DIGITAL FOOTPRINTS: TECHNOLOGY, RACE, AND JUSTICE

Cassandra Jones Havard†

Data aggregation is ubiquitous. To widen credit access, lenders now use nonconventional sources of personal technological information to measure borrower creditworthiness. Alternative data credit scoring is touted as a useful solution for borrowers with little or no credit history or "thin credit files." The supposedly neutral algorithm provides a predictive analysis of the borrower's ability to repay, thus allowing the borrower to obtain credit within the formal banking network.

Alternative data has the potential to expand access to financial services for underserved populations and make credit markets more competitive. Machine learning, or predictive behavioral analytics, collects and sorts the borrower's social and business network identity data to assess their risk. The lender's perception (and assessment) of tangible factors such as, educational level, internet browsing history, social media associations, health status, past and current employment, or even movies, all become relevant in assessing the borrower's repayment of credit. Proponents argue that this digital footprint of an individual's personal networks, choices, and habits is fairer, more transparent, or even "color-blind," reducing discrimination in the decision-making process.

The algorithmic formulae and machine learning models that produce digital footprint technologies are protected as trade secrets. Regulators routinely evaluate lenders for compliance with fair lending laws. Lenders, however, assert trade secret protections to shield their algorithmic scoring models. By directing, adopting, and using technologies created in-house or purchased from private companies, these lenders may intentionally or unintentionally obscure discriminatory conduct. How, then, do regulators determine if digital footprint algorithms fairly assess creditworthiness? What if the underlying data of the algorithm is incomplete or implicitly biased? What if a lender impermissibly designs the digital footprint

† Professor of Law, University of South Carolina School of Law. I am grateful to the University of South Carolina School of Law for summer research funding and to Ashley Long '24 for research assistance. Finally, thank you to the editors at *Cardozo Law Review* for participation in the Symposium and for their diligence and hard work in preparing this Article for publication.

algorithm to segment markets in legally prohibited ways, thereby perpetuating credit inequality? This Article questions lenders' broad ability to keep secret the alternative data relied upon and offers policy recommendations for regulating this new world of digital footprint algorithmic scoring.

This Article makes three contributions to the existing literature. First, it shows how the lender's choice of alterative data and its interpretation of that data may result in technological redlining in violation of existing fair lending laws. Second, the Article participates in the ongoing critical race theory debate about algorithmic bias and how law and technology must combine to create algorithmic justice. Specifically, it posits that failure to police algorithms for bias prior to their use can contribute to systemic discrimination in lending.

Third, the Article proposes protections for consumers' algorithmic network identities and recommends policies that regulators are uniquely positioned to implement. It recommends a specific, transparent lending disclosure when lenders use algorithmic network identity data. Unlike in the European Union, American consumers may be unaware that a lender has used alterative data in its creditworthiness evaluation. Lenders should disclose when alternative data has been used and how it is used. Similar to other areas of law where prophylactic antidiscrimination measures are appropriate, the history of redlining and sub-prime lending in minority communities dictates a similar control in this context. Consequently, a lender will bear the burden of ensuring fairness before the lending process begins instead of providing the individual with an ineffectual post hoc remedy.

TABLE OF CONTENTS

INTRODUCTION 1179

I. BIG DATA ALGORITHMS....................1181

  A. *Algorithmic Network Identity*....................1182

  B. *Trade Secret Protections for Algorithms* 1185

II. CRITICAL RACE AND ALGORITHMS....................1189

  A. *Antisubordination Theory*....................1190

  B. *Algorithmic Injustice*....................1191

    1. Criminal Justice System....................1192

    2. Employment and Hiring....................1194

III. PRICE DISCRIMINATION....................1197

  A. *The Technology of Lending*....................1197

    1. Stereotypes and Perception of Creditworthiness Based on Identity....................1197

    2. Fintech Lending....................1199

B. *Algorithmic Redlining* 1201
IV. TOWARDS ALGORITHMIC FAIRNESS IN BIG DATA LENDING 1204
A. *Algorithmic Inclusion* 1205
B. *Algorithmic Oversight* 1206
C. *Algorithmic Accountability* 1209
CONCLUSION 1211

INTRODUCTION

Data aggregation is ubiquitous.[1] To widen credit access, lenders now use nonconventional sources of personal technological information to measure borrower creditworthiness. Alternative data credit scoring is touted as a useful solution for borrowers with little or no credit history or "thin credit files."[2] The supposedly neutral algorithm provides a predictive analysis of the borrower's ability to repay, thus allowing the borrower to obtain credit within the formal banking network.[3]

Alternative data has the potential to expand access to financial services for underserved populations and make credit markets more competitive. Machine learning, or predictive behavioral analytics, collects and sorts the borrower's social and business network identity data to assess their risk. This algorithmic network identity becomes the basis for

[1] Data aggregation takes numerous forms. This Article discusses the use of alternative data in credit underwriting. "[A]lternative data means information not typically found in the consumer's credit files of the nationwide consumer reporting agencies or customarily provided by consumers as part of applications for credit." BD. OF GOVERNORS OF THE FED. RSRV. SYS., CONSUMER FIN. PROT. BUREAU, FED. DEPOSIT INS. CORP., NAT'L CREDIT UNION ADMIN. & OFF. OF THE COMPTROLLER OF THE CURRENCY, INTERAGENCY STATEMENT ON THE USE OF ALTERNATIVE DATA IN CREDIT UNDERWRITING n.1 (2019). The Consumer Financial Protection Bureau (CFPB) is responsible for ensuring the safe use of consumers' data and regulates how banks allow data aggregators to access consumers' bank account transactions and other account data in connection with a variety of financial products and services. *See* CONSUMER FIN. PROT. BUREAU, CONSUMER PROTECTION PRINCIPLES: CONSUMER-AUTHORIZED FINANCIAL DATA SHARING AND AGGREGATION 3 (2017) (issuing a regulatory statement on consumers' ability, "upon request, to obtain information about their ownership or use of a financial product or service from their product or service provider").

[2] In 2015, approximately 26 million American consumers were "credit invisible." These consumers do not have credit scores because they have little or no credit history, leading to the label that they have "thin [credit] files." CONSUMER FIN. PROT. BUREAU, WHO ARE THE CREDIT INVISIBLES? 2 (2016), https://files.consumerfinance.gov/f/documents/201612_cfpb_credit_invisible_policy_report.pdf [https://perma.cc/B7WQ-DE78].

[3] *See* Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 10–16 (2014) (explaining algorithms or automated systems).

evaluation of creditworthiness and risk assessment, although the borrower may be unaware.

The lender's perception and assessment of tangible factors, such as, educational level, internet browsing history, social media associations, health status, past and current employment, or even downloaded or watched movies, all become relevant in assessing the borrower's repayment of credit. Proponents argue that these digital footprints of an individual's personal networks, choices, and habits are fairer, more transparent, and reduce discrimination in the decision-making process.[4]

Algorithmic formulae and machine learning models that produce digital footprint technologies are protected as trade secrets. Regulators routinely evaluate lenders for compliance with fair lending laws.[5] Lenders, however, assert trade secrecy protections to shield their algorithmic scoring models.[6] By directing, adopting, and using technologies created in-house or purchased from private companies, these lenders may intentionally or unintentionally obscure discriminatory conduct. How, then, do regulators determine if digital footprint algorithms fairly assess creditworthiness? What if the underlying data of the algorithm is incomplete or implicitly biased? What if a lender impermissibly designs the digital footprint algorithm to segment markets in legally prohibited ways, thereby perpetuating credit inequality? This Article questions a lender's broad ability to keep secret the alternative data relied upon and offers policy recommendations for regulating this new world of digital footprint algorithmic scoring.

Part I describes the impact of "big data" and machine learning in using digital footprints to create an algorithmic network identity. The algorithmic network identity becomes the basis for a lending algorithm, which trade secret protections shield from disclosure. These technologies, created in-house or purchased from private companies, may intentionally

[4] Robert Bartlett, Adair Morse, Richard Stanton & Nancy Wallace, Consumer-Lending Discrimination in the FinTech Era 29 (Nov. 2019) (unpublished manuscript), https://faculty.haas.berkeley.edu/morse/research/papers/discrim.pdf?_ga= 2.161360829.1884437453.1583517590-1779220203.1578413279 [https://perma.cc/5657-R3KW] (finding approximately one-third less discrimination using fintech algorithms in loan pricing).

[5] Federal banking regulators and the CFPB share responsibility for conducting fair lending examinations and referring violations of fair lending laws to the Department of Justice. *See* Martha J. Svoboda, *The Evolution of Redlining Post-Financial Crisis and Best Practices for Financial Institutions*, 22 N.C. BANKING INST. 67, 78–80 (2018); Cassandra Jones Havard, *"On the Take": The Black Box of Credit Scoring and Mortgage Discrimination*, 20 B.U. PUB. INT. L.J. 241, 280–83 (2011).

[6] David Stein, *AI In Lending: Key Challenges and Practical Considerations*, LAW360 (Aug. 9, 2018, 1:03 PM), https://www.cov.com/-/media/files/corporate/publications/2018/08/ai_in_lending_key_challenges_and_practical_considerations.pdf ("[T]here is a long history of making credit decisions based on the output of proprietary 'black box' algorithms, where the underlying computer logic—the secret sauce—is shielded from regulatory and public scrutiny.").

or unintentionally obscure discriminatory conduct. Part I further provides the historical basis for protecting automated decision-making as trade secrets.

Part II begins by introducing the antisubordination theory and explores the use of algorithms in other contexts. Applying antisubordination principles to algorithmic network identity demonstrates how digital footprint algorithms, when unpoliced, effortlessly reproduce structural racism. The resulting big data algorithms can result in algorithmic redlining, leading to high cost and destabilizing financial products, arguably reducing borrowers' creditworthiness and, in general, narrowing rather than expanding access to credit.[7]

Part III argues that lenders should disclose when they use big data algorithms. It also recommends that regulators should require preclearance of big data algorithms to ensure fair lending. Comparable to other areas of law where prophylactic antidiscrimination measures are appropriate, the history of redlining and sub-prime lending in minority communities dictates that a similar control is needed in this context. If required, the lender will bear the burden of ensuring fairness before the lending process begins instead of providing the individual with an ineffectual post hoc remedy. This lending framework provides inclusion, oversight, accountability, and ultimately fairness to consumers.

I. BIG DATA ALGORITHMS

The online economy has spawned a new data industry. Technology provides voluminous amounts of data which, when extracted, creates a profile of individuals' preferences.[8] "Big data," referencing the amount and variety of information that is processed to reveal unknown patterns or trends, arguably provides valuable insights into human behavior.[9] Using this alternative data is beneficial for examining the

[7] One author identifies "algorithmic redlining" as using algorithms, or any type of computational lending, in discriminatory lending that prohibits access to housing. James A. Allen, *The Color of Algorithms: An Analysis and Proposed Research Agenda for Deterring Algorithmic Redlining*, 46 FORDHAM URB. L.J. 219, 223 (2019).

[8] When data becomes a commodity, privacy concerns arise with the collection of the alternative data, or "big data," and its use in machine learning algorithms. Jack Balkin, *Three Questions: Prof. Jack Balkin on Facebook and the Risks of 'Data Capitalism,'* YALE INSIGHTS (May 8, 2018), https://insights.som.yale.edu/insights/three-questions-prof-jack-balkin-on-facebook-and-the-risks-of-data-capitalism [https://perma.cc/6WBQ-4RPA] (advocating that online companies should have an obligation to protect consumers' data, becoming "information fiduciaries").

[9] *See* Nate Cullerton, Note, *Behavioral Credit Scoring*, 101 GEO. L.J. 807, 821–22 (2013) (discussing the use of behavioral and geo-demographic data in developing alternative credit models).

creditworthiness of borrowers with "thin" or no credit history. As massive amounts of data are collected on individuals, the social context that creates the technology requires examination.

A. *Algorithmic Network Identity*

Algorithmic network identity is a complex notion.[10] Technology documents daily life in unprecedented ways. In everyday activities, individuals interact with digital platforms to create information spaces. Mining the data of consumers' buying habits, social relationships, political preferences, lifestyles, hobbies, health, and personalities, businesses gather an immense amount of information.[11] Data aggregators use a wide range of sources, from "public records, web browsing activity, emails, banking activity, social media, store loyalty cards, online quizzes, license plate readers, app usage, smart devices (such as fitness watches and internet-connected doorbells), and geo-location tracking on . . . smartphones."[12] The interactions with others and the machine-learning algorithms that aggregate the data left in cyberspace create and define identity, self-identity, and cultural categories. These information spaces and the use of technology present complex issues about data, identity, and agency.[13]

The information flow of the internet creates a perpetual state of surveillance of its users. Using algorithms, companies follow consumers' data to observe and analyze users' online behavior. Specifically, machine learning algorithms sift through and aggregate the trace data that users create by posting and clicking in the virtual landscape. Based upon an individual's web-surfing activities, data is generated, compiled, and grouped to establish a digital identity.[14]

[10] One group of researchers describe network identity as the "algorithmically produced position of an individual." Zahra Stardust et al., *High Risk Hustling: Payment Processors, Sexual Proxies, and Discrimination by Design*, 26 CUNY L. REV. 57, 131 (2023) (quoting danah boyd, Karen Levy & Alice Marwick, *The Networked Nature of Algorithmic Discrimination*, *in* DATA AND DISCRIMINATION: COLLECTED ESSAYS 53, 56 (Seeta Peña Gangadharan, Virginia Eubanks & Solon Barocas eds., 2014), https://timlibert.me/pdf/2014-Data_Discrimination_Collected_Essays.pdf [https://perma.cc/73Y7-H3HU] (discussing discrimination based on personal networks)).

[11] Michele E. Gilman, *Five Privacy Principles (from the GDPR) the United States Should Adopt to Advance Economic Justice*, 52 ARIZ. ST. L.J. 368, 375 (2020).

[12] *Id.*

[13] *See* Dan L. Burk, *Algorithmic Legal Metrics*, 96 NOTRE DAME L. REV. 1147, 1185 (2021) (discussing use of data and the possibility of bias in developing individual identities).

[14] The lack of privacy regulations means that most consumers are not aware of the information that is collected about them. Ignacio N. Cofone, *The Dynamic Effect of Information Privacy Law*,

Algorithms comb the gathered data searches for correlations among thousands of individuals to form categories. Drawing inferences, algorithms use users' digital histories to create groups and sub-groups within categories. Matching similarities, correlations among the data result in categories.

Creating a recognizable digital person is hardly a voluntary decision.[15] The social construction of categories varies based on characteristics such as gender, race, and occupation. Yet, the social media identity that most users create is an idealized identity.[16] Using "likes" and "follows," users cultivate the person that others will find acceptable and lead to more connections and an expanded network. These shared identities infer categories of identity upon users based largely on their web-surfing habits. The interpretation of the aggregated data varies depending on who is using it. The correlations create categories which are further subdivided.

Users' digital definitions of identities and self-identities are complex. Often the creator is unaware of how the identity might be used or interpreted. Creators' self-generated profiles' interactions with algorithms are hidden from consumers. The algorithms analyze users' online behaviors, making predictive inferences about users' decision-making. In effect, the digital self is based on the infrastructures that computer code and algorithms create.[17] Rather than individuals choosing identities, identities are created based on the analysis of the aggregated data.[18] The design choice of the selected algorithm may underscore patterns in the data, obscure them, make assumptions, and draw

18 MINN. J.L. SCI. & TECH. 517, 533 (2017) (distinguishing between active and passive digital footprints).

[15] While consumers often "opt in" to data analytic notifications on a company's website and voluntarily visit or post on social media sites, the average consumer is unaware of how that information is produced and consumed across various platforms. *See* Rebecca Lipman, *Online Privacy and the Invisible Market for Our Data*, 120 PENN ST. L. REV. 777, 796–97 (2016) (discussing the lack of standardization among company websites).

[16] Individuals create identities based on both authentic and idealized versions of themselves. Erica R. Bailey, Sandra C. Matz, Wu Youyou & Sheena S. Iyengar, *Authentic Self-Expression on Social Media is Associated With Greater Subjective Well-Being*, 11 NATURE COMMC'NS, Oct. 6, 2020, at 2.

[17] As one author states,

> Codes are cultural objects embedded and integrated within a social system whose logic, rules, and explicit functioning work to determine the new conditions of possibilities of users' lives. How a variable like X comes to be defined, then, is not the result of objective fact but is rather a technologically-mediated and culturally-situated consequence of statistics and computer science.

John Cheney-Lippold, *A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control*, 28 THEORY, CULTURE & SOC'Y 164, 167 (2011).

[18] *Id.*

conclusions. The quantitative analytics of the data and the various methods of available interpretation are independent of consumer knowledge or choice.

Essentially, the defined identity of users is neither self-regulated nor self-managed. Computer code and statistics construct cultural categories within populations, and thus exercise control according to users' surveilled internet history. Most consumers are oblivious to being a co-participant in creating an algorithmic network identity.[19] Users are unaware that the categories also are a compilation of current and past classifications. Further, because the digital "self" involves idealized and actual notions, and virtual interactions with others, creating the algorithmic identity is not solely an individual act. Web-surfing and media interfacing produces complex networks of information that create individual identity, cultural categories, and meaning.

Furthermore, social network data sources present challenges in terms of transparency and discrimination. Given that thousands of data points are collected without consumers' knowledge,[20] consumers are probably unaware of what information will be used for credit decision-making and therefore are unable to challenge unfair credit decisions.[21] Or consumers may be strategically constructing social networks and identities to improve their credit quality.[22] Lenders may be targeting underserved consumers for profit-maximization, although doing so is explicitly prohibited under current law. Also, when creditworthiness is determined by familial, religious, or social associations, algorithms may treat facially neutral data points as proxies for immutable characteristics, such as race and gender.

The complexity of data algorithms can lead to opaqueness such that consumers are unable to identify the source of the harmful data's inputs. The question becomes whether consumers understand the power of the

[19] A "'passive digital footprint' is a data trail you unintentionally leave online." *Digital Footprint*, TECHTERMS.COM, https://techterms.com/definition/digital_footprint#:~:text=A%20digital%C20footprint%C20is%C20a,trail%C20you%C20unintentionally%C20leave%20online [https://perma.cc/SNS5-FVQ4].

[20] Uri Gal, *How the Shady World of The Data Industry Strips Away Our Freedoms*, THE CONVERSATION (Aug. 14, 2020, 2:12 AM), https://theconversation.com/how-the-shady-world-of-the-data-industry-strips-away-our-freedoms-143823 [https://perma.cc/S3PH-AYAG] (discussing how data brokers gather and harvest information about individual).

[21] Consumers have limited ability to identify and contest unfair credit decisions and have little chance to understand what steps they should take to improve their credit. Recent studies have also questioned the accuracy of the data used by these tools, in some cases identifying serious flaws that have a substantial bearing on lending decisions. Havard, *supra* note 5, at 280–83.

[22] Indeed, they may be incentivized to do so artificially or maliciously. *The Surprising Ways that Social Media Can Be Used for Credit Scoring*, KNOWLEDGE AT WHARTON (Nov. 5, 2014), https://knowledge.wharton.upenn.edu/article/using-social-media-for-credit-scoring [https://perma.cc/2EA6-W2CS] (interview with Pinar Yildirim).

algorithmic network identity that they create based on their social behavior. Knowing how algorithmic decisions are made and what specific data and principles shape them is crucial to protect all borrowers, but especially marginalized borrowers for whom fintech lending is designed to provide greater access to credit.[23]

As discussed below, the use of big data in the financial services sector intersects with trade secret protections for algorithmic lending models. The protection of intellectual property is, however, often at the expense of fair lending and privacy protections for consumers.

B. *Trade Secret Protections for Algorithms*

The personal data economy provides businesses with customer insights and market predictions. The proprietary analytics of big data create algorithms that are valuable assets based on a company's use of data.[24] These methods for analyzing and processing the data are competitive, intangible assets that need legal protection. The resulting decision-making clashes with the demands on consumer privacy and the legal protections of technology.

Law protects technologies as a way of promoting innovation.[25] The legal protections also allow inventors control over their original work. Intellectual property law safeguards technology. Patent and copyright protections promote scientific and artistic progress. In exchange for publicly releasing information about the invention or work, inventors and authors are awarded a limited monopoly on their work.[26]

Software and its related inventions have limited patent protections. Initially, the Patent and Trademark Office announced severe limitations on providing patents for software, describing it as the automation of

[23] Michael Griffith, Note, *AI Lending and the ECOA: Avoiding Accidental Discrimination*, 27 N.C. BANKING INST. 349, 363–64 (2023) (discussing how AI can be used to target vulnerable consumers). *See generally* Loretta J. Mester, *What's the Point of Credit Scoring?*, FED. RSRV. BANK PHILA. BUS. REV., Sept.–Oct. 1997, at 3 (outlining the history, use, and methods of credit scoring).

[24] *See* FRANK J. OHLHORST, BIG DATA ANALYTICS: TURNING BIG DATA INTO BIG MONEY 2–4 (2012).

[25] Irene Kosturakis, *Intellectual Property 101*, TEX. J. BUS. L., Fall 2014, at 37, 40.

[26] Federal laws protect authors' and inventors' works under patent, copyright, trademark, and trade secret rules. *See, e.g.*, 17 U.S.C. § 102; 18 U.S.C. §§ 1831–1839. Patent law protects inventions and provides a twenty-year term of exclusivity on an invention. 35 U.S.C. §§ 101, 154(a)(2). Copyright law protects work fixed in a tangible form providing exclusive protection of an author's work during the author's lifetime plus seventy years thereafter. 17 U.S.C. §§ 102, 302(a). Patent law also grants rights to inventors in exchange for public disclosure. Lanham Act, ch. 540, 60 Stat. 427 (1946) (codified as amended in scattered sections of 15 U.S.C.).

mental steps.[27] Those limitations continued through a Supreme Court ruling in which the question raised was whether the algorithm contained a particular machine or transformation.[28]

Algorithms, as mathematical formulae based on abstract ideas, were not considered original works worthy of legal protection. In *Gottschalk v. Benson*, the Court was unpersuaded that somehow a computer made the thought process of creating the algorithm less abstract.[29] Instead, the Court indicated that there should be "[t]ransformation and reduction of an article 'to a different state or thing,'" suggesting that patentability cannot result from the use of generic computer, but must be based on the use of a particular machine.[30] Under the Court's reasoning, patents were eligible on algorithm-based inventions only if those inventions themselves were "new and useful."[31] Whether a computer could create a concrete, original thought was the issue that the Court struggled to resolve. Accordingly, the Court determined that the software could not be patented.[32]

Yet, the rules on the patentability of computer software expanded under the interpretations of the Federal Circuit. Distinguishing "disembodied" mathematical algorithms from an invention containing a mathematical algorithm, the Federal Circuit established precedents that evaluated the composite invention and its practical application.[33] In this way, the abstract idea when tied to a physical invention was patentable subject matter.[34] Later clarifications of this line of reasoning solidified the

[27] Meghan J. Ryan, *Secret Algorithms, IP Rights, and the Public Interest*, 21 NEV. L.J. 61, 76 (2020).

[28] Gottschalk v. Benson, 409 U.S. 63, 64–65, 67–68 (1972).

[29] *Id.* at 64–65, 72. As the *Benson* Court found, "The mathematical formula involved here ha[d] no substantial practical application except in connection with a digital computer . . . ." *Id.* at 71–72. Later, in *Parker v. Flook*, 437 U.S. 584, 591 (1978), the Supreme Court, following the decision in *Benson*, affirmed that disembodied algorithms are unpatentable unless those inventions are "new and useful."

[30] *Benson*, 409 U.S. at 69–70 (quoting Cochrane v. Deener, 94 U.S. 780, 780 (1876)).

[31] *Id.* at 67 (quoting Funk Bros. Seed Co. v. Kalo Inoculant Co., 333 U.S. 127, 130 (1948)).

[32] *Id.* at 72–73. This was an inapposite result from the Supreme Court's reasoning in *Diamond v. Chakrabarty*, 447 U.S. 303, 309 (1980), that "anything under the sun that is made by man" is patentable. *Id.* at 309 (quoting S. REP. NO. 1979, at 4 (1952)).

[33] *In re* Alappat, 33 F.3d 1526, 1544 (Fed. Cir. 1994). In *Alappat*, the court determined that if a claim includes something more than just an algorithm or other vague mathematical concept, then it would not be subject to the exceptions outlined in the previous line of cases limiting the software's patentability. *Id.* at 1542–44. The court described the patentable subject matter as a mathematical concept used on a "specific machine to produce a useful, concrete, and tangible result." *Id.* at 1544.

[34] *See id.* at 1544–45.

concept that an algorithm with a computer constituted a patentable "machine" under 35 U.S.C. § 101.[35]

The Supreme Court rolled back the permissiveness of the federal circuit courts when it examined what constitutes a patentable process under § 101. Expanding the *Benson* Court's "machine-or-transformation" test, the Court examined whether the questioned software was a patentable "process."[36] Subsequently, in *Mayo Collaborative Services v. Prometheus Laboratories, Inc.*, the Court reaffirmed that principle, stating that "simply implementing a mathematical principle on a physical machine, namely, a computer, [is] not a patentable application of that principle."[37]

The most recent and defining decision about the § 101 exception to computer software is *Alice Corp. Proprietary Ltd. v. CLS Bank International*.[38] Patent law rarely protects algorithms. The *Alice* decision underscores why trade secret rules historically provide protection for lending algorithms. To be "patent-eligible," the algorithm must be an "inventive concept."[39] In *Alice*, the Supreme Court invalidated a software patent, deciding that the proprietary technologies of computerized algorithms are too abstract to be patented.[40] Under *Alice*, an algorithm is patentable only when it produces a method that is unique, novel, non-obvious, and useful.[41]

Traditional lending algorithms are considered proprietary secrets.[42] Patent law presents other obstacles for protecting algorithms, especially those based on artificial intelligence (AI).[43] AI algorithms differ from traditional algorithms. AI algorithms use machine learning–enabling

[35] 35 U.S.C. § 101; *see* State St. Bank & Tr. Co. v. Signature Fin. Grp., Inc., 149 F.3d 1368, 1373 (Fed. Cir. 1998); AT&T Corp. v. Excel Commc'ns, Inc., 172 F.3d 1352, 1361 (Fed. Cir. 1999).

[36] Bilski v. Kappos, 561 U.S. 593, 603–04 (2010) (declaring that *Benson* cannot be the sole test for assessing whether something constitutes a "process" under § 101).

[37] 566 U.S. 66, 84 (2012).

[38] 573 U.S. 208 (2014).

[39] *See Mayo*, 566 U.S. at 72 (quoting Parker v. Flook, 437 U.S. 584, 594 (1978)).

[40] *See Alice*, 573 U.S at 227.

[41] *See id.* at 223–24.

[42] Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54, 59 (2019) ("Because many algorithms are proprietary, they are resistant to discovery and scrutiny.").

[43] One author opines that the effect of the ruling is that companies are finding other legal ways to maintain the secrecy of their technology. *See* Ryan, *supra* note 27, at 84–87 (discussing how the Supreme Court is limiting both the patent protection for software by changing the rules regarding principles of non-obviousness, definiteness, and equivalents as well as making defenses more difficult and attorneys' fees more accessible for patent challengers). Furthermore, as Professor Ryan points out, the Defend Trade Secrets Act, which provides federal trade secret protection, increases the viability of trade secrets as a means of protecting algorithms. *Id.* at 86–87.

software to update and "learn" from previous outcomes. Unlike algorithms based on programmable software, machine learning algorithms continuously update without human intervention and therefore are considered proprietary analytics.[44] Protecting them as trade secrets provides more confidentiality for the inventor than the patent protections, which require public disclosure of information in the application.[45] Proving infringement of an AI algorithm may be difficult.[46] The process is lengthy, with the technology possibly changing during the process. Enforcement of a patent requires significant disclosure of the technology. While the raw data is not patentable, the information that must be disclosed to secure the patent may need to be kept secret to protect its proprietary nature.

Trade secrets are more commonly used to protect technology that is not clearly patent eligible. Trade secret laws provide broader protection over AI algorithms.[47] The information and data sets used for machine-based learning or training models are also not protectable under patent law. The law of trade secrets has several advantages for protecting this information. The trade secret serves as a license activated with a nondisclosure, noncompete agreement, or other means of shielding disclosure. The license also shields from public scrutiny the technology, resulting in risks and harms to consumers. For consumers alleging bias or error in an algorithmic result, there is little to no basis for challenge.[48]

To the extent that trade secret laws hide decision-making that results in discrimination, civil rights laws are implicated. As discussed below,

[44] *See* Hyunjong Ryan Jin, *Think Big! The Need for Patent Rights in the Era of Big Data and Machine Learning*, 7 NYU J. INTELL. PROP. & ENT. L. 78, 97 (2018).

[45] J. Jonas Anderson, *Secret Inventions*, 26 BERKELEY TECH. L.J. 917, 928 (2011) ("Because it requires disclosure, patent law precludes simultaneous protection of an invention as both a patent and a trade secret."). One author identifies three conditions that will impact change in the data industry, giving consumers more proprietary data rights. Those conditions are consumer mistrust, governmental regulation, and market competition. Hossein Rahnama & Alex "Sandy" Pentland, *The New Rules of Data Privacy*, HARV. BUS. REV. (Feb. 25, 2022), https://hbr.org/2022/02/the-new-rules-of-data-privacy [https://perma.cc/N6WJ-X2C2].

[46] *See* Ryan, *supra* note 27, at 84–85.

[47] *Id.* at 62–64.

[48] The Consumer Privacy Bill of Rights was a proposal during the Obama administration to regulate the processing of electronic personal data by providing consumers with privacy and control over personal data. This proposed legislation, which would have allowed consumers to challenge and correct data that algorithms use to make decisions about credit or insurance, was never enacted by Congress. *See generally* CONSUMER DATA PRIVACY IN A NETWORKED WORLD: A FRAMEWORK FOR PROTECTING PRIVACY AND PROMOTING INNOVATION IN THE GLOBAL ECONOMY, WHITE HOUSE (Feb. 2012) https://obamawhitehouse.archives.gov/sites/default/files/privacy-final.pdf. *See also* Ryan, *supra* note 27, at 88–89 (arguing that algorithm secrecy prevents the examination of outcomes based on their use and the examination of whether they are accurate and fair).

accountability requires that trade secret laws consider the intersection of technology and fairness.

II. Critical Race and Algorithms

Critical Race Theory (CRT) is a legal perspective that evaluates the intersection of race and law.[49] Algorithmic bias is often viewed as a technical problem. In general, bias in lending often receives deference under the business justification of creditworthiness.[50] As I have argued previously, the "black box" protects lenders and the creators of algorithms from the discriminatory harm that may ensue.[51] The burgeoning technology requires evaluating not just the resulting lending tools, but also the data aggregators and the algorithmic creator, for discriminatory results. This Part looks first at the use of algorithms in areas outside of lending, then at historical race-based economic discrimination, and outlines how critical race theoretical precepts expose their present-day vestiges. It argues that antisubordination principles, a subset of CRT, expose how mathematical principles can imbed racial subordination in lending decisions. It also argues that present antidiscrimination law is inadequate to protect consumers, especially given lending algorithms based on aggregated social media. This sets the background for arguments about why lending applicants need expanded legal protections.

[49] The term "intersectionality" was originally developed by critical race theorist Kimberlé Crenshaw. David Gillborn, *Intersectionality, Critical Race Theory, and the Primacy of Racism: Race, Class, Gender, and Disability in Education*, 21 QUALITATIVE INQUIRY 277, 278 (2015). Intersectionality is "the complex, cumulative way in which the effects of multiple forms of discrimination (such as racism, sexism, and classism) combine, overlap, or intersect especially in the experiences of marginalized individuals or groups." *Intersectionality*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/intersectionality [https://perma.cc/Q5XD-Q33M]; *see also* RICHARD DELGADO & JEAN STEFANCIC, CRITICAL RACE THEORY: AN INTRODUCTION 8–9 (2001) ("Closely related to differential racialization—the idea that each race has its own origins and ever evolving history—is the notion of intersectionality and anti-essentialism. No person has a single, easily stated, unitary identity."); Tukufu Zuberi, *Critical Race Theory of Society*, 43 CONN. L. REV. 1573, 1587–89 (2011); Kristin Brandser Kalsem, *Bankruptcy Reform and the Financial Well-Being of Women: How Intersectionality Matters in Money Matters*, 71 BROOK. L. REV. 1181, 1186 (2006) ("An intersectional analysis explores 'the way power has clustered around certain categories and is exercised against others' and identifies 'particular values attached to [such categories] and the way those values foster and create social hierarchies.'").

[50] Francesca Lina Procaccini, *Stemming the Rising Risk of Credit Inequality: The Fair and Faithful Interpretation of the Equal Credit Opportunity Act's Disparate Impact Prohibition*, 9 HARV. L. & POL'Y REV. S43, S62–63 (2015).

[51] Cassandra Jones Havard, *"On the Take": The Black Box of Credit Scoring and Mortgage Discrimination*, 20 B.U. PUB. INT. L.J. 241, 278–83 (2011).

A. *Antisubordination Theory*

Two distinct frameworks exist for understanding the racial implications of contemporary assessment of creditworthiness. The dominant framework is concerned with the assessment and management of risk, whereas an antisubordination framework is concerned with the subordinating effects of creditworthiness assessments, algorithmic risk assessment, and the borrower's ability to pay. What the dominant framework often fails to do in evaluating the use of technology generally, and algorithms specifically, is address how this type of credit risk assessment has emerged from, or sustained, prevailing racist ideologies. The antisubordination framework not only encourages such a critique but also facilitates a more inclusive analysis of race in relation to a creditworthiness assessment.

An antisubordination framework can supply a new and sharper lens for interpreting the ways in which algorithmic fairness can be achieved. This analysis points out how supposedly neutral credit risk assessments do not account for the spatial segregation of the country based on race and the profit-making motivations of lenders. Even in using nonracial categories, the classifications can reinforce perceptions of minorities as exceptional credit risks. This type of substitution functions in tandem with the law and nullifies the effects of fair lending's presumption of an individual assessment of risk. By uncovering and addressing this dynamic, law can monitor technology to develop and pursue algorithmic fairness strategies that ultimately achieve fair lending objectives.

Creditworthiness is a measure of a borrower's ability to manage debt. The lender's assessment is individual to the borrower. Before risk-based pricing became acceptable, lenders did not extend credit to borrowers with low or no credit scores.[52] Risk-based pricing methodologies allow lenders to tailor borrower interest rates based on the credit profile characteristics that vary by credit quality. Historically, delinquencies and bankruptcy were adverse credit events resulting in denial.[53] In risk-based pricing, lenders justify a higher interest rate based

[52] Michael Staten, *Risk-Based Pricing in Consumer Lending*, 11 J.L. ECON. & POL'Y 33, 50 (2015) (discussing how risk-based pricing makes credit available to higher-risk consumers); *see also* Alan M. White, *Borrowing While Black: Applying Fair Lending Laws to Risk-Based Mortgage Pricing*, 60 S.C. L. REV. 677, 689 (2009) (describing how lenders use either a FICO score or one based on their own model to determine the borrower's interest rate with the variation in price depending on the loan product and borrower characteristics).

[53] Susan Block-Lieb & Edward J. Janger, *The Myth of the Rational Borrower: Rationality, Behavioralism, and the Misguided "Reform" of Bankruptcy Law*, 84 TEX. L. REV. 1481, 1516 (2006) (discussing how lenders use risk-based pricing to grant loans to consumers who have filed for bankruptcy).

on the likelihood of default.[54] According to the dominant framework, the solution to making credit available is to use higher interest rates and uniform standards for the assessment of risk.

The dominant framework recognizes that lending has the dual purpose of assessing borrower likelihood of default and loan profitability and supports a framework that melds these two objectives. Proponents of this framework contend that risk-based pricing and inaccurate assessments of risk, together, result in higher loan defaults and undermine the sustainability of the consumer debt market. This is partially because, in a system that conditions credit availability on traditional credit risk factors, the risk of default is extremely high for persons who are outside of traditional credit markets. While that person may have the ability to repay the credit obligation, the label of being a high credit risk assigns a high interest rate. Whether these rates accurately use reliable metrics or tools for assessment, or are in fact arbitrary, is beyond examination. Lenders rely on the scoring mechanisms which in fact may be more subjective than objective in making the assessments of risk and setting the interest rate determinations. The probability that there is disproportionate representation of Black borrowers is unsurprising in this context, given the percentage of unbanked consumers within these communities.[55] Furthermore, the subjective assessment of risk permits implicit biases to influence perceptions of how risky a person is, encouraging higher interest rates for African American borrowers.

B. *Algorithmic Injustice*

Algorithms using aggregated social network data are widely used in many contexts.[56] The discussion that follows provides examples of the civil rights concerns in the criminal justice, employment, and price

[54] Robert G. Schwemm & Jeffrey L. Taren, *Discretionary Pricing, Mortgage Discrimination, and the Fair Housing Act*, 45 HARV. C.R.-C.L. L. REV. 375, 395–97 (2010) (describing the "discretionary pricing" policy of Countywide Mortgage Company); *see also* Adam J. Levitin, *Rate-Jacking: Risk-Based & Opportunistic Pricing in Credit Cards*, 2011 UTAH L. REV. 339 (defining the phenomenon of "rate-jacking" in the credit-card industry).

[55] Cassandra Jones Havard, *Doin' Banks*, 5 U. PA. J.L. & PUB. AFFS. 317, 327 (2020).

[56] *See* Nina I. Brown, *Regulatory Goldilocks: Finding the Just and Right Fit for Content Moderation on Social Platforms*, 8 TEX. A&M L. REV. 451, 458 (2021) (discussing content regulation for social platforms).

discrimination contexts when using social network algorithms in decision-making.[57]

1. Criminal Justice System

Criminal justice systems use predictive algorithms in a range of activities from policing to bail and sentencing to parole.[58] Commonly used to predict recidivism and assess a criminal defendant's risk of reoffending, supporters of criminal justice algorithms claim that they are objective, computer-driven calculations.[59] They argue that the possibility

[57] Algorithms are pervasive throughout society, including within government agencies that use them extensively in making determinations. *See* Katyal, *supra* note 42, at 56–57 (discussing how government agencies use algorithmic decision-making). A comprehensive list of recent "Examples of Discriminatory Bias by AI Systems" is available in Barry E. Hill, *Environmental Justice and the Transition from Fossil Fuels to Renewable Energy*, 53 ENV'T. L. REP. 10317, 10333 (2023).

[58] Mass incarceration has disproportionately impacted racial minorities in the United States. *See generally* MICHELLE ALEXANDER, THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS (2010). The impact of algorithms on minorities in the criminal justice system is the subject of much legal scholarship. *See, e.g.*, Sean Allan Hill II, *Bail Reform and the (False) Racial Promise of Algorithmic Assessment*, 68 UCLA L. REV. 910 (2021) (applying an antisubordination framework in the criminal context); Jessica M. Eaglin, *Technologically Distorted Conceptions of Punishment*, 97 WASH. U. L. REV. 483 (2019) (arguing for abolition of algorithms in criminal legal reform); Aziz Z. Huq, *Racial Equity in Algorithmic Criminal Justice*, 68 DUKE L.J. 1043 (2019); Aziz Z. Huq, *The Consequences of Disparate Policing: Evaluating Stop and Frisk as a Modality of Urban Policing*, 101 MINN. L. REV. 2397 (2017); Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327 (2015).

[59] The criminal justice system began using Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), a statistically based algorithm, in the 1990s. COMPAS generates a recidivism-risk score based on the defendant's responses to an in-depth questionnaire, which in turn determines the defendant's sentence. Data proved COMPAS to be biased against African Americans. White defendants that were equally likely to reoffend received a lower score and, therefore, a more favorable sentence, while African Americans were more likely to be assigned a higher score. Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, *Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [https://perma.cc/74XC-STXP]; Andrew Lee Park, *Injustice Ex Machina: Predictive Algorithms in Criminal Sentencing*, UCLA L. REV. (Feb. 19, 2019), https://www.uclalawreview.org/injustice-ex-machina-predictive-algorithms-in-criminal-sentencing [https://perma.cc/ACS6-L8KR]. Another predictive algorithm, Prisoner Assessment Tool Targeting Estimated Risk and Needs (PATTERN), uses factors such as criminal history, education level, disciplinary incidents while incarcerated, and whether an inmate has completed any programs aimed at reducing recidivism to predict a score. The Department of Justice found that PATTERN overpredicts recidivism among minority inmates by between two-to-eight percent as compared to white inmates and also estimates an inmate's risk of committing a crime after release. Duncan Purves & Jeremy Davis, *Criminal Justice Algorithms: Being Race-Neutral Doesn't Mean Race-Blind*, THE CONVERSATION (Mar. 31, 2022, 8:44 AM), https://theconversation.com/criminal-justice-algorithms-being-race-neutral-doesnt-mean-race-blind-177120 [https://perma.cc/WGH9-

of bias in assessment appeared greater in an individual decisionmaker who has access to personal data about an individual defendant. Prior methods of assessing recidivism involved personal interviews. The probation officer's personal, fallible judgments were the basis for assessing why an offense had been committed, determining the type of basis for needed rehabilitation, and predicting further harm to the community.

Supporters of algorithms argue that decisions are more uniform.[60] The analytical tool is more objective because it combines different types of information about risk, crime, and recidivism and thus offers a better assessment of the condition and communities into which a convicted person might be returning and the social effects on the ex-criminal's behavior.[61] Police also use social media for surveillance and to monitor potential criminal activity.[62] Social media data provides detailed, easily accessible information from its users and their associates. Surveillance has always occurred more in minority neighborhoods.[63] Yet communities may be unaware about how the criminal justice system uses the algorithms.

Opponents contend that criminal justice algorithms are constructed in a way that disproportionately harm marginalized communities.[64] The

TNTV]; *see also* Itay Ravid & Amit Haim, *Progressive Algorithms*, 12 U.C. IRVINE L. REV. 527, 560–61 (2022) (explaining that proprietary criminal justice algorithms may not be objective). *See generally* Virginia Eubanks, AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR (2018).

[60] Pauline T. Kim, *Race-Aware Algorithms: Fairness, Nondiscrimination and Affirmative Action*, 110 CALIF. L. REV. 1539, 1592–93 (2022) (discussing "the mistaken belief, common among non-technical people, that algorithms are objective and neutral").

[61] The predictive analytics in a criminal justice setting may use various types of information, including police records, personal data, and social networks to forecast future criminal activity. *See, e.g.*, John Buntin, *Social Media Transforms the Way Chicago Fights Gang Violence*, GOVERNING (Sept. 26, 2013), https://www.governing.com/archive/gov-social-media-transforms-chicago-policing.html [https://perma.cc/3EQA-7PDZ].

[62] Rachel Levinson-Waldman, *Government Access to and Manipulation of Social Media: Legal and Policy Challenges*, 61 HOW. L.J. 523, 524 (2018).

[63] *See* LUKE SCRIVENER, ALLIE MEIZLISH, ERICA BOND & PREETI CHAUHAN, TRACKING ENFORCEMENT TRENDS IN NEW YORK CITY: 2003–2018 (2020), https://datacollaborativeforjustice.org/wp-content/uploads/2020/09/2020_08_31_Enforcement.pdf [https://perma.cc/NF79-F54F] (discussing overenforcement in minority communities in New York City); *see also* Christopher Thomas & Antonio Pontón-Núñez, *Automating Judicial Discretion: How Algorithmic Risk Assessments in Pretrial Adjudications Violate Equal Protection Rights on the Basis of Race*, 40 MINN. J.L. & INEQ. 371, 407 (2022) (discussing the unconstitutionality of using algorithms in risk assessments because the risk assessments are not narrowly tailored); I. Bennett Capers, *Policing, Race, and Place*, 44 HARV. C.R.-C.L. L. REV. 43, 69–70 (2009) (describing the maintenance of racialized borders in minority neighborhoods).

[64] Ngozi Okidegbe, *The Democratizing Potential of Algorithms?*, 53 CONN. L. REV. 739, 742–46 (2022).

predictive analytics are based on qualitative and quantitative data.[65] When the results are based on "carceral knowledge sources," such as police records and court data, without an examination of the underlying data, the predictive results could be discriminatory.[66] A further criticism is that the experiences of incarcerated persons and their communities are often not included in the training data on which predictive algorithms are built.[67] Expanding the training information base would produce fairer results. The opaqueness about the data used contributes to the further marginalization of oppressed communities.[68]

Incomplete and inaccurate data affects the algorithm's decision-making power. The lack of data and other crucial information required for an independent evaluation can lead to bias. A separate issue is the inherent biases that result from using data from criminal justice institutions, such as the police, that are usually unfavorable towards minorities, and the resulting intentional discrimination and disparate treatment that may occur in using algorithms as early as in the pretrial context.[69]

2. Employment and Hiring

Employers use algorithms to assist in recruitment and hiring.[70] Algorithms improve the process by screening applicants through mining and matching processes.[71] The algorithms can determine everything from shaping the candidate pool by determining who views job postings to who

[65] Jessica M. Eaglin, *Predictive Analytics' Punishment Mismatch*, 14 I/S: J.L. & POL'Y FOR INFO. SOC'Y 87, 103–04 (2017) (discussing how information is used in predictive analytics).

[66] Ngozi Okidegbe, *Discredited Data*, 107 CORNELL L. REV. 2007, 2014 (2022) (recommending that a new category of information, "community knowledge sources," be used in pretrial algorithmic training data to produce criminal justice algorithms that have more racially and socioeconomically just outcomes).

[67] *Id.* at 2052–56.

[68] Okidegbe, *supra* note 64, at 743–44.

[69] Melissa Hamilton & Pamela Ugwudike, *A 'Black Box' AI System Has Been Influencing Criminal Justice Decisions For Over Two Decades—It's Time To Open It Up*, PHYS.ORG (July 26, 2023), https://phys.org/news/2023-07-black-ai-criminal-justice-decisions.html [https://perma.cc/5EYJ-36M8]; Thomas & Pontón-Núñez, *supra* note 63, at 376–77, 393. Proposals to improve outcomes for Black inmates that include race in the algorithms as a way of avoiding bias likely violate the Fourteenth Amendment of the Constitution. *Id.* at 402–05 (arguing that the use of algorithmic risk assessments is not narrowly tailored, and in many pretrial contexts, the opaqueness of the algorithms is not narrowly tailored and therefore cannot meet burden of proof standards).

[70] Alexandra N. Marlowe, *Robot Recruiters: How Employers & Governments Must Confront the Discriminatory Effects of AI Hiring*, 22 J. HIGH TECH. L. 274, 275 (2022) (discussing algorithms in hiring and potential bias).

[71] Stephanie Bornstein, *Antidiscriminatory Algorithms*, 70 ALA. L. REV. 519, 531 (2018) (discussing the use of algorithms throughout the hiring process).

will interview and receive an offer.[72] As the algorithm learns the employer's preferences, it solicits similar applicants.[73] The algorithms shape the applicant pool by searching for and identifying potential candidates to receive the hiring solicitation, effectively determining who has access to the hiring opportunity.[74]

Algorithms can also select the pool of candidates by ranking the applicants.[75] In determining the strongest candidates for consideration, algorithms may sort the files for minimal qualifications.[76] If a machine learning algorithm is used, the sorting process may use the employers' past screening decisions.[77] Predictive analytics sort and score resumes and assess competencies and personality types.[78]

Finally, algorithms also collect data on potential applicants and use it to predict job performance.[79] Applicants' digital footprints of social and

[72] For example, a study found that targeted ads on Facebook sorted the audience based on traditional gender roles for the advertised positions. The selected audience for supermarket cashier positions was eighty-five percent women while the audience for taxi drivers was seventy-five percent Black. Miranda Bogen, *All the Ways Hiring Algorithms Can Introduce Bias*, HARV. BUS. REV. (May 6, 2019), https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias [https://perma.cc/6WAC-V85D] ("This is a quintessential case of an algorithm reproducing bias from the real world, without human intervention."). Hiring algorithms may also artificially reduce the talent pool. JOSEPH B. FULLER, MANJARI RAMAN & FRANCIS HINTERMANN, HIDDEN WORKERS: UNTAPPED TALENT 3 (2021), https://www.hbs.edu/managing-the-future-of-work/Documents/Hidden%20Workers—Part%20Time%20Potential%2003.13.23.pdf [https://perma.cc/T9DH-M2GB] (concluding that hiring algorithms unfairly eliminated otherwise qualified persons who were previously incarcerated persons, veterans, refugees, immigrants, or those with mental or physical disabilities for not matching specific criteria, such that eighty-eight percent of these individuals were shown to be fully qualified for the position).

[73] *See* LEARNING COLLIDER, HIDDEN BIAS IN HIRING 6 (2021), https://static1.squarespace.com/static/60d0c05ace34212ef5a1131b/t/62ab8039e3a4642b49f2f730/1655406650864/Learning+Collider%27s+White+Paper+-+Hidden+Bias+in+Hiring+-+2022+Master.pdf [https://perma.cc/MY2H-B676]. For example, a training algorithm that looked for applicants named Jared who played lacrosse provides a good example of how training algorithms can impute bias. *Id.*

[74] Pauline T. Kim & Sharion Scott, *Discrimination in Online Employment Recruiting*, 63 ST. LOUIS U. L.J. 93, 114 (2018) ("Not informing people of a job opportunity is a highly effective barrier."); *see also* Bogen, *supra* note 72.

[75] Kelly Cahill Timmons, *Pre-Employment Personality Tests, Algorithmic Bias, and the Americans with Disabilities Act,* 125 PENN ST. L. REV. 389, 400–04 (2021) (discussing how algorithms are used in hiring).

[76] Manish Raghavan & Solon Barocas, *Challenges for Mitigating Bias in Algorithmic Hiring*, BROOKINGS (Dec. 6, 2019), https://www.brookings.edu/articles/challenges-for-mitigating-bias-in-algorithmic-hiring [https://perma.cc/XCB7-74KH] (discussing algorithmic use during the sourcing, screening, interviewing, and selection phases of employee interview hiring).

[77] LEARNING COLLIDER, *supra* note 73, at 4–6.

[78] Raghavan & Barocas, *supra* note 76.

[79] Chien-Chun Chen, Chiu-Chi Wei, Su-Hui Chen, Lun-Meng Sun & Hsien-Hong Lin, *AI Predicted Competency Model to Maximize Job Performance*, 53 CYBERNETICS & SYS.: AN INT'L J. 298, 316 (2022) (discussing how AI can assist with performance predictors).

professional information become the basis of determining their qualifications and even predicting job performance.[80] Predicting job success may include variables, such as tenure, productivity, or performance, as well as positive signals, such as a lack of tardiness or disciplinary action.[81]

Whether and how this incorporates human prejudice into the screening process is an issue of concern.[82] Arguably, while employers use algorithms to root out bias and increase an applicant pool, machine learning algorithms may perpetuate structural bias. Opponents of hiring algorithms argue subjectivity is readily present in many aspects of algorithmic hiring and protected by law.[83] The analytics can result in bias depending on how the algorithm defines the criteria and training variables.[84] Bias in training criteria and variables or a small, weak data set can produce bias.[85] If the underlying data is unfiltered, including past hiring practices, subjective decision-making and structural bias may exist.[86] Although employers have a duty to regularly audit the algorithms for compliance with antidiscrimination laws, current law justifies employers' use of subtle bias when using machine learning-enhanced

[80] Timothy M. Snyder, *You're Fired! A Case for Agency Moderation of Machine Data in the Employment Context*, 24 GEO. MASON L. REV. 243, 263–64 (2016).

[81] Lori Andrews & Hannah Bucher, *Automating Discrimination: AI Hiring Practices and Gender Inequality*, 44 CARDOZO L. REV. 145, 173–74 (2022) (discussing various AI hiring tools to predict workplace performance).

[82] *See* Alex P. Miller, *Want Less-Biased Decisions? Use Algorithms.*, HARV. BUS. REV. (July 26, 2018), https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms [https://perma.cc/RX6F-6QDB] (citing examples of how algorithms are less biased).

[83] Predictive analytics in employment include services that mine and collect data for employers with behavioral and performance data on potential job applicants. *See* Bornstein, *supra* note 71, at 530.

[84] Soojin Jeong, Margaret Sturtevant & Karis Stephen, *Countering Bias in Algorithmic Hiring Tools*, REGUL. REV. (Sept. 11, 2021), https://www.theregreview.org/2021/09/11/saturday-seminar-countering-bias-algorithmic-hiring-tools [https://perma.cc/4L8F-XDJG]. An example of this is Amazon creating a hiring algorithm in 2015 to screen resumes for top talent. *Id.* Amazon trained its algorithm using a decade of resumes from mostly male applicants, which caused the algorithm to replicate historical patterns in discrimination against female applicants. *Id.*

[85] *See* Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 678–87 (2016) (explaining different patterns of algorithmic bias).

[86] *See* Kit Ramgopal, *Using Algorithms and Artificial Intelligence for Hiring Risks Violating the Americans with Disabilities Act, Biden Admin Says*, NBC NEWS (May 12, 2022, 11:00 AM), https://www.nbcnews.com/tech/tech-news/hiring-algorithms-artificial-intelligence-risk-violating-americans-dis-rcna28481 [https://perma.cc/BF5N-RGS2]. Another example comes from a company called iTutorGroup, which hires U.S.-based tutors to provide English-language services to students in China; iTutorGroup allegedly programmed application software to automatically reject female applicants over the age of fifty-five and male applicants over the age of sixty. *Id.*

predictive hiring tools if the employer has a legitimate business justification.[87]

III. PRICE DISCRIMINATION

The criminal justice system, employment, and price discrimination[88] contexts raise issues of race that demonstrate how, even if unintentional, machine learning model algorithms can contain bias. Algorithmic accountability examines the underlying structure to ensure that technology, though innovative, does not trample legal rights.[89] The discussion in this Part turns to the technology of lending before addressing the nexus of algorithmic accountability in policing social networks' algorithms in lending.

A. *The Technology of Lending*

Technology has the potential to make lending more widely available to the financially underserved and significantly improve the way individuals make financial decisions. Using readily available personal social networks, lenders can access highly granular information about consumers. When detailed information about consumers is combined with big data analysis and machine learning, consumers' needs, preferences, and behavior implicate how law should react to this technology going forward.

1. Stereotypes and Perception of Creditworthiness Based on Identity

Historically, minorities have had difficulty in obtaining credit.[90] The negative perceptions of specific consumer identities have fostered

[87] Bias can also enter hiring decisions through the algorithm targeting other factors in resumes such as gender, Black-sounding names, women's colleges, or mention of a disability. *See* Barocas & Selbst, *supra* note 85, at 706–12 (arguing that neither Title VII's disparate treatment nor disparate impact standards specifically remedy algorithmic discrimination).

[88] Michal S. Gal, *Algorithms as Illegal Agreements*, 34 BERKELEY TECH. L.J. 67, 91–92 (2019).

[89] *See* Katyal, *supra* note 42.

[90] Sheila D. Ards & Samuel L. Myers, Jr., *The Color of Money: Bad Credit, Wealth, and Race*, 45 AM. BEHAV. SCIENTIST 223 (2001); Andrea Freeman, *Racism in the Credit Card Industry*, 95 N.C. L. REV. 1071 (2017); *see also* Lena Felton, *Can a Credit Card Be Sexist?*, THE LILY (Nov. 12, 2019), https://www.thelily.com/can-a-credit-card-be-sexist [https://perma.cc/MH7L-JF8L].

systemic obstacles to credit and lending for certain groups.[91] Credit-based discrimination has forced African Americans to use undesired forms of credit to fulfill basic needs.[92]

Several studies have exposed the discriminatory and differential treatment that African Americans endure when seeking to obtain access to credit.[93] In part, minorities are the targeted prey of predatory lending practices because "they are overrepresented in the lower-income levels, due to historical and present structural inequalities, and because of racial discrimination."[94] Lenders also target vulnerable immigrant Latino populations, who are susceptible to predatory lending practices due to language barriers, cultural differences in finances, and an unfamiliarity with U.S. financial products.[95] Knowing these vulnerabilities, lenders

[91] Since the early twentieth century, African Americans have been denied credit because they were deemed uncreditworthy based on the personal beliefs of lenders. *See* Ards & Myers, *supra* note 90, at 227–30 (describing the discrimination against African Americans resulting in their denial of access to preferred forms of credit).

[92] *Id.* at 228 ("The devastating consequence of this historical legacy of discrimination in credit is that Blacks have been overrepresented among those using the worst type of credit available . . . the concentration of Blacks in the bottom of the credit market has contributed to a tainted perception of Blacks' credit risk."). Perceptions of African American consumers are further tarnished by the consistently negative narratives pushed by the mass media. In *Racism in the Credit Card Industry*, Professor Freeman discusses how the creditworthiness of lower-class Black women and Black single mothers was negatively impacted by the "Welfare Queen" narrative that became popular in the late twentieth century. Freeman, *supra* note 90, at 1111–13. Similarly, the creditworthiness of Black men was negatively impacted by the creation of the "thug." *Id.* at 1113–14. Black men are regularly portrayed as thugs through music, television, and social discussion. Several scholars feel that these narratives and discriminatory actions have resulted in the internalization of Black creditworthiness stereotypes within the Black community. *Id.* at 1116–17.

[93] *See, e.g.*, Ethan Cohen-Cole, *Credit Card Redlining* 1–3, 6–7 (Fed. Rsrv. Bank of Bos., Working Paper No. QAU08-1, 2008), https://www.bostonfed.org/publications/risk-and-policy-analysis/2008/credit-card-redlining.aspx [https://perma.cc/2FV3-ZN9V] (comparing the terms and availability of credit card agreements entered into by credit card owners with identical risk profiles and payment histories living in different geographical locations); Chi-Jack Lin, Racial Discrimination in the Consumer Credit Market (2010) (Ph.D. dissertation, The Ohio State University) (OhioLINK) (revealing that being an African American consumer negatively affects the probability of owning a credit card and the amount of credit given); Andrea Freeman, *Payback: A Structural Analysis of the Credit Card Problem*, 55 ARIZ. L. REV. 151, 180–81 (2013) (discussing lenders predatory practices regarding minority borrowers) .

[94] Freeman, *supra* note 93, at 180–81; *id.* at 181 ("Deeply entrenched structural inequality, originating in slavery and reinforced by policy, cultural stereotypes, and segregation, creates the circumstances that allow credit card companies to exploit vulnerabilities in African American households for profit."). The wealth gap, a twenty-to-one difference between the wealth accumulation of African Americans and whites, is one of the primary systemic factors in the negative lending behaviors towards and perceptions of African Americans and credit. *Id.* at 181–86.

[95] Freeman, *supra* note 93, at 186–88.

often invest in targeted advertisements to attract Latino customers.[96] Similarly, women have historically dealt with and currently experience disparate treatment from credit lenders.[97] Despite the existence of the Equal Credit Opportunity Act, lending disparities continue to persist.[98]

2. Fintech Lending

Fintech lending is growing expansively and holds promise for both lenders and borrowers. Lenders use credit scores as screening devices to determine borrower likelihood of repayment and loans' interest rate.[99] Lenders' use of algorithms to assess creditworthiness is not new.

[96] *Id.* at 186–88; *see id.* at 187 n.258; Jeremy M. Simon, *Study: Credit Card Use and Revolving Debt Rising Among Hispanics*, CREDITCARDS.COM (Mar. 9, 2007), https://www.creditcards.com/credit-card-news/credit-card-study-shows-rise-in-hispanics-usage-1276.php [https://web.archive.org/web/20160415200734/https://www.creditcards.com/credit-card-news/credit-card-study-shows-rise-in-hispanics-usage-1276.php].

[97] *See* Taylor Telford, *Apple Card Algorithm Sparks Gender Bias Allegations Against Goldman Sachs*, WASH. POST (Nov. 11, 2019, 10:44 AM), https://www.washingtonpost.com/business/2019/11/11/apple-card-algorithm-sparks-gender-bias-allegations-against-goldman-sachs [https://perma.cc/Q6B2-4C2L]; *see also* Tom Simonite, *The Best Algorithms Struggle to Recognize Black Faces Equally*, WIRED (July 22, 2019, 7:00 AM), https://www.wired.com/story/best-algorithms-struggle-recognize-black-faces-equally [https://perma.cc/QHJ6-ESDX]. The 1960s and 1970s saw a dramatic influx of women entering the work force and gender-based legislation to ensure equal pay. Lenders often denied women from obtaining their own line of credit based on their gender alone. Allen Abraham, Note, *Credit Discrimination Based on Gender: The Need to Expand the Rights of a Spousal Guarantor Under the Equal Credit Opportunity Act*, 10 BROOK. J. CORP. FIN. & COM. L. 473, 477–79 (2016). In 1972, five primary patterns of credit-based discrimination against women were found:

> (1) Single women have more trouble obtaining credit than single men. (2) Creditors generally require a woman upon marriage to reapply for credit, usually in her husband's name. Similar reapplication is not asked of men when they marry. (3) Creditors are unwilling to extend credit to a married woman in her own name. (4) Creditors are usually unwilling to count the wife's income when a married couple applies for credit. (5) Women who are divorced or widowed have trouble reestablishing credit. Women who are separated have a particularly difficult time, since the accounts may still be in the husband's name.

Id. at 477–78. "Further discrimination 'evolved out of the widely-held presumption directed at the probability of pregnancy, the subsequent termination of employment upon childbirth, and the general instability and inability of women to control their personal affairs.'" *Id.* at 478 (quoting Gail R. Reizenstein, Note, *A Fresh Look at the Equal Credit Opportunity Act*, 14 AKRON L. REV. 215, 219 (1981)).

[98] Schwemm & Taren, *supra* note 54, at 405 (discussing discretionary pricing during the subprime lending crisis). Congress passed the Equal Credit Opportunity Act (ECOA) in 1974 in response to the race-based discrimination. *See* 15 U.S.C. § 1691. In 1976, Congress expanded the ECOA to forbid discrimination "on the basis of race, color, religion, national origin, sex or marital status, or age," creating a legal right to have equal access to credit. *Id.*

[99] *See* Mester, *supra* note 23 (outlining the history, use, and methods of credit scoring).

Mainstream lenders, such as banks and credit unions, use credit scores, which are based on a mathematical formula, to measure creditworthiness, which in turn determines interest rates and risk. The credit score itself is a composite of specific categories with varying assigned weights. The score evaluates an individual's use of credit cards and bank services, including negative credit performance such as bankruptcies and foreclosures. These pre-determined categories disadvantage borrowers who are not already in the credit system.

Big data credit scoring, which identifies risks based on detailed inquiries of public and private information about borrowers, fills the gap. Fintech lending relies on machine learning algorithms to evaluate the data. An algorithm based on correlations between multiple variables assigns statistically derived weights among the consumer data points, the digital footprints.[100] The datasets used for lending decisions are voluminous based on the many available sources that create an individual's profile. Big data allows lenders to make lending more profitable.[101] Lenders use big data analytics to check customers' credit and assess the ability to repay. Based on past behavior, machine learning algorithms process available information about potential borrowers to evaluate future risk.[102]

Big data analytics are useful in targeting products to consumers. Using social media platforms, lenders become aware of consumers' spending interests, lifestyles, and preferences. These provide the lender with a broader picture of the consumer and, when combined with machine learning algorithms, accurately predict consumer behavior.[103] Digital lending creates a fuller profile of the loan applicant. A digital credit score bases the loan terms on the data gathered from the individuals' personal networks.[104] Fintech lenders argue that the data

[100] Andreas Tsamados et al., *The Ethics of Algorithms: Key Problems and Solutions*, 37 AI & SOC'Y 215, 223–25 (2022).

[101] Julia F. Hollreiser, Note, *Closing the Racial Gap in Financial Services: Balancing Algorithmic Opportunity with Legal Limitations*, 105 CORNELL L. REV. 1233, 1234 (2020) (describing lenders using big data as "profit-oriented").

[102] *See* LEARNING COLLIDER, *supra* note 73. Similarly, big data assists with fraud detection by denoting a sudden change in borrowers' behavior. Fraud will undoubtedly increase as online technology evolves. Big data analytics can alert a lender instantaneously, increasing the ability to stop deceptive activity. *See* FED. TRADE COMM'N, REPORT TO CONGRESS: COMBATTING ONLINE HARMS THROUGH INNOVATION, 20220922A NYCBAR 154 (discussing AI and antifraud measures in the credit card context).

[103] Griffith, *supra* note 23, at 358 (discussing how lenders use data in AI algorithms).

[104] Cullerton, *supra* note 9, at 814–15 (discussing how social media is used to create credit scores).

helps them make better underwriting decisions, which reduces borrower defaults and lower interest rates.[105]

Big data analytics provide lenders with the insight to handle large, complex pieces of information efficiently. Generating digital credit scores is not limited to financial interactions. The analytics evaluate literally hundreds and thousands of data, such as sleeping and messaging habits, geo-locations, and purchasing history.[106] The predictive behavioral analytics of machine learning collect and sort the borrowers' social and business network identity data. The lenders' perception of tangible factors, such as shopping history, education, internet browsing history, social media associations, health history, employment, or movies, all become relevant.[107] The distinctive properties of big data—large size, high dimension, and complex structure—give lenders more accurate insights into loan applicants.[108]

B. *Algorithmic Redlining*

The proprietary algorithms and advanced scoring methods of big data are far more advanced when evaluating borrowers' credit risk and have proven more accurate in predicting lending risk than traditional scoring methods.[109] As technology evolves, using machine learning to

[105] *See* Christophe Croux, Julapa Jagtiani, Tarunsai Korivi & Milos Vulanovic, *Important Factors Determining Fintech Loan Default: Evidence from a Lending Club Consumer Platform,* 173 J. ECON. BEHAV. & ORG. 270 (2020).

[106] Tsamados et al., *supra* note 100, at 16.

[107] *See* Elizabeth Fernandez, *Will Machine Learning Algorithms Erase the Progress of the Fair Housing Act?*, FORBES (Nov. 17, 2019, 8:30 AM), https://www.forbes.com/sites/fernandezelizabeth/2019/11/17/will-machine-learning-algorithms-erase-the-progress-of-the-fair-housing-act/#779c1f1b1d7c [https://perma.cc/MD54-JL2J] (describing a proposed Department of Housing and Urban Development ruling that "landlords, lenders, and property sellers who use third-party machine learning algorithms to decide who gets approved for a loan or who can purchase or rent a property would not be held responsible for any discrimination resulting from these algorithms").

[108] *See* Itay Goldstein, Chester S. Spatt & Mao Ye, *Big Data in Finance*, 34 REV. FIN. STUD. 3213, 3215–17 (2021). Big data can provide more nuanced interpretation of data when certain features are present. First, large size datasets overcome the common problem in data analytics of sample bias. *Id.* at 3215. The data is broken into subsets and compartmentalized based on varying characteristics including activities, time, or any specified distinction. *Id.* Data with "high dimension" refers to information that can be extracted from the data's variables. *Id.* This assesses both the ways in which variables interact among other variables and the efficacy of the predictions based on them. *Id.* Data complexity refers to the amorphous nature and the variety of the data. *Id.* at 3216. When the underlying features of text, pictures, videos, and audio are extracted, there is deeper interpretation. *Id.*

[109] Si Shi, Rita Tse, Wuman Luo, Stefano D'Addona & Giovanni Pau, *Machine Learning-Driven Credit Risk: A Systemic Review,* 34 NEURAL COMPUTING & APPLICATIONS 14327, 14332 (2022).

understand human behavior will expand even more, with machines becoming the decision-makers.[110] The financial regulatory structure is designed currently to monitor human behavior. Regulators have not monitored algorithms for biases.

Using alternative data algorithms to assess creditworthiness is not innocuous. Although there are countless claims that algorithms are objective and even "colorblind" for minorities,[111] the risks of discrimination are ever-present. These types of concerns have been raised repeatedly in other sectors. For example, in the criminal justice arena, the predictive analyses are labeled as perpetuators and sources of continuing inequality.[112] The central problem is two-fold—data sources based on systemic racism and discrimination, and creators who are oblivious to bias. Specifically, code source and machine learning in the financial sector, like their use in other areas, raise intrinsic fairness issues and other challenges.[113]

What requires examination is how existing regulations need to be adapted to monitor machines.[114] The fairness of algorithms overlaps with privacy issues. Data is a significant source of revenue for lenders.[115] Specifically, it is problematic that in lending, the underlying algorithm affects transactions that are not yet executed or realized. It is difficult to measure the impact of unfairness or detect the potential discriminatory aspect under the existing regulatory scheme.[116] How big data is collected, analyzed, and used is an issue of both consumer privacy and ethics.

[110] Bob Lambrechts, *May It Please the Algorithm: The Future of A.I. in the Practice of Law*, J. KAN. BAR. ASS'N, Jan. 2020, at 36, 40 (arguing that systems will eventually replace humans as decision-makers).

[111] Kim, *supra* note 60, at 1592–93 ("This reinforces the mistaken belief, common among non-technical people, that algorithms are objective and neutral . . . .").

[112] Griffith, *supra* note 23, at 360 (citing studies evidencing that AI algorithms accurately predict default risks).

[113] Alice Xiang, *Reconciling Legal and Technical Approaches to Algorithmic Bias*, 88 TENN. L. REV. 649, 659–60 (2021) (discussing the "negative side effects" of algorithms).

[114] Nydia Remolina, *Open Finance: Regulatory Challenges of the Evolution of Data Sharing Arrangements in the Financial Sector* 23 (Oct. 24, 2019) (Sing. Mgmt. Univ. Ctr. for AI & Data Governance Rsch. Paper No. 2019/05), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3475019 [https://perma.cc/PDZ5-38JF] (discussing the regulatory concerns that need to be addressed with the advent of open banking technology).

[115] John L. Douglas, *New Wine into Old Bottles: Fintech Meets the Bank Regulatory World*, 20 N.C. BANKING INST. 17, 27–28 (2016) (discussing how data provides links to potential customers).

[116] Xiang, *supra* note 113, at 705 (discussing the importance and difficulty of establishing causality in algorithmic bias).

Lenders may have unfair access to data, raising issues about accountability in algorithmic systems.[117]

For the credit-invisible, many of whom are minorities, there is a particular threat that using the data gathered from the individual's social network identity will lead to an undetectable discriminatory or biased assessment. Digital footprint algorithms, if left unmonitored, can be harmful to borrowers—just as redlining and subprime lending have been—but are much less obvious.[118]

Social context is foundational in examining how technology is created and adopted. In the financial sector, big data may close the information asymmetry between lenders and nontraditional borrowers. The borrowers that banks or credit unions consider too risk-averse are sought after by lenders who are willing to use soft information gained through profiling an individual's social network. The traditional credit scoring models do not include an assessment of data that could be helpful in assessing risk for this group. Yet, just as traditional credit assessments use proxies for race (for example, wealth, higher education, and savings accounts), there is concern that fringe data assessments are opportunities for predation.[119]

While the law appears passive to the dilemmas that technologies are often offered to resolve, the threat of racial and economic biases by individual actors is ever-present.[120] Current legal rules imagine technologies as "code" simply achieving its stated objectives.[121] Such a narrow view of the law's intersection with technology ignores underlying, preexisting values. Law creates social forces and conditions just as much as it regulates the technologies created to address them. As one scholar posits: "Technologies enter a world shaped by law just as much as they

[117] Michael Akinwumi, John Merrill, Lisa Rice, Kareem Saleh & Maureen Yap, *An AI Fair Lending Policy Agenda for the Federal Financial Regulators*, BROOKINGS (Dec. 2, 2021), https://www.brookings.edu/articles/an-ai-fair-lending-policy-agenda-for-the-federal-financial-regulators/ [https://perma.cc/5FYH-6R4H] ("In some respects, the U.S. federal financial regulators are behind in advancing non-discriminatory and equitable technology for financial services.").

[118] *See* Allen, *infra* note 125. Redlining is the illegal practice which bases lending on the geographical location of the property. Sub-prime lending, the centerpiece of the global financial crisis, arbitrarily elevated the risk level of minority borrowers by providing high interest rate loans to minority borrowers, regardless of their credit status.

[119] Jason Jia-Xi Wu, *Algorithmic Fairness in Consumer Credit Underwriting: Towards a "Harm-Based" Framework for AI Fair Lending*, 21 BERKELEY BUS. L.J. (forthcoming 2024) (manuscript at 76), https://ssrn.com/abstract=4320444 [https://perma.cc/S8AR-4VD2] (discussing proxies for race); *see also* Saule T. Omarova, *New Tech v. New Deal: Fintech as a Systemic Phenomenon*, 36 YALE J. ON REG. 735, 745 (2019).

[120] *See* Ifeoma Ajunwa, *The Paradox of Automation as Anti-Bias Intervention*, 41 CARDOZO L. REV. 1671 (2020).

[121] Jay P. Kesan & Rajiv C. Shah, *Shaping Code*, 18 HARV. J.L. & TECH. 319, 372–73 (2005) (discussing governmental policy in shaping code).

shape that world. As such, a broader orientation in legal scholarship is necessary to begin to fully grapple with the issues that technology raises in society in this historical moment."[122] The use of alternative data and predictive analytics to "help" damaged or otherwise credit-impaired borrowers cannot be allowed to do more harm than good. To the extent that these analytics embed and identify groups within markets, objective factors are in fact subjective. These obscure analytics can easily produce hidden discriminatory lending practices. Antisubordination algorithmic analysis is critical to assuring that the consumer benefits from this type of fintech lending.

As discussed below in Part IV, the implicit legal protections of computer code need to be eased to reveal the complex set of relationships and structures that impact lending based on network identity. Access to credit based on machine learning and predictive analytics requires unique legal protections to be equitable.

IV. TOWARDS ALGORITHMIC FAIRNESS IN BIG DATA LENDING

Machine learning algorithms have led to many positive developments in finance and lending. Fundamental questions remain, not the least of which is how to address discrimination. As argued above, the existing fair lending regime is ineffectual in preventing the use of predatory scoring techniques.[123] Arguments for transparency and disclosure abound.[124] Yet, transparency and disclosure alone will not address the impact of big data and discrimination in the financial sector.[125] The antisubordination lens calls for remedies that address the

[122] Jessica M. Eaglin, *When Critical Race Theory Enters the Law & Technology Frame*, 26 MICH. J. RACE & L. 151, 158 (2021).

[123] Barocas & Selbst, *supra* note 85 (discussing how big data proxies can introduce implicit bias, which is not actionable under the disparate impact theory).

[124] *See id.* at 714–28; *see* Stephen Buranyi, *Rise of the Racist Robots—How AI Is Learning All Our Worst Impulses*, THE GUARDIAN (Aug. 8, 2017, 2:00 PM), https://www.theguardian.com/inequality/2017/aug/08/rise-of-the-racist-robots-how-ai-is-learning-all-our-worst-impulses [https://perma.cc/G32B-Q5PZ] ("It's unclear how existing laws to protect against discrimination and to regulate algorithmic decision-making apply in this new landscape. Often the technology moves faster than governments can address its effects. . . . Sinyangwe recently worked with the ACLU to try to pass city-level policies requiring police to disclose any technology they adopt, including AI. But the process is complicated by the fact that public institutions adopt technology sold by private companies, whose inner workings may not be transparent. 'We don't want to deputise these companies to regulate themselves,' says Barocas.").

[125] Solutions involving transparency of algorithms include amending the Computer Fraud and Abuse Act to allow third parties to audit algorithmic processes, including source code, for discrimination, and requiring private companies to fully disclose to the affected parties when algorithms are used in decision-making. *See* Peggy Bruner, *A Case Against Bad Math*, 22 J. TECH. L. & POL'Y 1, 17–19 (2018); Allen, *supra* note 7, at 256–58.

intersection between race, technology, and the law, as applied in the lending context.

A. *Algorithmic Inclusion*

The promise of fintech is access to financial markets, particularly credit, for those who are underserved.[126] Fintech also aids in economic development by broadening and deepening credit markets. Filling the gap in financial services improves the functioning of a financial system by expanding its reach. The ability to participate fully in a market economy requires that all individuals have beneficial and affordable financial products and services. The imperfections in financial markets constrain credit. Individuals are excluded from the formal banking sector due to informational asymmetries.[127]

Fintech undoubtedly *can* reduce the information and transaction costs that exclude participation. Public policies designed to address inclusion in financial services development must address both the inadequate capacity of financial institutions to deliver the products and services appropriate to these segments and regulate these special strategies. Developing the concept of "inclusion" is paramount.

Inclusion by necessity connotes diversity.[128] Whether and how banking institutions welcome underserved individuals into the financial mainstream is also a function of how public policy both supports and protects that participation. The responsibility of government is to make banking and credit access inclusive and available to all who qualify. Similarly, the regulation of technology is a governmental function.[129] The intersection and overlap between inclusiveness and access to financial services acknowledges that established distinctions need to be modified to accommodate different credit circumstances. In this regard, big data algorithms represent the latest needed modification. Thus, regulations should ensure their inclusiveness and fairness. The structural barriers that have created institutional inequities may perpetuate discrimination if not identified and addressed. Democratic principles require that innovation

[126] Kristin Johnson, Frank Pasquale & Jennifer Chapman, *Artificial Intelligence, Machine Learning, and Bias in Finance: Toward Responsible Innovation*, 88 FORDHAM L. REV. 499, 504 (2019) (discussing potential fintech customers).

[127] Havard, *supra* note 55.

[128] Tara Nair, *In Pursuit of an Inclusive Finance-scape in India: Changing Course, Shifting Goals*, 2016 INDIA: SOC. DEV. REP. 278.

[129] Kim Vu-Dinh, *Black Livelihoods Matter: Access to Credit as a Civil Right and Striving for a More Perfect Capitalism Through Inclusive Economics*, 22 HOUS. BUS. & TAX L. J. 1, 22 (2021) (discussing the availability of affordable banking).

not be rampant and untethered to governmental policy, resulting in harm to consumers.

B. *Algorithmic Oversight*

When lenders use digital footprint algorithms, consumers should know what underlying information the lender is using. Only then will the borrower be able to challenge the use of any incorrect or inaccurate information. Algorithmic oversight requires recognizing the human involvement in the decision-making and exercising oversight over it.

Congress is considering establishing regulatory guardrails on emerging AI technology.[130] Presently, American consumers do not have control over their financial data.[131] By contrast, the European Union established the consumer's "right to know," explicitly recognizing that data gathering and use often occurs without consumers' knowledge or consent.[132] Congress should separately and purposefully regulate this sphere, including how financial data is controlled by third-party data aggregators.[133]

The transformation of big-data and machine-learning techniques present challenges to the credit-scoring industry, regulators, and consumers.[134] Of significant concern are the interdependent variables

[130] Karoun Demirjian, *Schumer Lays Out Process to Tackle A.I., Without Endorsing Specific Plans*, N.Y. TIMES (June 21, 2023), https://www.nytimes.com/2023/06/21/us/ai-regulation-schumer-congress.html [https://web.archive.org/web/20230727222912/https://www.nytimes.com/2023/06/21/us/ai-regulation-schumer-congress.html].

[131] Nizan Geslevich Packin, *Show Me the (Data About the) Money!*, 2020 UTAH L. REV. 1277, 1316 (contrasting the U.S. market-based approach to protecting consumer financial data); Asress Adimi Gikay, *The American Way-Until Machine Learning Algorithm Beats the Law?*, 12 CASE W. RSRV. J.L. TECH. & INTERNET, no. 2, 2021, at i, 5 (discussing the "weakness" of consumer data protection in the United States).

[132] The European Union enacted the General Data Protection Regulation, which gives European citizens control over their personal information, including the right to know when automated decision-making is used. Specifically, under Article 22, individuals "have the right not to be subject to a decision based solely on automated processing." *See* Regulation 2016/679, of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119) 1.

[133] Data aggregators, who collect the data, including digital footprints, do not have to comply with a regulatory structure, which means that there are no mandated procedures for transparency or correction of errors. The CFPB is currently collecting information as it considers how to account for and regulate data errors. *See* Small Business Lending Data Collection Under the Equal Credit Opportunity Act (Regulation B), 86 Fed. Reg. 56356 (proposed Oct. 8. 2021) (to be codified at 12 C.F.R. pt. 1002).

[134] While both the Fair Credit Reporting Act and the Equal Credit Opportunity Act would make discriminatory algorithms actionable, neither is tailored to effectively prohibit potential violations. Griffith, *supra* note 23, at 367–68.

that present risks to financially excluded consumers. Beyond the issue of biased and discriminatory scoring is the possibility that the very vulnerable consumers who need access to credit are in fact targeted for high-risk loans.[135]

Moreover, digital footprint algorithms raise unique legal issues requiring a different type of remedy when policing for lending discrimination. First, discrimination may be undetectable. The potential for proxy-based discrimination being undiscoverable by consumers is significant. Because the algorithms may use data points or combinations of data points correlated to immutable characteristics, discrimination may be obscured, and consumers oblivious to a possible claim. Second, consumers do not have defined privacy rights regarding social network identities. Specifically, unlike in the European Union, lenders are not required to disclose the use of digital footprint algorithms. Consumers are unable to challenge any personal misinformation and are not allowed to prohibit or limit its use. Finally, under current trademark protections, a lender's code selections are proprietary.[136] Exactly how the code design employs social media data to identify borrowers, assess creditworthiness, and offer loan terms and products can only be examined after becoming operational. A lender's decision to target vulnerable consumers for predatory products cannot be discovered until serious harm has occurred. Lenders' accountability for developing an injurious algorithm is untimely deferred.[137]

As in the civil rights context, preclearance of digital algorithms is necessary to avert the harm of persistent discrimination in lending. The unconventional solution of section 5 of the Voting Rights Act required jurisdictions with a history of voting law violations, principally, but not entirely located in the South, to seek preclearance changes to voting procedures before implementation.[138] At that time, Congress recognized

[135] Mikella Hurley & Julius Adebayo, *Credit Scoring in the Era of Big Data*, 18 YALE J.L. & TECH. 148, 149 (2016).

[136] *See* Katyal, *supra* note 42, at 59 ("Because many algorithms are proprietary, they are resistant to discovery and scrutiny.").

[137] This result is similar to what occurred during the subprime lending crisis. During the height of the subprime lending crisis, African American consumers "were almost four times more likely to have a subprime loan than white consumers . . . , and Hispanics were almost three times more likely." Andre K. Gray, *Caveat Emptor: Let the Borrower Beware of the Subprime Mortgage Market*, 11 U. PA. J.L. & SOC. CHANGE 195, 224 (2008).

[138] Voting Rights Act, 42 U.S.C. § 1973 (2000). In *Allen v. State Board of Elections*, the Supreme Court acknowledged that section 5 preclearance was aimed at "subtle as well as the obvious" regulations that denied citizens the right to vote because of their race, starting that "voting includes 'all action necessary to make a vote effective.'" 393 U.S. 544, 565–66 (1969) (quoting § 1973). In 2013, the Supreme Court determined that preclearance of new voting changes was not required. *See* Shelby County v. Holder, 570 U.S. 529, 529–30, 557 (2013). The Court found section

that states with a history of denying African American citizens the right to vote continued to engage in obstructionist tactics.[139] Congress's approach to section 5 was proactive and recognized that to remedy discrimination in voting the harm of denying fair access to participate in elections could not be undone after an election had occurred.[140] Preclearance struck the balance between the denial of voting rights and expeditiously enjoining an election for potential discriminatory changes, by requiring advanced approval so the rights of persons traditionally discriminated against could be protected.[141]

Having a lender take responsibility for an algorithm before it becomes operational is critical to ensuring that big data lending does not become predatory. The risk that these sophisticated tools will be used to identify vulnerable individuals, who will be most susceptible to predatory loan products, outweighs allowing the developers and lenders who use them to defer responsibility for their harm. To both ensure fairness in lending algorithms and avert their potential harm, Congress should require that these lenders submit these tools for monitoring and auditing prior to their use.[142]

4(b) of the Voting Rights Act, which sets out a coverage formula identifying which jurisdictions are subject to the preclearance requirement, invalid. Although the Court refused to determine the constitutionality of section 5, it has nevertheless been rendered ineffective due to the lack of a proper coverage formula.

[139] Kareem Crayton & Terry Smith, *Unteachable: Shelby County, Canonical Apostasies, and A Way Forward for the Voting Rights Act*, 67 SMU L. REV. 3, 49 (2014); *id.* at 15 n.77 ("The legislative history reveals that the basic purpose of Congress in enacting the Voting Rights Act was 'to rid the country of racial discrimination in voting.' Section 5 was intended to play an important role in achieving that goal: 'Section 5 was a response to a common practice in some jurisdictions of staying one step ahead of the federal courts by passing new discriminatory voting laws as soon as the old ones had been struck down.'" (quoting Beer v. United States, 425 U.S. 130, 140 (1976))).

[140] Gilda R. Daniels, *Unfinished Business: Protecting Voting Rights in the Twenty-First Century*, 81 GEO. WASH. L. REV. 1928, 1936 (2013) (discussing the history of the Voting Rights Act).

[141] In *South Carolina v. Katzenbach*, the state of South Carolina challenged section 5 as unconstitutional. 383 U.S. 301, 307 (1966). In upholding the constitutionality of the provision and rejecting a case-by-case approach, Chief Justice Warren cited the historical necessity of prompt and effective action in eradicating racial voting discrimination. *Id.* at 308. Additionally, Warren found "that exceptional conditions can justify legislative measures not otherwise appropriate." *Id.* at 334.

[142] Andrew Tutt, *An FDA for Algorithms*, 69 ADMIN. L. REV. 83, 90–91, 122–23 (2017); *see also* Oren Bar-Gill & Elizabeth Warren, *Making Credit Safer*, 157 U. PA. L. REV. 1, 98–99, 98 nn.323 & 325 (2008).

C. *Algorithmic Accountability*

A regulatory agency with enforcement authority provides the type of safety that consumers need in this context.[143] The agency would be charged with reviewing algorithms before they are released on the market. That pre-market approval process could provide an opportunity for the agency to require that companies substantiate the safety performance of their algorithms. The agency, because it would be steeped in technological expertise, can determine the need for and extent of an audit of an algorithm. The regulator would have the capability to stay current with the changes in technology and to monitor algorithmic products.

Lenders, out of necessity, should bear the burden of authenticating scoring algorithms. By their very nature, these formulae are opaque. Consumers are unable to ascertain whether the offered financial products based on digital footprint scoring are safe and responsible or unsafe and possibly predatory. The pre-market approval process could provide an opportunity for the agency to require that companies substantiate the financial viability and performance of the algorithms. The agency could create a performance standard developed with industry input. Algorithms could also be approved with use conditions, making use beyond the established criteria subject to sanctions.[144]

The ex ante enforcement power would be the most significant power.[145] The primary reason is mitigation of harm. Unlike the present fair lending scheme which is ex post and puts the burden of proving discrimination on the plaintiff, this approach would be more deterrent and protective.[146] The agency would have the expertise to evaluate the technology, requiring lenders to reveal the code and training data, the exact metrics or data points that generate scores or determine borrower interest rates, to determine that the algorithms are safe. Only then can the agency properly exercise its authority to limit or prohibit the implementation of any algorithm that is hazardous and wealth-stripping.

[143] Arguably, the CFPB has the authority to implement regulations to require this type of review. The structural and institutional questions that surround creating the agency are beyond the scope of this Article.

[144] The regulatory agency would work with developers to set performance and safety standards, and could establish safe harbor and conditional approval standards. Unapproved algorithms should be subject to the highest enforcement sanctions and fines, including severe penalties and allowing consumers a private right of action.

[145] Jenigh J. Garrett, *The Continued Need for the Voting Rights Act: Examining Second-Generation Discrimination*, 30 ST. LOUIS U. PUB. L. REV. 77, 77–78 (2010) (discussing congressional findings leading to the reauthorization of section 5 of the Voting Rights Act).

[146] Plaintiffs using a disparate impact theory of lending discrimination have the initial burden of proving the lender's wrongdoing. *See* Schwemm & Taren, *supra* note 54, at 415–22.

Arguments against preclearance are numerous. Repeated most often, after the copyright protections discussed above, are arguments about stifling innovation and subjectivity of review.[147]

First is the conflict between the rapid development of technology and the review of it. The argument is that audits and assessments are slowed down and may be counterproductive if approval is not equally swift. The possibility of the auditing process becoming bureaucratic may either thwart the process or encourage developers to circumvent review.[148]

The second most popular point of dispute concerns algorithmic design. Creating an algorithm is by necessity a trade-off among variables. Any given system balances among others, issues of privacy, safety, security, objectivity, and accuracy.[149] Opponents argue therefore that the evaluation of any design scheme is subjective.[150] The solution from opponents of disclosure and review is market efficiency and self-regulation.[151] Opponents posit that companies respond to demand by consumers for safe and affordable products and services. A pro-innovation approach argues against any regulatory control.[152] The alternative is for companies to follow industry best practices to remain competitive.

Disclosure is only a first step to solving the problems raised by a lack of privacy for financial data. Transparency alone will not address the harms that digital footprint lending creates. The unregulated, complex safety risks—some of which lenders are aware of and hide in the "black boxes" others, unintended—pose safety issues for consumers. The arguments opposing disclosure protect algorithms and their profit-

[147] *See* Yafit Lev-Aretz & Katherine J. Strandburg, *Privacy Regulation and Innovation Policy*, 22 YALE J.L. & TECH. 256, 309–10, 317 (2020); *see also* Toni Lester & Dessislava Pachamanova, *The Dilemma of False Positives: Making Content ID Algorithms More Conducive to Fostering Innovative Fair Use in Music Creation*, 24 UCLA ENT. L. REV. 51, 68 (2017) (discussing the subjectivity of assessing copyright violations in the music industry using algorithms).

[148] Thierer & Chilson, *supra* note 137.

[149] *Id.*

[150] *Id.*

[151] *See generally* Dirk A. Zetzsche, William A. Birdthistle, Douglas W. Arner & Ross P. Buckely, *Digital Finance Platforms: Toward a New Regulatory Paradigm*, 23 U. PA. J. BUS. L. 273, 333 (2020) ("The downside of self-regulation is the dependency of the 'self-regulated constituency' on adopting rules. Where the collective private and public interests collide, we might expect few serious efforts at self-regulation.").

[152] Dirk A. Zetzsche, Ross P. Buckley, Janos N. Barberis & Douglas W. Arner, *Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation*, 23 FORDHAM J. CORP. & FIN. L. 31, 52 (2017) ("The proponents of free markets often characterize regulation as simply an unnecessary cost to business.").

making capabilities. Eventually, the consuming public loses its confidence in the ability of the lending algorithms to be effective.

CONCLUSION

AI algorithms have led to many positive developments in finance and lending. Digital footprint credit scoring is an innovative way to ensure access to credit for those who are underserved. However, fundamental questions remain. Unaddressed are the practices and regulations that will result in the responsible use of personal, network data, algorithms, and machine learning. Also unaddressed is the issue of race in lending algorithms. A common presumption is that there is no human decision-making involved in algorithmic design and that all the data is neutral. Algorithmic design may indeed embed stereotypes, immutable characteristics, cultural and political assumptions, or proxies that negatively impact the result. As argued above, the existing fair lending regime is ineffectual in preventing training data from replicating bias and discrimination. Race inevitably intersects with law and technology with what is included or excluded. Left unpoliced, an algorithm can begin a continuous cycle of discrimination. For law to cloak these technological advances in neutrality and provide the protection of secrecy denies the power of algorithms to harm and ensures that those who use them for profit-making will not be answerable for the resultant injury. The antisubordination lens calls for remedies that address the intersection between race, technology, and the law as applied in the lending context. Algorithmic justice requires incorporating principles of inclusion, oversight, and accountability into law and regulation.