1-2021

Beyond Transparency and Accountability: Three Additional Features Algorithm Designers Should Build into Intelligent Platforms

Peter K. Yu

BEYOND TRANSPARENCY AND ACCOUNTABILITY: THREE ADDITIONAL FEATURES ALGORITHM DESIGNERS SHOULD BUILD INTO INTELLIGENT PLATFORMS

*By Peter K. Yu**

 Regents Professor of Law and Communication, and Director, Center for Law and Intellectual Property, Texas A&M University. This article was commissioned for the 2020 *Northeastern University Law Review* Symposium entitled "Eyes on Me: Innovation and Technology in Contemporary Times" at Northeastern University School of Law, which was canceled due to the COVID-19 pandemic. The author would like to thank the editors of the *Review*, in particular Amy Hahn and Sarah Odion Esene, for their hard work in preparing for the Symposium and their professionalism in handling the challenging situation. He is also grateful to Ari Waldman for valuable comments and Somer Brown, Mark Hochberg, and Rohan Vakil for helpful editorial suggestions. The article draws on research the author conducted for earlier or forthcoming articles in the *Alabama Law Review* and the *Florida Law Review*.

Table of Contents

Editors' Note

The following article is written by Professor Peter K. Yu, an intended panelist for *Northeastern University Law Review*'s March 2020 Symposium, "Eyes on Me: Innovation and Technology in Contemporary Times," which was unfortunately canceled due to the COVID-19 pandemic. The Symposium was intended to host discussions about the impact of innovation and technology on contemporary legal society; following the event, the editors of the *Law Review* intended to publish three related Symposium pieces. This article, a Symposium piece that has been lengthened for clarity, was written based on Professor Yu's intended remarks at his panel on innovation.

INTRODUCTION

In the age of artificial intelligence (AI), innovative businesses are eager to deploy intelligent platforms to detect and recognize patterns, predict customer choices, and shape user preferences.[1] Yet such deployment has brought along the widely documented problems of automated systems, including coding errors, corrupt data, algorithmic biases, accountability deficits, and dehumanizing tendencies.[2] In response to these problems, policymakers, commentators, and consumer advocates have increasingly called on businesses seeking to ride the artificial intelligence wave to build transparency and accountability into algorithmic designs.[3]

While I am sympathetic to these calls for action and appreciate

1 *See* U.S. Pub. Policy Council, Ass'n for Computing Machinery, *Statement on Algorithmic Transparency and Accountability* 1, ASS'N FOR COMPUTING MACHINERY (Jan. 12, 2017), https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf [hereinafter *ACM Statement*] ("Computer algorithms are [now] widely employed throughout our economy and society to make decisions that have far-reaching impacts, including their applications for education, access to credit, healthcare, and employment."); VIRGINIA EUBANKS, AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR 9 (2017) ("Digital tracking and decision-making systems have become routine in policing, political forecasting, marketing, credit reporting, criminal sentencing, business management, finance, and the administration of public programs."); Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54, 56 (2019) [hereinafter Katyal, *Private Accountability*] ("Today, algorithms determine the optimal way to produce and ship goods, the prices we pay for those goods, the money we can borrow, the people who teach our children, and the books and articles we read—reducing each activity to an actuarial risk or score."); Peter K. Yu, *The Algorithmic Divide and Equality in the Age of Artificial Intelligence*, 72 FLA. L. REV. 331, 332–33 (2020) [hereinafter Yu, *Algorithmic Divide*] ("In the age of artificial intelligence . . . , highly sophisticated algorithms have been deployed to provide analysis, detect patterns, optimize solutions, accelerate operations, facilitate self-learning, minimize human errors and biases, and foster improvements in technological products and services.").

2 *See* ANDREW MCAFEE & ERIK BRYNJOLFSSON, MACHINE, PLATFORM, CROWD: HARNESSING OUR DIGITAL FUTURE 53 (2017) (noting the "biases and bugs" in intelligent machines); Dan L. Burk, *Algorithmic Fair Use*, 86 U. CHI. L. REV. 283, 285 (2019) (listing "ersatz objectivity, diminished decisional transparency, and design biases" among the inherent pitfalls in reliance on algorithmic regulation); Richard M. Re & Alicia Solow-Niederman, *Developing Artificially Intelligent Justice*, 22 STAN. TECH. L. REV. 242, 275 (2019) ("As AI adjudicators play a larger role in the legal system, human participation will change and, in some respects, decrease. Those developments raise the prospect of alienation . . . ."); Peter K. Yu, *Can Algorithms Promote Fair Use?*, 14 FIU L. REV. 329, 335 (2020) [hereinafter Yu, *Fair Use*] (noting "the biases, bugs, and other documented problems now found in automated systems"); Yu, *Algorithmic Divide*, *supra* note 1, at 354–61 (discussing algorithmic discrimination and distortion).

3 *See infra* text accompanying notes 30–39.

the benefits and urgency of building transparency and accountability into algorithmic designs, this article highlights the complications the growing use of artificial intelligence and intelligent platforms has brought to this area. Drawing inspiration from the title "Eyes on Innovation" of my intended panel in the 2020 *Northeastern University Law Review* Symposium,[4] this article argues that owners of intelligent platforms should pay greater attention to three I's: inclusivity, intervenability, and interoperability.

Part I of this article sets the stage with a brief background on the black box designs that have now dominated intelligent platforms. Part II explains why the I in AI has greatly complicated the ongoing efforts to build transparency and accountability into algorithmic designs. Part III identifies three additional I's that owners of intelligent platforms should build into these designs: inclusivity, intervenability, and interoperability. These in-built design features will achieve win-win outcomes that help innovative businesses to be both socially responsible and commercially successful.

4 The canceled 2020 symposium was titled "Eyes on Me: Innovation and Technology in Contemporary Times." *Eyes on Me: Innovation and Technology in Contemporary Times*, Ne. U. L. Rev. (Mar. 21, 2020), http://nulawreview.org/2020-symposium.

I. THE DAMN BLACK BOX

In the age of artificial intelligence, algorithms and machine learning drive the operation of online platforms. Although the term "algorithms" has multiple definitions, ranging from arithmetic methods to computer-based instructions,[5] most discussions in the artificial intelligence context broadly define the term to cover those "self-contained step-by-step set[s] of operations that computers and other 'smart' devices carry out to perform calculation, data processing, and automated reasoning tasks."[6] Whether we

5 As Rob Kitchin observed:

> [Shintaro] Miyazaki traces the term "algorithm" to twelfth-century Spain when the scripts of the Arabian mathematician Muḥammad ibn Mūsā al-Khwārizmī were translated into Latin. These scripts describe methods of addition, subtraction, multiplication and division using numbers. Thereafter, "algorism" meant "the specific step-by-step method of performing written elementary arithmetic" and "came to describe any method of systematic or automatic calculation." By the mid-twentieth century and the development of scientific computation and early high level programming languages, such as Algol 58 and its derivatives (short for ALGOrithmic Language), an algorithm was understood to be a set of defined steps that if followed in the correct order will computationally process input (instructions and/or data) to produce a desired outcome.
>
> From a computational and programming perspective an "Algorithm = Logic + Control"; where the logic is the problem domain-specific component and specifies the abstract formulation and expression of a solution (what is to be done) and the control component is the problem-solving strategy and the instructions for processing the logic under different scenarios (how it should be done). The efficiency of an algorithm can be enhanced by either refining the logic component or by improving the control over its use, including altering data structures (input) to improve efficiency. As reasoned logic, the formulation of an algorithm is, in theory at least, independent of programming languages and the machines that execute them; "it has an autonomous existence independent of 'implementation details.'"

Rob Kitchin, *Thinking Critically About and Researching Algorithms*, 20 INFO. COMM. & SOC'Y 14, 16–17 (2017) (citations omitted); *see also* CHRISTOPHER STEINER, AUTOMATE THIS: HOW ALGORITHMS CAME TO RULE OUR WORLD 53–74 (2012) (providing a brief history of man and algorithms).

6 As the U.S. Public Policy Council of the Association for Computing Machinery defined:

> An algorithm is a self-contained step-by-step set of operations that computers and other "smart" devices carry out to perform calculation, data processing, and automated reasoning tasks. Increasingly, algorithms implement institutional decision-making based on analytics, which involves the discovery, interpretation, and communication of meaningful patterns in data. Especially valuable in areas rich with

notice them or not, algorithms are ubiquitous and have far-reaching impacts on our daily lives. As Pedro Domingos observed in the opening of his best-selling book, *The Master Algorithm*:

> You may not know it, but machine learning is all around you. When you type a query into a search engine, it's how the engine figures out which results to show you (and which ads, as well). When you read your e-mail, you don't see most of the spam, because machine learning filtered it out. Go to Amazon.com to buy a book or Netflix to watch a video, and a machine-learning system helpfully recommends some you might like. Facebook uses machine learning to decide which updates to show you, and Twitter does the same for tweets. Whenever you use a computer, chances are machine learning is involved somewhere.[7]

Thus far, platform owners have carefully protected information concerning algorithmic designs and operations, for reasons such as privacy protection, intellectual property, and platform security and integrity.[8] Frustrated by the "black box" designs that have now dominated intelligent platforms, commentators have widely condemned the continuous lack of algorithmic disclosure.[9] In his widely-cited book, *The Black Box Society*, Frank Pasquale described a black box system as one "whose workings are mysterious; we can observe its inputs and outputs, but we cannot tell how one becomes the other."[10] To him, these "[b]lack boxes embody a paradox

---

> recorded information, analytics relies on the simultaneous application of statistics, computer programming, and operations research to quantify performance.

ACM Statement, *supra* note 1, at 1. For discussions of the transformation provided by the deployment of algorithms, see generally Pedro Domingos, The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World (2015); Steiner, *supra* note 5.

7 Domingos, *supra* note 6, at xi.

8 *See* discussion *infra* text accompanying notes 100–102.

9 *See* Domingos, *supra* note 6, at xvi ("When a new technology is as pervasive and game changing as machine learning, it's not wise to let it remain a black box."); Lee Rainie & Janna Anderson, Code-Dependent: Pros and Cons of the Algorithm Age 19 (2017), https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2017/02/PI_2017.02.08_Algorithms_FINAL.pdf ("There is a larger problem with the increase of algorithm-based outcomes beyond the risk of error or discrimination – the increasing opacity of decision-making and the growing lack of human accountability." (quoting Marc Rotenberg, Exec. Dir., Elec. Privacy Info. Ctr.)). *See generally* Frank Pasquale, The Black Box Society: The Secret Algorithms That Control Money and Information (2015) (providing an excellent and comprehensive discussion of black box systems).

10 Pasquale, *supra* note 9, at 3. For Professor Pasquale, the term "black box" has a second meaning. That meaning focuses on the recording or tracking function, a function that

of the so-called information age: Data is becoming staggering in its breadth and depth, yet often the information most important to us is out of our reach, available only to insiders."[11] Likewise, Virginia Eubanks lamented the suffering of "being targeted by an algorithm: you get a sense of a pattern in the digital noise, an electronic eye turned toward *you*, but you can't put your finger on exactly what's amiss."[12]

As if the inscrutability of these black boxes were not disturbing enough, Kate Crawford and Ryan Calo highlighted their tendency to "disproportionately affect groups that are already disadvantaged by factors such as race, gender and socio-economic background."[13] Cathy O'Neil, who dubbed black box systems "weapons of math destruction," concurred: "[These systems] tend to punish the poor . . . because they are engineered to evaluate large numbers of people. They specialize in bulk, and they're cheap."[14] Even worse, "black box" designs "hid[e] us from the harms they inflict upon our neighbors near and far."[15]

Consider, for example, the following scenario, which has happened to many of us during the COVID-19 pandemic. When you encountered price surges on the platform on which you shopped for food and other basic necessities, you could not tell whether those surges were caused by supply and demand, a pricing algorithm, or other factors.[16] Likewise, when that

is often identified with "the [black boxes or] data-monitoring systems in planes, trains, and cars." *Id.*

11 *Id.* at 191.

12 EUBANKS, *supra* note 1, at 5.

13 Kate Crawford & Ryan Calo, *There Is a Blind Spot in AI Research*, NATURE (Oct. 13, 2016), https://www.nature.com/news/there-is-a-blind-spot-in-ai-research-1.20805; *see also* EUBANKS, *supra* note 1, at 12 ("Automated decision-making shatters the social safety net, criminalizes the poor, intensifies discrimination, and compromises our deepest national values."); ARI EZRA WALDMAN, PRIVACY AS TRUST: INFORMATION PRIVACY FOR AN INFORMATION AGE 28 (2018) ("Black box algorithms . . . discriminate against marginalized groups. Google shows ads for higher paying, more prestigious jobs to men and not to women, ads for arrest records show up more often when searching names associated with persons of color than other names, image searches for 'CEO' massively underrepresent women; and search autocomplete features send discriminatory messages, as when completing the search 'are transgender people' with 'going to hell.'" (footnotes omitted)). *See generally* SAFIYA UMOJA NOBLE, ALGORITHMS OF OPPRESSION: HOW SEARCH ENGINES REINFORCE RACISM (2018) (discussing how search engines promote racism and sexism).

14 CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY 8 (2016) (noting that algorithm-driven automated systems "tend to punish the poor . . . because they are engineered to evaluate large numbers of people").

15 *Id.* at 200.

16 *See* Danielle Wiener-Bronner, *How Grocery Stores Restock Shelves in the Age of Coronavirus*, CNN BUSINESS (Mar. 20, 2020, 3:24 PM), https://www.cnn.com/2020/03/20/

platform informed you about a delay in delivery, you wondered whether the delay was the result of increased shopping orders or an algorithm that prioritized customers in high-spending neighborhoods.

As technology continues to improve and as platforms become more intelligent, online shopping will only become more complicated in the future. The next time you face a pandemic, the platform may automatically deliver food and other basic necessities to you based on your preferences and prior purchases and the behavior of other customers. As part of this delivery, the platform may also include hand sanitizers, household disinfectants, and toilet paper, even if you have not purchased them before. After all, the platform may be intelligent enough to notice the growing demand for those items in your area and therefore make a proactive decision to take care of the platform's repeat customers.

Since the mid-1990s, when the Internet entered the mainstream and online shopping became commonplace, governments introduced a wide array of legislation to protect consumers and their personal data.[17] Although the protection in the United States lagged behind what the European Union offered,[18] policymakers, legislators, and consumer advocates made efforts to ensure that the protection on this side of the Atlantic did not lag too far behind.[19] When the European Union introduced the General

business/panic-buying-how-stores-restock-coronavirus/index.html (reporting about panic shopping and hoarding in the early days of the COVID-19 pandemic).

17 *See generally* WALDMAN, *supra* note 13, at 80–85 (discussing the "notice and choice" regime for privacy protection); Symposium, *Data Protection Law and the European Union's Directive: The Challenge for the United States*, 80 IOWA L. REV. 431 (1995) (providing an excellent collection of articles on data protection and the 1995 EU Data Protection Directive).

18 *See* Council Directive 95/46, art. 12, 1995 O.J. (L 281) 31, 42 (EC) (mandating EU-wide protection of personal data).

19 In response to the 1995 EU Directive and to enable EU-compliant data transfers, the United States negotiated with the European Union for the development of a "safe harbor" privacy framework. *See* Peter K. Yu, *Toward a Nonzero-Sum Approach to Resolving Global Intellectual Property Disputes: What We Can Learn from Mediators, Business Strategists, and International Relations Theorists*, 70 U. CIN. L. REV. 569, 628–34 (2002) (discussing the EU-U.S. negotiation). This framework lasted for more than a decade until October 2015, when the Court of Justice of the European Union found it noncompliant with the Directive. *See* Case C-362/14, Maximillian Schrems v. Data Prot. Comm'r, 2015 EUR-Lex CELEX LEXIS 650 (Oct. 6, 2015) (Grand Chamber) (invalidating the Commission Decision 2000/520 that had found adequate the protection afforded by the safe harbor privacy principles and related frequently asked questions issued by the U.S. Department of Commerce). Since then, the U.S. Department of Commerce introduced the EU-U.S. Privacy Shield Framework, which was designed in conjunction with the European Commission to replace the old "safe harbor" privacy framework. *See* Int'l Trade Admin., U.S. Dep't of Commerce, *EU-U.S. and Swiss-U.S. Privacy Shield Frameworks*, PRIVACY SHIELD FRAMEWORK, https://www.privacyshield.gov/servlet/

Data Protection Regulation (GDPR),[20] which took effect in May 2018, U.S. companies quickly scrambled to respond, fearing that their collection, storage, processing, and utilization of EU-originated data would violate the new regulation.[21]

In the artificial intelligence context, Recital 71 of the GDPR states that the automated processing of personal data "should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision."[22] Articles 13.2(f) and 14.2(g) further require data controllers to provide the data subject with information about "the existence of automated decision-making, including profiling, . . . and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject."[23] Although the nature and coverage of what commentators have called "the right to explanation" remain debatable,[24] the GDPR's emphasis on explainability shows its drafters' keen awareness of the complications

servlet.FileDownload?file=015t0000000QJdg (last visited July 11, 2020). In July 2020, the Court of Justice of the European Union once again invalidated the United States' privacy framework. *See* Case C-311/18, Facebook Ireland Ltd v. Maximillian Schrems, 2020 EUR-Lex CELEX LEXIS 559 (July 16, 2020) (Grand Chamber) (invalidating the Commission Implementing Decision 2016/1250 that had deemed the privacy shield framework to be adequate while leaving intact the Commission Decision 2010/87 on standard contractual clauses for the transfer of personal data to third-country processors). It remains to be seen what new framework the United States will institute.

20 Council Regulation 2016/679, art. 35(1), 2016 O.J. (L 119) 1.

21 *See* Sarah Jeong, *No One's Ready for GDPR*, VERGE (May 22, 2018, 3:28 PM), https://www.theverge.com/2018/5/22/17378688/gdpr-general-data-protection-regulation-eu (reporting that few companies were ready for full compliance with the GDPR); Steven Norton & Sara Castellanos, *Companies Scramble to Cope with New EU Privacy Rules*, CIO J. (Feb. 26, 2018, 6:14 PM), https://blogs.wsj.com/cio/2018/02/26/companies-scramble-to-cope-with-new-eu-privacy-rules/ (reporting the companies' intensive drive to comply with the GDPR).

22 Council Regulation 2016/679, *supra* note 20, recital 71.

23 *Id.* arts. 13.2(f), 14.2(g).

24 For discussions of the so-called right to explanation, see generally Isak Mendoza & Lee A. Bygrave, *The Right Not to Be Subject to Automated Decisions Based on Profiling*, *in* EU INTERNET LAW: REGULATION AND ENFORCEMENT 77 (Tatiani-Eleni Synodinou et al. eds., 2017); Lilian Edwards & Michael Veale, *Slave to the Algorithm: Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking for*, 16 DUKE L. & TECH. REV. 18 (2017); Margot E. Kaminski, *The Right to Explanation, Explained*, 34 BERKELEY TECH. L.J. 189 (2019); Andrew D. Selbst & Julia Powles, *Meaningful Information and the Right to Explanation*, 7 INT'L DATA PRIVACY L. 233 (2017); Bryce Goodman & Seth Flaxman, *European Union Regulations on Algorithmic Decision Making and a "Right to Explanation,"* AI MAG., Fall 2017, at 50.

brought about by the growing use of artificial intelligence and intelligent platforms.[25]

In the United States, the recent years have seen a growing volume of class action lawsuits targeting the unauthorized use of personal data in the artificial intelligence and machine learning contexts,[26] including the use of such data to train algorithms.[27] The Federal Trade Commission has also undertaken investigations or initiated lawsuits in cases involving artificial

25 The GDPR's right to explanation can be traced back to the 1995 EU Data Protection Directive. *See* Edwards & Veale, *supra* note 24, at 20 (noting that a remedy similar to the right to explanation "had existed in the EU Data Protection Directive . . . which preceded the GDPR" (footnote omitted)). Nevertheless, "commentators have now devoted greater energy and effort to understanding this emerging right, due in large part to the increasing need to explain how data are being collected and used in technological platforms that are heavily driven by algorithms." Yu, *Algorithmic Divide*, *supra* note 1, at 377.

26 *See, e.g.*, *In re* Google Assistant Privacy Litig., No. 19-CV-04286-BLF, 2020 WL 2219022 (N.D. Cal. May 6, 2020) (a class action lawsuit users of smart devices brought against Google for the unauthorized recording of conversations by its virtual assistant software and for further disclosure of such conversations); Davey Alba, *A.C.L.U. Accuses Clearview AI of Privacy 'Nightmare Scenario,'* N.Y. Times (June 3, 2020), https://www.nytimes.com/2020/05/28/technology/clearview-ai-privacy-lawsuit.html (reporting the American Civil Liberties Union's privacy lawsuit in Illinois against the facial recognition start-up Clearview AI for the unauthorized collection and use of personal photos found online and on social media); Brian Higgins, *Will "Leaky" Machine Learning Usher in a New Wave of Lawsuits?*, Artificial Intelligence Tech. & L. (Aug. 20, 2018), http://aitechnologylaw.com/2018/08/leaky-machine-learning-models-lawsuits/ (discussing the potential litigation involving the developers of customer-facing artificial intelligence systems that utilized flawed or "leaky" machine learning models).

27 As Amanda Levendowski explained:

> Good training data is crucial for creating accurate AI systems. The AI system tasked with identifying cats must be able [to] abstract out the right features, or heuristics, of a cat from training data. To do so, the training data must be well-selected by humans—training data infused with implicit bias can result in skewed datasets that fuel both false positives and false negatives. For example, a dataset that features only cats with tortoiseshell markings runs the risk that the AI system will "learn" that a mélange of black, orange, and cream markings [is] a heuristic for identifying a cat and mistakenly identify other creatures, like brindle-colored dogs, as cats. Similarly, a dataset that features only mainstream domestic cats could create an AI system that "learns" that cats have fluffy fur, pointy ears, and long tails and fail to identify cats of outlier breeds, like a Devon Rex, Scottish Fold, or Manx. And, in both examples, all manner of wildcats are excluded from the training data.

Amanda Levendowski, *How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem*, 93 Wash. L. Rev. 579, 592 (2018) (footnotes omitted).

intelligence and automated decision-making.[28] In addition, state government officials and legislators have stepped in to enhance consumer and privacy protections in this fast-changing technological environment when they find federal legislation inadequate.[29]

Apart from these efforts, legal commentators have advanced a plethora of promising proposals to address challenges posed by the growing use of artificial intelligence and intelligent platforms. In response to the problems precipitated by black box systems, Professor Pasquale outlined various legal strategies to provide checks against some of the systems' worst abuses while "mak[ing] the case for a new politics and economics of reputation, search, and finance, based on the ideal of an intelligible society."[30] In his new book, *Privacy's Blueprint*, Woodrow Hartzog also advanced "a design agenda for privacy law," explaining why "the design of popular technologies is critical to privacy, and the law should take it more seriously."[31] This agenda is built on the "privacy by design" approach the

28 As the director of the Bureau of Consumer Protection of the Federal Trade Commission (FTC) stated:

> Over the years, the FTC has brought many cases alleging violations of the laws we enforce involving AI and automated decision-making, and have investigated numerous companies in this space. For example, the Fair Credit Reporting Act . . . , enacted in 1970, and the Equal Credit Opportunity Act . . . , enacted in 1974, both address automated decision-making, and financial services companies have been applying these laws to machine-based credit underwriting models for decades. We also have used our FTC Act authority to prohibit unfair and deceptive practices to address consumer injury arising from the use of AI and automated decision-making.

Andrew Smith, *Using Artificial Intelligence and Algorithms*, FED. TRADE COMM'N (Apr. 8, 2020, 9:58 AM), https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms.

29 *See, e.g.*, Rebecca Heilweil, *Illinois Says You Should Know If AI Is Grading Your Online Job Interviews*, VOX (Jan. 1, 2020, 9:50 AM), https://www.vox.com/recode/2020/1/1/21043000/artificial-intelligence-job-applications-illinios-video-interivew-act (reporting the adoption in Illinois of a first-of-its-kind law for regulating the use of certain artificial intelligence tools in video job interviews); Jon Porter, *Vermont Attorney General Is Suing Clearview AI Over Its Controversial Facial Recognition App*, VERGE (Mar. 11, 2020, 8:45 AM), https://www.theverge.com/2020/3/11/21174613/clearview-ai-sued-vermont-attorney-general-facial-recognition-app-database (reporting the State of Vermont Attorney General's litigation against Clearview AI for its unauthorized collection of Vermonters' photos and facial recognition data); *State Artificial Intelligence Policy*, ELECTRONIC PRIVACY INFO. CTR., https://epic.org/state-policy/ai/ (last visited July 11, 2020) (providing information about state artificial intelligence law and policy).

30 PASQUALE, *supra* note 9, at 15; *see also id.* at 140–218 (outlining the legal strategies to curb "black box" abuses and calling for the development of "an intelligible society").

31 WOODROW HARTZOG, PRIVACY'S BLUEPRINT: THE BATTLE TO CONTROL THE DESIGN OF

Federal Trade Commission and other commentators have advocated for since the early 2010s.[32]

Finally, many commentators have underscored the need for greater transparency and accountability in the design and use of algorithms,[33] including the disclosure of technological choices made by algorithm designers.[34] As a group of legal and computer science researchers emphatically stated, "in order for a computer system to function in an accountable way—either while operating an important civic process or merely engaging in routine commerce—accountability must be part of the system's design from the start."[35] Some experts and professional associations have gone even further to call on businesses and organizations deploying automated systems to provide social impact statements[36] or be subject to

---

New Technologies 7 (2018).

32 *See* Fed. Trade Comm'n, Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers 22–34 (2012) (discussing "privacy by design" and the need for companies to "promote consumer privacy throughout their organizations and at every stage of the development of their products and services").

33 *See generally* Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. Pa. L. Rev. 633 (2017) (calling for the development of accountable algorithms); Frank Pasquale, *The Second Wave of Algorithmic Accountability*, L. & Pol. Econ. (Nov. 25, 2019), https://lpeblog.org/2019/11/25/the-second-wave-of-algorithmic-accountability/ (discussing the first and second waves of research on algorithmic accountability).

34 *See* Hannah Bloch-Wehba, *Access to Algorithms*, 88 Fordham L. Rev. 1265, 1295–1306 (2020) (calling for the use of the First Amendment, the Freedom of Information Act, and state equivalents to promote algorithmic transparency and accountability in the public sector); Danielle Keats Citron, *Open Code Governance*, 2008 U. Chi. Legal F. 355, 371–81 (discussing how open code governance would enhance the transparency, democratic legitimacy, and expert quality of automated decisions made by administrative agencies); Sonia K. Katyal, *The Paradox of Source Code Secrecy*, 104 Cornell L. Rev. 1183, 1250–79 (2019) (calling for the controlled disclosure of source code).

35 Kroll et al., *supra* note 33, at 640.

36 *See* Katyal, *Private Accountability*, *supra* note 1, at 111–17 (discussing human impact statements in the artificial intelligence context); Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 Ga. L. Rev. 109, 168–93 (2017) (advancing a regulatory proposal based on the requirement of algorithmic impact statements); Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 Fordham L. Rev. 1085, 1133–38 (2018) (discussing algorithmic impact statements); Nicholas Diakopoulos et al., *Principles for Accountable Algorithms and a Social Impact Statement for Algorithms*, Fairness Accountability & Transparency in Machine Learning, https://www.fatml.org/resources/principles-for-accountable-algorithms (last visited June 13, 2020) (proposing that "algorithm creators develop a Social Impact Statement using the [listed] principles as a guiding structure").

periodic assessments[37] or algorithmic audits.[38] The calls for periodic analyses underscore the need for evaluations at different stages of the design and development process.[39]

37 *See* Council Regulation 2016/679, *supra* note 20, art. 35(1) ("Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data."); *see also* INST. ELEC. & ELEC. ENG'RS, ETHICALLY ALIGNED DESIGN: A VISION FOR PRIORITIZING HUMAN WELL-BEING WITH AUTONOMOUS AND INTELLIGENT SYSTEMS 98 (2017) ("A system to assess privacy impacts related to [autonomous and intelligent systems] needs to be developed, along with best practice recommendations, especially as automated decision systems spread into industries that are not traditionally data-rich."); Lorna McGregor et al., *International Human Rights Law as a Framework for Algorithmic Accountability*, 68 INT'L & COMP. L.Q. 309, 330 (2019) (discussing impact assessments in an algorithmic context); Diakopoulos et al., *supra* note 36 (calling for assessment "(at least) three times during the design and development process: design stage, pre-launch, and post-launch").

38 *See* Deven R. Desai & Joshua A. Kroll, *Trust but Verify: A Guide to Algorithms and the Law*, 31 HARV. J.L. & TECH. 1, 36–42 (2017) (discussing ways to test and evaluate algorithms); Pauline T. Kim, *Auditing Algorithms for Discrimination*, 166 U. PA. L. REV. ONLINE 189, 190–91 (2017) [hereinafter Kim, *Auditing Algorithms*] (discussing the use of audits as a check against discrimination); Yu, *Algorithmic Divide*, *supra* note 1, at 380–82 (discussing the need for algorithmic audits); *Digital Decisions* 11, CTR. FOR DEMOCRACY & TECH., https://cdt.org/files/2018/09/Digital-Decisions-Library-Printer-Friendly-as-of-20180927.pdf ("While explanations can help individuals understand algorithmic decision making, audits are necessary for systemic and long-term detection of unfair outcomes. They also make it possible to fix problems when they arise.").

39 As Lorna McGregor, Daragh Murray, and Vivian Ng explained:

> During the design and development stage, impact assessments should evaluate how an algorithm is likely to work, ensure that it functions as intended and identify any problematic processes or assumptions. This provides an opportunity to modify the design of an algorithm at an early stage, to build in human rights compliance—including monitoring mechanisms—from the outset, or to halt development if human rights concerns cannot be addressed. Impact assessments should also be conducted at the deployment stage, in order to monitor effects during operation . . . .
>
> [T]his requires that, during design and development, the focus should not only be on testing but steps should also be taken to build in effective oversight and monitoring processes that will be able to identify and respond to human rights violations once the algorithm is deployed. This ability to respond to violations is key as [international human rights law] requires that problematic processes must be capable of being reconsidered, revised or adjusted.

McGregor et al., *supra* note 37, at 330; *see also* Diakopoulos et al., *supra* note 36.

II. Transparency and Accountability

Although transparency and accountability remain crucial to consumer and privacy protections—bringing to mind Justice Louis Brandeis's century-old adage that "[s]unlight is said to be the best of disinfectants"[40]—building these features into an environment involving artificial intelligence and machine learning has been difficult. To begin with, algorithmic transparency requires the disclosure of not only the algorithms involved (and the accompanying source code) but also training data and algorithmic outcomes.[41] The disclosure of these outcomes is particularly important because many of them will reenter the intelligent platforms as training or feedback data.[42] The continuous provision of these data will create a self-reinforcing feedback loop that amplifies the "garbage in, garbage out" problem, turning inaccurate, biased, or otherwise inappropriate inputs into faulty outputs.[43] As time passes, the biases generated through these loops

40 Louis D. Brandeis, *What Publicity Can Do*, Harper's Wkly., Dec. 20, 1913, at 10, *reprinted in* Louis D. Brandeis, Other People's Money and How the Bankers Use It 92 (1st ed. 1914).

41 *See* O'Neil, *supra* note 14, at 229 ("We have to learn to interrogate our data collection process, not just our algorithms."); Anupam Chander, *The Racist Algorithm?*, 115 Mich. L. Rev. 1023, 1024–25 (2017) ("What we need . . . is a *transparency of inputs and results*, which allows us to see that the algorithm is generating discriminatory impact."); Kroll et al., *supra* note 33, at 641 ("[W]ithout full transparency—including source code, input data, and the full operating environment of the software—even the disclosure of audit logs showing what a program did while it was running provides no guarantee that the disclosed information actually reflects a computer system's behavior.").

42 Ajay Agrawal, Joshua Gans, and Avi Goldfarb distinguished between three types of data that enter artificial intelligence systems: "Input data is used to power [the machine] to produce predictions. Feedback data is used to improve it . . . . Training data is used at the beginning to train an algorithm, but once the prediction machine is running, it is not useful anymore." Ajay Agrawal et al., Prediction Machines: The Simple Economics of Artificial Intelligence 163 (2018).

43 *See* Sofia Grafanaki, *Autonomy Challenges in the Age of Big Data*, 27 Fordham Intell. Prop. Media & Ent. L.J. 803, 827 (2017) ("[A]lgorithmic self-reinforcing loops are now present across many spheres of our daily life (e.g., retail contexts, career contexts, credit decisions, insurance, Google search results, news feeds) . . . ."); Katyal, *Private Accountability*, *supra* note 1, at 69 ("Bad data . . . can perpetuate inequalities through machine learning, leading to a feedback loop that replicates existing forms of bias, potentially impacting minorities as a result."); Ronald Yu & Gabriele Spina Alì, *What's Inside the Black Box? AI Challenges for Lawyers and Researchers*, 19 Legal Info. Mgmt. 2, 4 (2019) ("[T]here is a strong risk that AI may reiterate and even amplify the biases and flaws in datasets, even when these are unknown to humans. In this sense, AI has a self-reinforcing nature, due to the fact that the machine's outputs will be used as data for future algorithmic operations."); *Digital Decisions*, *supra* note 38, at 8 ("Unreliable or unfair decisions that go unchallenged can contribute to bad feedback loops, which can make algorithms even more likely to marginalize vulnerable populations.").

will become much worse than the biases found in the original algorithmic designs or initial training data.

Worse still, it remains unclear if the full disclosure of all the information involved in the algorithmic designs and operations will allow users or consumer advocates to identify the problem. For example, such disclosure may result in an unmanageable deluge of information, making it very difficult, if not impossible, for the public to understand how data are used and how intelligent platforms generate outcomes.[44] Many commentators have also lamented how the public often finds source code and training data incomprehensible.[45] How many platform users or consumer advocates can actually understand algorithmic designs and operations by scrutinizing the source code and datasets involved? Even for those with the requisite skills to handle computer code and technical data, analyzing all the disclosed information will require considerable time, effort, resources, and energy.[46]

---

44 *See* JULIE E. COHEN, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF INFORMATIONAL CAPITALISM 180 (2019) ("In the era of information overload, . . . more comprehensive disclosures do not necessarily enhance understanding."); Maayan Perel & Niva Elkin-Koren, *Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement*, 69 FLA. L. REV. 181, 194–96 (2017) (discussing the problem of having too much information about algorithmic designs and operations). *See generally* OMRI BEN-SHAHAR & CARL E. SCHNEIDER, MORE THAN YOU WANTED TO KNOW: THE FAILURE OF MANDATED DISCLOSURE (2017) (discussing the limitations of mandatory disclosure requirements).

45 *See* RAINIE & ANDERSON, *supra* note 9, at 19 ("Only the programmers are in a position to know for sure what the algorithm does, and even they might not be clear about what's going on. In some cases there is no way to tell exactly why or how a decision by an algorithm is reached." (quoting Doc Searls, Dir., Project VRM, Berkman Klein Ctr. for Internet & Soc'y, Harv. Univ.)); Chander, *supra* note 41, at 1040 ("[T]he algorithm may be too complicated for many others to understand, or even if it is understandable, too demanding, timewise, to comprehend fully."); Kroll et al., *supra* note 33, at 638 ("The source code of computer systems is illegible to nonexperts. In fact, even experts often struggle to understand what software code will do, as inspecting source code is a very limited way of predicting how a computer program will behave."); Guido Noto La Diega, *Against the Dehumanisation of Decision-Making—Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information*, 9 J. INTELL. PROP. INFO. TECH. & ELECTRONIC COM. L. 3, 23 (2018) (suggesting that "a technical document which includes the algorithm used and the mere explanation of the logic in mathematical terms will not in itself meet the legal requirement [for the right to explanation]" and that this requirement "should be interpreted as the disclosure of the algorithm with an explanation in non-technical terms of the rationale of the decision and criteria relied upon").

46 *See* Yu, *Algorithmic Divide*, *supra* note 1, at 375 ("[I]t can be cost-prohibitive to collect or disclose all algorithmic outcomes, not to mention the lack of incentives for technology developers to reveal the algorithms used or to make algorithmic outcomes available for public scrutiny."); *see also* Perel & Elkin-Koren, *supra* note 44, at 195–96 ("[A]nalyzing th[e] overflow of disclosed data in itself requires algorithmic processing that is capable

Such analysis will therefore be cost-prohibitive and difficult to conduct, except for individual projects.

Nevertheless, some target investigations have provided revealing analyses. One such analysis concerns ProPublica's widely cited exposé on COMPAS, the highly controversial scoring software used by law enforcement and correction personnel to determine risks of recidivism.[47] This investigatory report showed, shockingly, that "black defendants were far more likely than white defendants to be incorrectly judged [by the software] to be at a higher risk of recidivism, while white defendants were more likely than black defendants to be incorrectly flagged as low risk."[48]

For intelligent platforms using learning algorithms or neural networks,[49] it has become even more challenging to analyze the algorithmic operations. Because key parts of these operations come from what the platforms have learned on their own, a careful analysis of the original source code is unlikely to provide the explanations needed to fully understand the operations. As Kartik Hosanagar and Vivian Jair observed:

> [M]achine learning algorithms – and deep learning algorithms in particular – are usually built on just a few hundred lines of code. The algorithms['] logic is mostly learned from training data and is

---

of turning the data into meaningful information. Yet this creates a vicious cycle: More transparency only strengthens users' dependence on algorithms, which further increases the need to ensure adequate accountability of the algorithms themselves." (footnote omitted)).

47 *See* Jeff Larson et al., *How We Analyzed the COMPAS Recidivism Algorithm*, ProPublica (May 23, 2016), https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. COMPAS stands for "Correctional Offender Management Profiling for Alternative Sanctions." *Id.*

48 *Id.*

49 As a government report on artificial intelligence explained:

> Deep learning uses structures loosely inspired by the human brain, consisting of a set of units (or "neurons"). Each unit combines a set of input values to produce an output value, which in turn is passed on to other neurons downstream. For example, in an image recognition application, a first layer of units might combine the raw data of the image to recognize simple patterns in the image; a second layer of units might combine the results of the first layer to recognize patterns-of-patterns; a third layer might combine the results of the second layer; and so on.

Comm. on Tech., Nat'l Sci. & Tech. Council, Preparing for the Future of Artificial Intelligence 9 (2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf. For discussions of deep learning, learning algorithms, and neural networks, see generally Ethem Alpaydin, Machine Learning: The New AI 104–09 (2016); John D. Kelleher, Deep Learning (2019); John D. Kelleher & Brendan Tierney, Data Science 121–30 (2018); Thierry Poibeau, Machine Translation 181–95 (2017).

> rarely reflected in its source code. Which is to say, some of today's best-performing algorithms are often the most opaque.[50]

Anupam Chander concurred: "[I]n the era of self-enhancing algorithms, the algorithm's human designers may not fully understand their own creation: even Google engineers may no longer understand what some of their algorithms do."[51]

Given these disclosure challenges, it is no surprise that many commentators, technology experts, and professional organizations have advocated the more active development of explainable artificial intelligence to help document algorithmic analyses and training processes.[52] As Pauline Kim explained:

> When a model is interpretable, debate may ensue over whether its use is justified, but it is at least possible to have a conversation about whether relying on the behaviors or attributes that drive the outcomes is normatively acceptable. When a model is not interpretable, however, it is not even possible to have the conversation.[53]

In sum, the myriad challenges identified in this Part highlight the difficulty in promoting transparency and accountability in the age of artificial intelligence. While building these features into intelligent platforms remains highly important and urgently needed, it will take time and require additional support. The next Part therefore calls on innovative businesses to build additional, and often complementary, features into algorithmic designs if they are to better protect consumers and be more socially responsible.

---

50 Kartik Hosanagar & Vivian Jair, *We Need Transparency in Algorithms, but Too Much Can Backfire*, HARV. BUS. REV. (July 23, 2018), https://hbr.org/2018/07/we-need-transparency-in-algorithms-but-too-much-can-backfire.

51 Chander, *supra* note 41, at 1040 (citing Barry Schwartz, *Google's Paul Haahr: We Don't Fully Understand RankBrain*, SEARCH ENGINE ROUNDTABLE (Mar. 8, 2016, 7:55 AM), https://www.seroundtable.com/google-dont-understand-rankbrain-21744.html).

52 *See ACM Statement*, *supra* note 1, Princ. 4, at 2 ("Systems and institutions that use algorithmic decision-making are encouraged to produce explanations regarding both the procedures followed by the algorithm and the specific decisions that are made."); INST. ELEC. & ELEC. ENG'RS, *supra* note 37, at 68 (calling on software engineers to "document all of their systems and related data flows, their performance, limitations, and risks," with emphases on "auditability, accessibility, meaningfulness, and readability"); Diakopoulos et al., *supra* note 36 ("Ensure that algorithmic decisions as well as any data driving those decisions can be explained to end-users and other stakeholders in non-technical terms.").

53 Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857, 922–23 (2017) [hereinafter Kim, *Data-Driven Discrimination*].

III. The Three I's

In view of the ongoing challenge of building transparency and accountability into intelligent platforms, this Part calls on innovative businesses to build three additional design features into their platforms. This Part discusses each feature in turn and explains why these features can operate in tandem to enhance consumer protection while enabling the fulfilment of corporate social responsibility. Even better, the features can help platform owners achieve win-win outcomes that make good business sense.

Although these design features are usually classified as technology-based alternatives, extra-legal measures, or private self-regulation, they complement those legal and regulatory measures and proposals explored in Part I.[54] In the age of artificial intelligence, legal and technological measures will go hand-in-hand,[55] similar to how privacy designs and practices have not only been required by laws and regulations but have also informed and inspired new legal and regulatory developments.[56]

A. *Inclusivity*

The first proposed design feature targets the biases and discrimination—usually unintentional—found in algorithmic designs and

54 *See supra* text accompanying notes 22–39.

55 As Roger Brownsword observed:

> To the extent that technological management coexists with legal rules, while some rules will be redirected, others will need to be refined and revised. Accordingly, . . . the destiny of legal rules is to be found somewhere in the range of redundancy, replacement, redirection, revision and refinement.

Roger Brownsword, Law, Technology and Society: Re-Imagining the Regulatory Environment 181 (2019).

56 *See* Fed. Trade Comm'n, *supra* note 32, at i (calling on Congress "to consider enacting baseline privacy legislation and . . . data security legislation" while also "urg[ing] industry to accelerate the pace of self-regulation"); Hartzog, *supra* note 31, at 8 ("At base, the design of information technologies can have as much impact on privacy as any tort, regulation, or statute regulating the collection, use, or disclosure of information."); Waldman, *supra* note 13, at 4–5 ("[W]e should conceptualize information privacy in terms of relationships of trust and leverage law to protect those relationships."); *see also* Peter K. Yu, *Teaching International Intellectual Property Law*, 52 St. Louis U. L.J. 923, 939 (2008) ("As [technological and legal protections] interact with each other, and improve over time, they result in a technolegal combination that is often greater than the sum of its parts. It is therefore important to understand not only law and technology, but also the interface between the two.").

operations. In an environment involving artificial intelligence and machine learning, fostering inclusivity will require efforts to promote diversity in not only product choices and platform experiences but also training data.[57] Unless businesses deploying intelligent platforms have utilized sufficiently diverse datasets to train these platforms, the training and subsequent feedback will likely perpetuate the many historical biases found in the offline world[58] and will thereby generate what Sandra Mayson has termed the "bias in, bias out" phenomenon.[59]

Although different ways exist to make algorithmic designs and operations inclusive, commentators have widely underscored the desperate need to address the lack of diversity in the technology workforce.[60] As Justin

57 As I noted in an earlier article:

> [A]ddressing algorithmic distortion—and, to an equal extent, algorithmic discrimination—requires the development of a more inclusive environment. Such an environment needs to be diverse not only in terms of those designing algorithms and related technological products and services but also in terms of the training and feedback data that are being fed into the algorithms. The lack of diversity in either direction will likely perpetuate the many historical biases that originate in the offline world.

Yu, *Algorithmic Divide*, *supra* note 1, at 367–68 (footnote omitted); *see also* U.N. SEC'Y-GEN.'S HIGH-LEVEL PANEL ON DIGITAL COOPERATION, THE AGE OF DIGITAL INTERDEPENDENCE, 29–30 (2019) (underscoring the importance of developing "[a]n inclusive digital economy and society"); MEREDITH BROUSSARD, ARTIFICIAL UNINTELLIGENCE: HOW COMPUTERS MISUNDERSTAND THE WORLD 154 (2018) ("Th[e] willful blindness on the part of some technology creators is why we need inclusive technology . . . .").

58 *See* Katyal, *Private Accountability*, *supra* note 1, at 79 ("[W]hen algorithms train on imperfect data, or are designed by individuals who may be unconsciously biased in some manner, the results often reflect these biases, often to the detriment of certain groups."); Kim, *Data-Driven Discrimination*, *supra* note 53, at 861 ("Algorithms that are built on inaccurate, biased, or unrepresentative data can in turn produce outcomes biased along lines of race, sex, or other protected characteristics."); Tal Z. Zarsky, *Understanding Discrimination in the Scored Society*, 89 WASH. L. REV. 1375, 1392–94 (2014) (discussing the reliance on tainted datasets and data collection methods).

59 Sandra G. Mayson, *Bias In, Bias Out*, 128 YALE L.J. 2218, 2218 (2019).

60 As Amy Webb, CEO of the Future Today Institute, declared:

> The only way to address algorithmic discrimination in the future is to invest in the present. The overwhelming majority of coders are white and male. Corporations must do more than publish transparency reports about their staff – they must actively invest in women and people of color, who will soon be the next generation of workers. And when the day comes, they must choose new hires both for their skills and their worldview. Universities must redouble their efforts not only to recruit a diverse body of students – administrators and faculty must support them through to graduation. And not just students. Universities must diversify their faculties, to ensure that students see themselves reflected in their

Reich, the executive director of the MIT Teaching Systems Lab, reminded us, "[t]he algorithms will be primarily designed by white and Asian men – with data selected by these same privileged actors – for the benefit of consumers like themselves."[61] Likewise, Andrea Matwyshyn lamented, "[s]oftware reflects the biases of its creators, and tends to be biased in favor of what are perceived by many to be boys' interests."[62] Indeed, the gender and minority gap in the technology community has been so enormous and notorious that Kate Crawford referred to this gap as artificial intelligence's "white guy problem."[63] As she explained:

> Like all technologies before it, artificial intelligence will reflect the values of its creators. So inclusivity matters – from who designs it to who sits on the company boards and which ethical perspectives are included. Otherwise, we risk constructing machine intelligence that mirrors a narrow and privileged vision of society, with its old, familiar biases and stereotypes.[64]

Given the lack of inclusivity in the technology community, businesses deploying intelligent platforms should actively promote diversity in their workforce. Such promotion will provide at least two benefits. First, a more diverse workforce will enable businesses to come up with new products and

---

teachers.

Rainie & Anderson, *supra* note 9, at 23 (quoting Amy Webb, Chief Exec. Officer, Future Today Inst.).

61 Rainie & Anderson, *supra* note 9, at 12; *see also* Brad Smith & Carol Ann Browne, Tools and Weapons: The Promise and the Peril of the Digital Age 184–85 (2019) ("At most tech companies, women still represent less than 30 percent of the workforce, and an even lower percentage of technical roles. Similarly, African Americans, Hispanics, and Latinos typically account for less than half of what one would expect based on their representation in the American population."); Mariya Yao, *Fighting Algorithmic Bias and Homogenous Thinking in A.I.*, Forbes (May 1, 2017), https://www.forbes.com/sites/mariyayao/2017/05/01/dangers-algorithmic-bias-homogenous-thinking-ai ("When Timnit Gebru attended a prestigious AI research conference last year, she counted 6 black people in the audience out of an estimated 8,500. And only one black woman: herself.").

62 Andrea M. Matwyshyn, *Silicon Ceilings: Information Technology Equity, the Digital Divide and the Gender Gap Among Information Technology Professionals*, 2 Nw. J. Tech. & Intell. Prop. 35, 55 (2003) (footnote omitted).

63 Kate Crawford, *Artificial Intelligence's White Guy Problem*, N.Y. Times (June 25, 2016), https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html.

64 *Id.*; *see also* Katyal, *Private Accountability*, *supra* note 1, at 59 ("[A]lgorithmic models are . . . the product of their fallible creators, who may miss evidence of systemic bias or structural discrimination in data or may simply make mistakes. These errors of omission—innocent by nature—risk reifying past prejudices, thereby reproducing an image of an infinitely unjust world." (footnote omitted)).

services that improve platform experiences while expanding the customer base.[65] Because customers make different use of communication technologies, intelligent platforms, and smart devices,[66] having algorithm designers who understand, or are sensitive to, varied usage patterns will ensure the development of a wider array of products, services, and experiences.

Second, a more diverse workforce will enable algorithm designers to quickly spot problems that may seem odd from an engineering standpoint but are quite obvious when viewed through a social or socioeconomic lens. A case in point is the problem Amazon encountered when it rolled out same-day delivery services for its Prime members in several major cities.[67] Because the tech giant had deployed an algorithm that prioritized areas with "high concentration[s] of Prime members," its new service became unavailable in ZIP codes that had predominantly Black or Hispanic neighborhoods.[68] Anybody familiar with those neighborhoods would be quick to point out the different demographics involved and how an algorithmic focus on member concentration would ignore many current and potential customers living in

65 *See* RAINIE & ANDERSON, *supra* note 9, at 57–60 (collecting views on how "algorithms reflect the biases of programmers and datasets").

66 *See* GEOFFREY G. PARKER ET AL., PLATFORM REVOLUTION: HOW NETWORKED MARKETS ARE TRANSFORMING THE ECONOMY AND HOW TO MAKE THEM WORK FOR YOU 35 (2016) ("Platforms are complex, multisided systems that must support large networks of users who play different roles and interact in a wide variety of ways.").

67 As a *Bloomberg* report described:

> In Atlanta, Chicago, Dallas, and Washington, cities still struggling to overcome generations of racial segregation and economic inequality, black citizens are about half as likely to live in neighborhoods with access to Amazon same-day delivery as white residents.
>
> The disparity in two other big cities is significant, too. In New York City, same-day delivery is available throughout Manhattan, Staten Island, and Brooklyn, but not in the Bronx and some majority-black neighborhoods in Queens. In some cities, Amazon same-day delivery extends many miles into the surrounding suburbs but isn't available in some ZIP codes within the city limits.
>
> The most striking gap in Amazon's same-day service is in Boston, where three ZIP codes encompassing the primarily black neighborhood of Roxbury are excluded from same-day service, while the neighborhoods that surround it on all sides are eligible.

David Ingold & Spencer Soper, *Amazon Doesn't Consider the Race of Its Customers. Should It?*, BLOOMBERG (Apr. 21, 2016), https://www.bloomberg.com/graphics/2016-amazon-same-day/; *see also* Omer Tene & Jules Polonetsky, *Taming the Golem: Challenges of Ethical Algorithmic Decision-Making*, 19 N.C. J.L. & TECH. 125, 155–56 (2017) (discussing the problem with Amazon Prime).

68 Ingold & Soper, *supra* note 67.

the excluded neighborhoods. While a less inclusive but observant group of algorithm designers might still reach the same conclusion in the end, doing so would take more time, not to mention the group members' more limited ability to draw on their own personal experiences to develop appropriate solutions.

To be sure, it will take time to develop a technology workforce that is sufficiently diverse to push for products and services that would accommodate the needs and interests of a wide variety of platform users. Factors such as workplace hierarchy, peer pressure, inertia, and cost-effectiveness will not only continue to affect platform decisions but may also militate against the pro-diversity efforts.[69] Nevertheless, building inclusivity into algorithmic designs and operations will remain highly important, especially when the user base continues to grow and diversify.[70] As Microsoft President Brad Smith and his colleague rightly reminded us: "[I]n a world where today's hits quickly become yesterday's memories, a tech company is only as good as its next product. And its next product will only be as good as the people who make it."[71]

---

69 As Ruha Benjamin illustrated with an example concerning the decision not to focus on African Americans in the development of a speech recognition app for Siri, a virtual assistant program:

> [T]he Siri example helps to highlight how just having a more diverse team is an inadequate solution to discriminatory design practices that grow out of the interplay of racism and capitalism. Jason Mars, a Black computer scientist, expressed his frustration saying, "There's a kind of pressure to conform to the prejudices of the world . . . It would be interesting to have a black guy talk [as the voice for his app], but we don't want to create friction, either. First we need to sell products."

RUHA BENJAMIN, RACE AFTER TECHNOLOGY: ABOLITIONIST TOOLS FOR THE NEW JIM CODE 28–29 (2019) (alteration in original). As she continued: "by focusing mainly on individuals' identities and overlooking the norms and structures of the tech industry, many diversity initiatives offer little more than cosmetic change, demographic percentages on a company pie chart, concealing rather than undoing the racist status quo." *Id.* at 61–62.

70 *See Black Impact: Consumer Categories Where African Americans Move Markets*, NIELSEN (Feb. 15, 2018), https://www.nielsen.com/us/en/insights/article/2018/black-impact-consumer-categories-where-african-americans-move-markets/ ("Black consumers are speaking directly to brands in unprecedented ways and achieving headline-making results."); Sarah Cavill, *The Spending and Digital Habits of Black Consumers Present Opportunities for Marketers*, DIGITAL MEDIA SOLUTIONS (Feb. 27, 2019), https://insights.digitalmediasolutions.com/articles/black-consumers-digital (discussing the changing spending and digital habits of African American customers and how these changes have created new market opportunities).

71 SMITH & BROWNE, *supra* note 61, at 169.

B. *Intervenability*

The second proposed design feature responds to ill-advised decisions generated by algorithms and intelligent platforms. Part II underscored the importance of conducting periodic assessments and making public disclosure of relevant information, including algorithms, training data, and algorithmic outcomes.[72] This Section turns to the need for operators of intelligent platforms to be ready to intervene when things go wrong.[73] Such intervention is particularly important considering that humans are known to have made better decisions than machines in many situations, especially unprecedented ones.[74] As Anthony Casey and Anthony Niblett reminded us:

> Algorithmic decision-making does not mean that humans are shut out of the process. Even after the objective has been set, there is much human work to be done. Indeed, humans are involved in all stages of setting up, training, coding, and assessing the merits of the algorithm. If the objectives of the algorithm and the objective of the law are perfectly aligned at the ex ante stage, one must ask: Under what circumstances should a human ignore the algorithm's suggestions and intervene *after* the algorithm has made the decision?[75]

72 *See* discussion *supra* Part II.

73 *See* Council Regulation 2016/679, *supra* note 20, art. 22(3) (requiring data controllers to "implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest [a decision based solely on automated processing, including profiling]"); SARAH T. ROBERTS, BEHIND THE SCREEN: CONTENT MODERATION IN THE SHADOWS OF SOCIAL MEDIA 61 (2019) (noting that human intervention is a "key . . . part of the production chain in sites that rely on user-generated uploaded content requiring screening"); Yu, *Fair Use*, *supra* note 2, at 356 ("Although automation enhances efficiency and effectiveness, human intervention can be highly beneficial."). *See generally* Aziz Z. Huq, *A Right to a Human Decision*, 106 VA. L. REV. 611 (2020) (discussing whether individuals have a "right to a human decision"); Meg Leta Jones, *The Right to a Human in the Loop: Political Constructions of Computer Automation and Personhood*, 47 SOC. STUD. SCI. 216 (2017) (tracing the historical roots of "the right to a human in the loop" back to rights that protect the dignity of data subjects).

74 *See* AGRAWAL ET AL., *supra* note 42, at 59 (noting the weaknesses of machines in making predictions "when there is too little data" and concerning "events that are not captured by past experience"); Frank Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, 87 GEO. WASH. L. REV. 1, 53 (2019) ("Many past efforts to rationalize and algorithmatize the law have failed, for good reason: there is no way to fairly extrapolate the thought processes of some body of past decisionmaking to *all* new scenarios.").

75 Anthony J. Casey & Anthony Niblett, *A Framework for the New Personalization of Law*, 86 U. CHI. L. REV. 333, 354 (2019).

A case in point is the debacle confronting Uber when a gunman took seventeen hostages at the Lindt Chocolate Café in Sydney, Australia in December 2014.[76] Because many people were trying to simultaneously flee the Central Business District, the sudden increase in demand for rideshares caused the platform to "impose[] surge pricing in the city, charging passengers a minimum of [AU]$100 for a ride, four times the normal fare."[77] Unfortunately, the pricing algorithm was unable to connect the dots the same way a human operator could,[78] especially after the tragic news about the hostages had begun pouring in.

Even worse for Uber, charging higher prices in such an emergency situation created bad public relations—not that different from our reactions to price surges during the COVID-19 pandemic.[79] Following the unfortunate

76 *See* Michael Pearson et al., *With Two Hostages and Gunman Dead, Grim Investigation Starts in Sydney*, CNN (Dec. 15, 2014, 10:27 PM), https://www.cnn.com/2014/12/15/world/asia/australia-sydney-hostage-situation/index.html (reporting the "deadly siege" of the Sydney café and the hostage situation). As a *Wired* report recounted:

> On Sunday in Sydney, Australia, a hostage crisis caused extreme panic in the city's Central Business District, and ultimately, it left two hostages and one gunman dead. . . . As the crisis unfolded on Sunday and so many people were trying to flee Sydney's Business District, some noticed that Uber had imposed surge pricing in the city, charging passengers a minimum of [AU]$100 for a ride, four times the normal fare. Uber has always imposed surge pricing when demand for rides is highest, and it's not always popular, but in an emergency situation such as this one, the sky-high prices looked like yet another incredibly callous move by a company that's beginning to gain a reputation for putting profits before people. The public outcry was fierce.

Issie Lapowsky, *What Uber's Sydney Surge Pricing Debacle Says About Its Public Image*, Wired (Dec. 15, 2014, 12:30 PM), https://www.wired.com/2014/12/uber-surge-sydney/.

77 Lapowsky, *supra* note 76.

78 If one asks both a human and a computer to find the telephone number of classical music composer Ludwig van Beethoven, the former will likely respond more quickly than the latter. *See* Don Norman, The Design of Everyday Things 46 (rev. & expanded ed. 2013) ("What about Beethoven's phone number? If I asked my computer, it would take a long time, because it would have to search all the people I know to see whether any one of them was Beethoven. But you immediately discarded the question as nonsensical.").

79 *See AG Paxton Warns of Price Gouging as Texans Prepare to Prevent the Spread of Coronavirus*, Tex. Att'y Gen. (Mar. 13, 2020), https://www.texasattorneygeneral.gov/news/releases/ag-paxton-warns-price-gouging-texans-prepare-prevent-spread-coronavirus (providing a reminder that "state law prohibits price gouging in the wake of a declared disaster"); Memorandum for All Heads of Department Components and Law Enforcement Agencies from the Office of the Attorney General (Mar. 24, 2020), https://www.justice.gov/file/1262776/download (providing warnings against hoarding and price gouging).

episode in Sydney, Uber quickly issued an apology, offered refunds to the affected customers, and "put in place the ability to override automatic surge pricing in some circumstances."[80] By the time a series of terrorist attacks occurred in Paris a year later, Uber was able to "cancel[] surge pricing in the city [within half an hour of the first attack] and alerted all of its users to the emergency."[81] This drastically different outcome shows the importance and wisdom of increasing the platform's readiness for human intervention.

While platform operators often intervene based on internal data, they can also utilize external information to determine their courses of action. In a recent article, I advocated the development of a notice-and-correct mechanism to rectify problems generated by automated systems.[82] Inspired by the notice-and-takedown arrangements in copyright law,[83] my proposed mechanism underscored the need for platform operators to take expedited actions after they have been notified of problems generated by platform algorithms.[84] As I noted in that article, "as technology becomes

80 MCAFEE & BRYNJOLFSSON, *supra* note 2, at 55.

81 *Id.*

82 *See* Yu, *Algorithmic Divide*, *supra* note 1, at 379–80 (proposing the mechanism).

83 *See* 17 U.S.C. § 512(c)(1)(C) (2018) (requiring online service providers to "respond[] expeditiously to remove, or disable access to, the material that is claimed to be infringing or to be the subject of infringing activity" once these providers have been notified of copyright infringement or obtained knowledge or awareness of such infringement); *see also* Peter K. Yu, *Digital Copyright Reform and Legal Transplants in Hong Kong*, 48 U. LOUISVILLE L. REV. 693, 709–13 (2010) (providing an overview of the notice-and-takedown procedure in copyright law).

84 *See* Yu, *Algorithmic Divide*, *supra* note 1, at 379–80 ("[R]emediation-based accountability will require technology developers to quickly correct the problems once they have been notified of these problems—similar, perhaps, to the 'notice and takedown' arrangements now found in copyright law." (footnote omitted)); *see also ACM Statement*, *supra* note 1, Princ. 2, at 2 ("Regulators should encourage the adoption of mechanisms that enable questioning and redress for individuals and groups that are adversely affected by algorithmically informed decisions."); BROWNSWORD, *supra* note 55, at 297 (calling for "the regulatory framework [to] provide for the correction of the malfunction" in the technology); Chander, *supra* note 41, at 1025 ("[I]f we believe that the real-world facts, on which algorithms are trained and operate, are deeply suffused with invidious discrimination, then our prescription to the problem of racist or sexist algorithms is *algorithmic affirmative action*." (footnote omitted)); Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 126–27 (2014) ("Once notice is available, the question then becomes how one might challenge the fairness of the predictive process employed. We believe that the most robust mechanism for this is the opportunity to be heard and, if necessary, correct the record."); Diakopoulos et al., *supra* note 36 ("Make available externally visible avenues of redress for adverse individual or societal effects of an algorithmic decision system, and designate an internal role for the person who is responsible for the timely remedy of such issues.").

increasingly complicated and inscrutable, ensuring quick correction of the problem will likely be more constructive than punishing those who have allowed the problems to surface in the first place, often unintentionally."[85]

This call for human intervention is nothing new; such intervention has already been built into many existing platforms, including those that utilize artificial intelligence and learning algorithms. In *Ghost Work*, Mary Gray and Siddharth Suri documented the large pool of human workers performing on-demand tasks in the shadow to advance the development of artificial intelligence and automated systems.[86] Among their tedious but indispensable tasks are content classification, image tagging, photo comparison, video screening, and data cleaning.[87] Sarah Roberts also provided an important ethnographic study of human commercial content moderators, who work behind the scenes to screen and remove content and enforce policies on online platforms.[88] Although policymakers and industry leaders have pushed aggressively for greater automation,[89] it will remain important for algorithm designers to build human intervenability into intelligent platforms. Better still, because decisions made by human intervenors can be fed back into the algorithms as training and feedback data, such intervention will help make the platforms even more "intelligent" in the future.[90]

---

85 Yu, *Algorithmic Divide*, *supra* note 1, at 380; *see also* Wendell Wallach & Colin Allen, Moral Machines 208 (2009) ("If you are convinced that artificial agents will never satisfy the conditions for real punishment, the idea of holding them directly accountable for their actions is a nonstarter.").

86 Mary L. Gray & Siddharth Suri, Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass (2019). They went further to note the critical role this shadow workforce has played in advancing the field of artificial intelligence:

> Beyond some basic decisions, today's artificial intelligence can't function without humans in the loop. Whether it's delivering a relevant newsfeed or carrying out a complicated texted-in pizza order, when the artificial intelligence . . . trips up or can't finish the job, thousands of businesses call on people to quietly complete the project. This new digital assembly line aggregates the collective input of distributed workers, ships pieces of projects rather than products, and operates across a host of economic sectors at all times of the day and night.

Id. at ix–x.

87 *See id.* at x–xxiii.

88 Sarah T. Roberts, Behind the Screen: Content Moderation in the Shadows of Social Media (2019). As she observed, "[i]ssues of scale aside, the complex process of sorting user-uploaded material into either the acceptable or the rejected pile is far beyond the capabilities of software or algorithms alone." *Id.* at 34.

89 *See generally* Hannah Bloch-Wehba, *Automation in Moderation*, 52 Cornell Int'l L.J. (forthcoming 2020) (discussing the growing use of automation in content moderation and its impact on free speech, privacy, and other civil liberties).

90 *Cf.* Gray & Suri, *supra* note 86, at 6–8 (discussing the need for human workers to develop

Notwithstanding the need for and benefits of human intervention, deciding whether and when to intervene is not always easy, especially in an environment involving artificial intelligence and machine learning. While platform owners can set up monitoring procedures to ensure that the algorithm-generated outcomes match human intuition, such procedures may undermine a key advantage of intelligent platforms. Because humans and machines "think" differently,[91] these platforms can generate seemingly counterintuitive decisions that are superior to human decisions.[92] Even more complicated, human operators, due to cognitive barriers, may not always be able to fully appreciate the merits of those counterintuitive decisions. As Professors Casey and Niblett observed:

> Algorithms will often identify counterintuitive connections that may appear erroneous to humans even when accurate. Humans should be careful in those cases not to undo the very value that was added by the algorithm's ability to recognize these connections. This is especially true when the benefit of the algorithm was that it reduced human bias and behavioral errors.[93]

Thus, as important as it is for platform operators to intervene, they should be careful not to quickly reject counterintuitive algorithm-generated

---

datasets that are used for training artificial intelligence and how the new advances have generated new cycles that require even more human workers to complete intervening tasks).

91 *See generally* Jason Millar & Ian Kerr, *Delegation, Relinquishment and Responsibility: The Prospect of Expert Robots*, *in* ROBOT LAW 102, 117–24 (Ryan Calo et al. eds., 2016) (discussing human–robot disagreement).

92 *See* RAY KURZWEIL, THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY 261 (2005) ("Machines can pool their resources in ways that humans cannot. Although teams of humans can accomplish both physical and mental feats that individual humans cannot achieve, machines can more easily and readily aggregate their computational, memory, and communications resources."); ERIC J. TOPOL, DEEP MEDICINE: HOW ARTIFICIAL INTELLIGENCE CAN MAKE HEALTHCARE HUMAN AGAIN 117–18 (2019) (discussing the impressive progress in algorithmic image processing); Jonathan Guo & Li Bin, *The Application of Medical Artificial Intelligence Technology in Rural Areas of Developing Countries*, 2 HEALTH EQUITY 174, 175 (2018) (noting research that shows that systems using deep convolutional neural networks are "able to classify skin cancer at a comparable level to dermatologists" and "could improve the speed, accuracy, and consistency of diagnosis [of breast cancer metastasis in lymph nodes], as well as reduce the false negative rate to a quarter of the rate experienced by human pathologists"); Peter K. Yu, *Artificial Intelligence, the Law–Machine Interface, and Fair Use Automation*, 72 ALA. L. REV. 187, 215–16 (2020) (discussing the growing evidence concerning the machines' ability to outperform humans in select areas); *Digital Decisions*, *supra* note 38, at 2 ("Algorithms . . . are better and faster than humans at detecting credit card fraud.").

93 Casey & Niblett, *supra* note 75, at 354.

outcomes.[94] What looks counterintuitive at first glance may make more sense with hindsight.

C. *Interoperability*

The final proposed design feature aims to improve the quality of predictive analyses generated by algorithms and intelligent platforms. Compared with inclusivity and intervenability, interoperability, at first glance, seems to be more about the platforms than about the consumers they serve. In reality, customers will likely have more accurate predictions and better platform experiences if greater interoperability and portability exist for data collected, stored, processed, or utilized by intelligent platforms.

In the age of big data, intelligent platforms need to amass, aggregate, and analyze vast troves of data to detect and recognize patterns, predict customer choices, and shape user preferences.[95] The more data the platforms have, the better their analyses and predictions will become. As Professor Hartzog boldly declared, "[i]n the world of big data, more is always better."[96] Viktor Mayer-Schönberger and Kenneth Cukier concurred: "[B]ig data relies on all the information, or at least as much as possible . . . ."[97] The converse is also true. In a recent article, I discussed how the lack of data from a large segment of the population can result in algorithmic distortion, which will harm not only the excluded population but also other segments of the population.[98] Even worse, in an environment involving artificial intelligence and machine learning, such distortion can amplify over time when the algorithmic outcomes are fed back into the algorithms as training and feedback data.[99]

---

94 *See* Rainie & Anderson, *supra* note 9, at 40 ("People often confuse a biased algorithm for an algorithm that doesn't confirm their biases. If Facebook shows more liberal stories than conservative, that doesn't mean something is wrong. It could be a reflection of their user base, or of their media sources, or just random chance." (quoting an anonymous principal consultant of a consulting firm)); Harry Surden & Mary-Anne Williams, *Technological Opacity, Predictability, and Self-Driving Cars*, 38 Cardozo L. Rev. 121, 158 (2016) ("[I]t is not uncommon for pilots in the cockpit to be surprised or confused by an automated activity undertaken by an autopilot system."). *See generally* Selbst & Barocas, *supra* note 36 (documenting the limitations of intuition while noting the need to address inscrutability).

95 *See* sources cited *supra* note 1.

96 Hartzog, *supra* note 31, at 51.

97 Viktor Mayer-Schönberger & Kenneth Cukier, Big Data: A Revolution That Will Transform How We Live, Work, and Think 30 (2013).

98 *See* Yu, *Algorithmic Divide*, *supra* note 1, at 354–61 (discussing algorithmic distortion).

99 *See id.* at 360 ("Because biases in machine-generated analyses can amplify themselves by feeding these biases into future analyses, the unreliability of those analyses that omit

Thus far, business leaders have found the sharing of source code, training data, or other proprietary information highly unappealing.[100] As the U.S. Public Policy Council of the authoritative Association for Computing Machinery acknowledged in its *Statement on Algorithmic Transparency and Accountability*, "concerns over privacy, protecting trade secrets, or revelation of analytics that might allow malicious actors to game the system can justify restricting access to qualified and authorized individuals."[101] Likewise, Pauline Kim lamented, "transparency is often in tension with other important interests, such as protecting trade secrets, ensuring the privacy of sensitive personal information, and preventing strategic gaming of automated decision systems."[102]

In fact, the increased use of artificial intelligence and machine learning in recent years has led policymakers, commentators, and industry leaders to push for greater protection of data generated by intelligent platforms, smart devices, and networked sensors. In October 2017, for example, the European Commission proposed a new sui generis data producer's right for nonpersonal, anonymized machine-generated data.[103] This proposal "aim[ed] at clarifying the legal situation and giving more choice to the data producer, by opening up the possibility for users to utilise their data and thereby contribute to unlocking machine-generated data."[104] Had this proposal been adopted,[105] data producers would have greater proprietary control over nonpersonal, anonymized data generated by intelligent platforms, smart devices, and networked sensors.[106]

[relevant] data . . . will increase over time. Such analyses will eventually become much more unreliable than the initial skewing caused by a lack of training data concerning [the excluded population].").

100 *See* Yu & Spina Alì, *supra* note 43, at 6 ("Commercial providers could be reluctant to share information on their models or have their systems openly compared to their competitors."); *see also* Kim, *Auditing Algorithms*, *supra* note 38, at 191–92 ("[T]ransparency is often in tension with other important interests, such as protecting trade secrets, ensuring the privacy of sensitive personal information, and preventing strategic gaming of automated decision systems.").

101 *ACM Statement*, *supra* note 1, Princ. 5, at 2.

102 Kim, *Auditing Algorithms*, *supra* note 38, at 191–92.

103 *See* EUR. COMM'N, BUILDING A EUROPEAN DATA ECONOMY 13 (2017), https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017DC0009&from=EN.

104 *Id.*

105 This proposal has not received much traction lately. *See* Mark Davison, *Databases and Copyright Protection*, *in* RESEARCH HANDBOOK ON INTELLECTUAL PROPERTY AND DIGITAL TECHNOLOGIES 63, 77 (Tanya Aplin ed., 2020) ("[I]t appears that the EU has now abandoned the idea [of creating a new data producer's right].").

106 *See generally* Peter K. Yu, *Data Producer's Right and the Protection of Machine-Generated Data*, 93 TUL. L. REV. 859 (2019) [hereinafter Yu, *Data Producer's Right*] (providing a comprehensive analysis and critique of the proposed EU data producer's right).

Despite the business leaders' eagerness to obtain stronger protection for data and their continued reluctance to share these data with competitors or third-party platforms, data interoperability and portability will be important to both businesses and consumers. From a business standpoint, greater data sharing—through voluntary transfer, pooling, licensing, or other arrangements—will allow businesses to undertake the much-needed big data analyses even when they do not have all the data needed for those analyses.[107] Unless the businesses involved have achieved a certain size (think about Google or Facebook) or have come up with ways to quickly collect a lot of data (think about Netflix[108]), they will need to actively share data to compete with those businesses that have already amassed prodigious quantities of data[109] and to provide customers with more accurate predictions and better platform experiences. Moreover, because accurate big data analyses sometimes require information not collected by the implicated platforms, even platforms with vast troves of data may still need to obtain data from others.[110]

From a societal standpoint, greater data interoperability and portability will also benefit consumers by making competition viable in the big data environment. Such competition will protect consumers from

107 *See id.* at 888–89 (discussing the business needs for large, comprehensive datasets to conduct big data analyses).

108 *See* Kal Raustiala & Christopher Jon Sprigman, *The Second Digital Disruption: Streaming and the Dawn of Data-Driven Creativity*, 94 N.Y.U. L. Rev. 1555, 1587 (2019) ("Some parameters that Netflix tracks include, but are likely not limited to, pause/rewind/fast-forward behavior; day of the week; date of viewing; time of viewing; zip code; preferred devices; completion rate; user ratings; user search behavior; and browsing and scrolling behavior."); Yu, *Fair Use*, *supra* note 2, at 345 ("Netflix . . . keeps track of the parts of a movie or TV program that its subscribers have paused or viewed repeatedly.").

109 *See* Viktor Mayer-Schönberger & Thomas Ramge, Reinventing Capitalism in the Age of Big Data 199 (2018) (calling for data to be made "available to small firms, especially start-ups, so that they can compete against the big players"). Data sharing is equally important to large technology companies. *See* Smith & Browne, *supra* note 61, at 282 ("Organizations need to decide whether and how to share data, and if so, on what terms.").

110 *See* Mayer-Schönberger & Cukier, *supra* note 97, at 153 ("[I]n a big-data age, most innovative secondary uses haven't been imagined when the data is first collected."); *see also* Mark Burdon & Mark Andrejevic, *Big Data in the Sensor Society*, *in* Big Data Is Not a Monolith 61, 69 (Cassidy R. Sugimoto et al. eds., 2016) (noting that the value in data "is provided by the fact that personal data can be aggregated with that of countless other users (and things) in order to unearth unanticipated but actionable research findings"); Margaret Foster Riley, *Big Data, HIPAA, and the Common Rule: Time for Big Change?*, *in* Big Data, Health Law, and Bioethics 251, 251 (I. Glenn Cohen et al. eds., 2018) ("The analysis of Big Data related to healthcare is often for a different purpose than the purpose for which the data were originally collected.").

monopoly pricing[111] while increasing the diversity of technological products and services.[112] As Ajay Agrawal, Joshua Gans, and Avi Goldfarb reminded us, "[t]here is often no single right answer to the question of which is the best AI strategy or the best set of AI tools, because AIs involve trade-offs: more speed, less accuracy; more autonomy, less control; more data, less privacy."[113] Because of the possibility for multiple technological solutions, competition will be badly needed to accommodate the different trade-offs consumers prefer. Such competition will also help identify problems in intelligent platforms, especially when those platforms utilize similar algorithms or training data.[114]

111 As Lee Kai-fu observed:

> As a technology and an industry, AI naturally gravitates toward monopolies. Its reliance on data for improvement creates a self-perpetuating cycle: better products lead to more users, those users lead to more data, and that data leads to even better products, and thus more users and data. Once a company has jumped out to an early lead, this kind of ongoing repeating cycle can turn that lead into an insurmountable barrier to entry for other firms.

LEE KAI-FU, AI SUPERPOWERS: CHINA, SILICON VALLEY, AND THE NEW WORLD ORDER 168–69 (2018); *see also* MAYER-SCHÖNBERGER & CUKIER, *supra* note 97, at 183 (expressing concern about "the rise of twenty-first-century data barons").

112 As I noted in a recent article:

> Competition is imperative if society is to develop more efficient, more effective, and less biased algorithms. Such competition is particularly needed when algorithmic choices are increasingly difficult, or time consuming, to explain. Indeed, without competition, it would be hard to identify problems within an algorithm or to determine whether that algorithm has provided the best solution in light of the existing technological conditions and constraints.

Yu, *Algorithmic Divide*, *supra* note 1, at 382–83 (footnotes omitted); *see also* Annie Lee, Note, *Algorithmic Auditing and Competition Under the CFAA: The Revocation Paradigm of Interpreting Access and Authorization*, 33 BERKELEY TECH. L.J. 1307, 1310 (2018) ("Online competitors . . . promote fair online practices by providing users with a choice between competitive products . . . .").

113 AGRAWAL ET AL., *supra* note 42, at 5; *see also* PAUL R. DAUGHERTY & H. JAMES WILSON, HUMAN + MACHINE: REIMAGINING WORK IN THE AGE OF AI 126 (2018) ("A deep-learning system . . . provides a high level of prediction accuracy, but companies may have difficulty explaining how those results were derived. In contrast, a decision tree may not lead to results with high prediction accuracy but will enable a significantly greater explainability.").

114 As Rob Kitchin observed:

> [R]esearchers might search Google using the same terms on multiple computers in multiple jurisdictions to get a sense of how its PageRank algorithm is constructed and works in practice, or they might experiment with posting and interacting with posts on Facebook to try and determine

Notwithstanding the benefits of competition, we cannot overlook the platforms' continuous need for large, comprehensive datasets for big data analyses, which has been frequently offered as a primary justification for the data-hoarding approach embraced by tech giants.[115] Indeed, the more competition there is, the more fragmentary datasets will become, and the more difficult it will be to realize the full potential of artificial intelligence.[116] Thus, if society is eager to develop a more competitive business environment—as many governments, policymakers, and commentators have strongly advocated[117]—businesses deploying intelligent platforms will need greater data interoperability and portability to achieve optimal performance in the big data environment.[118]

> how its EdgeRank algorithm positions and prioritises posts in user time lines, or they might use proxy servers and feed dummy user profiles into e-commerce systems to see how prices might vary across users and locales.

Kitchin, *supra* note 5, at 24 (citations omitted); *see also* Yu & Spina Alì, *supra* note 43, at 7 (calling on legal researchers to "compare outputs from different programs to detect flaws in the AI utilized and increase research accuracy").

115 *See* Shoshana Zuboff, The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power 13 (2019) (noting "a decisive turn toward a new logic of accumulation").

116 *Cf.* James Manyika et al., Mckinsey Glob. Inst., Big Data: The Next Frontier for Innovation, Competition, and Productivity 12 (2011) ("To enable transformative opportunities, companies will increasingly need to integrate information from multiple data sources."); Riley, *supra* note 110, at 254 ("One of the biggest challenges for Big Data [in the healthcare space] is linking data from multiple sources so that data describing an individual located in one source are linked with data about the same individual in other sources.").

117 *See* Tim Wu, The Curse of Bigness: Antitrust in the New Gilded Age 132–33 (2018) (discussing the benefits of the breakups and the blocking of mergers of large technology companies); Maurice E. Stucke, *Should We Be Concerned About Data-opolies?*, 2 Geo. L. Tech. Rev. 275, 275–77 (2018) (discussing the actions taken by the European competition authorities against Google, Apple, Facebook, and Amazon); Matt Stevens, *Elizabeth Warren on Breaking Up Big Tech*, N.Y. Times (June 26, 2019), https://www.nytimes.com/2019/06/26/us/politics/elizabeth-warren-break-up-amazon-facebook.html (discussing Senator Elizabeth Warren's call for the breakup of Amazon, Apple, Facebook, and Google); Peter K. Yu & John Cross, *Why Are the Europeans Going After Google?*, Newsweek (May 18, 2015), https://www.newsweek.com/why-are-europeans-going-after-google-332775 (discussing the EU antitrust probe of Google).

118 *See, e.g.*, Council Regulation 2016/679, *supra* note 20, art. 20, at 45 (introducing the right to data portability). *See also* Mayer-Schönberger & Cukier, *supra* note 97, at 183 ("We should enable data transactions, such as through licensing and interoperability."); Josef Drexl, *Designing Competitive Markets for Industrial Data: Between Propertisation and Access*, 8 J. Intell. Prop. Info. Tech. & Electronic Com. L. 257, 292 (2017) ("The functioning of the data economy will . . . depend on the interoperability of digital formats and the tools of data collecting and processing."); Wolfgang Kerber, *A New*

Finally, a growing volume of research and commentary has emerged to challenge the fundamental premise of big data analysis—that the use of more data will always lead to more accurate predictions.[119] For example, Matthew Salganik and his collaborators showed recently that, "[d]espite using a rich dataset and applying machine-learning methods optimized for prediction, the best predictions were not very accurate and were only slightly better than those from a simple benchmark model."[120] Brett Frischmann and Evan Selinger argued passionately that businesses and organizations do not need all the data they collect and ever-increasing data-driven "techno-social engineering" could ultimately threaten humanity.[121] In the business context, commentators have further explained why lean data can be just as effective as, if not better than, big data.[122] Given this line of research and commentary, what constitutes an optimal level of data collection, processing, and sharing will likely remain the subject of a continuous debate.

(Intellectual) Property Right for Non-Personal Data? An Economic Analysis, 65 GEWERBLICHER RECHTSSCHUTZ UND URHEBERRECHT INTERNATIONALER TEIL [GRUR INT] 989, 997 (2016) (Ger.) ("[S]upporting portability, interoperability and standardization in regard to data is seen as pivotal policy measures for improving the governance of data in the digital economy."); Yu, *Data Producer's Right*, *supra* note 106, at 889 ("[I]f we are to maximize our ability to undertake big data analyses, such analyses may require greater sharing of data—which, in turn, calls for greater data portability and interoperability.").

119 Thanks to Ari Waldman for pushing me on this point and offering reference suggestions.

120 Matthew J. Salganika et al., *Measuring the Predictability of Life Outcomes with a Scientific Mass Collaboration*, 117 PROC. NAT'L ACAD. SCI. U.S. 8398, 8398 (2020).

121 *See* BRETT FRISCHMANN & EVAN SELINGER, RE-ENGINEERING HUMANITY 17–28, 115–17, 166–72 (2018). Professors Frischmann and Selinger defined "techno-social engineering" as "processes where technologies and social forces align and impact how we think, perceive, and act." *Id.* at 4.

122 *See, e.g.*, Matti Keltanen, *Why "Lean Data" Beats Big Data*, GUARDIAN (Apr. 16, 2013), https://www.theguardian.com/media-network/media-network-blog/2013/apr/16/big-data-lean-strategy-business (offering four reasons why businesses may prefer lean data to big data); Daniel Newman, *Bigger Isn't Always Better: It's All About Lean Data*, FORBES (Dec. 4, 2019), https://www.forbes.com/sites/danielnewman/2019/12/04/bigger-isnt-always-better-its-all-about-lean-data/#7c84d0838940 (noting that many businesses "[a]re collecting a lot more data than [they] need or use" and that "being agile and lean in digital transformation doesn't require more data—it requires smarter data"); *Separating Better Data from Big Data: Where Analytics Is Headed*, KNOWLEDGE@ WHARTON (May 10, 2018), https://knowledge.wharton.upenn.edu/article/where-analytics-is-headed-next/ (providing interviews with marketing professors at the Wharton School of the University of Pennsylvania who call on businesses to focus on better data, rather than big data).

Conclusion

In the age of artificial intelligence, innovative businesses will need to think carefully and proactively about the different features that algorithm designers can build into intelligent platforms. Although policymakers, commentators, and consumer advocates have placed transparency and accountability high on their lists, they should pay greater attention to three additional design features: inclusivity, intervenability, and interoperability. Building these features into intelligent platforms will not only protect consumers in an increasingly data-pervasive, algorithm-driven world, but it will also achieve win-win outcomes that will benefit both consumers and platform owners.