The Death of the Legal Subject: How Predictive Algorithms Are (Re)constructing Legal Subjectivity

Katrina Geddes

ABSTRACT

This paper explores the epistemological differences between the socio-political legal subject of Western liberalism, and the algorithmic subject of informational capitalism. It argues that the increasing use of predictive algorithms in judicial decision-making is reconstructing both the nature and experience of legal subjectivity in a manner that is incompatible with law's normative commitments to individualized justice. Whereas algorithmic subjectivity derives its epistemic authority from population-level insights, legal subjectivity has historically derived credibility from its close approximation of the underlying individual, through careful evaluation of their mental and physical autonomy, prior to any assignment of legal liability. With the introduction of predictive algorithms in judicial decision-making, knowledge about the legal subject is increasingly algorithmically produced, in a manner that discounts, and effectively displaces, qualitative knowledge about the legal subject's intentions, motivations, and moral capabilities. This results in the death of the legal subject, or the emergence of new, algorithmic practices of signification that no longer require the input of the underlying individual. As algorithms increasingly guide judicial decision-making, the shifting epistemology of legal subjectivity has long-term consequences for the legitimacy of legal institutions.

CCS CONCEPTS

• **Theory of computation**; • **Computing methodologies - Artificial intelligence - Philosophical/theoretical foundations of artificial intelligence**; • **Social and professional topics - Computing / technology policy**; • **Applied computing - Law, social and behavioral sciences**; • **Theory of computation - Design and analysis of algorithms**;

KEYWORDS

algorithmic fairness, law, legal subjectivity, expressive power, epistemology, STS, critical algorithm studies, inequalitySchool of Law, New York University, kgg283@nyu.edu

ACM Reference Format:
Katrina Geddes. 2022. The Death of the Legal Subject: How Predictive Algorithms Are (Re)constructing Legal Subjectivity. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), June 21–24, 2022, Seoul, Republic of Korea.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3531146.3533134

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
FAccT '22, June 21–24, 2022, Seoul, Republic of Korea
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9352-2/22/06.
https://doi.org/10.1145/3531146.3533134

1 1 INTRODUCTION

Across a range of settings, legal decision-making relies increasingly on predictive algorithms to determine individual rights and interests. Scholarship on algorithmic decision-making has focused on the pernicious effects of algorithmic bias, and opacity [1-6]. This literature assumes that if algorithmic models can be disclosed, and de-biased, that their use in legal contexts is otherwise permissible. This perspective overlooks the ways in which algorithmic epistemology is reconfiguring legal subjectivity. As judges turn to algorithms to guide their decision-making, knowledge about the legal subject is increasingly algorithmically produced. Statistical predictions about the legal subject displace qualitative knowledge about their intentions, motivations, and moral capabilities. The reasons why a particular defendant might refrain from recidivism, for example, become less important than the statistical features they share with historical recidivists. This displacement of individual knowledge with algorithmic predictions diminishes the participation of the legal subject in the epistemic processes that determine their fundamental liberties. Given the impenetrability of algorithmic models, and the strength of automation bias, it is difficult for legal subjects to counter the prejudicial effect of algorithmic predictions. The resulting exclusion of the legal subject from the production of knowledge about themselves has participatory, dignitary, and expressive effects, as power over self-articulation is transferred from the legal subject to the data capitalist. Outsourcing the production of legal knowledge to private corporations effectively instantiates the power of private capital over the conditions of human freedom [41].

Using algorithmic knowledge to construct the legal subject also destabilizes a core epistemological foundation of law, namely, the morphological resemblance between the analog legal subject, and the corresponding flesh-and-blood individual. Contemporary legal subjectivity derives credibility from its close approximation of the underlying individual, through careful examination of their mental and physical autonomy, prior to any assignment of legal liability. This effort to paint a more complete, and accurate portrait of the underlying individual (using coherent causal explanations for their behavior) often legitimates the coercive power of the state in circumstances where such power may be fiercely resisted.

In contrast, the purpose of algorithmic subjectivity is not to faithfully portray the underlying individual, but to identify patterns of behavior at the population level that will facilitate their classification for "stochastic governance"[7]. Whereas morphological resemblance imposes constraints on the fragmentation of the analog legal subject, the algorithmic subject is designed to emerge from the unstable dynamics of fragmentation, and recombination; to manifest global behaviors that will facilitate classification with almost any sub-population based on shared statistical features [8-11]. For this purpose, an algorithm does not need to interrogate an

individual's subjective intentions, because their desires and preferences are statistically pre-empted, just as their future autonomy is inferred from the historical behavior of their statistical peers. The correlations extracted from this "impartial" metadata dispense with the need to develop coherent causal explanations of individual behavior. This produces the death of the legal subject, or the emergence of new, algorithmic practices of signification that no longer require the input of the underlying individual.

The observation that algorithmic and legal subjectivities derive from, and participate in, different epistemologies, is not a recommendation to prohibit the use of algorithms in judicial decision-making, or to uncritically reinstate the traditional paradigm of legal subjecthood. The utility of algorithmic guidance in modern jurisprudence is a complex question that this paper does not attempt to resolve. Even in the narrow context of criminal law, it is difficult to reach a firm conclusion about the net utility of algorithmic guidance. The potential benefits of risk assessment tools at one decision node (for example, diverting low-risk offenders from pretrial detention) are difficult to weigh against the harms generated by their use at other decision nodes (for example, the exacerbation of racial and socioeconomic disparities in sentencing). And this analysis is further complicated by shifting empirical evidence, and the conceptual incoherence of criminal law. Instead, this Paper contributes an observation about the underexplored epistemic effect of algorithmic knowledge on legal subjectivity. How is the basic unit of liberal society transformed by the elevation of algorithmic knowledge? How are they differently represented? Is the actuarial project of algorithmic governance fundamentally at odds with the law's normative commitment to individualized justice?

In answering these questions, it is important to resist nostalgic and short-sighted treatment of the traditional liberal paradigm of legal personhood. The traditional conception of the bounded, rational, self-determining legal subject (occupying a sphere of autonomy constructed by individual rights) is increasingly incompatible with contemporary understandings of systemic injustice, and evolving norms of collective responsibility based on mutual interdependence [12]. Crime, for example, cannot be explained by individual pathology, and cannot be solved through individual-level intervention. It is a deeply social phenomenon, sustained by social, cultural, political, and economic relations, that exist beyond the control of any individual. But predictive algorithms reflect persistent optimism that individual-level interventions can overcome the structural forces that sustain patterns of criminality [13]. Risk assessment tools target the "criminogenic" features of individuals, rather than the circumstances that shape and constrain their behavior. The baseline conditions of inequality that sustain patterns of criminality are illegible to the algorithm, so it treats these structural features as fixed constraints, choosing instead to engage in discriminatory profiling practices that require the persistence of existing disparities in order to be effective [63, 78]. This unrelenting scrutiny of individual behavior demands a critical re-evaluation of liberal individualism, as manifested through the discourse of individual rights, and the "self-determining" legal subject. A more nuanced and realistic understanding of legal responsibility (including responsibility for crime) requires a more nuanced and realistic conception of the legal subject, that pays greater attention to the relations that constitute individuals, and shape their behavior [12]. These are not the data relations of algorithmic subjectivity, but the social, cultural, and political relations that are meaningful to the individual. A more relational understanding of legal subjectivity will help us to recognize not just the limits of our power as individuals, but the kinds of power we can wield as a collective [12].

1.1 1.1 The Traditional Legal Subject

Law has always recognized legal subjecthood on the basis of highly specific metaphysical, empirical, and axiological beliefs about the state of the world, and the kind of subject the law should serve [14]. Slaves, for example, were once considered "property" rather than persons, and married women were denied separate legal personhood from their husbands [15]. Both the construction and experience of legal subjectivity are historically contingent, reflecting social norms, and evidentiary technologies. In pre-modern societies, the attribution of criminal responsibility turned on assumptions about character, based on the standing and reputation of the accused [16]. Trials were carried out on the basis of character evidence collected by the local community, and no effort was made to inquire into the defendant's mental state [17]. The contemporary legal subject, entitled to be judged in terms of their particular intentions, evolved gradually over time, in response to Enlightenment theories of agency, and utilitarian beliefs in the capacity of rational actors to be deterred from wrongdoing. Modern conceptions of autonomy as "free will" had to be rescued from premodern beliefs in determinism, including theological determinism (divine providence), and ancient philosophies of fate [18]. The developing fields of psychology and psychiatry led to increased optimism about the susceptibility of the mind to both evaluation and treatment, resulting in a more psychological view of personal responsibility [17]. Meanwhile, the social fragmentation of industrialization and urbanization (increased mobility and anonymity) made it difficult to collect the kinds of character evidence that had sufficed in pre-modern societies. The powers and resources of the modern administrative state were also expanding, with the birth of prisons, and the professionalization of policing and prosecution [17]. Statutory provisions affirming the right of the accused to testify on their own behalf reflected this emerging conception of "criminal responsibility as residing in psychological states of mind" [17]. Today, the conception of criminal responsibility as requiring mental autonomy is so deeply engrained in the moral legitimacy of the criminal law that strict liability offenses are "mentioned in hushed tones as an embarrassing and uncivilized exception" to the principles of criminal law [16].

It is important to note that legal subjectivity occupies a spectrum of abstraction. At one end of the spectrum, the stripped-down, highly abstracted subject of constitutional texts bears no more identifying features than natural personhood, and the possession of fundamental rights. This faceless rights holder expresses liberal society's commitment to equality, and the equal moral worth of all individuals as interchangeable rights-bearers [12]. Further along the spectrum is the reasonable person in tort law, who is sufficiently well delineated to capture variation in norms of reasonable behavior across different fields and disciplines. This legal subject provides the standard against which the behavior of tort defendants is measured. At the other end of the spectrum is the

criminal legal subject, whose mental and physical autonomy receive the most careful evaluation due to the severity of criminal sanctions. This legal subject expresses the law's commitment to the presumption of innocence, and the value of individual liberty. This Part will examine the traditional features of the contemporary criminal subject, given the growing use of predictive algorithms at almost every decision node in the criminal justice system [19]. The contemporary criminal subject is attributed with mental autonomy (the ability to form individual thoughts, intentions, and interpretations of the law); physical autonomy (the ability to act on one's intentions with causal efficacy); and future potentiality (in possession of an open, and undetermined future). Historically, the attribution of these traits has justified the imposition of severe legal sanctions [20].

1.2 1.2 Mental Autonomy

At a very basic level, the provision of textual instructions for moral behavior assumes that individuals possess the mental autonomy required to interpret and apply the law to their particular circumstances. When a legal subject approaches the law, they do not "discover" the preexisting properties of a fully-formed object; rather, they construct that object to suit their specific needs and preferences [21]. For example, the need for coherence, or principled consistency, will often compel legal subjects to understand legal doctrines as part of a reasonable, consistent, and non-arbitrary scheme of human regulation. Jack Balkin calls this "rational reconstruction," or the attempt to find normative coherence within the law [21]. This process is inherently subjective because different legal subjects will form different views about the substantive rationality of the law, depending on their moral and political beliefs, their knowledge of the legal system, and the extent of their cognitive exertion on its normative consistency [21]. In this sense, the coherence of the law is constituted by individual processes of subjectification – the hermeneutic interaction between a legal subject and a legal text. Mental autonomy is thus embedded in legal subjecthood because legal interpretation is a deeply subjective, socially-situated process in which the legal subject both constructs, and is constructed by, the legal text [21].

The importance of mental autonomy for the criminal legal subject also rests on fundamental conceptions of the minimum conditions for the attribution of blame. The mental state requirement of criminal liability (*mens rea*) is used to distinguish the blameworthy from the blameless, for example, the act of perjury from an innocent misstatement. And it reflects society's normative commitment to individual autonomy [22, 23]. Mental conditions that excuse criminal responsibility are tolerated for the same reason that civil transactions are invalidated upon proof of coercion or undue influence, which is that actions performed under those circumstances do not represent genuine choice [20]. Law is designed to enable individuals to exert control over their futures by giving effect to their informed and considered choices. A legal system that considers an individual's mental state maximizes the efficacy of those choices within the coercive framework of the law. Individuals are better able to predict whether and when the sanctions of the law will apply to them, because their individual choices will be a determining factor in those outcomes. As a result, they can identify in advance the space left open to them, "free from the law's interference" [20]. In contrast, under a system of strict liability, individual exposure to interference could neither be predicted nor controlled; every blow, even if accidental or careless, could give rise to liability. Accordingly, a legal system that requires mental conditions of responsibility reflects a normative commitment to autonomy. Although this approach bears more risk (not interfering until harm has occurred), that risk is "the price we pay for general recognition that a man's fate should depend upon his choice" [24].

1.3 1.3 Physical Autonomy

After the collapse of medieval society and its rigid class hierarchy, the economic opportunities of the marketplace enhanced the potentiality of individual choice. Social status was "decollectivized and relocated in the personal projects of free individuals undertaken within the protected space created for them by the law" [25]. The introduction of universal schooling also enhanced the efficacy of individual choice; educated individuals were regarded as more competent actors within society, prompting greater emphasis on agency, rather than fate. The individual was imbued with moral sovereignty, and new rights and responsibilities consistent with their newfound competencies. Individualism was viewed as a complement to the expansion of modern Western society [25].

As part of this liberal project, Western legal systems emphasized the autonomy of the legal subject. Without autonomy to choose between different courses of action, an individual could not reasonably be held responsible for the consequences of those choices [26]. In the words of Lon Fuller, "[t]o embark on the enterprise of subjecting human conduct to the governance of rules involves of necessity a commitment to the view that man is, or can become, a responsible agent, capable of understanding and following rules, and answerable for his defaults" [27]. Only an autonomous individual could "respond through acts of volition to the requirements of normative order" [28]. Naturally, the range of autonomy available to an individual will always be constrained by the normative order; our choices are always constrained by the environment in which they are made [26]. But this does not alter the fact that some minimum amount of autonomy underwrites our assumptions about the reasonableness of imposing sanctions on the choices made by individuals [29].

Law's normative commitment to individual autonomy partially explains judicial reluctance to adjudicate liability exclusively on the basis of statistical evidence [30]. Statistical evidence, because it is probative in aggregate, sacrifices interests in individual accuracy, and thereby undermines the efficacy of individual choice [31]. For example, if you purchase a ticket to a concert, but ninety-nine per cent of concert attendees do *not* purchase a ticket, and you are later prosecuted on the basis that, statistically speaking, you are *unlikely* to have purchased a ticket, the absence of any connection between your liability and your personal conduct strips the latter of causal efficacy. Your decision to buy a ticket has been rendered meaningless by the prioritization of statistical evidence, because you are being punished for the actions of others, over whom you exert no control. This loss of causal efficacy diminishes your individual autonomy [32, 33]. And individuals have little incentive

to obey the law if they will be punished for the lawlessness of their statistical peers [34]. The promotion of law-abiding behavior, then, is an instrumental reason to adjudicate liability on the basis of individualized rather than statistical evidence [31, 35]. The incentive-corrupting effect of reliance on statistical evidence does not occur with individual evidence, even if it is probabilistically equivalent [35]. For example, consider a pedestrian who has been injured by a ride-sharing vehicle, where there is eyewitness testimony (shown to be 80% reliable) that the vehicle was an Uber. If, instead, there is no eyewitness testimony, but we have statistical evidence that 80% of the ride-sharing vehicles operating in the area are Ubers, is that a sufficient basis on which to ground liability? Although both forms of evidence are probabilistically equivalent, the individual (eyewitness) evidence will be viewed as a more legitimate basis for liability than the statistical (market share) evidence [35]. If this were not the case, other ride-sharing operators (like Lyft) would have little incentive to improve their individual safety records, because Uber would act as their insurer. Thus, judicial reluctance to rely exclusively on statistical evidence, except in a handful of contexts [36], is partially explained by the desire to preserve the value of individual autonomy through the preservation of incentives for law-abiding behavior.

1.4 1.4 Future Potentiality

By providing textual instructions for moral behavior, systems of law have implicitly attributed to legal subjects the mental and physical autonomy required to interpret and apply the law to their individual circumstances. A necessary precondition for the exercise of this autonomy is the openness and indeterminacy of the future [37]. Without the possibility of influencing an undetermined future, individuals would struggle to exercise meaningful choice in the present; the very concept of choice (between action *a* and action *b*) would be farcical because the events of the future would be impervious to individual action [38]. As a result, the legal subject would struggle to apply the law with a view to influencing future outcomes [39].

1.5 1.5 The Criminalization of Status

In times of relative peace and stability, states may feel comfortable committing to a system of proof of individualized responsibility. But during periods of insecurity, when the costs of determining individual capacity-based responsibility seem intolerably high, and there is optimism about the ability of technology to control risk, perceptions of criminality as located in "stable" individual pathology can increase in popularity [17]. The first parole prediction tool, developed in 1927, included the nationality or race of the father as one of twenty-one factors "predictive" of parole violation [40]. In the 1970s and 1980s, the practice of selective incapacitation used predictions of recidivism to identify and incarcerate the small subset of the population that was believed to be responsible for the majority of crimes (so-called "career criminals") [19]. Today, the construction of the criminal legal subject again reflects prevailing social anxieties, and scientific theories [17]. The mythology of Big Data reassures an anxious public that criminal sanctions can be effectively and efficiently distributed using risk assessment tools. As a result, criminal legal subjectivity is increasingly algorithmically constructed. Predictive algorithms paint the criminal subject as "knowable" and "predictable" through statistical correlation, thereby dispensing with the need to consult the underlying individual, and ushering in a new generation of status criminalization [17].

2 2 THE ALGORITHMIC SUBJECT

The algorithmic subject is a "probabilistically determined behavioral profile" constructed from correlations identified in population-level data [41]. The descriptive term *algorithmic* refers to the high-speed, computational processes that collect and compare physical, transactional, and behavioral data from the digital surveillance technologies of networked information architectures [41]. The algorithmic subject is descended from the statistical subject of the eighteenth and nineteenth centuries, when developments in statistical modeling produced new tools for measuring and managing populations. Through the universalization of birth certificates, Social Security numbers, and other types of persistent formatting, individuals became fastened to, and made legible by, predefined categories of data [42]. As markets and states developed an affinity for processes that rendered individuals measurable, traceable, and manipulable, the statistical subject became the target of actuarial interventions [43, 44]. After the Civil War, many American families purchased life insurance as a means of preserving social status: keeping widows out of the workforce, or avoiding the embarrassment of a pauper's burial [45]. At the turn of the twentieth century, life insurers began to standardize their methods of risk classification. They retained medical examiners to identify the healthy, and reject the sick, in order to maintain low premiums, attract more customers, and expand capital reserves for investment [45]. This created tension between medical examiners, who tended to reject applicants, and insurance agents, who tended to accept applicants. To overcome this tension, New York Life introduced "sub-standard" life insurance policies: applicants who would ordinarily be rejected from insurance were instead approved at higher premiums. Actuaries were encouraged to *classify* rather than to aggregate; to "personalize" risk ratings, and construct risk classes [45]. Risk classification was further standardized by the adoption of numerical methods. An individual's "build" (height to weight ratio) provided the base value, which would then be adjusted upwards or downwards by insurance clerks depending on the "impact" of isolated factors (for example, add five points for height, or subtract five points for family history) [45]. Numerical methods offered a cheap and efficient means of "predicting" relative mortality; insurance clerks could use mortality tables to "calculate" risk ratings from paper applications, rather than consulting medical professionals. Meanwhile, similar developments were taking place in the consumer reporting industry, as credit bureaus began to develop more sophisticated tools for evaluating creditworthiness. Actuarial risk models replaced character interviews as credit scoring became the primary means of distributing financial credit [44]. Fair, Isaac & Company promoted the concept of "statistical objectivity" to shield their credit scorecards from anti-discrimination regulation [44].

Caley Horan describes the second half of the twentieth century as America's "actuarial age," in which the ideology of actuarial science normalized the risk classification of credit applicants, and

insurance holders [46]. Insurers framed economic security as an individual responsibility rather than a right of citizenship, justifying a reduced role for the state, and securing the indispensability of their own services. Promotional materials for private insurance enlisted citizens to "defend" distinctly American values of individualism and free enterprise against the normative threat of communism [46]. Older, more inclusive forms of social security founded on solidarity, interdependence, and mutual aid, were replaced by the segregationist logic of actuarialism, which emphasized differences, rather than mutuality, as a means of refining risk pools, and "shielding" individuals from the costs of Others. In this way, the anti-redistributive normative foundation of private insurance ("actuarial fairness," or the principle that each person should only pay for their own risk) was preserved through anti-communist Cold War rhetoric [46].

Today, the ideology of neoliberal self-governance is amplified by informational capitalism, and its unrelenting scrutiny of individual behavior as the sole determinant of socioeconomic status. By conditioning access to economic resources on the strength of algorithmic profiles, neoliberal markets encourage consumers to embrace algorithmic subjectivity as a legitimate mode of individuation [47]. Individuals assemble themselves as responsible algorithmic subjects, wearing fitness trackers and refreshing credit scores, constantly engaged in "self-surveillant algorithmic adjustment" before a vast, unblinking audience of data brokers [48]. Biometric information is sold as self-awareness, even as it decontextualizes data from lived experience in order to generate numeric homogeneity for quantitative data processing [48]. Debtors perform specific data-generating behaviors in order to appear "creditworthy," and thereby secure their imprisonment within the subjectivizing apparatuses of surveillance capitalism [49]. When individuals perform the classifications required for algorithmic measurement (rather than the behaviors that are meaningful for their own existence), they internalize the disciplining effects of datafication [49]. And so the actuarial subject of the twentieth century is reborn as the algorithmic subject of the twenty-first. Like its actuarial ancestors, the algorithmic subject is not designed to faithfully represent the underlying individual, but to facilitate the construction, management of, and trade in statistical populations, as a form of biopower [41].

The epistemic authority of the algorithmic subject is propagated by three legitimating narratives. The first is the construct of personal data as a "raw" resource, freely available for extraction, refinement, and "productive appropriation" by data capitalists [41]. This romantic narrative of "data prospecting" legitimates intrusive forms of surveillance as the "discovery" of natural resources, and normalizes efforts to manipulate user engagement in order to maximize opportunities for data extraction [50]. The framing of personal data as "raw" also obscures the normative choices that influence what kind of information is collected as data, and how it is measured, labeled, classified, and stored [41]. The second narrative frames the behavioral patterns, predictions, and forecasts derived from personal data as "new forms of datafied and depoliticized truth," previously invisible to the human eye [41]. This narrative legitimates the exclusion of the individual from the production of knowledge about themselves on the basis that "unmediated" behavioral data (not self-reported or otherwise subject to conscious manipulation by data subjects) offers unrivaled predictive accuracy [41]. Platforms are framed as "impartial" aggregators of unconscious traces of human behavior, or metadata [51]. This narrative of Big Data "objectivity" shields data extraction from regulatory scrutiny, on the basis that data capitalists are simply harvesting "depoliticized truth" [41]. The third narrative claims "personalized" knowledge about individual subjects, despite the fact that the exclusion of the individual from the knowledge production process forms the very basis of Big Data's claim to objectivity [52]. The algorithm's statistical "knowledge" is generated automatically, without any underlying theory, and with minimal human intervention, in the same way that Google translates sentences into Chinese with no underlying linguistic knowledge, just large datasets [52, 53]. The identification of correlations in categorical preference data (what individuals like, search, purchase, and share, relative to others) allows the algorithm to address not *the* you, but *a* you, that is refracted through multiple layers of relational data [47]. Big Data thereby "avoid[s] all forms of subjectivity," even as it claims to possess "personalized" knowledge about the very individual it ignores, and to whom it is entirely indifferent [47]. Bernard Stiegler describes this as "absolute non-knowledge" in the form of reticulated, and mimetic correlationist mythology [54]. Living knowledge, or noetic consciousness ("dreaming, wanting, reflecting, and deciding") is replaced by a closed loop of self-referential digital traces that construct a "personalized" simulation of consumerist drives, functionalized for the data economy [53].

3 3 THE DEATH OF THE LEGAL SUBJECT

The "death" of the legal subject refers to the emergence of new, algorithmic practices of signification that no longer require input from the underlying individual. Knowledge about the legal subject is increasingly algorithmically produced, in a way that discounts, and effectively displaces, qualitative knowledge about the underlying individual's subjective intentions, and motivations. Whereas legal subjectivity derives credibility from its close approximation of the underlying individual (through careful evaluation of their mental and physical autonomy), algorithmic subjectivity derives epistemic authority from population-level insights. The *relationality* of data, rather than its uniqueness, is what drives its economic value [55]. Predictive profiling apprehends data subjects as patterns of behavior, rather than unique individuals, so that the marginal cost of losing one person's data is relatively low [55]. In contrast, individualized knowledge about a particular person is necessary for judicial determination of their legal rights and interests. Evidence of subjective intent, for example, will inform an evaluation of mental autonomy, and the attribution of individual responsibility.

3.1 3.1 Mental Autonomy

Mental autonomy is not required for algorithmic subjectivity, and is, in fact, discouraged. Automated systems, approaching the speed of light, outstrip our sluggish capacities for individuation by preempting our desires and instincts with "personalized" recommendations for behavior [53, 54]. The algorithmic subject is not required to formulate, or express, individual desires or preferences because they are statistically pre-empted [52]. So the reasons why a criminal defendant may have behaved in a particular manner, and conversely,

the reasons why they might refrain from such behavior in the future, become less important than the statistical features they share with historical recidivists.

3.2 3.2 Physical Autonomy

A judge who chooses to sentence a criminal defendant for x years (the minimum sentence), or $x + n$ years (to incapacitate them from committing future crime) might take into consideration an algorithmic prediction that this particular defendant is a "high-risk" recidivist, and that will affect the defendant's autonomy in several ways. The act of prediction itself is autonomy-eroding because the behavior being predicted is subject to individual control, and the prediction materially affects the conditions of possibility for the exercise of that control [56]. Preventive incarceration based on expected future criminality (incapacitation) denies the defendant the opportunity to disprove the prediction through their autonomous behavior. And this outcome is exacerbated by the criminogenic effects of incarceration: individuals who are incarcerated for longer periods of time have greater difficulty re-integrating, and are more likely to recidivate [57, 58].

Incapacitation, as an approach to sentencing, also reveals the conceptual incoherence of the criminal law. Criminal law's primary justification for the imposition of severe legal sanctions (retribution for moral blameworthiness) has always sat uneasily with its desire to prevent the commission of future crime (deterrence). The same philosophical foundation that justifies severe deprivations of liberty, and informs almost every aspect of judicial due process (the presumption of innocence, the burden of proof, the right to trial), struggles to coexist with a system of preventive incarceration based on predictions of future behavior [59]. Incarceration for uncommitted crimes is normatively inconsistent with the contingency of punishment on proof of autonomous conduct [59]. If incapacitation of dangerous individuals was the sole objective of criminal law, there would be little reason to wait until a crime had been committed before imposing criminal sanctions; screening a population for factors predictive of crime would represent a more efficient use of resources [59]. The orthogonal tasks of assigning blame for past crime, and assessing risk for future crime, are difficult to integrate in a coherent system of punishment [60].

If incapacitation itself is autonomy-eroding, how do predictive algorithms alter that effect, if at all? Judicial reliance on predictive algorithms exacerbates the autonomy-eroding effects of incapacitation, in at least three ways. First, the reliance on statistical evidence treats the defendant as if their future conduct could reliably be inferred from the frequency of misconduct around them, or the dead hand of their own past – as if they were "determined rather than free" [61]. This ignores the defendant's capacity to diverge both from their past, and from their statistical peers [61]. Although predicting the future and determining the past represent distinct cognitive tasks, the same moral objection to the erosion of autonomy motivates judicial reluctance to adjudicate individual liability exclusively on the basis of group statistics. Such adjudication is inconsistent with the law's "commitment to treat the defendant as an autonomous individual" – as the author of their own destiny, rather than the object of statistical relations [24]. Because behavioral patterns at the population level may have no bearing on an individual's propensity towards recidivism, statistical inferences effectively punish the individual for the historical behavior of third parties.

Secondly, an algorithmic score cannot be controlled or altered by the individual it claims to represent because its correlationist design depends on the behavioral data of third parties [62]. Nor will the individual have a say in the statistical populations used to represent them, as the algorithm may choose variables that are fundamentally (e.g. race) or ethically (e.g. religion) unchangeable [61, 63]. As a result, the individual is punished for membership of a statistical group, where membership is neither voluntary nor causative of the outcome being predicted. It is not the case, for example, that being Black is "causative" of crime in the sense that its actuarial use implies, yet Equivant's risk assessment tool overpredicts recidivism for Black defendants at nearly twice the rate of Whites [64]. This outcome punishes the individual for society's history of racial discrimination. This is not to say that restricting the choice of predictive variables to those with a "plausible" causal connection to the predicted outcome would preserve the autonomy of the decision subject [65-67]. Rather, the autonomy-eroding effect of prediction is exacerbated by the use of variables that unfairly stigmatize particular groups through their attribution of causation [68].

Thirdly, a defendant may be unable to meaningfully counter an algorithmic prediction with qualitative information about their personal circumstances and intentions, due to the prejudicial effect of automation bias. Unlike standalone statistical evidence, algorithmic predictions carry the imprimatur of "datafied and depoliticized truth" [41]. Despite their undetermined accuracy at the individual level [69], algorithmic predictions are marketed as "personalized" assessments of an individual's propensity toward a specific behavioral outcome. The aggregation of vast amounts of data relating to an individual lends the algorithmic prediction greater epistemic authority than a standalone statistic. The disembodied "omniscience" of data surveillance amplifies the prejudicial effect of the algorithmic prediction, or the probability that it will be assigned greater weight than any qualitative evidence produced by the individual subject.

There is substantial empirical evidence that human decision-makers tend to accept, rather than challenge, quantitative assessments, and to assign greater weight, amongst a set of variables, to the variable that has been quantified [70-72]. This bias is especially likely when the algorithm's recommendation is expressed in unnuanced terms, and its calculations are opaque [73]. This means that the risk of future crime, because it has been quantified, may receive greater weight in a sentencing decision, resulting in the prioritization of deterrence over retribution [74]. An empirical study of the impact of risk assessment tools on judicial decision-making found that the introduction of algorithmic predictions reversed the effect of poverty from a mitigating factor that reduced the probability of incarceration, to a risk factor that increased the probability of incarceration, thereby exacerbating socioeconomic disparities in sentencing [75]. The same risk assessment information reduced the likelihood of incarceration for relatively affluent defendants, but increased the likelihood for relatively poor defendants (61.2% vs. 44.4%), after controlling for a judge's sex, race, politics, and jurisdiction [75].

3.3 3.3 Future Potentiality

A predictive algorithm does not perceive the future as undetermined; it views the future as entirely knowable, and predictable, through the lens of historical data. In this way, the algorithm constructs a specific temporal relation between past, present, and future, in which historical patterns recur throughout, thereby lending the algorithm its preemptive power [76]. The algorithm's focus on individual behavior as the sole determinant of the outcome being predicted also obscures the constraining conditions of circumstance. What the algorithm perceives as "innate" or inherently individual differences in behavioral propensity are actually the product of categorically differential access to housing, nutrition, education, and employment [77]. Because these structural features of the sociopolitical world are illegible to the algorithm, it treats these features as fixed constraints [78]. It faithfully reproduces what it considers to be "unchangeable" boundary conditions, rather than re-imagining what those conditions could be. Investments in education, housing, and healthcare, for example, would alter baseline conditions of inequality, and thus the "propensity" of individuals within certain groups toward specific behavioral outcomes. But the algorithm can only conceive of social possibilities in line with its technical capabilities, so it ignores such investments as targets of intervention, in favor of discriminatory profiling practices that require the persistence of existing disparities in order to be effective [79, 80]. In this way, the algorithm's "internalist, technologically determinist theory of social change" constrains the potentiality of the individuals whose behavior it claims to understand [78]. Consider, for example, the algorithm used to predict the grades of British high school students during the coronavirus pandemic. A high-performing student at an historically low-performing school could not receive a higher grade than had previously been achieved at their school, regardless of their individual performance [81]. Through its reproduction of historical boundary conditions, the algorithm restricted each student's future potentiality. When decision-makers uncritically accept the "likely future outcomes" predicted by algorithmic models, they narrow their decisional aperture to the permutations of the past, and thereby foreclose the future potentiality of the individual subject.

There are usually three rejoinders to the claim that predictive models foreclose future potentiality. The first is that it is the decisions themselves, rather than the algorithmic predictions, that foreclose future potentiality. It is certainly true that the decision to do "*x*" inherently forecloses the potentiality of "*not-x*" (to the extent that these options are mutually exclusive) but this foreclosure of outcome is different from the narrowing of decisional aperture that occurs with algorithmic predictions. For example, if you have three job applicants, A, B, and C, and you can only choose one, the decision to hire only one employee necessarily forecloses the possibility of hiring all three. But if in the process of deciding *which* applicant to hire, you rely on an algorithmic prediction of which employee is likely to perform the best, and the algorithm has been trained on historical data about the performance of employees like A and B, but it has no data on employees like C, then the algorithm will never recommend C, and that will narrow your own decisional aperture from A, B, and C, to just A and B. That narrowing of aperture will foreclose the unseen potential of applicant C in ways that the decision itself (to hire only one applicant) does not. In this sense, algorithmic predictions foreclose future potentiality in ways that future-oriented decision-making itself does not.

Another common rejoinder is that even if predictive algorithms lack imagination, humans do not, and they are the ultimate decision-makers. But this response ignores the influence of automation bias, or the tendency of human decision-makers to accept, rather than challenge, quantitative assessments [70]. An overburdened decision-maker, searching for ways to make efficient decisions under significant time pressure, may rely uncritically on an algorithmic prediction, treating it as a "fixed" attribute of the decision subject, without considering the ways in which they might disprove the prediction, or disrupt the social conditions underwriting it [56]. Repeated use of a predictive system encourages users to act as if the prediction were true, so that cases in which the prediction is proven wrong (for example, a "high-risk" recidivist never reoffends) are treated as sources of model error, rather than outcomes that should be encouraged [56]. Reliance on predictive tools creates a perverse incentive to support the "correctness" of the algorithm's prediction (and the "rightness" of the decision made upon it) when, in reality, we would prefer for the prediction to be wrong [82]. Human oversight often fails to perform the desired protective functions [83].

A third common rejoinder is that in some circumstances, the imaginative range of a human decision-maker may be just as narrow as that of a predictive algorithm. It is certainly true that all humans are limited by their own experiences. A community organizer, for example, living in an historically Black neighborhood, disproportionately affected by mass incarceration, might be able to imagine a future in which this neighborhood becomes a bustling metropolis, with investments in education, housing, and employment. A White judge, living in an affluent zip code thousands of miles away, might not. And their degrees of imaginative separation would likely influence their perception of the probability of recidivism within this community. But it is "never a reason for adding to injustice that we are already guilty of some" [24]. Not every human decision-maker will cabin an individual's future potentiality within the limits of their own experience in the way that an algorithm is structurally *bound* by the historical data on which it was trained. And so human decision-making (as flawed and inadequate as it may be) is at least capable of accommodating ruptures with the past in a way that algorithms are technically unable.

As future potentiality is algorithmically foreclosed to particular social groups, pattern-based discrimination produces a "seemingly permanent economic underclass," bound on all sides by historical data, and the self-reinforcing loop of predictive profiling [41]. Equivant's risk assessment tool, for example, over-predicts recidivism for Black defendants, and under-predicts recidivism for White defendants [84]. Over time, the extended incarceration of a particular social group generates disproportionality between their share of the offending population, and their share of the carceral population [85]. As their share of the carceral population is erroneously assumed to reflect their share of the offending population, more institutional resources are directed toward that group, increasing this disproportionality [85]. This is consistent with Stevenson and Doleac's finding that risk assessment use can increase racial disparities in sentencing [86]. Stevenson and Doleac observed the sentencing

practices of judges who were most responsive to risk assessment, and found that the probability of incarceration for Black defendants increased by 4 percentage points relative to Whites, and the length of the sentence increased by approximately 17 per cent. Judges were more likely to deviate downward for White defendants with high risk scores, than for Black defendants [86].

Judicial reliance on predictive algorithms has the potential to sustain the actuarial imprisonment of marginalized communities through "data determinism" [87]. Through its unequal distribution of future potentiality, the algorithm splits the future into two racially distinct times: a White time that is "futurally open" (indeterminate), and a non-White time that is "futurally closed" (predetermined) [76]. Charles Mills describes this as the "racialization of time," or the transfer of time from one set of lives to another [88]. The algorithmic administration of populations, or "stochastic governance,"[7] secures the (data) freedom of a minority of elites, while categorizing and disciplining the remaining "risky" majority, whose performance of everyday activities (as consumer, passenger, debtor, patient, guest) is subject to constant, quantitative evaluation [51]. This is how the apparatuses of algorithmic governmentality exert power, not in the present, but in the future, by controlling what we are "likely" to become, and thus who we are allowed to be [52].

3.4 3.4 The Epistemological Inferiority of the Algorithmic Subject

The commensuration of individuals along a statistical distribution erases important qualitative differences between them [89, 90]. These qualitative differences often correspond to aspects of an individual's character, history, and circumstances that resist datafication, and are therefore illegible to quantified systems, which overlook the ways in which they affect the outcome being predicted. Although almost any human experience or characteristic can be "quantified" in some form, many are physically or ethically difficult to observe or record, so that there is limited data from which to construct a sufficiently complex model. Instead, sites of individual "difference" are statistically constructed using labels, categories, and classifications that are meaningful to the model, but may not be meaningful to individuals. Consider, for example, an algorithmic prediction of parenting quality. The algorithm's assessment will incorporate various institutional data points – an eviction notice, a poor credit score, a brush with law enforcement – but it will miss all the other ways in which a parent cares for their child, for which no quantitative data exists [91]. This creates a lopsided situation in which negative institutional data is accorded greater epistemic weight than the unrecorded experience of parenting [92]. The algorithm creates its own metrics of parenting, based on what is accessible to measurement, and thereby invisibilizes the lived experience it claims to represent.

3.5 3.5 The Redistribution of Expressive Power

The datafication associated with algorithmic subjectivity implicates not only the legibility of the legal subject, but the distribution of power between citizen and state. Criminal defendants already suffer from significant expressive disempowerment. Algorithmic predictions exacerbate this asymmetry by excluding the legal subject from the production of knowledge about themselves. Although defendants have multiple opportunities to speak (the right to testify, the right to allocute, and the right to represent themselves), these expressive rights are rarely used [93]. Ninety-five per cent of defendants never go to trial, and of those who do, very few testify. The prevalence of plea bargains erodes the defendant's constitutional rights, and ritualized plea colloquies ("Do you understand the rights you are giving up?" "Yes") legitimate coerced speech [93]. The result is that defendant speech is ordinarily routed through legal counsel, who will convey the defendant's story in terms that are legible to the law [94]. And like any discourse of power, the limited discursive space constituted by legal scripts will constrain the range of subjectivity it can accommodate [95].

Given the infrequency of criminal trials, the right of allocution at sentencing represents a rare opportunity for a defendant to share their story in their own words [96]. Despite the changes that have occurred in criminal law since allocution was introduced over three hundred years ago (including the right to counsel, and the right to testify), courts have recognized that "[t]he most persuasive counsel may not be able to speak for a defendant as the defendant might ... speak for himself" [97]. Accordingly, many state and federal statutes explicitly provide for a right of allocution at sentencing [96]. Allocution is often framed as an opportunity for the defendant to share information that will reduce the severity of the sentence imposed, but it also bears significant non-instrumental value beyond sentence mitigation [96]. The opportunity to speak can have cognitive, dignitary, and participatory benefits for defendants who feel that they have had an opportunity to shape their legal destiny with their own words [93]. Individuals who participate expressively in their own cases may be more likely to accept the final outcome as a result [98].

Allocution may offer systemic benefits as well. When criminal defendants are silenced, and excluded from criminal justice discourse, the institution suffers the loss of their perspective [93]. Defendant silence maintains the ignorance of institutional actors "who never hear the full story about the individuals" they punish, nor the deficiencies of the system they serve [93]. As a result, judges and prosecutors never understand the "social circumstances that breed crime and violence from the perspectives of those who must survive under them" [93]. Where complex and contextualized narratives could illuminate the structural forces that shape individual behavior, public discourse on crime is instead sated with easy stereotypes [99]. This information deficit helps to sustain a democratically illegitimate institution that is shielded from, and unresponsive to, the voices of its subjects [93].

Judicial reliance on predictive algorithms throughout the criminal justice system must be understood as part of this systemic suppression of defendant speech. As judges turn to algorithms for "objective" predictions of individual behavior, the stories of defendants are further lost. Knowledge produced by the legal subject (about their intentions, motivations, and circumstances) is increasingly displaced by the epistemic authority of algorithms. Data "objectivity" is difficult to counter with a personal narrative shared in halting tones by an individual who may not understand, or trust, the judicial process. And predictive algorithms are generally inaccessible to the layperson, even if the veil of trade secrecy is pierced to expose their internal mechanics [100]. When Eric Loomis contested

his algorithmic classification as a high-risk recidivist, the Supreme Court of Wisconsin acknowledged its ignorance about how the classification had been calculated [101]. Nevertheless, the Court held that Loomis' ability to verify his responses to the algorithm's questionnaire, and to challenge the resulting risk score, provided sufficient protection of his due process right to be sentenced on the basis of accurate information [101]. The Court did not interrogate the variables selected by the algorithm, the weights assigned to them, the training data used to construct the model, or the population against which Loomis was compared. Similarly, when Willie Allen Lynch appealed his conviction for the sale of crack cocaine on the basis that he had been misidentified by a facial recognition algorithm, the District Court of Appeal of Florida affirmed his conviction on the basis that the trial result would not have been different if Lynch had had access to the other photographs in the facial recognition database [102]. Despite the impenetrability of many evidentiary technologies, courts continue to indulge them with uncritical deference, exacerbating the power imbalance between the defendant, and the prison industrial complex. Prosecutors have a longstanding duty, affirmed in *Brady v. Maryland*, to disclose potentially exculpatory evidence, but courts have been unwilling to recognize algorithmic tools as meeting the *Brady* standard [103], thereby instantiating the power of private capital over the conditions of human freedom [41]. Inscrutable evidentiary tools erode "the public's sense that the law's fact-finding apparatus is functioning in a somewhat comprehensible way, on the basis of evidence that speaks, at least in general terms, to the larger community that the processes of adjudication must ultimately serve" [82].

4 4 CONCLUSION

As algorithmic epistemology reshapes the legal arena, it is worth asking whether, and why, legal subjectivity matters. Is there a fundamental incompatibility between the algorithmic subjectivity normalized by data capitalism, and the legal subjectivity underpinning a system of coercive interference? Is the actuarial project of algorithmic governance fundamentally antithetical to individualized justice? In answering these questions, it is helpful to conceive of law not only as a system of coercive interference, but as a mechanism for regulating human behavior, and for communicating moral condemnation. Accordingly, the rituals of law, including legal subjecthood, matter not only as devices for achieving certain legal outcomes, but as affirmations of respect for the individual, as an end in themselves, and as a reminder about the kind of society we want to be [82]. To recognize someone as a subject in law is not merely to afford them certain rights and duties, but to communicate a message about their moral value, and to acknowledge their subjective interests as imposing legitimate constraints on their treatment by others [104]. Much of our commitment to democratic values turns on our view of the citizen as a "responsible agent entitled to be praised or blamed depending upon [their] free choice of conduct" [24]. A conception of citizens as alterable, predictable, or manipulable things "is the foundation of a very different social order indeed" [24]. When the basic unit of a liberal society is no longer an autonomous, unknowable individual, but an algorithmic subject, anticipating its own datafication, society is significantly altered. Individual behaviors become traceable and predictable components of surveillant disciplinary outcomes, and actuarial predictions foreclose opportunities for meaningful autonomy. The law is no longer seen as addressing free and equal subjects, but as "managing [the] threat posed by particular categories of subject" [17]. If the subject-as-assemblage is the atomistic social unit, what kind of society do we have?

The shifting epistemology of legal subjectivity presents a unique opportunity to re-examine the dominant liberal conception of the self-determining legal subject. Does this legal subject reflect contemporary understandings of systemic injustice, or evolving norms of collective responsibility based on mutual interdependence? Arguably, no. The bounded, rational, self-determining liberal subject occupies a sphere of autonomy constructed by individual rights [12]. Within this bounded sphere, the liberal subject is protected from threats to its autonomy from Others, including the state. As long as this boundary can be maintained, the liberal subject can remain isolated, and "in control" [12]. The liberal fantasy of autonomy-as-control fosters illusions of independence that can only be sustained through harmful practices of domination [12]. Social interactions necessarily involve affecting, and being affected by, other autonomous individuals. Autonomy cannot be unilaterally possessed or manifested, because it is a quality of human relations [12]. For people to enjoy autonomy, they need to exist in autonomous relations with others. Every individual is embedded within a web of nested relations (intimate, social, cultural, and political) that shape their capacity for autonomy [12]. In this sense, the traditional liberal conception of the self-determining legal subject fails to reflect the realities of interdependence.

The rhetoric of individual rights – Western liberalism's primary means of expressing selfhood – directs our attention to the individual rightholder, rather than the circumstances that shape the lived experience of rights [12]. Rights discourse perpetuates an "alienating and unrealistic individualism," in which the liberal subject engages in ostensibly individuated actions, disembedded from the social context in which those actions are occurring [12]. This directs us to view social problems (such as crime) as solvable through the exercise of individual entitlements. The reality, however, is that the pervasiveness of crime cannot be explained by individual pathology; it is a deeply social phenomenon, sustained by social, cultural, political, and economic relations, that exist beyond the control of any individual [12]. Without altering the structure of relations that sustain criminality, enforcing prohibitions on individual conduct will not reduce crime. Predictive algorithms, however, focus the deterrent gaze of the law on the "criminogenic" features of individuals, rather than the circumstances that shape and constrain their behavior. Algorithms insulate these political choices from scrutiny by reinforcing the narrative that criminal behavior can be both predicted, and prevented, based on assumptions about the ability of algorithms to understand them [13, 83]. This is the same approach used by data capitalists to justify predictive profiling, namely, that the target of intervention must always be individual behavior, rather than social context. This unrelenting examination of individual behavior obscures the effect of historical and structural forces [105], and minimizes the state's responsibility to its citizens on the basis that individuals can access essential resources through "self-surveillant algorithmic adjustment" [48, 106]. This is the neoliberal politics of algorithmic worldmaking [107].

The false security of rights, of limited personal responsibility, helps us to ignore the overwhelming nature of human interconnectedness. But an unrealistic understanding of autonomy – generated by the dominant liberal conception of the "self-determining" legal subject – will generate an unrealistic assignment of responsibility. It is time for our conception of the legal subject to evolve, as our understanding of human interconnectedness also evolves. We need a new paradigm of legal subjecthood, that retains and expresses our fundamental commitment to equality and autonomy, but pays greater attention to the social relations that constitute the legal subject. These are not the data relations of algorithmic subjectivity, but the social, cultural, and political relations that are meaningful to the individual. As algorithmic epistemology encroaches upon the legal arena, we are reminded once again that the greatest political act of law is the (un)making of the legal person [14].

ACKNOWLEDGMENTS

No funding declarations.

REFERENCES

[1] Danielle Citron, Technological Due Process, 85 Wash. U. L. Rev. 1249 (2008).
[2] Joshua Kroll, *et al.*, Accountable Algorithms, 165 University of Pennsylvania Law Review (2017).
[3] Danielle Citron and Frank Pasquale, The Scored Society: Due Process for Automated Predictions, 89 Wash. L. Rev. 1 (2014).
[4] Aziz Z. Huq, A Right to a Human Decision, 106 Va. L. Rev. 611 (2020).
[5] Andrew Tutt, An FDA For Algorithms, 69 Admin. L. Rev. 83 (2017).
[6] Andrew Selbst, An Institutional View Of Algorithmic Impact Assessments, Harvard Journal of Law and Technology (forthcoming), 2021.
[7] Carrie Sanders and James Sheptycki, Policing, crime and 'big data': towards a critique of the moral economy of stochastic governance, 68 Crime Law Soc Change 1-15 (2017).
[8] Lauren Wilcox, Embodying algorithmic war, 48(1) Security Dialogue 11-28 (2017).
[9] Haraway D (1991) A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In: Haraway D, Simians, Cyborgs and Women: The Reinvention of Nature. New York and London: Routledge, 149–181.
[10] Katherine Hayles (1999) How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics. Chicago, IL: University of Chicago Press.
[11] Hayles NK (2005) My Mother was a Computer: Digital Subjects and Literary Texts. Chicago, IL and London: University of Chicago Press.
[12] Jennifer Nedelsky, Law's Relations: A Relational Theory of Self, Autonomy, and Law, Oxford Scholarship Online (2012).
[13] Kelly Hannah-Moffat, The Uncertainties of Risk Assessment: Partiality, Transparency, and Just Decisions, 27 Fed. Sent'g Rep. 244 (2015).
[14] Ngaire Naffine, Who are Law's Persons - From Cheshire Cats to Responsible Subjects , 66 Mod. L. Rev. 346 (2003).
[15] Visa AJ Kurki, A Theory of Legal Personhood, Oxford Scholarship (2019).
[16] Nicola Lacey, In Search of the Responsible Criminal Subject: History, Philosophy and Social Sciences in Criminal Law Theory, 64:3 The Modern Law Review 350 (2001).
[17] Nicola Lacey, The Resurgence of Character: Responsibility in the Context of Criminalization, in Duff, R. A. and Green, Stuart, (eds.) Philosophical Foundations of Criminal Law, Oxford University Press (2011).
[18] Phillip Cary, A brief history of the concept of free will: issues that are and are not germane to legal reasoning, 25 BEHAV. Sci. & L. 165 (2007).
[19] Danielle Kehl, Priscilla Guo, and Samuel Kessler, Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing, Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School (2017).
[20] H. L. A. Hart, Punishment and Responsibility, Oxford: Oxford University Press (1968).
[21] J. M. Balkin, Understanding Legal Understanding: The Legal Subject and the Problem of Legal Coherence, The Yale Law Journal (1993) Vol. 103, No. 1, pp. 105-176.
[22] Claire Finkelstein, The Inefficiency of Mens Rea, 88 Calif. L. Rev. 895 (2000).
[23] Jeremy M. Miller, Mens Rea Quagmire: The Conscience or Consciousness of the Criminal Law, 29 W. St. U. L. Rev. 21 (2001).
[24] Sanford H. Kadish, The Decline of Innocence, 26 Cambridge L.J. 273 (1968).
[25] John W. Meyer, "Myths of socialization and of personality," in Reconstructing Individualism: Autonomy, Individuality, and the Self in Western Thought, ed. Thomas C. Heller, Morton Sosna, and David E. Wellbery (Stanford, CA: Stanford University Press, 1986), 208–221.
[26] Susanna Lindroos-Hovinheimo, Excavating Foundations of Legal Personhood: Fichte on Autonomy and Self-Consciousness. *Int J Semiot Law* 28, 687–702 (2015). https://doi.org/10.1007/s11196-015-9420-2.
[27] Lon L. Fuller, The Morality of Law, (1969) Yale University Press.
[28] Neil MacCormick, My philosophy of law. In The law in philosophical perspectives, (1999) ed. L.J. Wintgens, 121–145. Dordrecht: Springer.
[29] Jeremy Waldron, How Law Protects Dignity, 71(1) The Cambridge Law Journal 200, 208 (2012).
[30] Smith v. Rapid Transit, Inc., 317 Mass. 469, 58 N.E.2d 754 (1945).
[31] Craig Callen, Adjudication and the Appearance of Statistical Evidence, 65 Tulane Law Review 457 (1990-1991).
[32] Thomas Hurka, Why Value Autonomy? 13(3) Social Theory and Practice 361-382 (1987).
[33] Jiwei Ci, Evaluating Agency: A Fundamental Question for Social and Political Philosophy, 42(3) Metaphilosophy 261 (2011).
[34] David Enoch, Levi Spectre and Talia Fisher, Statistical Evidence, Sensitivity and the Legal Value of Knowledge, 40(3) Philosophy & Public Affairs 197 (2012).
[35] David Enoch and Talia Fisher, Sense and "Sensitivity": Epistemic and Instrumental Approaches to Statistical Evidence, 67(3) Stanford Law Review 557 (2015).
[36] George L. Priest, Market Share Liability in Personal Injury and Public Nuisance Litigation: An Economic Analysis, 18 Supreme Court Economic Review 109 (2010).
[37] Søren Kierkegaard, Kierkegaard's Writings, Volume V: Eighteen Upbuilding Discourses, Princeton University Press (2015).
[38] Werner Hamacher, "Disgregation of the Will," in Reconstructing Individualism: Autonomy, Individuality, and the Self in Western Thought, ed. Thomas C. Heller, Morton Sosna, and David E. Wellbery (Stanford, CA: Stanford University Press, 1986), 208–221.
[39] Joseph Raz, The Morality of Freedom, Oxford University Press (1988).
[40] Bernard E. Harcourt, Risk as a Proxy for Race: The Dangers of Risk Assessment, 27 Fed. Sent'g Rep. 237 (2015).
[41] Julie Cohen, Between Truth and Power, Oxford University Press (2019).
[42] Colin Koopman, How We Became Our Data, The University of Chicago Press (2019).
[43] Marion Fourcade and Kieran Healy, Seeing Like a Market, Socio-Economic Review Vol.15, 1, 9-29 (2017).
[44] Rodrigo Ochigame, The Long History of Algorithmic Fairness, Phenomenal World, January 30, 2020.
[45] Dan Bouk, How Our Days Became Numbered: Risk and the Rise of the Statistical Individual, The University of Chicago Press (2015).
[46] Caley Dawn Horan, Actuarial Age: Insurance and the Emergence of Neoliberalism in the Postwar United States, University of Minnesota Doctoral Dissertation, August 2011.
[47] Celia Lury & Sophie Day, Algorithmic Personalization as a Mode of Individuation, Theory, Culture & Society 2019, Vol. 36(2) 17–37.
[48] Dan L. Burk, Algorithmic Legal Metrics, Notre Dame Law Review (forthcoming) (2020), available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3537337.
[49] Mark Kear, Playing the credit score game: algorithms, 'positive' data and the personification of financial objects, Economy and Society 46:3-4 (2017) 346-368.
[50] Mark Ledwich and Anna Zaitsev, Algorithmic Extremism: Examining YouTube's Rabbit Hole of Radicalization, arXiv:1912.11211v1 [cs.SI] 24 December 2019.
[51] Van Dijck, J. (2014). Datafication, dataism and dataveillance: big data between scientific paradigm and ideology. Surveillance and Society, 12(2), 197–208.
[52] Antoinette Rouvroy and Thomas Berns, Algorithmic Governmentality and Prospects of Emancipation, trans. by Elizabeth Libbrecht, Réseaux 2013/1 (No. 177), pp. 163-196.
[53] Bernard Stiegler, Automatic Society, Vol. 1: The Future of Work (2017).
[54] Bernard Stiegler, The Neganthropocene, translated by Daniel Ross, Open Humanities Press (2018).
[55] Salome Viljoen, A Relational Theory of Data Governance, 131(2) The Yale Law Journal 370 (2021).
[56] Barbara D. Underwood, Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgment, The Yale Law Journal, Vol. 88, No. 7 (1979), pp. 1408-1448.
[57] Robert D. Crutchfield, Gregory A. Weeks, The Effects of Mass Incarceration on Communities of Color, 32(1) Issues in Science and Technology (2015).
[58] Mackenzie A (2015) The production of prediction: What does machine learning want? European Journal of Cultural Studies 18(4–5): 429–445.
[59] Paul Robinson, Punishing Dangerousness: Cloaking Preventive Detention as Criminal Justice, 114(5) Harvard Law Review (2001).
[60] John Monahan and Jennifer L. Skeem, Risk Assessment in Criminal Sentencing, 12 Annu. Rev. Clin. Psychol. 489 (2016).
[61] David T. Wasserman, The Morality of Statistical Proof and the Risk of Mistaken Liability, 13 Cardozo L. Rev. 935, 942-43 (1991).

[62] Susanna Lindroos-Hovinheimo, Excavating Foundations of Legal Personhood: Fichte on Autonomy and Self-Consciousness. Int. J. Semiot. Law 28, 687–702 (2015).
[63] Kasper Lippert-Rasmussen, Born Free and Equal? A Philosophical Inquiry into the Nature of Discrimination, Oxford University Press (2014).
[64] Julia Dressel and Hany Farid, The accuracy, fairness, and limits of predicting recidivism, 4 Science Advances (2018).
[65] Lily Hu, Disparate Causes, Phenomenal World, October 17th, 2019.
[66] Regina Austin, The Insurance Classification Controversy, University of Pennsylvania Law School, Faculty Scholarship Paper 1395 (1983).
[67] Barbara Kiviat, The Moral Limits of Predictive Practices: The Case of Credit-Based Insurance Scores, American Sociological Review 2019, Vol. 84(6) 1134–1158.
[68] Douglas Laycock & Teresa A. Sullivan, Sex Discrimination as Actuarial Equality: A Rejoinder to Kimball, 1981 Am. B. Found. Res. J. 221 (1981).
[69] Kristian Lum, David Dunson, and James Johndrow, "Closer than they appear: A Bayesian perspective on individual-level heterogeneity in risk assessment," arXiv: 2102.01135v1 (2021).
[70] Raja Parasuraman and Victor Riley, Humans and Automation: Use, Misuse, Disuse, Abuse, 39(2) Human Factors 230-253 (1997).
[71] Kathleen Mosier and Linda Skitka, "Human Decision Makers and Automated Decision Aids: Made for Each Other?", Chapter in Raja Parasuraman & Mustapha Mouloua (Eds), Automation and Human Performance: Theory and Applications, Lawrence Erlbaum Associates (1996).
[72] Andrew Selbst, *et al.*, Fairness and Abstraction in Sociotechnical Systems, FAT* '19: Proceedings of the Conference on Fairness, Accountability, and Transparency, January 2019.
[73] Karen Yeung, 'Hypernudge': Big Data as a Mode of Regulation by Design, Information, Communication & Society, 20(1) (2017).
[74] Sonja B. Starr, The New Profiling: Why Punishing Based on Poverty and Identity Is Unconstitutional and Wrong, 27 Fed. Sent'g Rep. 229 (2015).
[75] Jennifer Skeem, John Monahan, and Nicholas Scurich, Impact of Risk Assessment on Judges' Fairness in Sentencing Relatively Poor Defendants, 44(1) Law and Human Behavior 51 (2020).
[76] Bonnie Sheehey, Algorithmic Paranoia: The Temporal Governmentality of Predictive Policing, 21 Ethics and Information Technology 49-58 (2019).
[77] Charles Tilly, Durable Inequality, University of California Press (1998).
[78] Ben Green and Salomé Viljoen, Algorithmic Realism: Expanding the Boundaries of Algorithmic Thought, FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, January 2020.
[79] Jenna Burrell and Marion Fourcade, The Society of Algorithms, Annual Review of Sociology (2021) 47:23.1–23.25.
[80] Kasper Lippert-Rasmussen, Born Free and Equal? A Philosophical Inquiry into the Nature of Discrimination, Oxford University Press (2014).
[81] Melissa Fai, Jen Bradley, and Erin Kirker, Lessons in 'Ethics by Design' from Britain's A Level Algorithm, Gilbert + Tobin, September 11, 2020.
[82] Laurence H. Tribe, Trial by Mathematics: Precision and Ritual in the Legal Process, 84(6) Harvard Law Review 1329 (1971).
[83] Ben Green, The Flaws of Policies Requiring Human Oversight of Government Algorithms (September 10, 2021). Available at SSRN: https://ssrn.com/abstract=3921216; or http://dx.doi.org/10.2139/ssrn.3921216.
[84] Julia Angwin and Jeff Larson, "Machine Bias", ProPublica, May 23, 2016.
[85] Bernard Harcourt, Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age, The University of Chicago Press (2007).
[86] Megan T. Stevenson and Jennifer L. Doleac, Algorithmic Risk Assessment in the Hands of Humans (April 21, 2021). Available at SSRN: https://ssrn.com/abstract=3489440; or http://dx.doi.org/10.2139/ssrn.3489440.
[87] Rameriz E. 2013. The privacy challenges of big data: a view from the lifeguard's chair. Technology Policy Institute Aspen Forum, 19 August 2013.
[88] Charles Mills, White time: The chronic injustice of ideal theory, Du Bois Review (2014).
[89] Wendy Espeland, and Mitchell L. Stevens, Commensuration as a Social Process, Annual Review of Sociology 24:313–43 (1998).
[90] Cass Sunstein, Incommensurability and Valuation in Law, 92 Michigan Law Review 779 (1993).
[91] Data & Society and Public Books, Becoming Data Episode 1: Data & Humanity, with Mimi Onuoha, Lam Thuy Vo, and Annie Galvin (May 17, 2021).
[92] Simone Brown, Dark Matters: On The Surveillance of Blackness, Duke University Press (2015).
[93] Alexandra Natapoff, Speechless: The Silencing of Criminal Defendants, 80 N.Y.U. L. Rev. 1449 (2005).
[94] Dragan Milovanovic, Rebellious Lawyering: Lacan, Chaos, and the Development of Alternative Juridico-Semiotic Forms, 20 LEGAL Stud. F. 295 (1996).
[95] Terezie Smejkalová, Legal Performance: Translating into Law and Subjectivity in Law, Tilburg Law Review 22 (2017) 62-76.
[96] Kimberly A. Thomas, Beyond Mitigation: Towards a Theory of Allocution, 75 Fordham L. Rev. 2641 (2007).
[97] Green v. United States, 365 U.S. 301, 304 (1961).
[98] William M. O'Barr & John M. Conley, Litigant Satisfaction Versus Legal Adequacy in Small Claims Court, 19 Law & Soc'y Rev. 661, 662 (1985).
[99] Austin Sarat, Narrative Strategy Death Penalty Advocacy, 31 Harv. C.R.-C.L. L. Rev. 353 (1996).
[100] Jens Halterlein, Epistemologies of predictive policing: Mathematical social science, social physics, and machine learning, Big Data & Society, January-June: 1-13, 2021.
[101] State v. Loomis, 371 Wis.2d 235 (2016).
[102] Lynch v. State, 260 So.3d 1166 (2018).
[103] Brady v. Maryland, 373 U.S. 83 (1963).
[104] Tomasz Pietrzykowski, Personhood Beyond Humanism, Springer (2018).
[105] Martha McCluskey, "Countering Neoliberal Logic With The Vulnerable Human Subject," LPE Project, September 8, 2021.
[106] Lois McNay, Self as Enterprise: Dilemmas of Control and Resistance in Foucault's The Birth of Biopolitics, Theory, Culture & Society (2009), Vol. 26(6): 55-77.
[107] Louise Amoore and Rita Raley, Securing with Algorithms: Knowledge, decision, sovereignty, 48(1) Security Dialogue 3-10 (2017).