A Framework for the New Personalization of Law

Anthony J. Casey† & Anthony Niblett††

Law has always strived for accurate contextualization, but only with recent technological advances in data processing and communication has this goal become meaningfully achievable at the personal level. While the other essays in this Symposium explore the costs and benefits of personalizing particular areas of law, we present a general framework for thinking about the new personalization of law. We identify two fundamental questions that every personalization project must address: First, how do lawmakers set the objective of a personalized law? Second, how is the content of a personalized law communicated to the citizens who must follow it? We explore these questions and identify specific challenges they pose to any personalization project.

INTRODUCTION

Personalized law is an old concept. The idea that the law should be tailored to better fit the relevant context to which it applies is obvious and has been around as long as the idea of law itself.[1] Indeed, every law has some contextual parameters. The question is how specific—or how finely tailored—those parameters will be. In a world without any frictions, an ideal law would take all relevant (and no irrelevant) contextual factors into account. But frictions do exist, and so lawmakers face various trade-offs when determining the context of a law.

These trade-offs have been well rehearsed.[2] On the one hand, it is costly to add more context to law. Lawmakers must discover

---

† Professor of Law, The University of Chicago Law School. I am grateful for research support from the Richard Weil Faculty Research Fund and the Paul H. Leffman Fund. Stephanie Xiao and Courtney Block provided excellent research assistance.

†† Associate Professor and Canada Research Chair in Law, Economics, & Innovation, University of Toronto Faculty of Law. In the interests of full disclosure, Professor Niblett is also a cofounder of Blue J Legal, a start-up bringing machine learning to the law.

1 See Frederick Schauer, *Profiles, Probabilities, and Stereotypes* 27–54 (Harvard 2003) (discussing how Greek philosophers debated the appropriate breadth of the law and questioning how individualized the law should be).

2 For an economics perspective, see generally Louis Kaplow, *Rules versus Standards: An Economic Analysis*, 42 Duke L J 557 (1992). See also Anthony J. Casey and Anthony Niblett, *The Death of Rules and Standards*, 92 Ind L J 1401, 1402 n 2 (2017)

the relevant factors and then communicate those factors to the citizens. And there is always the risk of error. On the other hand, laws with less context can be crude and rigid. They will be over- or underinclusive.[3]

Often, the debate about adding context turns on timing. If the content of the law remains vague ex ante, a judge can fill in the contextual details ex post. This is how the law of negligence works. The judge contextualizes the rules after the alleged tort has occurred. Doing so reduces some of the costs of contextualization. An ex post adjudicator has the benefit of hindsight in determining which factors are relevant in a particular context. Moreover, the adjudicator needs to add content only for situations that actually did arise rather than those that might have arisen. Still, ex post adjudication imposes its own costs. For example, citizens live in uncertainty because no one has communicated the specific content of the law to them, and the ex post adjudicator may infect the process with noise, inconsistency, hindsight bias, or her own idiosyncratic views on what the law's objectives should be.[4] Again, this timing question has been explored for decades.[5]

(collecting sources that discuss the distinction between rules and standards from an economics perspective).

3 See Frederick Schauer, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* 31–34 (Oxford 1991). Indeed, the literature on the costs and benefits of using rules and standards is vast. See, for example, Kaplow, 42 Duke L J at 559–60 & n 1 (cited in note 2) (explaining the distinction between rules and standards and collecting sources); Frederick Schauer, *The Tyranny of Choice and the Rulification of Standards*, 14 J Contemp Legal Issues 803, 803 n 1 (2005) (collecting sources); Cass R. Sunstein, *Problems with Rules*, 83 Cal L Rev 953, 969–96 (1995) (describing the strengths and limitations of rules). See also generally Kathleen M. Sullivan, *The Supreme Court 1991 Term—Foreword: The Justices of Rules and Standards*, 106 Harv L Rev 22 (1992); Joseph Raz, *Practical Reason and Norms* (Princeton 1990); Isaac Ehrlich and Richard A. Posner, *An Economic Analysis of Legal Rulemaking*, 3 J Legal Stud 257 (1974).

4 See generally, for example, Raz, *Practical Reason* (cited in note 3). For further discussion of inconsistency, see generally Anthony Niblett, *Tracking Inconsistent Judicial Behavior*, 34 Intl Rev L & Econ 9 (2013). On hindsight bias, see generally Jeffrey J. Rachlinski, *A Positive Psychological Theory of Judging in Hindsight*, 65 U Chi L Rev 571 (1998). See also Christine Jolls, Cass R. Sunstein, and Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 Stan L Rev 1471, 1523–27 (1998). For further discussion on noise, see Daniel Kahneman, et al, *Noise: How to Overcome the High, Hidden Cost of Inconsistent Decision Making*, 94 Harv Bus Rev 38, 40–42 (Oct 2016). For further discussion of idiosyncrasy, see Robert C. Farrell, *Justice Kennedy's Idiosyncratic Understanding of Equal Protection and Due Process, and Its Costs*, 32 Quinnipiac L Rev 439, 502 (2014).

5 The question of timing of contextualization is often framed as one of rules (providing ex ante context) and standards (providing ex post context). See, for example, Kaplow, 42 Duke L J at 581–82 (cited in note 2). But not everyone agrees that the choice between rules and standards turns on the timing in this way. Indeed, as we have noted elsewhere,

The question for this Symposium then becomes: What exactly is *new* about personalized law? After all, a personalized law is just a law that is more contextualized or tailored to the relevant factors facing an individual. Again, the law of negligence in its theoretical form demonstrates the point. As usually stated, the law of negligence incorporates all relevant factors facing an individual. It asks how a reasonable person would have acted facing the *same situation*.[6] The personalized context comes in through the definition of the relevant "same" situation. When all relevant factors are taken into account—which the conventional view of tort law would seem to require—the law as applied by the judge is fully personalized.

But as Professors Omri Ben-Shahar and Ariel Porat have pointed out, that is not how the law of negligence works in the real world.[7] While the law has forever aspired to a high level of personalization, it has persistently remained quite far from that goal.[8] Things appear, however, to be changing. As technologies associated with big data, prediction algorithms, and instantaneous communication reduce the costs of discovering and communicating the relevant personal context for a law to achieve its purpose, the goal of a well-tailored, accurate, and highly contextualized law is becoming more achievable. And that is the "new" personalization of law that we explore in this Symposium.

As a starting point for this inquiry, this Essay proposes a general framework for thinking about the new personalization of law. Without such a framework, projects to personalize law run the risk of being ad hoc and unconnected. To avoid this, we identify two fundamental questions that lie at the heart of any move toward personalization through data and analytics.

there is wide disagreement in legal scholarship about what the words "rules" and "standards" even mean and what questions are implicated in the choice between the two forms of law. See Casey and Niblett, 92 Ind L J at 1405 n 9 (cited in note 2); Anthony J. Casey, *The Short Happy Life of Rules and Standards* 3:00 (Feb 21, 2017), online at http://www.youtube.com/watch?v=TnbRApMEumU (visited Aug 27, 2018) (Perma archive unavailable). This is striking given the ubiquity of the concept in legal scholarship.

6 See *Brown v Kendall*, 60 Mass 292, 296 (1850); Alan D. Miller and Ronen Perry, *The Reasonable Person,* 87 NYU L Rev 323, 329 (2012).

7 Omri Ben-Shahar and Ariel Porat, *Personalizing Negligence Law*, 91 NYU L Rev 627, 636–46 (2016).

8 Id at 637–46.

The first question is how to set law's objective. As we have suggested in prior work,[9] advances in data processing and communications technologies create the potential for law to migrate from the traditional forms of rules and standards to microdirectives that update and change in real time with the changing personal circumstances of the regulated citizen. But for this personalization to work, lawmakers must know and precisely state the objective of law up front in a way that has never before been required. Indeed, the entire purpose of personalization is to fit legal outcomes to relevant contextual factors. Big data can facilitate that fit through greater accuracy in determining the relevant factors for a law's application. But the relevance of any given factor is ascertainable only by reference to the objective the law seeks to achieve. And personalization technology, for all its promise, cannot provide that objective. Moreover, the use of big data and machine-driven analytics requires a statement of the objective in the most precise and accurate form as inaccuracy and imprecision can lead to perverse outcomes.[10]

The second question is how to communicate a law's content to citizens. The new personalization of law will lead to highly specific and complicated laws that must be translated and communicated to citizens in a form and at a time that makes compliance possible. Along these lines, technology will allow for the personalization not just of the substance of law but also of the means and timing by which a law's directive is provided to a citizen.

Thus, in Part I of this Essay, we explore the importance and challenges inherent in (1) identifying the objective of a personalized law and (2) deciding how and when to apply and communicate the content of a personalized law to citizens.

In Part II, we identify several specific logistical challenges that manifest in implementing these two fundamentals: source and quality of data, discrimination and bias, human intervention, transparency of data, and regulation of the providers of personalized law.[11]

---

9 See Casey and Niblett, 92 Ind L J at 1410–12 (cited in note 2); Anthony J. Casey and Anthony Niblett, *Self-Driving Laws*, 66 U Toronto L J 429, 431 (2016).

10 For discussion of perverse outcomes, see Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* 120–22 (Oxford 2014).

11 We have addressed this final question about regulating the providers of such laws in the public law context in *The Death of Rules and Standards* and in the private law context in *Self-Driving Contracts*. See Casey and Niblett, 92 Ind L J at 1417–23 (cited in note 2); Anthony J. Casey and Anthony Niblett, *Self-Driving Contracts*, 43 J Corp L 1,

I. TWO FUNDAMENTAL QUESTIONS

As a preliminary note, we observe that law collides with personalization in two prominent ways. First, many legal researchers have explored the issue of when and how the state should best regulate the way private actors use algorithms to personalize products. Some essays in this Symposium address this question by looking at how well equipped the law is for dealing with this type of personalization.[12] Our focus, though, has primarily been on the second issue: How can law be better personalized through the use of algorithms? That is, how should the state use big data and algorithms to personalize the production of law? These two research agendas are not mutually exclusive. Indeed, many of the issues raised—for example, data sources, privacy, discrimination, and bias—are common to both strands of the literature. And as the lines between private ordering and public law become blurred, so too will the lines between these inquiries.

The new personalization of law will pose two fundamental challenges. First, it requires lawmakers to be explicit about the objective that a law seeks to achieve. When laws are not highly personalized, lawmakers can and often do punt on this issue. They leave it to citizens and judges to discover or invent their own views about the purpose of a law. Big data and algorithms that translate large amounts of information into specific legal directives do not, however, permit such ambivalence or hand-waving. Instead, they require (and facilitate) an up-front and clear statement of an objective. Second, the new personalization poses new questions about the methods and timing of the application of a

26–31 (2017). Additionally, Professor Gillian Hadfield provides extensive analysis of related issues of how to regulate private providers of law in a world of increasing information technology. See Gillian K. Hadfield, *Rules for a Flat World: Why Humans Invented Law and How to Reinvent It for a Complex Global Economy* 249–59 (Oxford 2016). This Symposium also includes a deep dive into the question by Professor Andrew Verstein. See generally Andrew Verstein, *Privatizing Personalized Law*, 86 U Chi L Rev 551 (2019). As a result, we will do little in this Essay other than flag the issue as one major concern in virtually every personalization project.

12 See generally Oren Bar-Gill, *Algorithmic Price Discrimination When Demand Is a Function of Both Preferences and (Mis)perceptions*, 86 U Chi L Rev 217 (2019) (examining how to respond to data-driven price discrimination); Talia B. Gillis and Jann L. Spiess, *Big Data and Discrimination*, 86 U Chi L Rev 459 (2019) (examining possible legal responses to automated credit pricing); Gerhard Wagner and Horst Eidenmüller, *Down by Algorithms? Siphoning Rents, Exploiting Biases, and Shaping Preferences: Regulating the Dark Side of Personalized Transactions*, 86 U Chi L Rev 581 (2019) (examining possible negative effects of personalization on business to consumer transactions).

law and communication of its directive to citizens. This Part explores these two challenges.

In this Essay, for expository purposes, we use the extreme form of personalization—the microdirective—to demonstrate the concerns and questions raised by personalization. One can think of the microdirective as the idealized version of a personalized law.

With a microdirective, lawmakers create a law that is nothing more than a general objective. It looks like a standard. But a microdirective also provides for an algorithm to use big data to transform that objective into a specific rule-like and personalized directive that is communicated to the citizen when the citizen needs to know the content of the law. Early forms of microdirectives already exist. For example, smart traffic lights collect data inputs to personalize the directives provided to drivers at an intersection. One can think of a yield sign as a standard, a stop sign as a clunky rule, and traffic lights and smart traffic lights as progressions toward the more personalized microdirective.

A. Objectives

The use of big data and algorithms to provide the contours of law will force lawmakers to address the question of what, precisely, is the objective of a particular law. To what end are we creating this highly contextual and personalized law? Personalizing a law means taking into account the personal factors of the individual to whom it applies. But to take a factor into account, one must ask: For what purpose? A personal factor relevant in answering one legal question will be irrelevant in answering another. Another way to think about it is that personalization makes a law more accurate (less error prone).[13] But how does one define accuracy and error? Again, the answer comes from understanding the objectives of the law. An error is an application of the law that does not achieve its objective. To know whether an error has occurred, one must understand that objective.

Conventional law—with its vagaries and lack of personalization—often allows lawmakers to avoid this question. A speed limit may be imposed to achieve one of many goals, which might include reducing accidents, promoting efficient transportation, or

[13] See Ariel Porat and Lior Jacob Strahilevitz, *Personalizing Default Rules and Disclosure with Big Data*, 112 Mich L Rev 1417, 1458 (2014) (noting that the value of algorithms using big data to personalize default rules lies in the ability to make the law more accurate).

reducing pollution. But the lawmakers need not make their goal clear. And they may not even need to have a goal. They can use a crude rule (55 miles per hour) with no personalization and leave the objective unstated. Or they can use a standard (drive reasonably) that leaves it to the driver and the judge to figure out the objective.

Predictive algorithms and big data do not work that way.[14] The lawmaker has to tell the algorithm what to do with the data. She must specify an ex ante objective.[15] Such specification is not unique to the new personalization of law. All lawmakers and judges make implicit judgments about the objective of a law when they announce the content of that law. But the problem becomes crucial when using big data and automated technology to achieve a purpose. In the context of algorithms, the objective will be fixed once the program is initiated and must be stated with precision if it is to be translated into code. Thus, the new personalization of law forces two things with regard to objective setting: clarity and forethought.[16] The objective must be known and programmed *with clarity*, and this must be done at the time of the algorithm's creation.

There are different modes through which law can set an algorithm's objective. We will explore two to demonstrate the importance and challenges of objective setting, and then we will examine one way that lawmakers can use data to avoid explicitly setting an ex ante objective.

1. Algorithmic legislation.

First, algorithms could improve ex ante legislation.[17] This is the clearest avenue to add accuracy through personalization. Ex

---

14 See Jerry Kaplan, *Artificial Intelligence: What Everyone Needs to Know* 94–95 (Oxford 2016). See also Harry Surden, *Machine Learning and Law*, 89 Wash L Rev 87, 102–10 (2014).

15 As Professor Solon Barocas and Andrew Selbst put it, when using data to find relevant correlations, there is a first step of "problem specification": "translat[ing] some amorphous problem into a question that can be expressed in more formal terms that computers can parse." Solon Barocas and Andrew D. Selbst, *Big Data's Disparate Impact*, 104 Cal L Rev 671, 678 (2016).

16 See Michael Luca, Jon Kleinberg, and Sendhil Mullainathan, *Algorithms Need Managers, Too: Know How to Get the Most out of Your Predictive Tools*, 94 Harv Bus Rev 96, 99 (Jan–Feb 2016) (explaining the importance of having explicit, defined, and quantifiable goals for algorithms).

17 An alternative but related version of this is personalization through algorithmic administrative regulation. We have argued elsewhere that that avenue is the most likely

ante laws are often the least personalized and do not do a good job of achieving their objective. There are high ex ante costs of discovering and articulating all relevant factors.[18] As big data reduces those information costs, personalization can make law more accurate. Speed limits provide a canonical example. Consider the errors created by having one-size-fits-all speed limits. Lawmakers might use big data to create personalized speed limits (communicated to the driver's dashboard) to minimize car accidents. But minimizing car accidents is not the only objective involved in setting a speed limit. An algorithm programmed merely to minimize accidents would simply set the speed limit at zero.[19] One has to know and articulate all of the competing objectives of a personalized law and their relationship to each other to provide a means of balancing potential error reductions.

In the speed limit example[20]—which will become highly important in the regulation of the software behind self-driving cars—the lawmaker must make an ex ante judgment about the precise balance between speed of travel, the risk of accidents, pollution, the consumption of fuel, and so on.[21] Indeed, she must

---

public source of personalization. Casey and Niblett, 92 Ind L J at 1418 (cited in note 2). For the analysis here, the concerns are the same.

18 Examples of inaccurate impersonal rules abound. Indeed, in this Symposium alone, several are discussed. See Omri Ben-Shahar and Ariel Porat, *Personalizing Mandatory Rules in Contract Law*, 86 U Chi L Rev 255, 262–63 (2019) (discussing an example of inefficient mandatory contract rules that ignore relevant personal characteristics); Matthew B. Kugler and Lior Jacob Strahilevitz, *Assessing the Empirical Upside of Personalized Criminal Procedure*, 86 U Chi L Rev 489, 494–95 (2019) (discussing criminal procedure rules that may ignore personal expectations of privacy); Christoph Busch, *Implementing Personalized Law: Personalized Disclosures in Consumer Law and Data Privacy Law*, 86 U Chi L Rev 309, 314–24 (2019) (discussing consumer and privacy laws that ignore consumer heterogeneity); Adi Libson and Gideon Parchomovsky, *Toward the Personalization of Copyright Law*, 86 U Chi L Rev 527, 528 (2019) (discussing copyright laws that ignore the value of content to users).

19 See Ehrlich and Posner, 3 J Legal Stud at 260 (cited in note 3).

20 We use the speed limit example because the availability of the relevant data is obvious. There is much available data on driving; thus, it is an area in which we should expect some of the quickest moves toward personalization. This example demonstrates one key point about personalization: the supply of relevant data will be a major factor in locating the emergence of personalized law.

21 See generally Oliver Moore, *Toronto to Use Big Data to Help Reduce Traffic Congestion* (The Globe and Mail, Apr 7, 2015), archived at http://perma.cc/8R2E-AERU. See also generally Yuanfang Chen, et al, *When Traffic Flow Prediction Meets Wireless Big Data Analytics* (arXiv.org, Sept 23, 2017), archived at http://perma.cc/CWU4-CZYG. Beyond personalizing the speed itself, you might even personalize the fines for violations based on concepts of deterrence and social equality. See, for example, Alec Schierenbeck, *The Constitutionality of Income-Based Fines*, 85 U Chi L Rev 1869, 1876–79 (2018). In this way, personalized law could be a set of personalized fines or prices presented to citizens. There

insert specific weights for these measurable values. Without personalization, lawmakers can obscure those value judgments. But when one tries to use data about personalized factors to predict outcomes and set a rule, one needs to know and declare the specific outcome desired.

2. Algorithmic judging.

Second, algorithms could assist (or perhaps even replace) judges to improve the accuracy of decisions.[22] Arguably, conventional judicial application of law is more personalized than conventional legislative statements of law.[23] Judges can take more factors into account because they are applying the law after all the evidence is in. But there are costs to this. Judges may make mistakes in determining which factors are relevant, or they may—intentionally or not—seek to achieve the wrong objective in determining the relevance of various factors. The result is that a judge may deliver decisions that do not align with the objective of the law. That is problematic for two reasons: (1) there are errors in the fit of the application of the law, and (2) those errors create inconsistency and variance that make it harder for citizens to comply with the law.

The new personalization of law can reduce both of these errors. A recent study using big data and machine learning technologies in the context of granting and denying bail provides a telling example of how data can improve the personalization of law by

is much efficiency to be gained by such price discrimination. But as Professor Oren Bar-Gill points out, algorithmic price discrimination can also be used to take advantage of the one paying the price. Bar-Gill, 86 U Chi L Rev at 230–31 (cited in note 12). See also Wagner and Eidenmüller, 86 U Chi L Rev at 585–86 (cited in note 12) (describing "first degree price discrimination," under which individual consumers are charged different prices). The fact that the price is coming from the lawmaker may reduce our worry about advantage-taking, but that certainly will not always be true.

[22] Again, there is a close alternative option: personalization through algorithmic enforcement, in which enforcement officers or agencies use big data to personalize the application of laws and regulation in each case. See Benjamin Alarie, Anthony Niblett, and Albert H. Yoon, *Regulation by Machine* *2–3 (Conference on Neural Information Processing Systems, Dec 2, 2016), archived at http://perma.cc/L9LJ-T3RB.

[23] This personalization is less costly with ex post adjudication, as the adjudicator has to figure out only the context-specific applications for cases that actually arise, whereas an ex ante rule has to address all possible applications. See John O. McGinnis and Steven Wasick, *Law's Algorithm*, 66 Fla L Rev 991, 1030 (2014); Sunstein, 83 Cal L Rev at 1003–04 (cited in note 3); Kaplow, 42 Duke L J at 582 (cited in note 2).

judges.[24] In most jurisdictions, the stated objective of a bail decision requires a judge to balance the risk that the defendant will flee or commit another crime against the costs and burdens of incarceration.[25] Judges are not given specific directives, and so it is up to them to personalize the law when applying it to a specific case.

In the study, machine learning techniques were shown to provide more accurate assessments of risk, which, if used by the lawmaker, could allow for a reduction in both the detention rate of defendants and the rate of crimes committed by those who were released.[26] This suggests that the algorithm does a better job at eliminating irrelevant factors and at assessing those factors that are relevant to achieving the law's objectives.

The authors of the study do note one difficulty in their study: judges may be maximizing other objectives or preferences.[27] While judges may inject their own preferences when determining an objective of a law, algorithms will not inject new objectives. This raises an important concern. To the extent that lawmakers value their ability to delegate objective setting to judges, then algorithmic personalization would not be preferred unless the law delegates to the judges the power to create the algorithm as well.

It is worth noting that allowing an individual judge to set the objectives of law is highly problematic. It reduces democratic accountability and consistency in the law. Judges might introduce bad objectives without disclosing them, or they may simply fail to achieve the good objectives they set. Again, the bail study is illustrative. The study suggests that the algorithmic bail program can reduce crime and detention rates *while also* "*reducing racial disparities*."[28] This suggests that judges either were intentionally imposing discriminatory purpose on the law or, more likely, were taking into account discriminatory factors that were not relevant to the putative objective of the law.[29]

---

24 Jon Kleinberg, et al, *Human Decisions and Machine Predictions,* 133 Q J Econ 237, 240–45 (2018).

25 See, for example, the standard in Massachusetts, where bail is determined by examining the alleged crime, the likely penalty, the likely flight risk, history of defaults, family in the area, employment status, and previous criminal records, among other criteria. Mass Ann Laws ch 276, § 57. The Kleinberg study looks primarily at data from New York. Kleinberg, et al, 133 Q J Econ at 246 (cited in note 24).

26 See Kleinberg, et al, 133 Q J Econ at 285–86 (cited in note 24).

27 Id at 243 ("[D]ecisions that appear bad may simply reflect different goals.").

28 Id at 241 (emphasis added).

29 More on this below in Part II.B.

What makes personalization in the bail context work is that a measurable objective can be clearly stated.[30] The takeaway is that algorithmic personalization can improve the fit of personalized law and reduce errors if the data are available and lawmakers can agree on the definition of an error. The corollary is that personalizing some areas of law, through legislative or judicial algorithms, may be problematic because lawmakers or the general polity do not agree on the objective. Sentencing of criminal defendants provides a salient example. Competing objectives like retribution and deterrence will lead to different outcomes through personalization. The factors that are relevant and the data that are necessary will look very different based on which objectives the lawmakers choose. With conventional lawmaking, we tolerate (to some extent) a human judge, legislator, or police officer injecting her version of the law's objective after the fact. Algorithmic personalization does not tolerate such a wait-and-see approach. This is likely to be the biggest impediment to personalization. In areas of law for which no ex ante consensus exists as to the purpose of a law, conventional lawmaking may endure. On the other hand, one might question the legitimacy of a law whose purpose cannot be identified.[31]

3. Using judicial data.

There is a shortcut that may allow a lawmaker to avoid setting an objective when developing ex ante microdirectives. In some areas of law, lawmakers could use big data to predict what judges would do based on past decisions and implement those predictions rather than achieve a stated ex ante objective.[32] This

---

[30] This is not to take a position on whether this is the right objective. The point, however, is that, if one has a clearly stated objective function that values reducing crime, detention rates, and racial disparities, there is evidence that an algorithm can "improve" personalization by improving judicial outcomes on every dimension even though the algorithm was developed with only the first two objectives in mind.

[31] One might also argue that allowing human variation with regard to the purpose of a law is a feature rather than a bug because it provides for experimentation and the evolution of the law. Algorithms could, of course, be programmed to intentionally introduce arbitrary ex post variation. But that would have to be programmed as part of the ex ante objective. And it is likely that citizens will be more squeamish about arbitrary variation when it is intentional and automated than when it is dressed up in the "reasoning" of a judicial opinion.

[32] Researchers have illustrated the power of machine learning tools to predict outcomes of cases decided by courts. See, for example, Benjamin Alarie, Anthony Niblett, and Albert H. Yoon, *Using Machine Learning to Predict Outcomes in Tax Law*, 58 Can Bus L J 231, 235–36 (2016); Daniel Martin Katz, *Quantitative Legal Prediction—or—How I*

method has the result of incorporating the collective objective function of the population of past judges. To be sure, this does not add any new personalization to the law. Instead, this method simply entrenches the existing personalization that exists in judicial applications of law, but with the potential added benefit of consistency.

In areas that have generated a wealth of litigation, algorithms can be used to map judicial behavior and predict—or replicate—how judges in similar cases would have decided these cases. When using existing case data as the basis for microdirectives, the objective of the algorithm is to predict the outcome that would be reached by judges who decided prior cases.[33] The algorithm seeks to find hidden patterns in the data and weigh factors the way judges would.

The objective of such an algorithm is not to add personalization to the law but to improve it, primarily by reducing the inconsistency of judicial decisions and minimizing the likelihood of outlier decisions while still allowing for the law to take into account the personal factors of individuals. Thus, using the algorithm as the basis for the law provides citizens with greater certainty and consistency. An ancillary benefit that emerges from the use of algorithms is the reduction in ex post administration costs. There are fewer cases litigated in a world in which outcomes can be predicted.[34]

This mechanism of algorithmic judging will be particularly attractive when lawmakers cannot agree on a precise objective of a law but do agree that the approaches taken by judges are

---

Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry, 62 Emory L J 909, 936–42 (2013). See also Daniel Martin Katz, Michael J. Bommarito II, and Josh Blackman, *A General Approach for Predicting the Behavior of the Supreme Court of the United States*, 12 PLOS ONE 1, 7–15 (Apr 2017); John O. McGinnis and Russell G. Pearce, *The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services*, 82 Fordham L Rev 3041, 3046–53 (2014).

33 There are complicated questions of what this would mean for implementation. For example, if an algorithm suggests that 60 percent of judges would classify a defendant as liable, what should the algorithm do? If it simply says defendant is liable, that changes the probabilistic outcome a litigant faces. This changes settlement and deterrence calculations. An alternative approach would be to impose damages at 60 percent. But that would reflect a dramatic shift in legal doctrine. These are some of the difficult technical implementation questions posed by the use of algorithms in law.

34 See Part I.B for a discussion of how improvements in communication technology will allow predicted outcomes to be communicated to litigants.

generally satisfactory and that consistency is important. For example, it may be difficult to identify the exact policy considerations behind the distinctions between employees and independent contractors. What objectives should underpin the tax law or employment law on this question? How much weight should lawmakers put on the various considerations of revenue maximization, unemployment reduction, and equity concerns? These weights may be difficult to specify ex ante. It may be much easier, however, to simply observe the ways in which these distinctions manifest themselves in the case law when judges decide actual cases.

But there are, of course, important questions and concerns that will arise when replicating existing case law in algorithmic form. First, this method of personalization is not suitable when the underlying objective of the law shifts over time. Second, and relatedly, even though the biases of individual outlier judges will diminish, the new law will replicate any systematic biases that are baked into existing case law. Third, if we replace judicial decisions with algorithmic decisions, then there will be fewer (or perhaps no) new cases. This means that the system of law will not learn or evolve. Fourth, there may not be sufficient data to capture every future contingency. Ex ante personalization of this type relies on having enough judicial decisions to predict what will happen with other, hypothetical cases. Areas of law that are rarely litigated are not suitable candidates for this type of algorithmic judging.

B. Timing the Application and Communication of Law

The previous Section deals with the *content* of law. This Section deals with the *timing* of law. There are two questions that big data personalization raises about timing: (1) When and how is the legal directive communicated to the relevant citizen? (2) And when does the government commit itself to a specific application of the law?

Big data combined with advances in communication technology opens up new possibilities for timing by reducing the personalization trade-off that lawmakers otherwise face. Communication technology adds two innovations: (1) it enhances options about *when* to reveal that law's content and directive to a citizen, and (2) it simplifies the form in which that content and directive is revealed.

To see why these are important, consider the conventional views of the personalization trade-off. Traditionally, law is personalized through ex post adjudication of standards. Judges with the benefit of hindsight look at all the relevant factors and then personalize the law. Thus, more personalization comes at the expense of waiting.[35] This ex post method imposes two important costs: uncertainty[36] and the potential for government misbehavior.[37]

The uncertainty arises because citizens do not know how the law will apply to their specific situation. Judges are unpredictable, make mistakes, and may consider factors that do not seem relevant to the citizen. The risk of government (or judicial) misbehavior arises because the government (through judges or some other adjudicator or enforcer) can decide the content of the law after the individual has already acted. This discretion can be abused.[38] The government—because it is not precommitted to the law's outcome—could use its after-the-fact discretion to punish disfavored citizens,[39] to impose outcome or distributive preferences that are not part of the law's intended objective, or to change the objective based on hindsight bias.[40]

Personalization through big data, rather than through ex post adjudication, can reduce these problems. We start with uncertainty. When law is personalized through ex post adjudication, the communication of the law's effect on that action is delayed until after the action occurs.[41] With big data and advances in communication, the effect can be communicated to the citizen as soon as the relevant evidence about the citizen's personal characteristics and situation is available.[42] Thus, in some (but not all) cases,

---

35 See Eric A. Posner, *Standards, Rules, and Social Norms*, 21 Harv J L & Pub Pol 101, 101–03 (1997); Kaplow, 42 Duke L J at 585–86 (cited in note 2).

36 See Sunstein, 83 Cal L Rev at 974–77 (cited in note 3); Kaplow, 42 Duke L J at 569, 575 n 42, 587–88 (cited in note 2); Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 89 Harv L Rev 1685, 1689–1701 (1976). See also Richard Craswell and John E. Calfee, *Deterrence and Uncertain Legal Standards*, 2 J L Econ & Organization 279, 285–88 (1986).

37 See Saul Levmore, *Double Blind Lawmaking and Other Comments on Formalism in the Tax Law*, 66 U Chi L Rev 915, 919 (1999); Posner, 21 Harv J L & Pub Pol at 113 (cited in note 35); Sunstein, 83 Cal L Rev at 974–76 (cited in note 3).

38 See Levmore, 66 U Chi L Rev at 919 (cited in note 37).

39 See Posner, 21 Harv J L & Pub Pol at 113 (cited in note 35).

40 See Antonin Scalia, *The Rule of Law as a Law of Rules*, 56 U Chi L Rev 1175, 1179 (1989).

41 See Posner, 21 Harv J L & Pub Pol at 101–03 (cited in note 35); Kaplow, 42 Duke L J at 585–86 (cited in note 2).

42 The timing of content creation also changes when law is personalized. The more personalized a law is, the more likely it is that relevant factors will come into existence

the law's personalized application to a case can be communicated before the citizen has to act.

Importantly, technology can also transform the personalized law into a simplified form of communication. For an example, consider the smart traffic light, which uses data inputs to personalize traffic directives. In raw form, the personalized microdirective would be incomprehensible to the driver. It would contain a whole catalog of contingent directives that turn on traffic patterns, weather, time of day, proximity of other vehicles, and so on. The technologies at work translate that information into the simple form of a green or red light that is communicated to the driver.[43]

Now consider the government misbehavior problem. Even though an algorithmic microdirective could be programmed to take into account factors that are not known when the law is enacted, the government could precommit to the law's objective. Indeed, lawmakers could rely on an algorithm that prevents human intervention after the citizen has acted but still takes all relevant personal context into account. In this way, the government would bind itself against abusing ex post discretion even when the law is personalized and takes into account factors that become known only after the law is promulgated.[44] Thus, the use of an automated personalization algorithm can prevent the lawmaker from changing the law's content or directive based on illegitimate ex post factors.[45]

With traditional law, however, there is an additional cost to government precommitment. In some cases, precommitment facilitates evasion. In many regulatory areas, there is a strategic

only later in time, right before the moment that the law becomes applicable. But it still could become available before the action is required. On the topic of how advance rulings promote certainty, see Yehonatan Givati, *Resolving Legal Uncertainty: The Unfulfilled Promise of Advance Tax Rulings*, 29 Va Tax Rev 137, 144–49 (2009). See also Carlo Romano, *Advance Tax Rulings and Principles of Law: Towards a European Tax Rulings System?* 320–32 (IBFD 2002).

43 This demonstrates another key point about the new personalization of law. It is most likely to occur first in areas of law in which complex information about contextual factors can be easily gathered, processed in the relevant time frame, and then transformed into a simple communication to the citizen. Thus, we should expect to see personalization thrive in fields like the regulation of food and drug safety, securities law, tax law, workplace safety regulation, consumer law, and police accountability.

44 For a more detailed exploration of this idea, see Casey, *Short Happy Life* at 30:07 (cited in note 5).

45 In some sense, this is like a governmental version of Professor Lee Fennell's analysis of individual precommitment. See generally Lee Anne Fennell, *Personalizing Precommitment*, 86 U Chi L Rev 433 (2019). Just as individuals may desire to bind themselves in the future, governments may want to do the same thing.

game between the lawmakers and the regulated citizens. Regulating evasion in tax law provides a salient example. If you reveal the content of the law up front, the taxpayer can structure her behavior to evade the spirit of the law.[46] On the flip side, if you have a standard like "behave reasonably," the government can change the rules of the game after the fact and take advantage of those who acted reasonably in good faith.[47]

This problem can potentially be solved by delaying the revelation of the law's directive until after the citizen has committed to action. This idea of delayed content revelation[48] provides a partial solution to the evasion-commitment problem. The government commits to the law ex ante (so it functions like a rule). But the taxpayer does not know the content and thus must comply with the law as if it were a standard. This prevents the search for technicalities and loopholes that are invited by a preannounced rule. And the law is still personalized in a way that traditional precommitment would not have allowed. That said, uncertainty for the taxpayer remains. And so the use of delayed revelation will be optimal for areas of personalized law in which uncertainty is less of a concern than evasion.

All of this is to say that the timing of the communication and application of law can be tailored to optimally fit the context of the particular law in a particular situation. The law could communicate the personalized directive as soon as it is known, at the moment it becomes relevant to the citizen, whenever the citizen requests it, or even after the citizen has acted.[49] And it can take various forms (simple or complex).

This content-revelation question will be important for all of personalized law. Citizens will sometimes have an incentive to alter their personal characteristics to take advantage of the personalization. In some cases, this will be desirable. Often citizens desire to follow the spirit of the law. And real-time information about how to comply helps with that. Well-designed traffic lights

---

[46] See David A. Weisbach, *Formalism in the Tax Law*, 66 U Chi L Rev 860, 869–72 (1999).

[47] See id at 860; Levmore, 66 U Chi L Rev at 920 (cited in note 37); Scalia, 56 U Chi L Rev at 1180–82 (cited in note 40).

[48] Professor Saul Levmore explored the idea of delayed revelation as a means of combatting evasion in 1999. Levmore, 66 U Chi L Rev at 919–20 (cited in note 37).

[49] One additional limitation on timing is the ability of the algorithm to gather evidence about personal characteristics of the citizen in question. This suggests that personalized law, at least when ex ante revelation is important, will be successful in areas in which evidence about the relevant characteristics and factors is readily available.

have this characteristic. Most drivers want to obey traffic lights that prevent accidents. One can even envision automated traffic rules that provide rewards in the form of open lanes or green lights for drivers who opt into safer behavior or other favored characteristics.[50] Other areas of law, such as tax, are different because citizens have a greater incentive to try to evade the spirit of the law.[51]

The important takeaway here is that, in addition to setting clear objectives, lawmakers must also understand for each area of law the factors that determine the optimal timing and form of the application and communication of personalized law.

II. Specific Challenges

In this Part, we discuss five additional challenges, each of which relates to the fundamental questions we explore above: (1) source and quality of data, (2) discrimination and bias, (3) human intervention, (4) transparency of data, and (5) regulation of the providers of personalized law. The tasks of choosing data sources and preventing bias and discrimination are essentially projects in identifying the correct objective of personalized law and considering the costs of obtaining data relevant to that objective. Similarly, the issues of human intervention and transparency pose questions about how to identify and deal with errors in objective setting and how to use timing to precommit to a legal directive. And the choice and regulation of the provider of personalized law turn on questions about setting the objective and communicating law's content and directives to citizens.

A. Data

The question of whether an algorithm can achieve the objective of the law turns on the *quality* of the data that a lawmaker relies on. The concern here is whether the data actually measure the relevant factors and adequately predict the central objective of the law. This raises obvious procedural issues. How much relevant data are available? Who collected the data—are they impartial, and what biases do they bring? How and why did they collect

---

50 We already see early forms of this in high-occupancy-vehicle lanes.

51 Judicial replication is never the appropriate mechanism when evasion of this type is a concern.

the data? What potential biases are lurking in the data, and can those biases be corrected?[52]

The objective and the form of data are intrinsically linked. If measuring expectations or social objectives, one might turn to survey data to try to extract information. If replicating decisions of human decisionmakers, one might look to data sources of case law or regulatory decisions. Indeed, one might think of judicial decisions as surveys of past decisionmakers from which we extract broader principles. If seeking to minimize consequential errors, the consequences in the objective have to be measurable. In the bail example,[53] if the objective requires an algorithm to assess the risk of the accused fleeing or committing a crime, one would need data that accurately describe the behavior of accused persons who have been granted bail in the past. But observational data may reflect biases, not only of those who collect the data but also of those actors who influence particular variables of interest.[54] For example, data on whether accused persons commit a crime while out on bail may reflect the biases and resources of those who police and prosecute crimes.

Fundamentally, the data have to be well tailored to the objective. Researchers must account for potential asymmetries in the measurement of errors. The bail example again proves helpful. Here, the researchers are able to measure two types of error. First, they observe whether errors were made in *granting* bail by measuring consequences.[55] That is, some accused persons are granted bail and then commit crimes or flee. But to measure the other type of error—the error of *denying* bail to less risky individuals—the researchers exploit variation in the way that judges decide cases.[56] Some judges are more lenient than others. Only through this judicial variation could the researchers observe the error of denying bail. Creators of algorithms have to be conscious

---

52 See generally Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Crown 2016) (exploring biases inherent in big data in a number of contexts, such as employment, insurance, and criminal law).

53 See notes 24–26 and accompanying text.

54 See, for example, Danielle Keats Citron and Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 Wash L Rev 1, 13–14 (2014) ("The biases and values of system developers and software programmers are embedded into each and every step of [the] development [of credit-scoring software]."); Danielle Keats Citron, *Technological Due Process*, 85 Wash U L Rev 1249, 1262 (2008) (describing the possibility that programmers' individual biases, such as a preference for binary questions, could be embedded into the algorithms they create).

55 See Kleinberg, et al, 133 Q J Econ at 247 (cited in note 24).

56 See id at 261–62.

of this asymmetry. A law personalized through algorithms may eliminate variation. As a result, new information about denying bail would be unavailable. Resolving this problem may require artificial randomization in order for the evolutionary algorithm to "learn" about the different types of errors.[57]

Second, the need for expansive data raises privacy questions.[58] The data collected may infringe on the privacy expectations of individuals who may not know where and how their information is being used. This is frequently raised as a concern in the regulation of private companies using big data to personalize products.[59] But it is, of course, an issue for any public law that requires the government to have access to personal data.[60] This raises additional questions about the need for—and limits of—consent by citizens, whether the government should compensate citizens for the use of their data,[61] and who should collect the data on behalf of the government.[62]

These types of issues have attracted the attention of legal scholars. Notably, Professors Niva Elkin-Koren and Michal Gal,

---

57 Similarly, personalization that sorts people in the consumer context may reduce the availability of data in the future. Consider Professors Ben-Shahar and Porat's proposal of personalized mandatory contract rules. Ben-Shahar and Porat, 86 U Chi L Rev at 265 (cited in note 18). If all Type A consumers are subject to a certain mandatory contract rule, we may lose information about how their preferences change subsequent to implementation and how they would act in a world without a mandatory rule. See id at 255.

58 See Kate Crawford and Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 BC L Rev 93, 96–109 (2014); Omer Tene and Jules Polonetsky, *Big Data for All: Privacy and User Control in the Age of Analytics*, 11 Nw J Tech & Intell Prop 239, 251–52 (2013). See also Craig Konnoth, *Health Information Equity*, 165 Penn L Rev 1317, 1333–46 (2017) (exploring the various privacy trade-offs inherent in big data use in health care). Concerns about who provides the algorithm and privacy will be especially important for personalization projects in consumer law. See Ben-Shahar and Porat, 86 U Chi L Rev at 281–82 (cited in note 18).

59 See, for example, Mark Scott and Laurens Cerulus, *Facebook Data Scandal Opens New Era in Global Privacy Enforcement* (Politico, Mar 26, 2018), archived at http://perma.cc/9MMU-6YAA (discussing the reactions of global privacy regulators to the Cambridge Analytica–Facebook scandal in 2016).

60 See Porat and Strahilevitz, 112 Mich L Rev at 1467–69 (cited in note 13).

61 See Niva Elkin-Koren and Michal S. Gal, *The Chilling Effect of Governance-by-Data on Data Markets*, 86 U Chi L Rev 403, 414 (2019) (explaining that data collectors feel reduced pressure to pay data subjects when the government uses the data); Eduardo Porter, *Your Data Is Crucial to a Robotic Age. Shouldn't You Be Paid for It?* (NY Times, Mar 6, 2018), archived at http://perma.cc/RS3S-QGR8. See also Michael Pollack, *Taking Data*, 86 U Chi L Rev 77, 99–106 (2019) (arguing that the Takings Clause should apply to personal data used by the government).

62 This is part of the broader question about who provides the algorithm. See Part II.E.

in their piece for this Symposium, discuss the fundamental tensions generated by data collection by private companies for the purpose of commercial gain and by the government for the purpose of refining the law.[63]

B. Discrimination and Bias

Related to the quality-of-data issue is the question of whether the algorithm exhibits discriminatory behavior and bias. The new personalization of law, with its emphasis on algorithmic decision-making, can reduce or exacerbate existing biases in the law.

Certain biases currently exhibited by adjudicators can be tempered through the use of big data and algorithms. Most obviously, when an algorithm programmed in advance dictates the application of the law to a case, there will be no hindsight bias. Similarly, the impact of individual biases of judges should be reduced by the reliance on data rather than gut intuition. Humans often consciously or unconsciously assume correlations that do not exist. A human may have a biased view about what factors matter when writing or applying a factor-based rule. That is, they assume that some factors are relevant to an inquiry when they are not. The use of big data to predict outcomes will help reduce some of these biases.

In the bail example,[64] the authors illustrate how a machine-learning algorithm can reduce racial bias in decision-making by instructing the algorithm to focus on nonracial factors.[65] Importantly, their results suggest that the algorithm reduced racial disparity and was better at achieving the law's objectives when compared to judges.[66] More basically, the existence of their study demonstrates a key point: we can audit the effectiveness of big data personalization by auditing its outcomes just the same way that the legal academy audits the old personalization of law by human judges.[67] And it is likely easier to correct an algorithmic bias once detected than it is to correct a human bias.

---

63 See generally Elkin-Koren and Gal, 86 U Chi L Rev 403 (cited in note 61).

64 See notes 24–26 and accompanying text.

65 Kleinberg, et al, 133 Q J Econ at 275–78 (cited in note 24).

66 Id.

67 See, for example, Jeffrey A. Segal, *Judicial Behavior*, in Robert E. Goodin, ed, *The Oxford Handbook of Political Science* 275, 280–83 (Oxford 2009) (arguing that judges' personal ideologies affect their decisions); Thomas J. Miles and Cass R. Sunstein, *The New Legal Realism*, 75 U Chi L Rev 831, 835–41 (2008) (finding that political preference, race, gender, and other demographic characteristics sometimes have effects on judicial judgments).

On the other hand, and as noted above, the data itself might be biased.[68] Or discrimination may arise because of proxy variables and correlations that exist in society because of other systemic problems.[69] Personalization adds contextual factors to the law. Identifying and adding relevant contextual factors and removing irrelevant factors can increase accuracy and reduce pernicious biases. But there may be correlations in the data that suggest relevant factors—in the sense that they correlate with certain outcomes—that have a discriminatory effect. Hard questions arise as to whether it is appropriate to consider these factors. And even if the law does not consider them directly, disparities might appear in the results through indirect relationships and proxy variables. The question then is whether the law should do anything proactively to reverse those disparities.

At first, this appears as a trade-off between accuracy in achieving law's objective and reducing discrimination.[70] That is the wrong framing. The real question here is one of determining the appropriate objective in the first place. The law rarely functions on one dimension. Most laws have dynamic objectives. With personalized law, the various dimensions must be understood. There will often be arguments for sacrificing success on one dimension in service of success on another dimension. For example, an algorithm that reduces crime while exacerbating societal inequalities presents difficult policy questions about objective setting. If lawmakers are unwilling to answer those questions, the personalization of law will stall.[71]

---

68 See note 54 and accompanying text. See also Barocas and Selbst, 104 Cal L Rev at 684 (cited in note 15) ("[C]onclusions drawn from incorrect, partial, or nonrepresentative data may discriminate.").

69 For example, an employer may choose criteria for competency that happen to be less common with members of the protected group due to systemic inequalities. The employer's employment practices, made based on these criteria, will have a disparate impact on the members of the group. See Barocas and Selbst, 104 Cal L Rev at 691 (cited in note 15) (noting that systematic discrimination may result when "the criteria that are genuinely relevant in making rational and well-informed decisions also happen to serve as reliable proxies for class membership").

70 See Ya'acov Ritov, Yuekai Sun, and Ruofei Zhao, *On Conditional Parity as a Notion of Non-discrimination in Machine Learning* *16–19 (arXiv.org, June 26, 2017), archived at http://perma.cc/NLK9-GXGX (analyzing whether minority neighborhoods pay higher insurance premiums); Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan, *Inherent Trade-offs in the Fair Determination of Risk Scores* *17 (arXiv.org, Nov 17, 2016), archived at http://perma.cc/WWY7-7X7Q (concluding that no model of risk assignment can meet every goal of fairness).

71 Selbst and Professor Barocas make a similar point in a slightly different context. See Andrew D. Selbst and Solon Barocas, *The Intuitive Appeal of Explainable Machines*,

C. Human Intervention

Algorithmic decision-making does not mean that humans are shut out of the process. Even after the objective has been set, there is much human work to be done. Indeed, humans are involved in all stages of setting up, training, coding, and assessing the merits of the algorithm. If the objectives of the algorithm and the objective of the law are perfectly aligned at the ex ante stage, one must ask: Under what circumstances should a human ignore the algorithm's suggestions and intervene *after* the algorithm has made the decision?

To see how and when humans should intervene, consider the different ways lawmakers can use algorithms to personalize the law. Algorithms can merely provide human decisionmakers with more information about the context of the decision. Alternatively, algorithms can provide suggested decisions to a human or, in the most extreme version, they can be translated directly into automated legal directives. For suggestions to a human, the question will be: How much deference do the humans give to that suggestion? And for automated directives, the question is: When should humans intervene and have the ability to override that directive?

While algorithms may reduce errors, they cannot completely eradicate them. There will always be errors. And questions must be asked about how best to stomach those errors. But the types of errors made by an algorithm and the types of errors made by humans may be different. Some algorithmic errors will be obvious. One need only look at the errors made by algorithms that identify objects in pictures to determine that there is a mismatch of objectives.[72] Thus, humans can—in cases in which the type of error is clear—intervene. But in other cases, errors will be difficult to identify. Algorithms will often identify counterintuitive connections that may appear erroneous to humans even when accurate. Humans should be careful in those cases not to undo the very value that was added by the algorithm's ability to recognize these connections. This is especially true when the benefit of the algorithm was that it reduced human bias and behavioral errors.

---

87 Fordham L Rev 1085, 1133 (2018), archived at http://perma.cc/7T3M-D7LU ("Questions about justifying a model are often just questions about policy in disguise.").

72 Selbst and Barocas catalog some of the interesting mistakes—some obvious and some not so obvious—that experimental algorithms have made. Id at 1122–26.

Moreover, when predictability is paramount, we may not want humans interfering with the algorithm, injecting inconsistency into the system. Similarly, when a lawmaker's precommitment to the outcome is of high importance,[73] it will be important to limit the possibility of human intervention. When precommitment or overcoming human biases is important, human intervention should be kept at a minimum.[74]

D. Transparency

Closely tied to the issue of human intervention is the question of transparency. Some argue that algorithms must be transparent in their reasoning in order for us to be able to use them responsibly. One of the reasons presented in support of transparency is that it informs decisions about human intervention. While traditional statistical techniques enable users to understand the weights and interactions of different variables in the decision, more complex machine learning algorithms do not allow for such interpretation. This is true in part because the algorithms are recognizing nonintuitive connections that human intuition cannot recognize. As Andrew Selbst and Professor Solon Barocas point out, the transparency problem "is a particularly pronounced problem in the case of machine learning, as its value lies largely in finding patterns that go well beyond human intuition."[75]

Commentators have suggested that it is irresponsible for lawmakers to delegate duties to algorithms whose reasoning and decisions are not transparent. These concerns are misdirected.[76] First, it sets up a false comparison. Critics of algorithmic decision-making often emphasize the importance of human judges offering reasons for their opinions.[77] But the human brain is even more of

[73] See text accompanying notes 44–45.

[74] In the private context, as Professor Fennell points out, individuals often want to prevent themselves from intervening with a directive in the future to deal with behavioral self-control problems. See Fennell, 86 U Chi L Rev at 434–47 (cited in note 45). Similarly, when the government is using a microdirective to precommit itself, human intervention would be counterproductive.

[75] Selbst and Barocas, 87 Fordham L Rev at 1129 (cited in note 71); id at 1094 (noting that an algorithm's ability to learn things "that humans might overlook or cannot recognize . . . render[s] the models developed with machine learning exceedingly complex and, therefore, impossible for a human to parse").

[76] Selbst and Barocas provide a deep analysis of the various flaws in the transparency arguments. See id at 1089–93.

[77] This critique reveals a particularly academic bias about what judges do. While the appellate decisions taught in law schools are supported by written opinions, lower court

a black box than machine learning algorithms. Judges' written opinions may simply provide ex post justification for opinions that are actually driven by other factors.[78]

The fix here lies elsewhere. Algorithmic personalization requires transparency of the human process, not the computer reasoning. The relevant information to test the validity of an algorithm will be what objective it was given (and how that objective was developed), how the algorithm was programmed to achieve that objective, how the data was selected, and audit data on the algorithm's performance.[79]

E. Regulating Providers

Finally, one of the most important logistical questions for the personalization of law through big data is who provides (and who owns) the personalization. This is true across the various methods that we have discussed. We have made this point elsewhere,[80] as has Professor Gillian Hadfield, who suggests it will be one of the most important changes coming in all of law.[81] Moreover, Professor Andrew Verstein provides a deep analysis of the question in this Symposium.[82] So we will say only a few words on the topic.

---

judges are constantly ruling on motions and objections and entering orders without any written opinions that can be examined after the fact.

78 Indeed, this disconnect between the stated reasons in judicial opinions and the actual reasons is a major driver behind the robust existence of the field of judicial behavior. Scholars have long recognized that many factors not stated in an opinion might be driving outcomes. See generally, for example, Nicola Gennaioli and Andrei Shleifer, *Judicial Fact Discretion*, 37 J Legal Stud 1 (2008) (modeling how judges find facts and arguing that the summaries of facts in written opinions cannot be trusted); Jerome Frank, *Are Judges Human? Part One: The Effect on Legal Thinking of the Assumption That Judges Behave Like Human Beings*, 80 U Pa L Rev 17, 33–38 (1931) (explaining that even honest judges are influenced by extralegal factors, such as their own prejudices).

79 Selbst and Barocas make a similar proposal in their article. Selbst and Barocas, 87 Fordham L Rev at 1130–33 (cited in note 71) (noting that personalization will "require process, documentation, and access to that documentation" and proposing the idea of an algorithmic impact statement).

80 See Casey and Niblett, 43 J Corp L at 30–32 (cited in note 11).

81 Hadfield, *Rules for a Flat World* at 323–45 (cited in note 11) (describing the potential pitfalls and benefits of third-party provision of legal infrastructure).

82 See generally Verstein, 86 U Chi L Rev 551 (cited in note 11). Rebecca Wexler also provides a thoughtful analysis of the problems inherent in private ownership of sentencing algorithms. See generally Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 Stan L Rev 1343 (2018).

An automated algorithmic method would require a public or private software provider. A survey method would require a public or private survey process. The worry with a private provider is that it could manipulate or game the system. It may be allied with other private actors. Knowing the underlying code, running preemptive surveys, or having advance access to data could give asymmetric advantages to certain parties.

Again, an analysis of the objective of an algorithm is key. Privately created algorithms may serve to correct only one type of error. For example, as Professor Dan Burk notes in this Symposium, private enforcers of copyright may focus on reducing infringement.[83] How can the law best achieve outcomes that reflect the objectives of lawmakers if interested parties are creating algorithms that define the bounds of the law?

Even when the provider of the algorithm is a neutral party, the provider may not be forthcoming in revealing how the algorithm was created.[84] This is especially true when a private provider has financial incentives to obscure the reasons why particular results are generated in order to heighten barriers to competition or when it has reasons to favor one side because of repeat-player issues.[85]

In these cases, the law should require government providers or initiate regulation of these third-party providers.[86] On this note, the law must be proactive. The personalization of law is likely to develop in different ways and from varied sources, becoming a widespread phenomenon long before the legislature takes any action.[87]

CONCLUSION

The key theme of this Essay is that everything in personalization comes back to objectives. If one propounds the benefits of a personalized law through the use of an algorithm, one must ask

---

83 Dan L. Burk, *Algorithmic Fair Use*, 86 U Chi L Rev 283, 284 (2019).

84 See Wexler, 70 Stan L Rev at 1364–68 (cited in note 82) (illustrating how, in the case of pretrial suppression hearings, "third-party developers will try to use intellectual property law as a shield against judicial scrutiny").

85 See Casey and Niblett, 43 J Corp L at 28–29 (cited in note 11).

86 See Verstein, 86 U Chi L Rev at 567–72 (cited in note 11) (discussing when the law should be privatized); Gillian Hadfield, *Producing Law for Innovation*, in *Rules for Growth: Promoting Innovation and Growth through Legal Reform* 23, 39–44 (Kauffman 2011) (proposing that third parties be allowed to compete in the regulatory market).

87 See Hadfield, *Producing Law* at 52–53 (cited in note 86).

whether the algorithm is achieving the law's purpose. That will be a comparative analysis among the three choices: (1) big data personalization, (2) human judge personalization, and (3) no personalization. As long as big data provides better outcomes than humans and better outcomes than blunt ex ante rules, then one should not reject them simply because they are imperfect. One wants to make sure the law is achieving its desired outcome (without costly unintended consequences) better than the next-best alternative. But this presupposes that lawmakers have been able to identify a measurable objective. What does it mean for a law to be "better"? Can we agree on the normative content of a "better" outcome? Do we want a law based on consequentialism or one that reflects a more deontological approach?

The trend toward the new personalization, therefore, presents a challenge for the lawmaker that has hitherto been relatively easy to set aside for a later date: What is it, exactly, that we want the law to achieve? When this all plays out, it will not be technological infeasibility or lack of data standing in the way of personalized law. It will be the inability of humans to agree on what law is designed to do.