JOURNAL
OF
LAW & INNOVATION

VOL. 5 JANUARY 2023 NO. 1

ESSAY

Algorithmic Grey Holes

ALICIA G. SOLOW-NIEDERMAN†

It's almost a cliché to talk about algorithms as "black boxes" that resist human understanding. This frame emphasizes opacity, suggesting that the inability to see inside the algorithm is the problem. If a lack of transparency is the problem, then procedural measures to enhance access to the algorithm—whether by requiring audits, by adjusting the technological parameters of a tool to make it more "explainable," or by pushing back against proprietary claims made by private vendors—are the natural solution.

But the relentless pursuit of transparency blinks the reality that algorithmic accountability is more complicated than opening a box. Neither critics nor backers of algorithmic tools have reckoned with a related, yet

† Associate Professor, University of Iowa College of Law; Affiliated Fellow, Yale Law School Information Society Project; Faculty Associate, Berkman Klein Center for Internet & Society at Harvard University; Non-Resident Affiliate, Northeastern University School of Law Center for Law, Innovation and Creativity. Thank you to Hannah Bloch-Wehba, Thomas Kadri, Richard Re, Rory Van Loo, and Ari Ezra Waldman for helpful feedback and suggestions. I am grateful to the Journal of Law & Innovation student editors at the University of Pennsylvania Carey Law School for their assistance in preparing this piece for publication and to the 2022 Journal of Law & Innovation Symposium participants, particularly Orin Kerr, for their incisive questions and thoughtful engagement. Any remaining errors or omissions are my own. This essay is dedicated to my parents, who gifted me the middle name Grae because life isn't black or white.

distinct challenge that emerges when the state employs algorithmic methods: algorithmic grey holes. Algorithmic grey holes occur when layers of procedure offer a bare appearance of legality, without accounting for whether legal remedies are in fact available to affected populations. Although opacity about how an algorithm works may contribute to a grey hole, reckoning with a grey hole demands more than transparency. A myopic emphasis on transparency understates not only the consequences for an individual, but also how a lack of effective individual review and redress can have systemic consequences for rule of law itself. This class of potential costs has not been adequately recognized.

This Essay puts the challenge of algorithmic grey holes and the threat to rule of law values, particularly for criminal justice applications, front and center. It evaluates the individual and societal stakes not only for criminal justice, but also for front-line enforcement decisions and adjunction of benefits and burdens in civil settings. By forthrightly confronting these concerns, it becomes possible both to diagnose individual and societal algorithmic harms more effectively and to contemplate how technological tools might innovate in more helpful ways.

INTRODUCTION

It's almost a cliché to talk about algorithms as "black boxes" that resist human understanding.[1] This frame emphasizes opacity, suggesting that the inability to see inside the algorithm is the problem. Opacity-driven concerns

[1] *See, e.g.*, Jeff Ward, Foreword, *Black Box Artificial Intelligence and the Rule of Law*, 84 L. & CONTEMP. PROBS. i, ii (2021) ("[A]ll kinds of AI tools—even those using simple symbolic or handcrafted algorithms that should be intuitive and accessible to human inquiry—might lack transparency and thus be characterized as black boxes."). This foreword introduced a recent volume of the journal *Law & Contemporary Problems* that focused on "Black Box Artificial Intelligence and the Rule of Law."

have many sources, ranging from alarm about the lack of accountability when inscrutable or incomprehensible algorithms are used to make consequential decisions to worries about the fairness or validity of trade secrecy claims mounted by private developers of algorithmic tools.[2] If a lack of transparency is the problem, then access to the algorithm—whether by requiring audits, by adjusting the technological parameters of the tool to make it more "explainable," or by pushing back against proprietary claims made by private vendors—is the natural solution.

But the relentless pursuit of transparency blinks the reality that algorithmic accountability is more complicated than opening a box. Neither critics nor backers of algorithmic tools have reckoned with a related, yet distinct challenge: *algorithmic grey holes*.[3] Algorithmic grey holes occur when layers of procedure offer a bare appearance of legality, without accounting for whether legal remedies are in fact available to affected populations. Although opacity about how an algorithm works may contribute to a grey hole, reckoning with a grey hole demands more than transparency.

A black box framing conflates two questions. The first is whether anyone can access an algorithm, in theory. The second is whether an individual affected by an algorithm has the ability to obtain redress for how that algorithm is applied to them, in practice. A black box frame foregrounds the first issue at the expense of the second. A myopic emphasis on transparency understates not only the consequences for an individual, but also how a lack of effective individual review and redress can have systemic consequences for rule of law itself.[4] And this class of potential rule of law costs has not

---

[2] There are many, many excellent examples of each of these genres. On accountability, see, for example, Margot E. Kaminski, *Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability*, 92 S. CAL. L. REV. 1529, 1533 (2019); Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson & Harlan Yu, *Accountable Algorithms*, 165 U. PA. L. REV. 633, 636 (2017); Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. U. L. REV. 1, 6–8 (2014). On secrecy, see, for example, Natalie Ram, *Innovating Criminal Justice*, 112 NW. L. REV. 659, 663 (2018); Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1353 (2018).

[3] This concept builds from David Dyzenhaus's "grey holes" in the law. See discussion *infra* Part I.

[4] Mireille Hildebrandt has focused on how "smart technologies" may blur online and offline worlds—what she calls "onlife"—and affect rule of law. *See generally* MIREILLE HILDEBRANDT, SMART TECHNOLOGIES AND THE END(S) OF LAW (2015). Emily Berman has argued that "certain characteristics of predictive analytics inevitably bring them into tension with rule-of-law principles." Emily Berman, *A Government of Laws, and Not of Machines*, 98 B.U. L. REV. 1277, 1282–83 (2018). Ari Ezra Waldman has also contended that automated

been adequately recognized.[5]

This Essay puts the challenge of algorithmic grey holes and the threat to rule of law values, particularly for criminal justice applications, front and center. It proceeds in three parts. Part I reviews the non-algorithmic literature on legal holes, with an emphasis on David Dyzenhaus's "grey holes" and how they can undermine rule of law. Part II then analyzes the risks of algorithmic grey holes in criminal justice and evaluates the individual and societal stakes. Part III grapples with potential implications for front-line enforcement decisions and adjunction of benefits and burdens beyond the criminal justice system. By forthrightly confronting these concerns, it becomes possible both to diagnose individual and societal algorithmic harms more effectively and to contemplate how technological tools might innovate in more helpful ways.

I. BEYOND BLACK OR WHITE: CONFRONTING GREY HOLES

This Part summarizes how general theories of rule of law have addressed legal holes. As we will see in Part I, neither the potential for holes in the law nor the effects of any such holes have been squarely recognized when it comes to the state's use of algorithmic tools such as risk assessment instruments.

What, exactly, makes a legal system a good system of law, and how

---

decision-making systems are "problematic for the rule of law and legal legitimacy," underscoring the inadequacy of process-driven solutions. Ari Ezra Waldman, *Power, Process, and Automated Decision-Making*, 88 FORDHAM L. REV. 613, 617 (2019). Most recently, Margot Kaminski and Jennifer Urban have identified a relationship between an individual right to contest algorithmic decisions and "rule of law values." Margot E. Kaminski & Jennifer M. Urban, *The Right to Contest AI*, 121 COLUM. L. REV. 1957, 1991 (2021). This Essay offers a distinct "legal hole" framing to focus on the interactions among legality, procedural protections, and rule of law, and considers the ways that algorithmic tools interact with existing legal systems in the American context.

[5] For rare exceptions, see sources cited *supra* note 4. Even those who have thoughtfully critiqued the limits of transparency or who have stressed that transparency is an instrumental value have generally not engaged with rule of law concerns. For an incisive analysis of the limits of transparency, see, for example, Mike Ananny & Kate Crawford, *Seeing Without Knowing: Limitations of the Transparency Ideal and its Application to Algorithmic Accountability*, 20 NEW MEDIA & SOC'Y 973, 973 (2016). A rule of law lens underscores how there is an unrecognized cost to foregrounding transparency, even accepting that transparency can be instrumentally or intrinsically valuable. For a defense of transparency as "necessary, if not sufficient, for building and governing accountable algorithms," see Margot E. Kaminski, *Understanding Transparency in Algorithmic Accountability*, *in* THE CAMBRIDGE HANDBOOK OF THE LAW OF ALGORITHMS 121 (Woodrow Barfield ed., 2020).

should that system constrain government actors? At a minimum, a thin understanding of the role of law might call for rule *by* law, in the sense of compliance with duly enacted positive law. A thicker understanding might call for rule *of* law, in the sense of requiring the system of law-making to comport with a broader set of legal principles.[6] This understanding of rule of law is commonly called "formal," drawing from work such as Lon Fuller's classic account of eight factors that establish the "inner morality" of the law.[7]

From this thicker rule of law perspective, the constraints placed on government actors within a legal system are especially salient.[8] The legality of the government's actions may depend on public officials' ability to "show a legal warrant—an authorization in preexisting law" for conduct that "affect[s] individuals' rights and interests."[9] Without legality, the state does not abide by rule of law.[10] Instead, there exists a legal black hole.[11] Black holes in the law "arise when statutes or legal rules either explicitly exempt[] the executive from the requirements of the rule of law or explicitly exclude[] judicial review of executive action."[12] In such a black hole, a government

---

[6] *See* Evan J. Criddle, *Mending Holes in the Rule of (Administrative) Law*, 104 NW. U. L. REV. 1271, 1272–73 & n.11 (2010) (discussing Adrian Vermeule, *Our Schmittian Administrative Law*, 122 HARV. L. REV. 1095, 1096 (2009)). There is, to be sure, "a great deal of controversy about what the Rule of Law requires." Jeremy Waldron, *The Rule of Law*, STAN. ENCYCLOPEDIA PHIL. (June 22, 2016), https://plato.stanford.edu/entries/rule-of-law/ [https://perma.cc/VW7L-6J8U]. The summary here is a stylized one that lays a foundation for the significance of algorithmically generated legal holes.

[7] *See generally* LON FULLER, THE MORALITY OF LAW 41–42 (1964) (establishing eight principles with which a system of law-making must comply to establish an "inner morality of law"). There are many possible definitions of "rule of law" itself, including formal, substantive, and procedural versions. *See* Aziz Z. Huq, *Artificial Intelligence and the Rule of Law*, *in* THE ROUTLEDGE HANDBOOK OF THE RULE OF LAW (forthcoming 2022) (manuscript at 5–6), https://ssrn.com/abstract=3794777 [https://perma.cc/L67Q-A4TD] (providing an overview of these three understandings). This Essay emphasizes the formal understanding without foreclosing the possibility that related concerns arise for other rule of law understandings, too.

[8] Hence the saying that "rule of law means the government of laws, and not of men." *See* Berman, *supra* note 4, at 1309–10 (quoting Justice John Marshall's use of the phrase in *Marbury v. Madison* and discussing historic pedigree of the concept).

[9] David Dyzenhaus, *The Rule of Law Project*, 129 HARV. L. REV. F. 268, 270 (2016) (citing David Dyzenhaus, *The Compulsion of Legality*, *in* EMERGENCIES AND THE LIMITS OF LEGALITY 33 (Victor Ramraj ed., 2008)).

[10] *See* Berman, *supra* note 4, at 1310 & n.133; *see also* Waldron, *supra* note 6, at § 4.

[11] *See* David Dyzenhaus, *Schmitt v. Dicey: Are States of Emergency Inside or Outside the Legal Order*?, 27 CARDOZO L. REV. 2005, 2006 (2006) [hereinafter Dyzenhaus, *Schmitt v. Dicey*].

[12] Adrian Vermeule, *Our Schmittian Administrative Law*, 122 HARV. L. REV. 1095, 1096 (2009); *see also* Noa Ben-Asher, *Legal Holes*, 5 UNBOUND: HARV. J. LEGAL LEFT 1, 1 (2009) (quoting Vermeule) (citing David Dyzenhaus, THE CONSTITUTION OF LAW: LEGALITY IN A

actor's discretion is unconstrained, leaving an affected person without legal recourse to contest state action.[13] A prime example is the suspension of any right to contest an individual's detention by the state, such as the executive's suspension of the right of habeas corpus in a time of emergency. Akin to an astrological black hole that exerts a force so strong that not even light can escape it, the black hole in the law has engulfed the reality and the appearance of legality.

But perhaps a black hole that obviously swallows up the law is not the only bad outcome. Perhaps it is not even the worst outcome for individuals affected by the state. Despite the risk that a legal black hole can allow the state to operate without adequate constraints, some scholars have offered that there exists an even greater threat to rule of law: what David Dyzenhaus has termed a "grey hole" in the law.[14] Such a grey hole is "a legal space in which there are some legal constraints on [government] action—it is not a lawless void—but the constraints are so insubstantial that they pretty well permit government to do as it pleases."[15] Dyzenhaus contends that a grey hole is "in effect worse" than a black hole because it is "in substance black," yet it provides "the façade of legality."[16] In other words, by enacting procedures that create the appearance but not the reality of constraints on government action, a grey hole risks undermining the very fabric of our law.[17]

Dyzenhaus's understanding of legal grey holes arose in the context of national security legislation and the expansion of executive authority in the wake of the 9/11 terrorist attacks.[18] He emphasized a particular kind of grey hole arising from a detainee's inability to use their procedural rights "effectively to contest the [state's] case for [their] detention."[19] The risk was that legislation purporting to allow an individual to contest the conditions of their detention offered only a "thin veneer" of legality.[20] Dyzenhaus worried that this "little bit of legality can be more lethal to the rule of law than none."[21]

TIME OF EMERGENCY 3 (2006) and Johan Steyen, *Guantanamo Bay: The Legal Black Hole*, 53 INT'L & COMPAR. L.Q. 1, 1 (2004)).

[13] Dyzenhaus, *Schmitt v. Dicey*, *supra* note 11, at 2006.

[14] Dyzenhaus, *supra* note 9, at 268 (quoting Dyzenhaus, *Schmitt v. Dicey*, *supra* note 11, at 2006).

[15] Dyzenhaus, *Schmitt v. Dicey*, *supra* note 11, at 2018.

[16] *Id.* at 2039.

[17] Dyzenhaus, *supra* note 9, at 268–69 (quoting Dyzenhaus, *Schmitt v. Dicey*, *supra* note 11, at 2026).

[18] Dyzenhaus, *Schmitt v. Dicey*, *supra* note 11, at 2018.

[19] *Id.* at 2026.

[20] *Id.* at 2040.

[21] Dyzenhaus, *supra* note 9, at 268; *see also* Shirin Sinnar, *Rule of Law Tropes in National Security*, 129 HARV. L. REV. 1566, 1617–18 (2016) (adopting Dyzenhaus's black

This Essay takes up the argument that a façade of legality without effective redress is problematic from a rule of law perspective. Specifically, it offers that the concept of grey holes has traction beyond the acceptable scope of executive power in emergencies: it is a useful framework for thinking about risks posed by government deployment of algorithms in ordinary operations, too. By reframing concerns about black box algorithms as concerns about algorithmic grey holes that offer only the appearance of individual redress for state action, we can better understand the impact of the government's use of algorithmic tools at both the individual and the societal level.

II. FROM BLACK BOXES TO GREY HOLES

This Part contends that thinking about algorithms as black boxes, without attention to the related but distinct issue of algorithmic grey holes, is a mistake. It focuses on criminal adjudication as an especially acute illustration of the potential harms.[22]

Because algorithmic tools must be integrated within existing criminal law institutions, their use in the judiciary touches on bedrock dynamics of criminal adjudication. Constitutional criminal procedure centers on how to craft rules that allocate inevitable error. The Blackstone principle that it is better to let ten guilty individuals go free, than to convict one innocent person, is at bottom a directive about how to account for mistakes.[23] To implement

---

and grey holes schema and applying it to the idea of "rule of law tropes"). Others have since disputed the dangers posed by grey holes in other legal contexts. Adrian Vermeule, for instance, contends that administrative law is full of grey holes in which there are few real checks on executive action because judicial review "is more apparent than real" and concludes that these grey holes are unavoidable, endemic features of our legal system. Vermeule, *supra* note 12, at 1096 n.7 & 1096–98. The resolution of this dispute may turn on one's baseline understanding of rule of law and underlying "attitude[] towards law" as a "*theistic*-like" or "*scientific*-like" system. Ben-Asher, *supra* note 12, at 2. This Essay reserves further study of how algorithms interact with underlying values in particular legal settings, such as the administrative state, for future work. *Id*.

[22] This analysis addresses a stylized version of the criminal justice system, reserving practical questions such as how the proliferation of plea bargaining might change this understanding. *See, e.g.*, Laura I. Appleman, *A Tragedy of Errors: Blackstone, Procedural Asymmetry, and Criminal Justice*, 128 HARV. L. REV. F. 91, 92–93 (2015) (underscoring "plea bargaining's triumph" and its interaction with the Blackstone principle).

[23] *See generally* Daniel Epps, *The Consequences of Error in Criminal Justice*, 128 HARV. L. REV. 1065, 1072–73 (2015) (identifying the Blackstone principle as a "moral principle about the distribution of errors: we are obliged to design the rules of the criminal justice system to reduce the risk of false convictions—even at the expense of … [overall

this vision of justice, the American criminal justice system relies on rules that are known in advance and consistently applied to all individuals to determine guilt or innocence. These rules guarantee certain individual rights throughout the trial and appeals process, and they generally assume that procedural due process mechanisms amply check errors and substantiate those rights. These procedural rules of the road not only control the adjudicative process, but also affect how individuals understand and respond to the legal system and perceive state actors' use of force as legitimate (or not).[24] As Tom Tyler's work on procedural justice emphasizes, "people's reactions to legal authorities," including both courts and government officials such as the police, "are based to a striking degree on their assessments of the fairness of the processes by which legal authorities make decisions," as well as how those authorities "treat members of the public."[25]

When it comes to algorithmic accountability, though, it's not self-evident that procedural rules can afford the same rights, correct for the risk of error, and preserve the legitimacy of the system in quite the same way.[26] Consider a 2016 case, *State v. Loomis*.[27] In *Loomis*, a criminal defendant mounted a due process challenge to a trial court's use of a proprietary risk assessment tool.[28] The trial court had used the tool to assess the defendant's risk of

---

accuracy]").

[24] *See* Tom R. Tyler, *Procedural Justice, Legitimacy, and the Effective Rule of Law*, 30 CRIME & JUST. 283, 284 (2003) (connecting "procedural elements" to "process-based judgments" that can contribute to "supportive values" such as legitimacy); *see also* Hannah Bloch-Wehba, *Access to Algorithms*, 88 FORDHAM L. REV. 1265, 1294 & n.221 (2020) (citing Tracey L. Meares & Tom R. Tyler, *Justice Sotomayor and the Jurisprudence of Procedural Justice*, 123 YALE L.J.F. 525, 535 (2014)) (discussing procedural justice scholarship and the importance of the "appearance of fairness" in government decision-making). Legitimacy, like rule of law itself, is a slippery and multi-faceted concept. This Essay's working definition positions "legitimacy" as inherently connected to "the manner in which authorities exercise their authority—that is, to issues of procedure." Tyler, *supra*, at 286.

[25] Tyler, *supra* note 24, at 284.

[26] Ari Ezra Waldman has separately critiqued procedural approaches to algorithmic accountability. *See* Waldman, *supra* note 4, at 615. Waldman challenges Tyler's argument that the "legitimacy of legal authorities depends on process" as the best understanding of legitimacy in the context of automated decision-making. *Id*. This Essay takes a different tack: it adopts Tyler's thesis that process matters for legitimacy and contends that algorithmic tools scramble the assumed relationship between procedural rules, legitimacy, and rule of law. As discussed in detail *infra* Parts I-II, algorithmic grey holes generate the possibility that more layers of procedure create an external veneer of legality (and thus legitimacy), yet do not actually empower individuals internally affected by the system (and thus might actually diminish legitimacy for them).

[27] 881 N.W.2d 749 (Wis. 2016).

[28] Mr. Loomis brought three claims. He contended that the use of the risk assessment tool violated his "due process right to be sentenced based on accurate information" and

recidivism, and this risk assessment then entered the trial court's sentencing calculations.[29] On appeal, the Wisconsin Supreme Court found no constitutional problem with the use of the tool because the trial court's decision-making did not rely exclusively on the risk assessment instrument.[30] The *Loomis* Court focused on how the tool's actuarial prediction was only one source of information available to the trial court; emphasized that the court had discretion to discount or disregard the risk assessment; and stipulated certain further procedures and best practices for how this sort of information should be presented to a trial judge and how the trial judge should explain their ultimate decision.[31] These specific procedural requirements, coupled with generally applicable procedural due process principles, constrain what judicial and executive actors can do and thus place some limit on the bounds of state action. There is no "lawless void" in the legal system here.[32]

And yet this emphasis on procedural interventions fails to reckon with the broader rule of law stakes. In *Loomis*, the prescribed processes that constrain how and when the trial court can use an algorithmic tool in its sentencing determinations give the appearance of legality. Individuals can point to the law and attempt to mount a challenge. Like a detainee confronting secret state law, however, the letter of the law does not help them. There is no legal remedy to press in service of their case. There is, in short, an algorithmic grey hole.

Transparency might initially seem like the solution. The opacity of the tool is indeed troubling, for reasons well-canvassed in prior scholarship: it neither gives the defendant the ability to inspect the proprietary tool (or the individualized data used to render the decision), nor explains the inner workings of the tool itself.[33] It's hard to contest something that you can't see. The opacity thus becomes a problem insofar as it affects the substance of the arguments that the defendant can make and impedes their ability to obtain

---

deprived him of "an individualized sentence" based on his charges and "unique character of the defendant." *Loomis*, 881 N.W.2d at 760, 764. He further argued that the tool used gender in a way that violated his due process rights. *Id.* at 765.

[29] *Id.* at 753.

[30] *Id.* at 765–67.

[31] *Wisconsin Supreme Court Requires Warning Before Use Of Algorithmic Risk Assessments In Sentencing:* State v. Loomis, 130 HARV. L. REV. 1530, 1532-33 (2017) [hereinafter Recent Case: *State v. Loomis*].

[32] Dyzenhaus, *Schmitt v. Dicey*, *supra* note 11, at 2018.

[33] For critiques of trade secrecy claims in the context of criminal justice technologies, see, for example, Ram, *supra* note 2; Wexler, *supra* note 2.

meaningful review of potential algorithmic errors.[34]

A variation of this transparency problem emerges for judges, too. The *Loomis* Court's conclusion that it's acceptable for the trial court to use a risk assessment algorithm, so long as that trial court doesn't rely exclusively upon it, misses the point. Judges cannot make rational, informed choices about how to weigh the tool's input if they do not evaluate how the tool operates and then calibrate their own decision-making protocol accordingly.[35] And they cannot do so without, at a minimum, more data about the way that the tool operates.

These issues are natural extensions of the "black box" critique. Transparency matters in a situation like *Loomis* because it is instrumental to redress and review. It can empower actors in a position to respond to the tool's actual operation to access information that might inform their response.[36] In a state of opacity, only the creators of the tool have access to this necessary information. There is an informational asymmetry. Transparency can help to level it out. Transparency may thus be a necessary instrumental step to make procedures more effective.[37]

Transparency is not sufficient, however. And it can be dangerous, insofar as urging transparency as a cure-all diverts attention from the deeper issues.

---

[34] There is room to debate what is good enough and to contest what makes a remedial process "meaningful," whether that threshold is determined by cost-benefit analysis such as the *Mathews v. Eldridge* due process framework, or by another metric. Still, there must be some floor below which review cannot be said to be "meaningful." And common sense suggests that a degree of opacity that prevents a defendant from raising specific challenges under the controlling procedures falls below that floor. Thank you to Orin Kerr for pressing me on this point.

[35] *See* Recent Case: *State v. Loomis*, *supra* note 31, at 1534 (warning that it is not enough to "encourag[e] judicial skepticism of the value of risk assessments . . ." because that step "alone does little to tell judges how much to discount these assessments"); *see also* Katherine J. Strandburg, *Adjudicating with Inscrutable Decision Rules*, *in* MACHINES WE TRUST: PERSPECTIVES ON DEPENDABLE AI 61, 63, 81–83 (Marcello Pelillo & Teresa Scantamburlo eds., 2021) (examining impact of "inscrutable automated decision tools" on "explanatory flows to and from adjudicators").

[36] As Margot Kaminski and Jennifer Urban argue, procedural mechanisms such as an individual right to contest AI decision-making can "ameliorate individual harms and give life to broader rule of law values." Kaminiski & Urban, *supra* note 4, at 2003.

[37] On transparency's importance in allowing contestation of state action, see generally Bloch-Wehba, *supra* note 24. On transparency's potential role in reducing information asymmetries, see Emre Bayamlıoğlu, *Transparency of Automated Decisions in the GDPR: An Attempt for Systemisation* 17–22 (Jan. 16, 2018) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3097653 [https://perma.cc/583E-3WXP].

The underlying challenge is a structural and practical one: whether there are consistent legal rules that constrain state action, and which provide individuals with the opportunity to contest the state's use of algorithms, particularly for punitive measures.[38] The existence, or lack, of these rules will inform individuals' perceptions that the system itself is just.[39]

This dynamic matters not only for the affected individual, but also at the societal level. For one, focusing on a remedy in a single case is an unsatisfying resolution in the face of questions such as whether the algorithmic tool's validation protocol is problematic across the board. Because they are developed to be applied across many cases, algorithmic tools tend to involve these sorts of system-wide questions.[40] For another, it's hard to have confidence in a system of laws when the letter of the law does not facilitate review of errors. Just as Dyzenhaus worried that an outcome in which the procedures available to detainees did not allow them to obtain effective review of their conditions of detention, so too is there cause for concern when the procedures available to defendants do not allow them to obtain effective review of their sentencing determinations. Courts risk constructing a façade of legality when they prop up procedures that seem effective, without doing much at all.

In addition to these direct consequences, algorithmic grey holes are likely to lead to broader regulatory dysfunction. Recall that a grey hole requires some amount of procedural redress; otherwise, it would be a black hole. In the abstract, these procedures provide some minimal appearance of legality. They seem to provide actual relief. But the available forms of redress function poorly from the standpoint of individuals affected by the system. This gap

---

[38] This focus on a more foundational question, related to yet distinct from the issue of transparency itself, is similar in spirit to Ryan Calo and Danielle Citron's shift away from the "project of restoring rights and values displaced by technology," and towards the more systemic question of whether a government actor's embrace of a technology is in tension with the underlying justification for the government actor's conduct. Ryan Calo & Danielle Keats Citron, *The Automated Administrative State: A Crisis of Legitimacy*, 70 EMORY L.J. 797, 803 (2021). Rather than focus on federal agencies and administrative legitimacy, as Calo and Citron do, the present piece foregrounds state and local agencies and rule of law values.

[39] See *supra* notes 24-25 and accompanying text.

[40] *See, e.g.*, Aziz Z. Huq, *Constitutional Rights in the Machine-Learning State*, 105 CORNELL L. REV. 1875, 1909–10 (2020) (arguing that algorithmic instruments are likely to produce "systemic problems" that are ill-suited to individualized, "retail challenges"); Mariano-Florentino Cuéllar & Aziz Z. Huq, *Toward the Democratic Regulation of AI Systems: A Prolegomenon* 8 (July 14, 2020) (unpublished manuscript), https://ssrn.com/abstract=3671011 [https://perma.cc/3Y77-6WE7] ("[W]e must look at AI systems and not just instruments in splendid isolation.").

between the promise and the reality of redress creates only the appearance of constraints on state action: a grey hole.

Picture a policymaker who has, for whatever reason, failed to recognize the existence of such a grey hole. This oversight need not be malicious; it could arise because the policymaker focused on the overall operation of the system, and not on the lived reality of individuals seeking redress.[41] The risk is that such a legislator reflexively forges familiar procedural safeguards, such as disclosure requirements, that in fact create only the appearance of legality. As Julie Cohen has emphasized in the context of information privacy legislation, human beings tend to craft legislation that reflects the known universe of pre-existing challenges.[42] Here, such a result would fail to engage with the existence of grey holes and how they affect subjects of punitive government actions. It would also divert regulatory attention away from more effective interventions.

To avoid these consequences, what should a court or regulator confronting a situation like *Loomis* do instead? There is no easy answer, particularly insofar as the issues that come up in the context of algorithms were present before, too. For instance, well before the Wisconsin trial court used a proprietary risk assessment tool to inform sentencing, other forms of risk assessment entered judges' calculations. It is not possible to open the black box of a judge's brain and understand exactly how their own lived experiences inform their weighting of various risks. The sentencing enterprise has long been about predictions of future risk, whether that assessment is made by clinical judgements, non-automated actuarial judgements, or more advanced forms of artificial intelligence.[43] As former

---

[41] An observer could also willfully ignore this gap or seek to exploit it. In the national security context, for example, a hawkish legislator might find the façade of legality attractive if it permits the extension of surveillance or detention activities without generating judicial pushback or public outcry. *See* Dyzenhaus, *Schmitt v. Dicey*, *supra* note 11, at 2039 (suggesting that proposed legislation concerning detention at Guantanamo Bay "is such a thin veneer of legality that the grey hole sanctioned by the plurality [of the Supreme Court] in Hamdi will shade into blackness"); Sinnar, *supra* note 21, at 1618 (suggesting that rules announced by the Executive, "evoking familiar concepts from constitutional or international law, may offer false comfort in the Executive's ability to protect liberty and to police itself — hindering accountability"). This Essay gives would-be policymakers the benefit of the doubt and focuses on the potential for negative regulatory consequences from even well-intentioned actors.

[42] Julie E. Cohen, *How (Not) to Write a Privacy Law*, KNIGHT FIRST A. INSTIT. 1, 2–3 & n.1 (2021), https://s3.amazonaws.com/kfai-documents/documents/306f33954a/3.23.2021-Cohen.pdf [https://perma.cc/F24W-GTF9].

[43] *See* Alicia Solow-Niederman, YooJung Choi, & Guy Van den Broeck, *The Institutional Life of Algorithmic Risk Assessment*, 34 BERKELEY TECH. L.J. 705, 710–18

district court judge Katherine Forrest explains, a judge may "fe[el] obligated by the existing legal requirements to base [their] sentences in part on . . . [their] personal assessment of the individual's . . . likelihood of recidivating, and the extent to which any recidivism would harm the community or those around [the individual]," yet lack any way to measure the accuracy of their predictions.[44] Put sharply, a human being's "algorithm" for making a decision about a defendant is not transparent. It is therefore tempting to conclude that a technological algorithm should be no different: both situations involve opacity, and we can trust in criminal law's tried and true procedures and norms to control for the risk of errors as best as we can.

But there are several key distinctions between algorithmic systems and judges. One is subjective: we sense that an algorithmic instrument could be amenable to more transparency. Algorithms introduce the tantalizing prospect that we can get at the data and processes and check them, in a way that we cannot with a person. That urge may be even more pressing where the tools are created by private firms with profit motives. Another is objective: we have evidence that algorithmic instruments are not accurate, and that they display certain forms of statistical unfairness, such as higher false positive error rates for Black defendants and higher false negative rates for white defendants.[45] We are thus on alert that there are opportunities both to improve the administration of criminal justice and to avoid known harms.

Because algorithms interact with human beings as part of a complex

(2020) (providing a brief history of actuarial tools and, more recently, algorithmic risk assessment instruments).

[44] Jed S. Rakoff, *Sentenced by Algorithm*, N.Y. REV. BOOKS (June 10, 2021) (reviewing KATHERINE M. FORREST, WHEN MACHINES CAN BE JUDGE, JURY, AND EXECUTIONER: JUSTICE IN THE AGE OF ARTIFICIAL INTELLIGENCE (2021)).

[45] ProPublica published a much-disseminated piece exposing this disparate error rate in a tool known as COMPAS and called the tool unfair based on error rate bias. Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [https://perma.cc/KY8C-YK5S]. However, other researchers have disputed this conclusion, asserting that the tool is in fact fair because it exhibits predictive parity, meaning that an equal number of Black and white defendants are predicted to be "high risk" at a given quantitative level of predicted risk. *See* Sam Corbett-Davies, Emma Pierson, Avi Feller & Sharad Goel, *A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear.*, WASH. POST (Oct. 17, 2016), https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/ [https://perma.cc/8FYA-SZP8]. For further discussion of the COMPAS controversy, see Sandra G. Mayson, *Bias In, Bias Out*, 138 YALE L.J. 2218, 2238 (2019), and Laurel Eckhouse, Kristian Lum, Cynthia Conti-Cook & Julie Ciccolini, *Layers of Bias: A Unified Approach for Understanding Problems with Risk Assessment*, 46 CRIM. JUST. & BEHAV. 185, 190–92 (2019).

social and technical system, there could be alternate interventions that might provide more, or different, forms of judicial review. To move in this direction, a court might insist on even more robust forms of public justification in the context of a given case, or the legislature might require certain forms of public justification in cases that involve the use of algorithms.[46] For example, a specific court might require different forms of explanation from the parties to assess how a tool is operating and make an informed judgment.[47] Or a legislative body might specify across the board that certain kinds of particularly inscrutable algorithmic decision-making, like so-called "neural net" machine learning, are off the table in high-stakes arenas such as criminal justice. Now, we may be holding the algorithm to a super-human standard if we insist on different standards or procedural requirements for algorithmic decision-making. People can be biased, and the systems we have are far from perfect. But there is an underappreciated cost to the status quo, too: when these alternate procedures are not taken, the existing ones can generate an aura of legality, without affording meaningful review or redress, and give rise to algorithmic grey holes.

* * *

Such a threat to rule of law is not limited to sentencing, nor does it appear only in the context of criminal justice adjudication or in narrow windows that demand urgent state action. It is omnipresent. Algorithmic grey holes clash with thick rule of law expectations whenever added procedures endow available processes with the aura of legality, yet do not empower affected individuals to challenge the state's use of force against them. This result corrodes the legitimacy of the government's exercise of authority. Moreover, the use of algorithms is not limited to an emergency; to the contrary, the state's use of a tool like the *Loomis* risk assessment instrument emerges as an everyday development to improve ordinary criminal justice operations.[48]

---

[46] *See* Megan Stevenson, *Assessing Risk Assessment in Action*, 103 MINN. L. REV. 303, 373–74 (2018) (suggesting, in the pretrial context, that policymakers seeking to accomplish certain policy outcomes—such as lowering the jail population—may need to "establish[] clear guidelines for when [judicial] deviation is or is not allowed, [or] mak[e] deviation costly for the judge in some way (e.g. requiring a detailed written explanation of the reasons for deviation)").

[47] *See* Ashley Deeks, *The Judicial Demand for Explainable Artificial Intelligence*,119 COLUM. L. REV. 1829, 1831, 1838–42 (2019) (identifying areas in which "judges are likely to play a critical role in fleshing out whether [explainable AI] is required and, if so, what forms it should take").

[48] *See, e.g.*, Solow-Niederman, *supra* note 4, at 710–14 (discussing history of actuarial tools). This apparent ordinariness does not diminish the rule of law stakes; to the contrary, "the debate about the rule of law is a theoretical and normative one and as much about what is appropriate during ordinary or normal times" as it is one about emergencies. Dyzenhaus,

This omnipresence makes it even more important to recognize where grey holes exist. That's essential not only to avoid direct harms to individuals or to the rule of law, but also because a lack of adequate checks on the deployment of algorithms can undermine the case for potentially helpful applications of emerging technologies.

III. ALGORITHMIC GREY HOLES BEYOND CRIMINAL ADJUDICATION

This Part exposes algorithmic grey holes as a widespread phenomenon that reaches law enforcement targeting decisions as well as civil settings, such as the distribution of welfare benefits, and contends that this phenomenon implicates not only individual rights, but also structural concerns.

A. *Before Adjudication: Algorithmic Enforcement*

Loomis involved the unique setting of a criminal trial, yet the questions it raises are not unique. Algorithmic grey holes pose equally difficult challenges in earlier stages of law enforcement, too. Before a defendant arrives in court, there are multiple antecedent points when law officers make "street-level" decisions about how to interpret policy and enforce the law on the ground.[49] When technological instruments enter these decision points and themselves mediate officers' exercise of discretion, the governance of these tools and their applications becomes even more fraught because there is a risk of creating algorithmic grey holes.

One especially salient moment occurs at the entry point to the criminal justice system: when a law officer exercises their enforcement discretion and determines that a particular individual has committed a criminal offense. To see how algorithms might affect this process, consider a facial recognition tool that is sold to a state or local law enforcement agency by a private

Schmitt v. Dicey, *supra* note 11, at 2035.

[49] *See* MICHAEL LIPSKY, STREET-LEVEL BUREAUCRACY 14, 24–25 (1983); *see also* Joseph Goldstein, *Police Discretion Not to Invoke the Criminal Process: Low-Visibility Decisions in the Administration of Justice*, 69 YALE L.J. 543, 546–54 (1960) (discussing police decisions not to enforce substantive criminal laws and locating these choices within the "criminal law process"); *see also* SARAH BRAYNE, PREDICT AND SURVEIL: DATA, DISCRETION, AND THE FUTURE OF POLICING 13 (2020) (identifying police as the "prototypical 'street-level bureaucrat'" and arguing that big data can "either ossify or upset existing organizational and legal dynamics"); Ali Alkhatib & Michael Bernstein, *Street-Level Algorithms: A Theory at the Gaps Between Policy and Decisions*, *in* 2019 PROC. CHI CONF. ON HUM. FACTORS COMPUTING SYS. 1, 2–3 (noting that street-level bureaucrats often "make consequential decisions about what to do with a person").

vendor.[50] The matches made by such tools are known to be much less accurate for people of color.[51] Indeed, the racial disparities of these tools are now sufficiently well-recognized that many activists and scholars call for them to be banned, and big tech firms like Microsoft and Amazon have enacted moratoria on the sales of their own facial recognition tools to law enforcement.[52] But they remain available to government actors, and they are not universally critiqued; in fact, facial recognition tools were used by both law enforcement and private citizens to identify perpetrators of the January 6, 2021 riot at the U.S. Capitol.[53] Descriptively, these tools thus represent a technological mechanism that a law officer can use to bridge the gap between a high-level policy goal (prevent crime and detain those who commit offenses) and a specific, street-level enforcement choice (detain a particular individual on the grounds that they have committed a particular offense).

The ongoing use of these tools makes it essential to face their human impact, and to acknowledge what current procedural constraints do and do not accomplish. The first-order effects are at the level of the individual. When a Black man is improperly identified and arrested based on a facial recognition match, what is the relief available to him? The case of Robert

---

[50] *See* Drew Harwell, *Wrongfully arrested man sues Detroit police over false facial recognition match*, WASH. POST (Apr. 13, 2021) https://www.washingtonpost.com/technology/2021/04/13/facial-recognition-false-arrest-lawsuit/ [https://perma.cc/ZLZ2-6HTP] (reporting Detroit and Michigan state police contract with biometrics firm DataWorks Plus).

[51] *See* Press Release, NIST, NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software (Dec. 19, 2019), https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software [https://perma.cc/PRD2-46XF] (documenting high rates of false positives for Asians, African Americans and native groups in set of 189 facial recognition algorithms evaluated by NIST). *See generally* Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, PROC. OF MACH. LEARNING RSCH. (2018), https://proceedings.mlr.press/v81/buolamwini18a.html [https://perma.cc/3HG8-Z2RF] (revealing racial and gender disparities in facial recognition tools).

[52] *See* Jay Greene, *Microsoft won't sell police its facial-recognition technology, following similar moves by Amazon and IBM*, WASH. POST (June 11, 2020), https://www.washingtonpost.com/technology/2020/06/11/microsoft-facial-recognition/ [https://perma.cc/F37Y-8MMT].

[53] *See* Mark Harris, *How Facial Recognition Technology Is Helping Identify the U.S. Capitol Attackers*, IEEE SPECTRUM (Jan. 11, 2021), https://spectrum.ieee.org/facial-recognition-and-the-us-capitol-insurrection [https://perma.cc/YN7K-A59P] (describing law enforcement use of technology in response to the January 6 activity in Washington, D.C.); Alice Hines, *How Normal People Deployed Facial Recognition on Capitol Hill Protesters*, VICE (Feb. 2, 2021), https://www.vice.com/en/article/4ad5k3/how-normal-people-deployed-facial-recognition-on-capitol-hill-protesters [https://perma.cc/X6VQ-B8KD] (describing private citizens' use of facial recognition technology to identify Capitol Hill protesters).

Williams, a Black resident of Detroit who was falsely identified and wrongfully arrested as a shoplifting subject, illustrates the current system's deficiencies.[54] Mr. Williams was identified based on the application of facial recognition technology to security video at a store, even though the security officer who provided the video was not physically present at the time that the shoplifting occurred and no one who was at the scene was shown the photo line-up.[55] This method of identification, without an in-person witness, contravened established rules. As such, the case against Mr. Williams was dismissed at his probable cause conference—but he spent time in jail, and he and his family were subjected to the traumatic experience of his false identification and subsequent arrest and detention. He has brought a civil suit against local police in response, as have others who were similarly falsely identified by police relying on facial recognition.[56] He seeks not only damages, but also policy changes within the police department.[57]

This push for policy changes reaches the heart of the issue: in addition to problematic individual applications of the tool, the tool is embedded in existing practices in ways that may subvert accountability by creating a bare appearance of legality.[58] This rule of law threat is a structural problem that goes beyond the directly affected individual. In Mr. Williams' case, the policing procedures that were in place appeared to constrain official conduct, and thus appeared to sustain the legality of the system. But these procedures did not in fact constrain officials' conduct when it came to the application of algorithmic force. Law enforcement officers instead manipulated existing policies and used the technology to route around the analog protections that should have controlled their investigative process. And they were able to do

---

[54] *See* Harwell, *supra* note 50 (discussing the arrest of Robert Williams).

[55] *See* Press Release, Cnty. of Wayne Off. of the Prosecuting Att'y, WCPO Statement in Response to New York Times Article Wrongfully Accused by an Algorithm (June 24, 2020), https://www.waynecounty.com/elected/prosecutor/wcpo-statement-in-response-to-new-york-times-article.aspx [https://perma.cc/7LCN-W6FT].

[56] Complaint & Demand for Jury Trial, Williams v. Detroit, No. 2:21-cv-10827 (E.D. Mich. Apr. 13, 2021); *see also* Harwell, *supra* note 50 (discussing Mr. William's civil lawsuits).

[57] Complaint & Demand for Jury Trial, *supra* note 56; Harwell, *supra* note 50.

[58] Ari Ezra Waldman has offered a trenchant critique of similarly "performative" accountability in privacy law, drawing on Lauren Edelman's work on "legal endogeneity" and performative compliance in the employment context. *See* Waldman, *supra* note 4, at 628–29 (describing Edelman's theory of "mobilization of symbolic structures"). *See generally* Ari Ezra Waldman, *Privacy Law's False Promise*, 97 WASH. U. L. REV. 773 (2020) (arguing that privacy law incentivizes merely symbolic compliance). Rather than focus on corporate compliance and how courts perceive private compliance mechanisms in the privacy or employment context, this Essay emphasizes the public law stakes and threat to rule of law values when algorithmic procedures merely appear law-like.

so while still benefiting from the aura of legality created by the bare existence of controlling procedures.

The rules on the books neither grapple with, nor provide a direct path to redress, the broader harms that this kind of violation can cause. Instead, a defendant like Mr. Williams must use a civil lawsuit as a vehicle to push for reform of the entire system, as a policy matter. Unless an affected individual challenges the way in which the procedures themselves endow the use of the tool with the appearance of legality, there is no lasting relief. Furthermore, absent such a lawsuit, it is unlikely that would-be policymakers will even become aware of the problems. As Clare Garvie, an expert who has researched and critiqued law enforcement use of facial recognition, explains, "there's the burden of somebody after the fact, who's already been injured by a misidentification, to inform the public of what happened to them."[59] Meanwhile, the procedures evoke the "barest forms of rule by law," and thereby "seem to evoke the idea that the rule is legitimate because it is in accordance with the law, that is, the rule of law."[60] But in reality, there is an algorithmic grey hole.

Society pays the price. For affected individuals, there is an uphill battle to vindicate rights in the face of government officials' use of the technology, while government officials continue to take advantage of the appearance of legality generated by existing legal rules. Moreover, the net effect is not limited to those who are directly affected. Each of us may be continually subject to review by algorithms that operate without effective checks, at risk of stumbling into an algorithmic grey hole if we become the object of an investigation. Even if we are never swept up in a criminal investigation, this dynamic generates mistrust in government action in ways that undermine government legitimacy and corrode a thick conception of rule of law.

Nor are these dynamics limited to a criminal investigation determination, per se. That example is an especially acute representation of the issues, yet any algorithm that prioritizes individuals and singles them out as an object of government attention, without guaranteeing that the procedures available to them can lead to meaningful relief, risks creating an algorithmic grey hole. This risk extends to a phenomenon that is seemingly as banal as an airport search. Others have written at length about the prospect of government overreach in the wake of 9/11; notably, Shirin Sinnar has suggested that national security agents take advantage of "rule of law tropes" to expand

[59] Harwell, *supra* note 50.

[60] Dyzenhaus, *Schmitt v. Dicey*, *supra* note 11, at 2029*; see also* Ben-Asher, *supra* note 12, at 13–14 (discussing Dyzenhaus).

surveillance capacity without meaningful accountability.[61] The same concerns apply with renewed force in the algorithmic context, which adds layers of opacity and technological complexity. Imagine an algorithm that tells a TSA officer a risk score for a passenger and then empowers the officer to use that score to determine whether the person warrants enhanced surveillance.[62] If an individual cannot meaningfully contest the algorithm that identifies them for enhanced surveillance, then the procedures that are in place risk legitimizing a process that, at bottom, lacks much if any substantive content—an algorithmic grey hole.

B. *Beyond Criminal Justice: Civil Applications*

Algorithmic grey holes, moreover, are not limited to the criminal justice context. The criminal justice system is a pressing example, given its high-stakes decisions that affect life and liberty. Criminal adjudication, where judges must inevitably fill in the gaps left by open-ended standards, is the domain where grey holes emerge in starkest relief. In criminal adjudication, there is an especially acute, direct relationship between procedures that endow a process with apparent legality and the question of whether an affected person can use those procedures to obtain relief. But it is not entirely unique; as we have seen, similar dynamics arise with law enforcement choices made before cases arrive in court. And in civil domains, government officials use algorithmic tools to prioritize who should be subjected to particular burdens, who should receive particular benefits, or how to distribute limited state resources. Analyzing patterns across both criminal and civil law can help us to understand how algorithmic tools are mediating the relationship between the government and the governed.[63]

---

[61] *See generally* Sinnar, *supra* note 21, at 1569–81 (identifying "rule of law tropes" in executive national security lawmaking that "leave the public misinformed about the extent to which internal rules constrain power and protect rights").

[62] Little imagination is required. According to a 2020 report on the federal government's use of artificial intelligence, U.S. Customs and Border Protection uses an "Automated Targeting System" that "generates and assigns a rating to every entity that crosses U.S. borders, determining the potential threat a given [person, vehicle, or container] poses and the level and priority of screening it should receive." DAVID FREEMAN ENGSTROM, DANIEL E. HO, CATHERINE M. SHARKEY & MARIANO-FLORENTINO CUÉLLAR, GOVERNMENT BY ALGORITHM: ARTIFICIAL INTELLIGENCE IN FEDERAL ADMINISTRATIVE AGENCIES 32–33 (Feb. 2020), https://www.acus.gov/report/government-algorithm-artificial-intelligence-federal-administrative-agencies [https://perma.cc/5EMR-Y4MA].

[63] This Essay focuses exclusively on public actors, where the rule of law connection is most explicit. But the phenomena identified here may sweep beyond state action. Technological platforms, such as social media sites, craft and implement their own procedural rules that apply to their users. There is at least a family resemblance between algorithmic grey holes generated by the government and those generated by private actors, insofar as available procedures leave users feeling that they do not have meaningful redress

As in the criminal justice system, the precise nature of an algorithmic grey hole in the civil sector varies depending on the type of work that the algorithm is used to do and how the instrument relates to human decision-making processes and constraints. At one extreme lie entirely automated algorithmic systems. Although such algorithms have not yet entered criminal justice in the United States,[64] their deployment in at least some civil society settings suggests that the institutional configuration of such tools must account for the potential rule of law costs.

One frequently-criticized system, Michigan's welfare fraud detection algorithm, is an illustrative example. For nearly two years, Michigan deployed the Michigan Integrated Data Automated System, or "MiDAS" for short, to automate its review of potential unemployment insurance fraud. The system had a broad scope and included retroactive review of nearly six years of records.[65] Michigan's process relied on the tool to identify alleged fraud

---

in the face of punitive platform decisions. For instance, Meta (formerly Facebook) publishes meeting minutes for its Product Policy Forum convenings, see Product Policy Forum Minutes, META (Nov. 15, 2018), https://about.fb.com/news/2018/11/content-standards-forum-minutes/ [https://perma.cc/Q926-EK22], reports removal decisions in quarterly transparency reports, see, e.g., *Community Standards Enforcement Report*, META, https://transparency.fb.com/data/community-standards-enforcement/ [https://perma.cc/GCK8-VFGH] (last visited Jan. 7, 2022), and operates an Oversight Board that permits users to appeal content decisions, see OVERSIGHT BOARD, https://oversightboard.com [https://perma.cc/8N5W-8XY6] (last visited Jan. 7, 2022). If these procedural measures do not empower affected users to seek meaningful redress for decisions that affect their conduct on the platform, and if one believes that these platforms are governance actors on a par with state actors, then the extension of algorithmic grey holes to this context feels quite natural. For scholarship applying a governance lens in the context of private platforms, see, for example, Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598 (2018) (positioning platforms as "new governors"); Rory Van Loo, *The Corporation as Courthouse*, 33 YALE J. ON REG. 547 (2016) (assessing private adjudicatory mechanisms within firms). The rule of law connection, however, may be less forceful in the platform context if one does not find it appropriate to analogize platforms to states. *See generally, e.g.*, Morgan Weiland, Beyond Platforms: The Laws, Ideology, and Political Economy That Broke the Public Sphere, and Why Recuperating Listeners' Rights Can Fix It 4 (Feb. 3, 2022) (unpublished manuscript) (on file with author). This Essay reserves treatment of private action and algorithmic grey holes, particularly in the content moderation context, for another day.

[64] The picture is different internationally. For instance, China has reportedly developed a "system [that] can replace prosecutors in the decision-making process to a certain extent." Stephen Chen, *Chinese scientists develop AI 'prosecutor' that can press its own charges*, S. CHINA MORNING POST (Dec. 27, 2021), https://www.scmp.com/news/china/science/article/3160997/chinese-scientists-develop-ai-prosecutor-can-press-its-own [https://perma.cc/9JDY-9DQT].

[65] *See* Paul Egan, *Data glitch was apparent factor in false fraud charges against jobless*

and then immediately demanded repayment plus interest and substantial penalties from individuals whose claims were deemed fraudulent.[66] There was no human intervention in this process, which identified thousands of workers' claims as fraudulent and collected over $57 million in penalties for the state.[67] But MiDAS was deeply flawed, operating with a 93% error rate.[68] Because both the fraud determination and the resulting penalties were applied automatically, it left affected individuals struggling to reconstitute their livelihoods as they attempted to appeal the automated decision.[69]

The concept of algorithmic grey holes highlights how the MiDAS system was so pernicious, above and beyond its inaccuracy and human toll. On the surface, there were procedures that endowed the system with the aura of legality: claimants were permitted to respond to an eight-question survey within ten days.[70] In reality, there was the bare appearance of redress procedures, without the availability of effective review. For instance, having been "flagged" for fraud, "MiDAS did not inform the claimant about the basis for the Agency's suspicion or provide the claimant with any information to allow [them] to rebut the fraud charge."[71] MiDAS also further limited how claimants could pursue review because it "did not allow for a fact-based adjudication or give the claimant the opportunity to present evidence to prove that [they] did not engage in disqualifying conduct."[72] As Aziz Huq explains, the inscrutability of the MiDAS enforcement process made it difficult to challenge, causing the "automati[on of] enforcement discretion [to] squeeze[] out the possibility of oversight by adjudication."[73] Moreover, many claimants

claimants, DETROIT FREE PRESS (July 30, 2017), https://www.freep.com/story/news/local/michigan/2017/07/30/fraud-charges-unemployment-jobless-claimants/516332001/ [https://perma.cc/Z23X-X7KM] (investigating the origin and impact of the "glitch"); *see also* Cahoo v. SAS Analytics, 912 F.3d 887, 892–93 (6th Cir. 2019) (describing how the MiDAS system worked in practice).

[66] *See* Michele Gilman, *AI Algorithms Intended to Root Out Welfare Fraud Often End Up Punishing the Poor Instead*, THE CONVERSATION (Feb. 14, 2020), https://theconversation.com/ai-algorithms-intended-to-root-out-welfare-fraud-often-end-up-punishing-the-poor-instead-131625 [https://perma.cc/8SNU-FG7V] ("Without any human intervention, the state demanded repayments plus interest and civil penalties of four times the alleged amount owed.").

[67] Egan, *supra* note 65; Ryan Felton, *Lawsuit Challenging Michigan Unemployment Fraud Cases Moves Forward*, DETROIT METRO TIMES (Mar. 30, 2016), https://www.metrotimes.com/news/lawsuit-challenging-michigan-unemployment-fraud-cases-moves-forward-2435449 [https://perma.cc/J7LL-2S53].

[68] Gilman, *supra* note 66.

[69] *Id.*

[70] *SAS Analytics*, 912 F.3d at 893.

[71] *Id.*

[72] *Id.*

[73] Huq, *supra* note 7 (manuscript at 5).

who might have pursued adjudication were unaware of the fraud determination until after their window to appeal to a court had passed—after they had already paid a steep price.[74]

These cumulative barriers amount to an algorithmic grey hole. This grey hole embedded an automated tool within the welfare system and left individuals unable to obtain review of the algorithmic determinations, short of a challenge to the entire system. It fell on individual MiDAS claimants to manage to successfully file claims before administrative law judges and thereby draw attention to the flaws that affected how the government exercised its authority across the board. And only policy changes, such as eliminating the state's reliance on automated decisions without human review, eventually altered the situation by correcting the overt algorithmic errors and forcing the government to stop relying on an automated detection and punishment process.[75] When so many individuals are harmed without effective review, it fosters a situation in which those on the receiving end of state force are unlikely to perceive the relationship between the government and the governed as legitimate. This outcome is obviously bad for individuals. Additionally, it is bad for society if it becomes harder to craft institutional processes that permit government to use algorithmic tools in more nuanced and productive ways.

The size and scope of this example suggests, moreover, that some algorithmic grey holes may be addressed only if the problem becomes egregious enough that there is a substantial class of affected persons. This implication raises thorny questions about subtler forms of harm in situations where it may not be quite so obvious whether an algorithm is hurting or helping affected populations. For instance, the Allegheny Family Screening Tool (AFST) that Pennsylvania developed to screen child welfare calls containing allegations of mistreatment has been both touted for improving the system-wide accuracy of predictions and critiqued for unfairly targeting poorer individuals and thereby perpetuating socioeconomic inequities.[76] In

---

[74] *See SAS Analytics*, 912 F.3d at 894 (documenting how claimants who attempted to appeal were told that they could not appeal, even if those claimants never received notice).

[75] The agency eventually put a human "in the loop" by relying on human verification before making a final fraud determination with MiDAS. *See* Egan, *supra* note 65 ("The agency continues to use MiDAS, but with human verification required for fraud determinations.").

[76] *Compare, e.g.*, Frequently-Asked Questions, ALLEGHENY COUNTY DEPT. HUMAN SERVICES 11 (Apr. 2019) https://www.alleghenycountyanalytics.us/wp-content/uploads/2019/05/FAQs-from-16-ACDHS-26_PredictiveRisk_Package_050119_FINAL-8.pdf [https://perma.cc/CWF6-6K8R] ("Compared to the existing system, the AFST is expected to increase accuracy and

the context of a tool like the AFST, how do affected persons perceive their available means of redress, and how do legal processes endow the deployment of the algorithm with a sense of legality? This under-examined challenge will only grow in importance as algorithms proliferate across government domains as far-ranging as welfare determinations; other civil enforcement settings like immigration or financial regulation; hiring decisions; and beyond.

CONCLUSION

Algorithmic proliferation can be for the good: data analysis can permit the state to escape the limitations and biases of human cognition, model previously unseen patterns, and unlock more efficient uses of limited government resources.[77] But those positive outcomes are possible only if the human beings who are most affected by a tool perceive that tool's use as legitimate. Assessing whether that is the case requires moving beyond questions about whether an algorithm is a black box, and instead confronting the risk that an algorithm will be embedded in a system in ways that seem to permit review of its methods and seem to be controlled by legal standards—yet leave individuals in an algorithmic grey hole. This reality suggests that decisions about whether to adopt an algorithm in the first instance should receive more weight. Even when an algorithmic turn seems to do little more than update an analog process, as in the case of sentencing, it displaces some amount of human decision-making. And in so doing, it changes the nature of the relationship between the government and the people who are affected by government decisions.

When the state endorses the displacement of human discretion by adopting a tool, particularly in high-stakes settings such as criminal justice, it risks generating algorithmic grey holes. Policymakers should take care that they do not pay for the promise of innovation and positive outcomes in the form of diminished legitimacy with affected human beings. That question goes to the heart of what rule of law demands. In the algorithmic context, it may require thinking about which groups warrant input in design and deployment choices, and how to account for the voices of affected

---

consistency of decision-making, which means wrongful stigma is expected to be reduced."), *with* Virginia Eubanks, *A Child Abuse Prediction Model Fails Poor Families*, WIRED (Jan. 15, 2018), https://www.wired.com/story/excerpt-from-automating-inequality/ [https://perma.cc/JT7K-PZKB] (arguing that the AFTS builds in "[h]uman choices, biases, and discretion" and engages in "poverty profiling").

[77] *See, e.g.*, Ryan Calo, *Modeling Through*, 71 DUKE L.J. 1391, 1405–09 (2022) (discussing how modeling might improve policymaking).

populations. The questions of whether, and under what conditions, algorithmic deployments warrant local control, and what exactly individuals should be able to contest to preserve government legitimacy, are ripe for further research. So too, may it require some additional process, particularly in the form of procedures that help to increase systemic accountability, such as algorithmic audits.[78] The potential use of these tools as systemic accountability measures, above and beyond mechanisms for individual redress, makes it critical to think hard about the standards that should apply to algorithmic auditing processes.[79] Otherwise, such procedures may amount to little more than ethics-washing. To make progress with algorithmic tools, government actors must recognize that adopting an algorithm is a policy choice, and not an inevitable conclusion.

[78] Margot Kaminski has suggested that algorithmic governance may require both systemic accountability mechanisms and protection of individual rights. *See* Kaminski, *supra* note 2. Scholars such as Andrew Selbst have suggested that algorithmic impact assessments might help to produce better algorithmic accountability, notwithstanding the practical challenges of industry self-governance. *See* Andrew D. Selbst, *An Institutional View of Algorithmic Impact Assessments*, 35 HARV. J. L. & TECH. 78 (2021). *But see* Waldman, *Privacy Law's False Promise*, *supra* note 58, at 773 (cautioning that privacy law is suffering from its reliance on industry-driven compliance measures, like audits, that are "standing in for real privacy protections").

[79] *See generally* ADA LOVELACE INSTITUTE & DATA KIND UK, EXAMINING THE BLACK BOX: TOOLS FOR ASSESSING ALGORITHMIC SYSTEMS (Apr. 29, 2020), https://www.adalovelaceinstitute.org/wp-content/uploads/2020/04/Ada-Lovelace-Institute-DataKind-UK-Examining-the-Black-Box-Report-2020.pdf [https://perma.cc/H38S-N9CX] (evaluating the promise, and limits, of different kinds of algorithmic audits and impact assessments).