2021

Normative Dimensions of Consensual Application of Black Box Artificial Intelligence in Administrative Adjudication of Benefits Claims

Frank Pasquale

NORMATIVE DIMENSIONS OF CONSENSUAL APPLICATION OF BLACK BOX ARTIFICIAL INTELLIGENCE IN ADMINISTRATIVE ADJUDICATION OF BENEFITS CLAIMS

FRANK PASQUALE*

I

INTRODUCTION

Recent calls for administrative austerity have included demands that agencies do more with less as they make decisions about benefit eligibility.[1] This economic logic dovetails with a business case for automating consideration of disputes. The field of computational legal studies suggests ways of deploying natural language processing to triage case filings, or otherwise to find patterns in past adjudications in order to inform (or even complete) the resolution of disputes.[2] For example, a certain combination of medical records may have always led to an award of disability benefits in the past. Administrators may decide to fast track such claims or may even decide to award benefits on the basis of those medical records. Conversely, claims that look too unlike past, successful claims, may be rejected at the outset, ideally with some instructions as to how they may be improved.

In the longer term, more ambitious surveillance programs may feed into administrative adjudications. For example, there are calls in the United States to review the eligibility of benefits recipients via evidence that could include surveillance of their social media feeds.[3] However, using black box AI to deny

---

    1. JULIE COHEN, BETWEEN TRUTH AND POWER 145, 170 (2019) (discussing neoliberal managerialism in the judiciary and administrative state).
    2. *See, e.g.,* MICHAEL A. LIVERMORE AND DANIEL N. ROCKMORE, EDS., LAW AS DATA: COMPUTATION, TEXT, AND THE FUTURE OF LEGAL ANALYSIS (2019); RYAN WHALEN, ED., COMPUTATIONAL LEGAL STUDIES: THE PROMISE AND CHALLENGE OF DATA-DRIVEN RESEARCH (2020).
    3. *See* SOC. SEC. ADMIN., FISCAL YEAR 2019 BUDGET OVERVIEW 17–18 (2018), https://www.ssa.gov/budget/FY19Files/2019BO.pdf [https://perma.cc/7CJH-NZPQ] ("We will study and design successful strategies of our private sector counterparts to determine if a disability adjudicator

benefits is an untested and dangerous proposal. Even when algorithms are transparent, problems arise. For example, Australia's CentreLink agency used defective algorithms and data to mail thousands of letters to claimants demanding return of alleged overpayments. Many were inaccurate, causing a great deal of distress among those who received the demand letters.[4]

Nevertheless, there are promising avenues for automation of law. Tax scholars have argued that as many as 100 million filed tax returns in the United States each year may be unnecessary, wasting millions of hours of tedious form-filling and record-keeping. Instead, the government could simply automatically decide the tax liability of persons who take the standard deduction.[5] With more advanced data, and tax laws written to be machine-readable, even complex returns may be automated.[6] Thus, automated administration offers both promising possibilities and clear warning signs with respect to potential negative consequences.[7]

Black box AI may eventually play an important role in several of the use cases mentioned so far. For purposes of this Article, "black box AI" refers to any natural language processing, machine learning, textual analysis, or similar software which uses data not accessible to the data subject, and/or which deploys algorithms which are either similarly inaccessible, or so complex that they cannot be reduced to a series of rules and rule applications comprehensible to the data subject.[8]

There are many ways in which governments may deploy black box AI in administrative adjudications. For example, they may apply it to all claimants, or only those who consent to its use. Further, they may apply its algorithms only to materials submitted by claimants, or they may sweep a wider set of data into the

---

should access and use social media networks to evaluate disability allegations. Currently, agency adjudicators may use social media information to evaluate a beneficiary's symptoms only when there is an OIG CDI unit's Report of Investigation that contains social media data corroborating the investigative findings. Our study will determine whether the further expansion of social media networks in disability determinations will increase program integrity and expedite the identification of fraud.").

4. *See* Terry Carney, *The New Digital Future for Welfare: Debts Without Legal Proofs or Moral Authority?*, U. NEW S. WALES L.J. F., Mar. 2018, at 1–2. Carney has observed more recently that while automated implementation of welfare schemes may contribute to unfairness, deeper problems in the Australian welfare system may be the root cause of recent scandals. *See* Terry Carney, *Artificial Intelligence in Welfare: Striking the Vulnerability Balance?* 46 MONASH U. L. REV., no. 2, 2020, at 3.

5. *See, e.g.*, MICHAEL J. GRAETZ, 100 MILLION UNNECESSARY RETURNS 103–07 (2008) (arguing for the elimination of most deductions and credits in order to simplify tax liability). Automation is already in widespread use at the IRS. Danielle K. Citron & Ryan Calo, *The Automated Administrative State: A Crisis of Legitimacy*, EMORY L.J. (forthcoming 2021) (manuscript at 11) (on file with author) (describing how the IRS already is using algorithmic processes to decide whom to audit).

6. *See* Sarah B. Lawsky, *A Logic for Statutes*, 21 FLA. TAX REV. 60, 78 (2017) (arguing for default logic based statutes because "[a]rtificial intelligence based on default logic can more easily encode statutes and extract information from statutes than artificial intelligence based on standard logic.").

7. *See* Citron & Calo, *supra* note 5, at 3 (explaining that erroneous rules led to automated systems violating the due process guarantees of beneficiaries in the public benefits arena).

8. This definition is developed in Chapter 1 of FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2015).

black box AI's purview. This Article will typologize key normative concerns raised even in scenarios where black box AI is consented to by the claimant.[9] Part II examines the normative desirability of consent-based black box AI use cases involving materials submitted by the claimant. Part III analyzes use cases involving a wider universe of materials not submitted by the claimant. I conclude with reflections on the important function of reasoned explanation in even mass justice systems.

II

NORMATIVE CONCERNS REGARDING BLACK BOX AI USED TO PROCESS SUBMITTED MATERIALS

A deal between the litigant and the state to permit the state or its instrumentalities to use black box AI to analyze submitted materials will likely have more procedural legitimacy than a deal permitting the use of a wider universe of data. Litigants can control what they submit, taking on responsibility for "quality control" and accuracy. Nevertheless, there are important, often unexpected side effects of such agreements which need further illumination, mitigation, or elimination before they are embraced.

Simply treating a certain class of persons better than others on the basis of black box AI highlights a deep tension between utilitarian ethics and the rule of law. Despite the compelling utilitarian rationale for such automated adjudication of benefits, it is difficult to fit the "determination" at issue into classic administrative law categories of rule or order, fact or law. Is the hypothetical matching based on similarities between past and current applications a legal or factual determination? Perhaps the two can be split apart so that the relevant NLP is limited to only reviewing the facts section of the claimant's submission, or only the law section. Yet this possibility would undermine the legitimacy of the NLP, since there should be some connection between the facts and the law in even the most summary determination.

Further problems arise when considering the significance and purpose behind distinctions made by black box AI. Assume, for example, that the entire submission is reviewed via a comparison with similar, past submissions. The key question then becomes, what is the nature of the similarity discovered?[10] If the similarity is only based on something superficial (say, the number of periods and

---

9. I save for future work an assessment of the imposition of black box AI without consent, or the use of black box AI to preemptively award benefits before a case is actually adjudicated. For a typology of the general types of issues raised by the opacity of data and its processing, and the revisability of judgments, in adjudications, *see* Frank Pasquale & Danielle Keats Citron, *Promoting Innovation While Preventing Discrimination: Policy Goals for the Scored Society*, 89 WASH. L. REV. 1413, 1422 (2014).

10. Perhaps there is no similarity analogous to the types of reasoning done by extant forms of legal analysis. The NLP may be doing something *sui generis*, as Judge Geneviève Vanderstichele has suggested in her illuminating work. *See* Geneviève Vanderstichele, *The Normative Value of Legal Analytics. Is There a Case for Statistical Precedent?* (Aug. 30, 2020) (MPhil thesis, University of Oxford), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3474878 [https://perma.cc/FRY2-2K2W].

punctuation marks in the document, or the presence of certain words with ample synonyms used in other, unsuccessful submissions, *ceteris paribus*), the problem of arbitrariness rears its ugly head.[11] While arbitrary "givings" may not be nearly as problematic as arbitrary takings, they nevertheless raise serious questions about the specifically *legal* nature of the AI in question.[12]

Given these concerns about black box AI relying on arbitrary distinctions, why not simply use NLP to identify the neediest cases based on some algorithm of merit or urgency? This seems like a more plausible goal for NLP than the more complex question of applying law to facts (ramified, as it necessarily is, by theories of interpretation, precedent, and internal guidance on limits to the number of claims that can be granted).[13] An algorithm of neediness may assist as pure description (albeit value-laden) of some dimension of the claimant's case. However, those crafting such algorithms must be very attentive to the advice of members of the relevant communities (here, of poor and disabled persons), lest the system be biased.[14] Moreover, a "neediest case" black box AI developer should be attentive to long-standing debates over just principles of allocation of resources.[15]

---

11. Frank Pasquale & Glyn Cashwell, *Prediction, Persuasion, and the Jurisprudence of Behaviorism*, 68 U. TORONTO L.J. 63, 75 (2018) ("[T]he entire 'predictive' project . . . may be riddled with spurious correlations. As any student of statistics knows, if one tests enough data sets against one another, spurious correlations will emerge.").

12. I use the category "givings" as defined in Abraham Bell & Gideon Parchomovsky, *Givings*, 111 YALE L.J. 547, 549 n.2 (2001).

13. As Yoshua Bengio explains: "[a]nother big challenge is natural language understanding . . . [is] still not at the level where we would say the machine understands. That would be when we could read a paragraph and then ask any question about it, and the machine would basically answer in a reasonable way, as a human would. We are still far from that." Will Knight, *Will Machines Eliminate Us?*, MIT TECH. REV., Jan. 29, 2016, https://www.technologyreview.com/2016/01/29/162084/will-machines-eliminate-us/ [https://perma.cc/5W4R-8A5P].

14. VIRGINIA EUBANKS, AUTOMATING INEQUALITY (1st ed. 2018); Seeta Pe a Gangadharan, *Digital Inclusion and Data Profiling*, 17 FIRST MONDAY (2012), https://doi.org/10.5210/fm.v17i5.3821 [https://perma.cc/Z23L-8UXS]; Michele E. Gilman, *Welfare, Privacy, and Feminism*, 39 U. BALT. L.F. 25 (2008); Rashida Richardson, Jason M. Schultz, & Vincent M. Southerland, *Litigating Algorithms 2019 US Report: New Challenges to Government Use of Algorithmic Decision Systems*, A.I. NOW INST. (2019), https://ainowinstitute.org/litigatingalgorithms-2019-us.pdf [https://perma.cc/E778-JA5M]; *Litigating Algorithms: Challenging Government Use of Algorithmic Decision Systems*, A.I. NOW INST. (2018), https://ainowinstitute.org/litigatingalgorithms.pdf [https://perma.cc/4SKQ-BW9X]; Dillon Reisman, Jason Schultz, Kate Crawford, & Meredith Whittaker, *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*, A.I. NOW INST. (2018), https://ainowinstitute.org/aiareport2018.pdf [https://perma.cc/T2CQ-FMF9]. For an example of the perils apparent here, see K.W. v. Armstrong, 180 F. Supp. 3d 703 (D. Idaho 2016) (a case hinging on application of a widely criticized algorithm to benefits determinations).

15. The diversity of theories of distributive justice, and differences between them, matter, too. For instance, "Amartya Sen illustrates these differences with a story about three children and a flute. One child justifies her claim to the flute because she is the only one of the three who can play and would therefore receive the most pleasure from owning the flute. Another child claims the flute on the ground that he is impoverished and has no toys of his own; the flute would therefore increase his happiness and his share of economic goods. The third child demands the flute because she actually made the flute; she therefore has a right to the flute because it is the product of her own labor. How we resolve the question of which child receives the flute will likely depend on whether we favor utilitarian, economic egalitarian,

Including more "humans in the loop" of decisionmaking may also ameliorate the alienating effects of black box AI. The legal scholar Guido Noto La Diega has categorized several rationales for avoiding the dehumanization of decision-making in law generally, many of which apply here.[16] He gives "good reasons not to trust" algorithms, even when they meet some standard of accuracy with respect to a given dataset:

> First, design choices make the decisionmaking process or the factors it considers too opaque; these choices may also limit the control of the designer. Second, the output of the system may be affected by the biases in data collection. Third, unlike human beings, algorithms cannot balance biases in interpretation of data by a conscious attention to the redress of the bias. Fourth, there are biases in the ways that learning algorithms are tuned based on the testing users' behavior. Fifth, algorithms may be designed for a purpose, but then inserted into systems designed for other purposes. Lastly . . . another factor is the biases in the data used to train the decision-making systems.[17]

To be sure, there are many dedicated attorneys, computer scientists, philosophers, and social scientists now working to address concerns like these.[18] Indeed, the fields of AI ethics and algorithmic accountability will have much to contribute to law in coming decades. However, even if researchers and policymakers manage to address all the concerns mentioned by La Diega, there is still the "role reversibility" problem raised by Stephen Henderson and Kiel Brennan-Marquez.[19] Their argument is that "in a liberal democracy, there must be an aspect of 'role-reversibility' to judgment. Those who exercise judgment should be vulnerable, reciprocally, to its processes and effects."[20] The problem with AI case determination, or even some super-sophisticated robot judge, is that is cannot experience punishment the way that a human being would. Role-reversibility is necessary for "decision-makers to take the process seriously, respecting the gravity of decision-making from the perspective of affected parties." Brennan-Marquez & Henderson derive this ideal from basic principles of self-governance:

> In a democracy, citizens do not stand outside the process of judgment, as if responding, in awe or trepidation, to the proclamations of an oracle. Rather, we are collectively responsible for judgment. Thus, the party charged with exercising judgment—who could, after all, have been any of us—ought to be able to say: *This decision reflects*

---

or libertarian conceptions of justice." Shannon M. Roesler, *Addressing Environmental Injustices: A Capability Approach to Rulemaking*, 114 W. VA. L. REV. 49, 60 (2011) (citing AMARTYA SEN, THE IDEA OF JUSTICE 12–14 (2009)). Even within egalitarian conceptions of merit, which would include a "neediest cases" algorithm, the concept of need may be defined differently by different groups.

16. Guido Noto La Diega, *Against the Dehumanisation of Decision-Making — Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information*, J. INTELL. PROP., INFO. TECH. & E-COM. L., no. 9, 2018, https://www.jipitec.eu/issues/jipitec-9-1-2018/4677 [https://perma.cc/C6ST-PRJ7].

17. *Id.* at 8–9.

18. ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (Jan. 14, 2021) https://facctconference.org/ [https://perma.cc/S9CV-4ECK].

19. *See generally* Kiel Brennan-Marquez & Stephen E. Henderson, *Artificial Intelligence and Role-Reversible Judgment*, 109 J. CRIM. L. & CRIMINOLOGY 137 (2019).

20. *Id.* at 140.

constraints that we have decided to impose on ourselves, and in this case, it just so happens that another person, rather than I, must answer to them. And the judged party—who could likewise have been any of us—ought to be able to say: *This decision-making process is one that we exercise ourselves, and in this case, it just so happens that another person, rather than I, is executing it.*[21]

Thus, for Brennan-Marquez & Henderson, "even assuming role-reversibility will not improve the accuracy of decision-making, it still has intrinsic value."[22]

There are of course some challenges to this position from devotees of robot "personhood." They may claim that a robot programmed to be terrified of being confined or turned off or denied electric power, for example, would be able to "empathize" before it imposed imprisonment, the death penalty, or a benefits denial, on a person. But a silicon-based machine (that is replicable and replaceable) could offer only a simulation of terror (and, *a fortiori*, empathy) that carbon-based, irreplaceable, dependent, rational animals like ourselves actually feel at the prospect of serious disadvantage.[23] Brennan-Marquez and Henderson's recognition of this ontological divide should be foundational for future work on AI and law, as it underscores an ineliminable advantage of human over machine judgment.

Another line of critique would emphasize the distance of many jurists from the experience of the persons' they are passing judgment on. Judges and other legal decisionmakers often lead lives far removed from those subject to their decisions, and this problem is particularly acute in benefits cases, where the decisionmaker usually enjoys a secure, middle class job, and the claimant is financially insecure. Nevertheless, there is still some basic grounding of common experience of such judges and persons judged, which can never be attained by entities that do not share a common biological substrate, which underwrites the experiences of mortality and natality that Hannah Arendt described so well as foundational to the human condition.[24]

Brennan-Marquez & Henderson build on a long tradition of scholarship which focuses on the intrinsic value of legal and deliberative processes, rather than their instrumental value. Their focus runs against the grain of a utilitarian American legal tradition that, while immensely influential for decades, does not exhaust our ethical commitments—and often manages to entirely misconceive them. For example, applications of the U.S. Supreme Court's famous *Mathews v. Eldridge* calculus have frequently failed to take into account the effects of abbreviated procedures on claimants' dignity—what George Kateb describes as a foundational commitment to respect.[25]

---

21. *Id.*

22. *Id.* at 142.

23. FRANK PASQUALE, NEW LAWS OF ROBOTICS: DEFENDING HUMAN EXPERTISE IN THE AGE OF AI (2020), chapter 8; ALASDAIR MACINTYRE, DEPENDENT RATIONAL ANIMALS: WHY HUMAN BEINGS NEED THE VIRTUES (1999).

24. HANNAH ARENDT, THE HUMAN CONDITION (1958).

25. Jerry L. Mashaw, *The Supreme Court's Due Process Calculus for Administrative Adjudication in* Mathews v. Eldridge*: Three Factors in Search of a Theory of Value*, 44 U. CHI. L. REV. 28, 52 (1976)

As Brennan-Marquez and Henderson show, bureaucracies, including the judiciary, have enormous power. They owe litigants a chance to plead their case to someone who can understand and experience, on a visceral level, the boredom and violence portended by a prison stay, the brutal need resulting from the loss of benefits, the sense of shame that liability for drunk driving or pollution can give rise to. And as the classic administrative law case *Morgan v. U.S.* held, even in complex administrative processes, the one who hears must be the one who decides.[26] Brennan-Marquez and Henderson teach that it is not adequate for persons to play mere functionary roles, gathering data for more authoritative machines. Rather, persons must take responsibility for the transparency and explainability of adjudication. To forsake this is to compromise the dignity of claimants.

III

NORMATIVE CONCERNS REGARDING BLACK BOX AI USED TO PROCESS A WIDER UNIVERSE OF MATERIALS NOT SUBMITTED BY THE CLAIMANT

The concerns discussed in Part II apply to the use of black box AI in administrative adjudications of benefits when they are based on materials submitted by claimants. Black box AI that utilizes a wider universe of materials beyond those submitted by claimants poses additional and distinct normative considerations. Nevertheless, at some point in the future, this technology may become a feature of our legal system.

Consider a kind of automated benefit determination where a kiosk may take a claimant's picture and then use facial recognition technology to connect the claimant to all manner of databases of spreadsheet entries about, surveillance camera footage of, and audiorecordings of the claimant.[27] A voice like that of Google Assistant, Siri, or Alexa may state to the claimant that the case will be based not only on written submissions, but also on surveillance of what the claimant has been doing in public and on social media for the past several months. Given the intense scrutiny of factual matters that are part of the standard for claiming disability, a wide variety of data may be canvassed by such AI. It may assess whether the claimant been seen exercising vigorously, or standing and walking for more than four hours in a row. Does the claimant's phone

---

("Decisions with substantial 'moral worth' connotations are generally expected to be highly individualized and attentive to subjective evidence. The adjudication of such issues on the basis of documents submitted largely by third parties and by adjudicators who have never confronted the claimant seems inappropriate. Instead, a court approaching an analysis of the disability claims process from the dignitary perspective might emphasize those aspects of disability decisions that focus on a particular claimant's vocational characteristics, his unique response to his medical condition, and the ultimate predictive judgment of whether the claimant should be able to work."); GEORGE KATEB, HUMAN DIGNITY 1 (2011).

26.   Morgan v. United States, 298 U. S. 468 (1936).

27.   I offer this hypothetical in the spirit of a Dennett-ian "intuition pump." DANIEL C. DENNETT, INTUITION PUMPS AND OTHER TOOLS FOR THINKING 5–7 (W.W. Norton & Co. ed., 1st ed. 2013).

accelerometer and related tracking technology disclose a gait that correlates with the gait of those previously found disabled? How many applications for jobs has the claimant submitted? All these factors may be commensurated and compressed down to a single score to assess likelihood of disability. Should such assessments play a role in administrative adjudications? This hypothetical system, involving black box AI that uses a wide universe of data concerning many aspects of the claimant's life, raises considerations distinct from those raised by black box AI analysis of the claimant's own submissions alone.

The surveillance such a system entails is deeply alienating in two important senses of the term. First, there is a sense of powerlessness, in that this automation furthers "the alienation of man from man, and the degradation of men into commodities."[28] Persons are simply processed, like any other commodity, with observable behavior displacing experience and explanation. In other words, direct person-to-person communication and interpretation are displaced by behaviorist computation.[29] The black box AI may have been programmed by persons, but there is little to no chance of their interaction with the claimant—or the claimant being able to influence the selection and processing of the pivotal information.

The second sense of alienation here is that black box AI can create meaninglessness regarding choices and rules. As sociologist Melvin Seeman explains: "We may speak of high alienation, in the meaninglessness usage, when the individual is unclear as to what he ought to believe—when the individual's minimal standards for clarity in decision-making are not met."[30] Black box AI can create this condition, leaving the objects of its surveillance uncertain, menaced, and second-guessing themselves. "Am I walking too fast," one might ask, "demonstrating that I really am fit for work?" Or "is this website visit too long, some evidence of distraction from jobseeking?"[31] Policymakers should not encourage the development of such all-encompassing self-suspicion.[32]

Ethical problems proliferate as we consider the potential scope of surveillance. First, there is the normalization of data–gathering in spaces that

---

28. Melvin Seeman, *On the Meaning of Alienation,* 24 AM. SOC. REV. 783, 783–84 (1959).

29. On the deep connections between AI and behaviorism, *see* ERIK J. LARSON, THE MYTH OF ARTIFICIAL INTELLIGENCE (2021).

30. *Id.* at 786. *See also* Richard M. Re & Alicia Solow-Niederman, *Developing Artificially Intelligent Justice*, 22 STAN. TECH. L. REV. 242 (2019) (explaining the alienation resulting from AI decision-making and setting forth proposals for alleviating it).

31. For a preview of the type of unexpected data analysis that might occur, consider a recent study which predicted persons' likelihood of being in car accidents from Google Maps images of their homes. KINGA KITA-WOJCIECHOWSKA, & ŁUKASZ KIDZIŃSKI, GOOGLE STREET VIEW IMAGE OF A HOUSE PREDICTS CAR ACCIDENT RISK OF ITS RESIDENT (2019), https://arxiv.org/ftp/arxiv/papers/1904/1904.05270.pdf [https://perma.cc/UQM9-LP8M].

32. *See* MARK ANDREJEVIC, AUTOMATED MEDIA 77 (1st ed. 2020) (examining how threats of surveillance and future punishment subordinate individual rights); Julie E. Cohen, *What is Privacy For?*, 126 HARV. L. REV. 1904, 1912 (2013) (discussing modulation); NEIL M. RICHARDS, WHY PRIVACY MATTERS (2021).

have justifiably been considered improper for outside surveillance. Second, there is a power differential between the state and the claimant. Government should not take advantage of the vulnerability of claimants to upend long-established social expectations of privacy. Otherwise, it may create what is effectively a competition for exposure among claimants, each worried that if they fail to agree to certain forms of surveillance via the black box AI (or convincingly performing in areas where such surveillance is possible), they risk losing benefits.

All these objections may seem to melt before the universal solvent of consent. However, the validity of consent is in question when much of the data and analysis ostensibly consented to is opaque to the claimant. And even if that objection can be overcome by providing some generalized description of the data and analysis, other normative problems emerge at a societal level. The competitive dimensions of "consented" disclosure ensure that social conditions can easily render the proposed deal coercive over time. Once a critical mass of persons has agreed to the big data analysis, to resist surveillance is to risk stigmatizing oneself as a person with something to hide. Legislators need to ensure that there is a common and inalienable right against being surveilled by intrusive new technologies, particularly when some arms of the state may weaponize their existing delays and dysfunction to force claimants into competing for favorable treatment by sacrificing their privacy.[33]

Unfortunately, the law has been trending in the opposite direction in cognate areas. It has required residents of public housing to give up basic rights as a condition of benefits, as Rachel Hannaford has described in *Trading Due Process Rights for Shelter*.[34] When this unfair bargain was challenged in *Department of Housing and Urban Development v. Rucker*,[35] the Supreme Court did not even address the petitioners' unconstitutional conditions claims directly.[36] From a purely contractarian perspective, this perspective makes sense: let the petitioners find other shelter if they do not find the terms of public housing amenable. However, this formalistic view elides the many challenges to dignity such one-sided leases impose on some of the most vulnerable persons in our society. Their marginalization deserves remedial attention, rather than being treated as one more point of vulnerability to be exploited.

---

33. Privacy concerns are particularly relevant given the centrality of the applicant's body in U.S. disability determinations, which consider (inter alia*) whether a claimant has a "severe medically determinable physical or mental impairment" (or combination of impairments). 20 C.F.R. § 416.921 (2019). The centrality of the body in privacy law and theory is a normative dimension of data protection of long standing. *See* ANITA ALLEN, UNPOPULAR PRIVACY: WHAT MUST WE HIDE? (2011).

34. Rachel Hannaford, *Trading Due Process Rights for Shelter:* Rucker *and Unconstitutional Conditions in Public Housing Leases*, 6 U. PA. J. CONST. L. 139 (2003).

35. *See* Dep't. of Hous. and Urban Dev. v. Rucker, 535 U.S. 125 (2002).

36. *See id.* at 136 n.6 (shutting down respondents' attempts to raise constitutional challenges).

IV

CONCLUSION

Despite the threats to privacy it portends, black box AI has already garnered interest from some academics as a method for automating certain legal determinations. Indeed, a recent conference on personalized law at the University of Chicago showcased proposals to dynamically adapt legal requirements for individuals and corporations based on automated or semi-automated analysis of vast quantities of data collected about them.[37] Expectations about the increasing volume, variety, and velocity of even more data about states of the world (and persons) grounds more ambitious visions of a future of law driven by big data.[38] Advanced technology can also promote a combination of approaches, such as smart contracts and personalization. At its limit, this vision replaces rules and standards with "microdirectives," specific requirements fusing factual determinations and legal control of behavior.[39] For example, a traffic authority may alter the speed limit driver by driver, based on drivers' history of accidents and moving violations. Such personalization imports a utilitarian reasoning common in antitrust's measurement of consumer welfare, to law generally.[40] In other words, from this Benthamite perspective, law's primary goal is to optimize some objective function (such as speed and convenience traded off against risk of injury and death).[41] Such an objective function can be extrapolated into the future based on analysis of the past (say, when deaths and injuries reached an unacceptable level, based on the speed of driving of individuals like those who are now being regulated).

The problems with such an approach are twofold. Black box AI could be inaccurate or unfair, and resistant to the usual corrective procedures that have helped mitigate other sources of unfairness and inaccuracy. Moreover, its widespread imposition may be deeply alienating. We are only in the beginning stages of articulating proper channels for collective governance of such scenarios. Without such collective governance, we risk moving seamlessly from the

---

37. *See Symposium on Personalized Law*, 86 CHI. L. REV. 217 (2019) (including various articles on this topic).

38. *See* VICTOR MAYER-SCHÖNBERGER & KENNETH CUKIER, BIG DATA: A REVOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK, AND THINK 174–78 (2013) (discussing how big data can influence how laws are applied and legal decision-making); IAN AYRES, SUPER CRUNCHERS: WHY THINKING-BY-NUMBERS IS THE NEW WAY TO BE SMART 77–78 (2008) (highlighting the use of big data to reveal that longer sentences for criminal defendants do not have an impact on recidivism rates).

39. Anthony Case & Anthony Niblett, *The Death of Rules and Standards,* 92 IND. L.J. 1401, 1401 (2017) ("A microdirective, like a rule, provides a clear instruction to a citizen on how to comply with the law. But, like a standard, a microdirective is tailored to and adapts to each and every context.").

40. *See* WILLIAM DAVIES, THE LIMITS OF NEOLIBERALISM: AUTHORITY, SOVEREIGNTY AND THE LOGIC OF COMPETITION (2015).

41. *See generally,* Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (1789) https://www.econlib.org/library/Bentham/bnthPML.html [https://perma.cc/96MC-F7JD]. Bentham was a pioneer of Utilitarianism, proponents of his work support the development of law and policy that achieves an ideal balance of benefit versus harm as interpreted through measurable criteria.

"juridification of the lifeworld" critiqued by Habermas[42], to a "technification of the lifeworld"—no less alienating in its failure to put human accountability at the heart of state action. The use of black box AI to deny benefits is deeply suspect, especially given the numerous examples of bias now emerging in large language models.[43]

From a utilitarian perspective, it may seem strange to limit the state in this way, foregoing a chance for a judicial supercomputer to cross-correlate millions of variables. Language seems so weak a tool in comparison. However, its limits can be strengths.[44] The burden of writing or speaking word by word ensures a line of thought capable of being comprehended (and challenged) by hearers. This idea is beautifully conveyed in Matthew Lopez's *The Inheritance*, when an imagined E.M. Forster says to an aspiring writer: "All your ideas are at the starting post, ready to run. And yet they must all pass through a keyhole in order to begin the race."[45] Word by word, a text or talk can be understood, agreed with, or disputed. Until the ability to challenge machine learning methods is similarly accessible and democratized, we should be wary of entrusting AI with the evaluation of humans. And even then, hermeneutics are inescapable: much of what counts as genuine data of positive and negative outcomes will be up for debate, ensuring ineliminably human participation in the gathering of the data necessary for black box computation. All these challenges counsel against using black box AI in administrative adjudications of benefit determinations, especially when simpler and more transparent software is still in early stages of deployment.

---

42.   Habermas envisioned a dichotomy in 20th century society that split the conscious social acts and institutions driven by human consent in everyday life—the "lifeworld"—from that of acts and institutions that demand formalized conduct—the "system." Through mounting "juridification of the lifeworld," i.e., the increasing use of laws to govern everyday life, Habernas observed that decisions made for and by people in modern society were becoming increasingly distant from the volition of the people themselves. *See* Bohman, et al., *Jürgen Habermas,* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward Zalta ed., 2017) https://plato.stanford.edu/entries/habermas/ [https://perma.cc/4EFR-DGCS].

43.   EMILY M. BENDER, TIMNIT GEBRU, ANGELINA MCMILLAN-MAJOR, & MARGARET MITCHELL, ON THE DANGERS OF STOCHASTIC PARROTS: CAN LANGUAGE MODELS BE TOO BIG? (2021) https://faculty.washington.edu/ebender/papers/Stochastic_Parrots.pdf [https://perma.cc/6Z59-PS8A].

44.   *See* Frank Pasquale, *Foreword* to IS LAW COMPUTABLE?, at xv (Deakin & Markou, eds., 2020) (arguing that language allows for flexibility a computer code cannot recreate).

45.   MATTHEW LOPEZ, THE INHERITANCE 5–8 (2018).