[182] *Id.* at 685.

[183] *Id.*

Perhaps more important to recognizing difference as a central factor is the need to see clearly how it gets obscured in the general debate, especially when analyzing the abilities of machine calculators to make legal determinations. We have explained how notions of similarity of "cognition" and language like "artificial intelligence" and "machine learning" tend to blind us to difference.[184] It is thus unsurprising that even those who recognize difference still get captured by the narrative of similarity and superiority.[185]

Professor Huq, for example, falls victim to the narrative framework in his own article when he fails to consider how difference would actually influence his normative analysis.[186] Consider, for example, his analysis of normative concerns around the "accuracy" of human cognition versus machine calculation.[187] Huq argues that the issue is ultimately one of equity rather than accuracy but notes that it cannot be solved by creating a right to a human decision, which would "revert to a more error-prone human decision-making protocol[.]"[188] This analysis obscures the difference between human and machine. Let us again return to the discussion of facial recognition. The different processes used by humans and machines are likely to both result in mistakes but also are likely to be accurate in different ways.[189] The difference in how machines and humans work is of significance and cannot just be assumed away. Put simply, the discussion of accuracy ignores difference.[190] To ask which processor is more accurate is to assume that legal decision-making is only about making a probabilistic determination of a statutorily-defined outcome and to ignore other factors, like morality, that human judges use in their decision-making and machines don't.[191] Even if this weren't the case, such an analysis privileges certain outcomes—in this case, the outcomes of machine calculations of facial recognition rather than human ones.

Indeed, despite his thoughtful analysis, even Professor Huq falls back on a view of machines as pure and better processors.[192] As he notes in his article:

> An account of the right to a machine decision would begin with the observation that while machine-learning tools have

[184] *See* discussion *supra* Part III.
[185] *See* discussion *supra* notes 20, 38 and accompanying text.
[186] *See* Huq, *supra* note 20, at 653–56.
[187] *See generally id.*
[188] *Id.* at 655.
[189] *See supra* Part II (noting that sometimes people can make identifications that evade the machines: when people are fully masked, we can hear and not see them, we note their touch or the way they feel, or we recall the way they move in the corner of our eyes).
[190] *See* discussion *supra* Section II.B.
[191] *See* discussion *supra* Part III.
[192] *See* Huq, *supra* note 20, at 687.

> the capacity to improve on humans' accuracy and neutrality, many of those now implemented by government are highly flawed. Even if this does not impel a reversion to (equally flawed) human decision making, the legal system should incentivize the correction of such errors. Its dynamic goal should be a machine decision well-calibrated in light of constitutional concerns. Most basically, an algorithmic tool is well-calibrated if it does not rely on flawed training data and otherwise meets common standards of industry performance. More work, however, needs to be done to describe the circumstances in which algorithms are in compliance with due process, privacy, and equality norms.[193]

By this analysis, machines are superior in terms of accuracy and neutrality despite not being perfect processors.[194] The goal of law should not be to revert to flawed human rationality but to incentivize the correction of machine flaws in ways that allow them to pass legal muster. Lost in this conclusion is a sense of how positive human traits should be maximized or balanced with positive machine traits. Similarly, there is a claim to machine neutrality that ignores the bias that comes with mathematical reasoning.

The lack of clarity about difference can also be found in the work of others who are clearly cognizant of it. Again, with due respect to their insightful empirical analysis of how judges and machines differ, even Professors Stevenson and Doleac find themselves echoing claims of superiority.[195] Their data suggests judges may be abandoning their algorithmic assistance altogether.[196] That is, they began by using the algorithms but began to ignore them over time.[197] The authors suggest this may be because the algorithms did not provide useful information, or it may be because the judges demonstrated an aversion to "using superior but imperfect algorithms."[198] To say algorithms are superior but imperfect is to suggest that factors such as morals and situational sense are irrelevant. We don't doubt that people may be adverse to recognizing they are not the best predictors, but to suggest that the ability to predict particular results makes a machine algorithm "superior" is to miss the point. Being the best "predictor" and being the best decision-maker are two different things. As we discussed above, scholars need to "start" from difference.[199] If they don't, they may

---

[193] *Id.*
[194] *Id.*
[195] *See* Stevenson & Doleac, *supra* note 26.
[196] *See id.* at 16.
[197] *Id.*
[198] *Id.* (internal citations omitted).
[199] *See supra* Part III.

tend to fall under the influence of a powerful frame that describes machine calculation as pure and a superior means of determining outcomes.

2. (Dis)-missing Difference: How Assumptions About Human Reasoning Lead to False Claims of Superiority

As we discussed above,[200] there are scholars who have taken the implicit assumption about similarity and superiority and made it explicit. Professor Eugene Volokh's recent article, *Chief Justice Robots*, specifically addresses and dismisses claims that human traits are relevant to law.[201] We discussed this argument *supra* when we described how scholars had explicitly argued that machine decisions are superior to human ones.[202] To refresh the readers' memory, in his article, Volokh assumes the existence of a machine processing algorithm that can write persuasive judicial opinions.[203] In such a case, he argues, whether a machine has human traits or not should not matter because the result of a decision is more important than the factors that go into it.[204] As he notes:

> But here again what matters is the result, not the process. If a poetry-translation program reliably produces translations that are emotionally rewarding for us as readers, it should not matter to us that Robot Frost can't itself have emotions. If, in a blind test, we view an AI sentencing judge as producing wiser and more compassionate results— by our lights—than a human sentencing judge, it should not matter to us as evaluators that the judge can't have "wisdom" or "compassion."[205]

A similar argument is made by Coglianese and Lehr when they suggest that machine processing may be able to replace humans as long as we can mathematically capture the specific outcomes that humans desire (such as wealth maximization) and have the machines mimic human decisions.[206] Coglianese and Lai too have specifically argued that machines are better processors than humans.[207] We have dealt with this latter argument by noting that it puts the cart before the horse. That is, the issue is really about how we

[200] *See supra* Section II.B.
[201] Volokh, *supra* note 39, at 1167.
[202] *See supra* Section II.B.
[203] *See* Volokh, *supra* note 39, at 1152 (noting that the decision must be persuasive).
[204] *Id.* at 1167.
[205] *Id.* at 1189.
[206] *See* Coglianese & Lehr, *supra* note 38, at 1202.
[207] *See* Coglianese & Lai, *supra* note 20, at 1281.

perceive human cognition. If we assume humans as nothing more than endpoint maximizers, then a vision of machines replacing humans, and deciding better than humans, follows easily.[208]

Professor Volokh's argument goes one step further. He doesn't simply argue from similarity; he argues against the need for human traits altogether.[209] His argument, however, doesn't negate the problem between process and product.[210] Rather, it simply hides the problem among his assumptions. Volokh argues that to "get it right," a machine must make a determination that is persuasive to a human being.[211] Assuming a machine makes such a determination, he notes, the way it reaches that determination shouldn't matter. But for a machine to make a determination that satisfies human concerns regarding morality, dignity, and compassion requires the machine logic to code for morality, dignity, and compassion, which is something machines cannot do. Of course, even without coding for such traits, a machine may make a decision that is persuasive once in a while, just as millions of monkeys typing may one day produce Shakespeare.[212] However, most of the time monkeys typing will produce documents that make no sense. To have consistently moral and ethical judgments, one would need a process into which morals and ethics are built.

In sum, the current discussion of machine deployment in law reflects to a high degree the vision of machines as similar and superior to humans when it comes to legal decision-making. As a discursive lens, the impact of this view is various. In this section, we have demonstrated how this view marginalizes human traits that are not subject to quantifiable analysis and treats machines as pure, when, in truth, the choice of machines carries with it its own bias. The result of hiding difference is a fragmented and unclear literature that tends to see machines as better legal "decision-makers" than they are and that ignores completely potential machine bias while suggesting the broad-scale deployment of machines given their superior nature.

V. IMPLICATIONS: HUMAN TRAITS AND MACHINE BIAS

In this section, we consider the implications of the understanding of difference for a variety of legal concerns. We first turn to the issue of

[208] *See supra* Section III.B.

[209] Volokh, *supra* note 39, at 1167–69.

[210] *See id.*

[211] *Id.* at 1152–53.

[212] Émile Borel, *Mécanique Statistique et Irréversibilité*, 3 J. PHYSICS: THEORIES & APPLICATIONS 189, 189–96 (1913) (Fr.) (introducing the amusing thought experiment known in popular culture as the "infinite monkey theorem," which hypothesizes that a monkey hitting keys at random on a typewriter keyboard for an infinite amount of time will almost surely type a given text, such as the complete works of William Shakespeare).

difference generally, suggesting that it provides a stable and understandable framework for considering how and when to deploy machines in law. We then turn to concerns that arise from the notion of superiority. We argue that superiority adds additional impetus to a process of machine deployment that can negatively impact human agency and decrease human flourishing. We then turn to one major source of the notion of superiority: language. We argue that it is incumbent on legal scholars to change their use of words such as "artificial intelligence," "neural networks," and "machine learning" to better describe what machines actually do and because these words serve private rather than public interests.

A. Dealing with Difference

Machine computation is different from human cognition and cannot simply be assumed away. A discursive model built on language like "artificial intelligence" and "machine learning" and a parsimonious vision of cognition as well as law-making hides difference and marginalizes the influence of emotion, morality, fairness, compassion, and other human traits on legal decision-making. [213] Even those who recognize that difference should influence our discussion of how and when machine computation should be deployed in law can fall victim to this frame.[214] Once the marginalization of particularly human traits becomes clear, we can begin to look to law itself for hints about how to properly deploy machine calculators in the legal realm.

The law has, of course, long made space for the consideration of human traits like compassion and ethics, as well as for other particularly human forms of cognition that machine calculation cannot do. We suggest that the beginning of any discussion on deployment of machines in law should recognize these traditional spaces, located in places like the use of discretion, the use of equity, and the application of standards rather than rules. All of these are places where machine calculation should not take the lead. In terms of a "right to a human decision," all of these are places where machines alone cannot be deployed as deciders.

Of course, machines have long been tools in decision-making. However, as we discuss in the next section, concerns over giving too much deference to perceived machine superiority must be considered. A number of other professions, such as medicine and aviation, use technology to aid in decision-making.[215] Yet, despite knowing of the limits of the technology they are using, doctors and pilots continue to defer to machines in ways that lead to bad results. While law is certainly different from medicine and aviation,

---

[213] *See* discussion *supra* Part IV.
[214] *See* discussion *supra* Section IV.B.1.
[215] *See infra* Section V.B.1.

we suggest that the potential for machine bias needs to be much better understood before deploying machine calculators, even as support for human decision-makers.

To this point, we have discussed the narrative frame's implications for human traits, but similar concerns also apply to machine bias. Specifically, we note that the narrative treats machines as pure when they are not. We need to consider how math itself may be biased rather than thinking of the task solely as opening the black box of machine processing to uncover human bias. There is a large literature on the potential for machines to exacerbate human bias and the need to open up the "black box" of machine determinations to protect against harms to particular groups.[216] Our analysis, of course, suggests that we must also consider the potential for code itself to carry with it bias so as to make sure that particular groups are protected from the harms that may come from the use of math as a basis for making, or aiding, legal determinations.

B. Other Concerns Raised by the Narrative

1. The Impact of Superiority on Human Agency

There are particular concerns that come from an understanding of machine processing as superior. Specifically, concerns arise that legal decision-makers who rely on machine processing may be overly deferent to machine calculations. The same can be said of the populace in general.

Other fields that use technology, such as medicine and aviation, have documented the existence of an "automation bias."[217] The bias actually refers to "2 separate but closely related bodies of research" conceived of as automation bias or "automation-induced complacency."[218] Automation bias is described as "the tendency to use automated cues as a heuristic replacement for vigilant information seeking and processing,"[219] while automation-induced complacency has been described as "self-satisfaction which may

---

[216] *See, e.g.*, Arun Rai, *Explainable AI: From Black Box to Glass Box*, 48 J. ACAD. MKTG. SCI. 137 (2020); Robin Feldman & Kara Stein, *AI Governance in the Financial Industry*, 27 STAN. J.L. BUS. & FIN. 94, 104 (2022); Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J.L. & TECH. 889 (2018); *see also* Cynthia Rudin & Joanna Radin, *Why Are We Using Black Box Models in AI When We Don't Need to? A Lesson from an Explainable AI Competition*, 1.2 HARV. DATA SCI. REV. (2019).

[217] David Lyell & Enrico Coiera, *Automation Bias and Verification Complexity: A Systematic Review*, 24 J. AM. MED. INFORMATICS ASS'N 423, 423–24 (2017).

[218] *Id.* at 423.

[219] *Id*. at 423–24 (quoting Kathleen L. Mosier & Linda J. Skitka, *Human Decision Makers and Automated Decision Aids: Made for Each Other?*, *in* AUTOMATION AND HUMAN PERFORMANCE 201 (Raja Parasuraman & Mustapha Mouloua eds., 1996)).

result in non-vigilance based on an unjustified assumption of satisfactory system state."[220]

Despite heavy use of, and reliance upon, automation in medicine and aviation, much about the bias is still unknown,[221] but the impact of automation bias on decision-making is quite high.[222] One meta-study of decision-making in health care found a 26% increase in errors when technological support was used.[223] Another study of computer aided EKG diagnoses found that doctors failed to change wrong machine calculations of atrial fibrillation 24% of the time.[224] One might assume that knowledge of the bias's existence is enough to limit its effects,[225] yet evidence that shows its continued impact despite the fact that doctors and pilots are aware of it suggests caution.[226]

Medical decision-making, of course, is not legal decision-making, and empirical evidence does suggest that judges are skeptical of machine algorithms, at least in the case of sentencing decisions.[227] Moreover, legal decision-makers are often democratically accountable, creating incentives to avoid simply adopting a machine calculation when democratic factors may be implicated. Clearing away the potential differences, however, the undue deference given to machines in decision-making in medicine and aviation certainly suggests we should not simply dismiss the potential for automation bias to affect legal decision-making, especially given the normative, ethical, and other non-quantitative factors that law often deals with and a bias that leads us to construct machine calculations as both similar to and superior to human ones.

The greater concern regarding beliefs about superiority is their potential to impact the polity. Before beginning this analysis, however, it is necessary to explain that the polity conceives of technology in the same way as legal scholars. There is ample evidence that this is the case. And of course, why shouldn't it be? In particular, to the extent the vision of technology is based in part on language like "artificial intelligence" and "machine learning," the influence of these words falls not just on legal scholars but also on all

---

[220] *Id.* at 424. *See also* CHARLES E. BILLINGS ET AL., NAT'L AERONAUTICS & SPACE ADMIN., NASA AVIATION SAFETY REPORTING SYSTEM Q. REP. NO. 76-1, 8 (1976); Lyell & Coiera, *supra* note 217, at 424.

[221] Lyell & Coiera, *supra* note 217, at 425.

[222] *Id.* at 425–26, 430.

[223] Kate Goddard et al., *Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators*, 19 J. AM. MED. INFORMATICS ASS'N 121, 123 (2012).

[224] Frank Bogun et al., *Misdiagnosis of Atrial Fibrillation and its Clinical Consequences*, 117 AM. J. MED. 636, 637 (2004).

[225] *See* Huq, *supra* note 20, at 646–48.

[226] *See* Lyell & Coiera, *supra* note 217.

[227] *See* Stevenson & Doleac, *supra* note 26, at 19–20.

members of society.[228] Even visions of machines as thinking similarly to humans can be found in popular culture.[229] How could an article on perceptions of similarity not, for example, refer to the *Terminator* movies and their effects on popular culture? The Frameworks Institute, in conjunction with the MacArthur Foundation, has recently conducted a study that finds just this.[230] As the study summarizes: "the discussion about AI is shaped—and derailed—by cultural mindsets that push people to either see AI in a virtuous light, leaving biases unquestioned, or to assume that technology is competing with people, working its way toward a takeover of humankind."[231] The polity, it seems, also marches to a narrative of similarity, superiority and purity.

A number of scholars, primarily outside of law, have focused on the potential for machine calculations to impact human agency.[232] For most scholars, this concern is rooted in an understanding of the power of AI to shape or distort human preference. As one professor of computer science explains: "Many current AI systems . . . try to avoid information and choice overload by replacing our decision-making processes with algorithmic predictions."[233] Of course, because the majority of machine processing is used by private entities as a means of maximizing profit, these scholars recognize that AI will shape individual decisions to serve commercial interests.[234] Put simply, the more we rely on machines to decide for us, the less effort we will make to think about our own choices and whether those choices will maximize our own happiness.

The same can be said about the political sphere. As one early volume on the topic suggests:

> But once artificial entities become more autonomic, and less dependent on deliberate human intervention, criteria like agency, intentionality and self-determination, become too fragile to serve as defining criteria for human subjectivity, personality or identity, and for characterizing the processes

---

[228] *See* Tucker, Ctr. Priv. & Tech., *supra* note 17.

[229] *E.g.*, WALL-E (Walt Disney Pictures 2008); THE TERMINATOR (Orion Pictures 1984).

[230] CONKLIN ET AL., *supra* note 9.

[231] *Id.*

[232] Lee Rainie et al., *Experts Doubt Ethical AI Design Will Be Broadly Adopted as the Norm Within the Next Decade*, PEW RSCH. CTR. (June 16, 2021), https://www.pewresearch.org/internet/2021/06/16/experts-doubt-ethical-ai-design-will-be-broadly-adopted-as-the-norm-within-the-next-decade/.

[233] Janna Anderson & Lee Rainie, *Concerns About Human Agency, Evolution and Survival,* PEW RSCH. CTR. (Dec. 10, 2018), https://www.pewresearch.org/internet/2018/12/10/concerns-about-human-agency-evolution-and-survival/.

[234] *See* Tucker, Ctr. Priv. & Tech., *supra* note 17.

> through which individual citizens become moral and legal subjects. Are autonomic, yet artificial, systems shrinking the distance between (acting) subjects and (acted upon) objects? How distinctively human will agency be in a world of autonomic computing?[235]

Implicit in these critiques is a view of the mechanism by which machines will impact autonomy. First, as machines start making more and more decisions for us, humans will simply stop deciding things for themselves, and second, as machines manipulate preference, humans will become the objects of decisions, rather than the decision-maker.

The potential for machines to impact our autonomy is large even when we don't think of them as superior, but a belief that machines are superior is likely to increase our willingness to abandon our agency even more. And people do think of machines as better.[236] The Frameworks study, for example, found that people literally see AI as divination:[237]

> The public also think of AI prediction as the ability to see into the future and prevent bad things from happening based on information gleaned from the present. This concept of seeing as knowing creates a false sense of certainty about the future. When thinking in this way, the public sees AI less like a scientist or machine and more like a fortune teller or psychic, who they believe will be able to sense a change and predict what will happen in the future so that it can be prevented in the present. When applied to the domains of policing, child welfare, and health care, participants reasoned that the use of AI would instantaneously identify and address problems before they happened or got worse. . . .[238]

This level of belief in technology provides a strong incentive for humans to devolve decision-making to machines. That is, it suggests that giving up agency is the right thing to do. Of course, giving up agency also provides other benefits. Instead of having to face the consequences that come from

[235] *Description* to LAW, HUMAN AGENCY AND AUTONOMIC COMPUTING: THE PHILOSOPHY OF LAW MEETS THE PHILOSOPHY OF TECHNOLOGY (Mireille Hildebrandt & Antoinette Rouvroy eds., Routledge 2011).

[236] *See* discussion *supra* Part III.

[237] CONKLIN ET AL., *supra* note 9, at 11.

[238] *Id.*

making difficult political choices, we can defer to the machine,[239] thus insulating ourselves from the emotional consequences of difficult choices.[240] In essence, we give up wrestling with the consequences of our decisions by deferring to a superior decision-making process. In a world where decisions are politicized, hiding them behind a veil of objectivity or superiority only serves to make us less involved citizens.[241]

2. Changing the Language of Power

Throughout this article we have suggested that language matters.[242] The language we use in legal scholarship and the language used by courts and other legal decision-makers all play a role in how we think about the deployment of machine calculation in law.[243] Indeed, the language of legal scholarship and the language of legal decision are likely to be iterative, with one influencing the other and vice-versa.[244] We now turn to language both to understand who the narrative of similarity serves and to discuss how the law should use it. Words like "artificial intelligence" and "machine learning" carry with them a sense of machines "thinking" like humans. Yet, how did these words come to be used to describe a technology that is nothing like human thinking?

The Center on Privacy and Technology at Georgetown Law School provides insights. It describes how private capital co-opted the language of artificial intelligence and machine processing for purposes of expanding their power.[245] Citing to Meredith Whittaker, it notes that the turning point was when companies debuted an algorithm called "AlexNet":

---

[239] *See* Douglas A. Kysar, *Regulating from Nowhere: Environmental Law and the Search for Objectivity*, 98 ARCHIVES PHIL. L. & SOC. PHIL. 152 (2012) (suggesting the same for the use of cost-benefit analysis).

[240] *Id.*

[241] A similar phenomenon may also be in play regarding the rooting out of technology bias. As Michael Sandel has noted, "AI not only replicates human biases, it confers on these biases a kind of scientific creditability. It makes it seem that these predictions and judgments have an objective status." Christina Pazzanese, *Ethical Concerns Mount as AI Takes Bigger Decision-Making Role in More Industries*, HARV. GAZETTE (Oct. 26, 2020), https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/ (quoting Michael Sandel). Performances of Greek tragedies often did their killing off stage partly because the human imagination could conceive of the killing in much more gruesome terms than what could be depicted on stage. To the extent machine determinations are a "black box," leaving us to fill in our own understanding based on our own predispositions and understandings, one can understand how a common understanding of machines as "pure" or "objective" impacts our concerns about bias.

[242] *See* discussions *supra* Sections II.A, IV.A.

[243] *See* Tucker, Ctr. Priv. & Tech., *supra* note 17.

[244] *Id.*

[245] Tucker, Ctr. Priv. & Tech, *supra* note 17.

> AlexNet mapped a path forward for large tech companies seeking to cement and expand their power. . . . Tech companies quickly (re)branded machine learning and other data-dependent approaches as AI, framing them as the product of breakthrough scientific innovation. Companies acquired labs and start-ups, and worked to pitch AI as a multitool of efficiency and precision, suitable for nearly any purpose across countless domains.[246]

The notion of similarity and superiority serves particular interests, and these interests are not in human flourishing. Rather, it is the opposite. As the Center notes: "[i]nstead of pursuing the limits of computers' potential for simulated humanity, the hawkers of 'AI' are pursuing the limits of human beings' potential to be reduced to their calculability."[247]

In response, the Center has chosen to stop using words like "artificial intelligence" and "machine learning."[248] The recognition that such language contributes to a misperception of machine calculation as similar to human thinking is now starting to be echoed by computer scientists,[249] psychologists,[250] and the media.[251]

In keeping with the Center's suggestion, we have chosen to use different language in this article, describing machines as "processors" or "calculators" rather than as "learners" or "intelligent." The language of calculation is more accurate and reflects the realities of what machine algorithms actually can do in the process of legal decision-making.[252] Similarly, we suggest additional considerations in how scholars discuss

---

[246] *Id.*
[247] *Id.*
[248] *Id.*
[249] *See, e.g.*, Kathy Pretz, *Stop Calling Everything AI, Machine-Learning Pioneer Says*, IEEE SPECTRUM (Mar. 31, 2021), https://spectrum.ieee.org/stop-calling-everything-ai-machinelearning-pioneer-says ("Artificial-intelligence systems are nowhere near advanced enough to replace humans in many tasks involving reasoning, real-world knowledge, and social interaction. They are showing human-level competence in low-level pattern recognition skills, but at the cognitive level they are merely imitating human intelligence, not engaging deeply and creatively. . . ."); Jack Bandy, *Killing A.I.*, MEDIUM (Mar. 5, 2022) https://jackbandy.medium.com/killing-a-i-3f08ad5de6d9 (considering getting rid of the term "AI").
[250] Sasha Luccioni & Gary Marcus, *Stop Treating AI Models Like People*, SUBSTACK (Apr. 17, 2023), https://garymarcus.substack.com/p/stop-treating-ai-models-like-people.
[251] Jaron Lanier, *There is No A.I.*, NEW YORKER (Apr. 20, 2023), https://www.newyorker.com/science/annals-of-artificial-intelligence/there-is-no-ai ("I don't like the term 'AI.' In fact I think it's misleading—maybe even a little dangerous.")
[252] Tucker, Ctr. Priv. & Tech., *supra* note 17.

issues of machine calculation in law. When scholars choose to rely on a parsimonious vision of rationality or legal decision-making, they should be explicit about their choice. Similarly, the structure of scholarship often leaves machine limitations to be considered as an afterthought, rather than integrating the limitations into its analysis.[253] Finally, machines are not pure. While we can talk of rooting out human bias from machine calculations, we should also recognize that we are not then making machine decisions pure. Instead, we are choosing to root out human bias in exchange for the bias inherent in machine calculation.

As scholars, it is incumbent upon us to take whatever steps we can to stop advancing a narrative that supports a false view of machine processing and that generally serves private rather than public interests. Changing our language is a starting point. As we note supra, language is powerful. Indeed, it can shape our understanding of the use of machines in law more than our actual experience of machine use in law.[254] Thus, the words we use in our scholarship, our classrooms, our op-eds, and elsewhere matters.

Private interests have a clear advantage in the dissemination of information[255] and a clear interest in the use of language like "artificial intelligence," "machine learning," and "neural networks" that will continue to tantalize the population into thinking of machines as almost mystical superintelligence. However, an obligation to describe processes truthfully, as well as a concern for furthering the public interest, both mediate strongly in favor of changing the language we use. In law, that mysticism is bound into a narrative that fools us into thinking of machines as pure and better. By removing language that obscures the landscape, we can help ensure that the use of machines in law is done in a way that promotes the public interest rather than private interests.

VI. CONCLUSION

There are many different narratives of technology. They include images of technology both as an existential threat to society and as society's

---

[253] *See, e.g.*, Coglianese and Lai, *supra* note 20, at 1325 (mentioning machine limitations after a detailed analysis of their superiority); Huq, *supra* note 20, at 651 (analyzing the normative bases of a right to a human decision separately from considerations of difference). *See also supra* Section IV.B (explaining how difference should be an organizing principle of the debate over deploying machines in law).

[254] *See supra* Section IV.A.

[255] *See* Sylvia Lu, *Algorithmic Opacity, Private Accountability, and Corporate Social Disclosure in the Age of Artificial Intelligence*, 23 VAND D. ENT. & TECH. L. 99, 113, 119 (2020) (explaining how private companies benefit from collecting information for AI algorithms to increase profits by tailoring products, services, and advertisements to consumers).

savior. We do not desire for this article to play into this dichotomy. It is neither "pro" technology nor "anti" technology. Our purpose is to ensure that we see more clearly the consequences and problems of technology that arise from a particularly influential narrative currently embedded in scholarship and legal decision-making. To understand machines as efficient, superior, and pure, we need to first see them as similar.[256] With the blinders of superiority off, the problem set and solutions change; not completely, but they change.[257] The bias inherent in choosing machine processing becomes more apparent and so too do concerns over the loss of agency, autonomy, and human flourishing.[258] These goals have been marginalized by the superiority narrative, hidden by language and parsimonious visions of cognition.[259] While changing language and tempering our almost reflexive acceptance of machines as superior is part of the solution, we also urge a deeper focus on these other concerns, which have been hidden by language that serves private rather than public interests.

[256] *See* discussion *supra* Section II.B.
[257] *See* discussion *supra* Section IV.A.
[258] *See* discussion *supra* Section III.B.1.
[259] *See* discussion *supra* Section IV.A.