Written Statement of Brenda Leong
U.S. House of Representatives
Committee on Financial Services, Task Force on AI
Rayburn House Office Building
Washington, D.C. 20515
February 12, 2020

Thank you for the opportunity to provide a written statement for the record of the hearing on Equitable Algorithms: Examining Ways to Reduce AI Bias in Financial Services with The Task Force on Artificial Intelligence. My name is Brenda Leong, and I am Senior Counsel and Director of AI and Ethics at the Future of Privacy Forum (FPF). FPF thanks the Task Force Chair and Ranking Member for convening this hearing, and for working to address the privacy and civil liberties challenges of the use of artificial intelligence and machine learning-based applications in financial services products and services, and specifically how to protect those systems from the impacts of undesired or unintended bias.

We submit this statement to:

- Observe that automated decision-making is not new in the financial services sector, and that AI-powered programs and services remain subject to the regulatory and compliance structures in place to protect consumers,
- Describe beneficial ways that financial institutions are using AI to gain efficiencies or add capabilities: to combat fraud, extend credit to traditionally underserved individuals, improve internal research and analysis and customer service functions,
- Identify several factors that can present fairness and equity concerns that are unique or heightened by processing within an AI or Machine Learning-based system, and to
- Identify the technical, policy, regulatory and legislative actions that can help mitigate risk and bias from the use of these systems.

About Future of Privacy Forum:

FPF is a nonprofit organization that serves as a catalyst for privacy leadership and scholarship, advancing principled data practices in support of emerging technologies. We believe that the power of information technology is a net benefit to society, and that it can be well-managed to control risks and offer the best protections and empowerment to consumers and individuals.

FPF has a substantial portfolio of work regarding the privacy, bias, and fairness issues surrounding Artificial Intelligence (AI), across many industry applications. We analyze policy proposals and provide feedback to policymakers. We speak with stakeholders – including leaders from the corporate, public sector, and non-profit communities – to exchange best practices and knowledge regarding machine learning models. After an extensive development process, we published Privacy Expert's Guide to AI and Machine Learning,[1] and created a continuously updated set of resources for Ethics, Governance and Compliance news and guides,[2] and Artificial Intelligence and Robotics Publications.[3] These references comprise a compendium of information for those seeking guidance and updated analysis of the various challenges of machine learning applications in a variety of contexts, focusing on the challenges in common across industries.

[1] Future of Privacy Forum, Privacy Expert's Guide to AI and Machine Learning, September 2018, https://fpf.org/2018/10/18/fpf-release-the-privacy-experts-guide-to-ai-and-machine-learning/

[2] Future of Privacy Forum, Ethics, Governance and Compliance Resources, https://sites.google.com/fpf.org/futureofprivacyforumresources/ethics-governance-and-compliance-resources?authuser=1

[3] Future of Privacy Forum, AI and Robotics Academic Publications, https://sites.google.com/fpf.org/futureofprivacyforumresources/artificial-intelligence-and-robotics-academic-publications?authuser=1

I. Introduction

Artificial Intelligence technology continues to evolve and appear in new contexts in the financial services sector. There are several main uses and functions that benefit from AI, including: Trading Algorithms, Digital Identity Verification, Credit Scoring, Process Automation, Fraud Detection, and Anti-Money Laundering. New applications are being considered all the time for both "back office" functions and in consumer-facing opportunities.

There are, however, specific concerns about the privacy protections needed for the responsible use of this expanding technology, particularly in a highly regulated area such as financial services companies. In this sector more than any other, trust in the fair and equitable impacts of AI is critical to creating a foundation of protections for personal data. Concerns around bias must be carefully understood and managed to ensure appropriate policy and regulatory controls.

II. AI and Machine Learning Are Being Used and Considered for a Variety of Beneficial Applications in Financial Services

"Artificial Intelligence" has become a catch-all phrase used to describe automated systems of all kinds. But it is important when considering consumer risks, as well as regulatory approaches, that the technology be specific and defined.[4] Machine Learning (ML) is the primary type of AI in use or being considered for Financial Services Applications, but not every form of AI is based on Machine Learning. AI includes natural language processing, much robotic process automation, machine learning, and within ML, the use of neural networks.

In the financial services industry – including commercial banks, retail banks, stock brokers, insurance companies, and others – AI is being incorporated in a variety of products and

[4] G. Zhe Jin, Artificial Intelligence and Consumer Privacy, January 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3112040

services. These may include setting interest rates for mortgages, savings accounts, and student loans; recommendations for approving or rejecting credit card and loan applications; and offering or setting the terms for insurance policies. One common use across various parts of the Financial Services industry is fraud prediction and prevention. AI powers the "RegTech" or regulatory technology that allows banking firms to stay in compliance with "Know Your Customer" requirements and Anti-Money Laundering regulations.[5] AI is also used for identity verification (device fingerprinting; personal logins), and interacting with virtual assistants[6], such as "chat bots" that help consumers set up an account, access help, or even provide long term investment strategy advice.[7]

Despite this extensive list, however, AI is used to a more limited degree in the Financial Services sector than many people might expect. Even in back office functions (predicting server down times; tracking data usage and flow; and staff management) where AI is used for improved process, efficiency, and accuracy, financial institutions most commonly keep people in the loop, using the system recommendations as an input to a human's final review or decision. These organizations broadly realize that there is still much uncertainty as to impact of these models, in both practical, and legislative compliance related aspects. Many have determined that the

---

[5] Sudipto Ghosh, AiThority.com Primer on What is RegTech: Definitions, Stats and Tools, AiThority, February 3, 2020, https://www.aithority.com/ait-featured-posts/aithority-com-primer-on-what-is-regtech-definitions-stats-and-tools/

[6] Capitol One Finds Ways to Make Its Digital Assistants More Proactive, Donna Fuscaldo, Forbes, January 2020, https://www.google.com/url?q=https://www.forbes.com/sites/donnafuscaldo/2020/01/31/capital-one-finds-ways-to-make-its-digital-assistant-more-proactive/%235e7d97211706&sa=D&ust=1581457597308000&usg=AFQjCNEeXhD1Vr3u0DXi0Du2Gy_dRYB81g

[7] This $11 Billion Tech Investment Could Disrupt Banking, JP Morgan Chase, https://www.google.com/url?q=https://www.jpmorganchase.com/corporate/news/stories/tech-investment-could-disrupt-banking.htm&sa=D&ust=1581457597307000&usg=AFQjCNHR-qJ91Q2SLSnGH-zqnorScMT-ZQ

maturity of these systems is such that while much may be implemented internally,[8] client facing features must be adopted slowly and carefully.

Other fintech areas where AI is being tested or considered include:

- Character recognition systems for medium term note issuance – this allows analysis and sharing of output data faster, more reliably
- Wire transfer processing
- Contract review (language search and analysis) – where the system is trained for targeted language and then processed for faster review and more consistent products

And like any business in any industry, fintech organizations may employ ML-based systems for HR processing; employee monitoring; machine monitoring; facility access; and cyber security.

Much of this is not new. Financial service providers have long engaged statistical and probability models as well as predictive analytics to forecast performance and evaluate risk. Now, with the inclusion of larger and more complex databases, and the availability of new methods of analysis, many fintech firms deploy extremely complex algorithms to predict the ROI, profitability, and repayment risks. Automation may be able to provide objective analysis using model-based assessments of a borrower's creditworthiness with the ability to better control for bias than traditional reviews subject to the limits of the human reviewer(s). At scale, the application of learning algorithms in credit markets may allow firms to consider nontraditional data in assessing creditworthiness and potentially integrate historically excluded individuals, expanding access to credit to the unbanked in the United States, as well as individuals globally who lack access to financial services.[9]

---

[8] DerivativePath, https://www.derivativepath.com/ (as an example, the use of AI for foreign exchange and derivative management)

[9] Kristin Johnson et al., Artificial Intelligence, Machine Learning, and Bias in Finance: Toward Responsible Innovation, Fordham Law Review, Vol. 88, Issue 2, 2019, https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=5629&context=flr

Some early examples of fintech firms promising to better integrate underserved communities were those who introduced digital money transfer services, the equivalent of cash exchanges (most familiarly using app platforms such as Venmo) as well as platforms that offered digitally distributed credit application functions. Facilitating cash exchanges provides opportunities for those who lack access to conventional banks with personal checking and savings accounts. And expanding credit markets by using sophisticated algorithms may increase the opportunities to offer credit – a necessary step for financial growth.[10] These services do also raise related concerns, including transparency and accountability on the part of the fintech organizations, along with the social impacts of determining "fairness" in credit markets and interest rates, marketing techniques, and structuring of credit products.[11] Many consumer advocates remain cautious. Even though "exclusionary and predatory" credit market practices are legally prohibited, discriminatory processes and inequitable outcomes persist.[12] Given the fears of exploitation and abuse of unbanked communities and higher risk credit applicants, plans for market expansion based on automated decision-making should be carefully considered.[13]

Automated decision-making processes in the financial services sector are built upon the combination of massive data built on the past and the freshest data from today. This means that decision-making algorithms can be "adversely trained" or taught to make sub-optimal decisions

---

[10] Duke University, FinReg Blog, November 2018, https://www.google.com/url?q=https://sites.duke.edu/thefinregblog/2018/11/14/fintech-lending-risks-and-benefits/&sa=D&ust=1581457597364000&usg=AFQjCNH5l84hUXE3c4r99iBd3ohjy0oGdA

[11] Kristin Johnson et al., Artificial Intelligence, Machine Learning, and Bias in Finance: Toward Responsible Innovation, Fordham Law Review, Vol. 88, Issue 2, 2019, https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=5629&context=flr

[12] Matthew Adam Bruckner, The Promise and Perils of Algorithmic Lenders' Use of Big Data, Chicago Kent Law Review, Volume 93, Issue 1 FinTech's Promises and Perils, 2018, https://scholarship.kentlaw.iit.edu/cgi/viewcontent.cgi?article=4192&context=cklawreview

[13] Tanaya Macheel, PayPal-TIO Deal Could Increase Venmo Revenue, Utility, Tearsheet, February 20, 2017, https://tearsheet.co/modern-banking-experience/paypal-tio-deal-could-increase-venmo-revenue-utility/

based upon short term variations in macro-economic forecasts, micro-economic trends, and local consumer consumption patterns. Without constant and effective monitoring of the performance of automated decision systems, such a system for approval of mortgage applications could be "adversely trained" to recommend approvals for applicants from areas with a higher income than other areas, even if that area has not historically been an area of high wealth or credit potential. Offering differential mortgage approvals based upon trends that adversely train AI systems is one form of undesirable biases resulting in disparate impacts.

III. Bias in Machine Learning Algorithms is a Complicated Problem, Implicating Fairness and Equality of Opportunities and Outcomes

Systems historically run and managed by people demonstrate biases that are well documented. As recently as 2017, data from the Home Mortgage Disclosure Act showed that:

- 10.1 percent of Asian applicants were denied a conventional loan. By comparison, just 7.9 percent of white applicants were denied.
- 19.3 percent of Black borrowers and 13.5 percent of Hispanic borrowers were turned down for a conventional loan.[14]

Loan denial rates for some ethnic groups are far higher than the average denial rate of 9.6 percent. These results are from processes that did not rely on the use of AI.

For financial services institutions transitioning to digital systems, bias is a concern in almost every application, including algorithms to review loan applications, trade securities, predict financial markets, identify prospective employees, and assess potential customers. Addressing sources of system bias – that is, inequalities in either inputs, outputs, analysis processes, or settings and error rates that result in "unfair" recommendations – are an on-going

[14] Sray Agarwal et al., Fair AI: How to Detect and Remove Bias from Financial Services AI Models, Finextra, September 11, 2019, https://www.finextra.com/blogposting/17864/fair-ai-how-to-detect-and-remove-bias-from-financial-services-ai-models.

challenge in ML-based models in general.[15] The technology to evaluate models for system bias is advancing at the same time, but not always at the same pace, and so constant review and oversight is essential for any automated decision-making system, particularly those with legal or personally impactful outcomes.[16]

Discrimination regarding the impacts on any legally protected class is prohibited under the US Equal Credit Opportunity Act of 1974.[17] But bias clearly impacts outcomes in many systems, both those based on human decisions, and those relying on algorithmic recommendations. A 2018 study conducted at UC Berkeley found that both traditional face-to-face decisions and those made by machine learning systems charged Latinx/African-American borrowers interest rates that were 6-9 basis points higher.[18] The higher rate equates to these borrowers paying $250-$500 million per year in extra mortgage interest. However, the automated system did offer recommendations for loan approval to a broader percentage of minority applicants. The study concluded that algorithms had not fixed existing discrimination, but may have shifted the mode in the sense that more applicants were able to find financing at all.

Part of the challenge of automating these systems is that the biases from the patterns of the past are all too easily embedded in the automation of the present and future. While ML and AI are technologies thought of as completely "other" from human thinking, they are so far still always based on algorithms and models created by people. Thus, these algorithms are prone to

---

[15] S. Corbett-Davies and S. Goel, The Measure and Mismeasure of Fairness,Aug 2018, https://arxiv.org/abs/1808.00023

[16] Sarah Tan et al., Distill-and-Compare: Auditing Black-Box Models Using Transparent Model Distillation, 2017, https://www.semanticscholar.org/paper/Distill-and-Compare%3A-Auditing-Black-Box-Models-Tan-Caruana/752fd6f73c0840e5919180441c3c575da4a41124.

[17] 15 U.S. Code § 1691, available at: https://www.law.cornell.edu/uscode/text/15/1691.

[18] Robert Bartlett et al., Consumer-Lending Discrimination in the FinTech Era, November 2019, NBER Working Paper No. 25943, http://faculty.haas.berkeley.edu/morse/research/papers/discrim.pdf.

incorporating the biases of their designers, as well as the biases of the systems they're designed to serve, because the only data available to train them already reflects decades or even centuries of inequality. Because AI algorithms learn from data, any historical partiality in an organization's data can quickly create biased AI that bases decisions on inherently unfair datasets.[19]

These human biases exist in all industries and fields. Research has shown that judges' decisions are influenced by their own personal characteristics, while employers grant interviews at different rates to candidates with identical resumes but with names perceived to reflect different races.[20] Humans also routinely misinterpret information that they may identify as representing patterns of correlation.[21] Employment applications are sometimes reviewed to consider credit histories in ways that unfairly disadvantage minority groups, even though a link between credit history and job performance has not been established. Human-run processes are also difficult to review for consistency or reliability. People who self-report are frequently imprecise, whether deliberately or not, about the factors they considered, or may not even be aware of the various influences on their thinking.

Thus, training data is a part of the problem. Huge amounts of training data are required to train ML-based systems to any usable degree, but if this data comes from existing biased processes, the datasets created will reflect those inequities, and it will train the model such that

---

[19] AI Now, Race, Gender and Power in AI, April 2019, https://medium.com/@AINowInstitute/gender-race-and-power-in-ai-a-playlist-2d3a44e43d3b

[20] Marianne Bertrand & Sendhil Mullainathan, Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination, American Economic Review, Vol. 94, No 4, September 2004, https://www.aeaweb.org/articles?id=10.1257/0002828042002561

[21] Roel Dobbe et al., A broader view on bias in automated decision-making: Reflecting on epistemology and dynamics, July 6, 2018, arXiv preprint, https://arxiv.org/pdf/1807.00553.pdf

its recommendations will reflect those historical biases.[22] Consider if AI might be used by a company to set starting salaries for new hires. One of the inputs would certainly be salary history, but given the well-documented history of sexism in corporate compensation levels, that data could import gender bias into the calculations.[23]

A key problem for resolving the challenge presented by biased algorithms is identifying where the sources of bias arise. For simple algorithms based upon linear models, outcomes suggesting a disparate impact could be traced back to the sources of bias in the data, or the model components or the computations that led to that outcome. However, the greater complexity of ML models, which reflect thousands of variables and complex programming techniques like neural networks, make it unlikely that even the original programmers can say assuredly what the factors or interactions at fault might be. This is the problem of "explainability" and relates to the transparency of ML systems for review and evaluation.[24]

A further complication is that AI algorithms are by definition evolving. Unlike a static computer program, they "learn" and change over time. Initially, an algorithm creates recommendations using the process as refined on the training and testing datasets available at launch. Then based on the application of the model to real data, the system will continue to adapt its functioning, reflecting the continued processing of the increasing amounts of data. As the system gains experience in the form of more and more data, it further refines its connections and pattern analysis. These changes do not require human intervention to edit the code, but are

[22] Sally Ward-Foxton, Reducing Bias in AI Models for Credit and Loan Decisions, EE Times, April 30, 2019, https://www.eetimes.com/reducing-bias-in-ai-models-for-credit-and-loan-decisions/#

[23] John Billasenor, Artificial Intelligence and Bias: Four Key Challenges, Brookings Institute, January 3, 2019, https://www.brookings.edu/blog/techtank/2019/01/03/artificial-intelligence-and-bias-four-key-challenges/

[24] Andrew Burt et al., Beyond Explainability: A Practical Guide to Managing Risk in Machine Learning Models, July 2018, https://fpf.org/2018/06/26/beyond-explainability-a-practical-guide-to-managing-risk-in-machine-learning-models/

modifications made by the model to its own programming. In some cases, this evolution can introduce or strongly reinforce an undesired bias.[25]

Even in systems that do not collect or use data that includes sensitive or protected class fields such as race or gender, bias relative to these traits can occur. This is due to data "proxies" – fields that strongly correlate with other factors such that the patterns identified using them will result in outcomes that impact along those protected categories. The most commonly used example is the fact that zip codes frequently turn out to be a proxy for socio-economic status, race, and sometimes even general employment categories. Thus, if the system at issue is searching for patterns to define fraud scoring risk levels, it can end up scoring some racial groups at higher levels, despite never having had access to data about their race.

Organizations using these systems must continuously test the adoption of proxies within the model, that is, outputs that align along discriminatory lines, regardless of original design or intent. Not only test for but also be willing to discard models that exhibit proxies with disparate outcomes.

However, the way AI works for analysis and pattern recognition means algorithms can also be part of the solution. In some cases, AI can be applied to identify, and then reduce, humans' misinterpretation of patterns. Some experiments show that algorithms can impact decision making in a way that causes it to become fairer when measured across identified classes.[26] One study resulted in automated financial underwriting systems benefitting historically

---

[25] J. Dunkelau, et al, Fairness Aware Machine Learning, October 2019, https://www.phil-fak.uni-duesseldorf.de/fileadmin/Redaktion/Institute/Sozialwissenschaften/Kommunikations-_und_Medienwissenschaft/KMW_I/Working_Paper/Dunkelau___Leuschel__2019__Fairness-Aware_Machine_Learning.pdf

[26] Jon Kleinberg et al., Human Decisions and Machine Predictions Quarterly Journal of Economics, Volume 133, Issue 1, February 2018, https://academic.oup.com/qje/article-abstract/133/1/237/4095198

underserved applicants.[27] Recommendations made by AI could be analyzed and audited by other AI systems for more accurate understanding of their consistency, reliability, and potential bias.

IV. How Can Bias Be Managed in Financial Services AI Using Technological and Governance Solutions?

Artificial Intelligence can be a pain point for the Financial Services sector, but it can be managed through both conventional governance tools and by using other technological tools to expose and mitigate algorithmically driven biases. For example, governance tools can include careful use of contractors' expertise and managerial attention to employee's attitudes towards uses of AI.

A survey of professionals in the financial services industry sought to identify the primary areas they felt could be, or had been, improved with AI systems. Higher accuracy, greater consistency, and reduced processing times were some of the most significant benefits of AI technology across the backoffice applications.[28] In the same study, most individuals preferred a contracted model where the financial services provider would partner with an outside provider to manage their AI technology systems. This reflects the recognition that AI technology requires particular expertise to implement and manage. In-house capabilities are unlikely to be sufficiently sophisticated for the maturity of increasingly complex AI platforms and models.

This perspective is likely to be correct, as any AI algorithm can have bias: in the data, in the model design, or creeping into it "in the wild" (i.e. in applications with real life data and situations). Trained AI programmers and designers are likely necessary to proactively look for

---

[27] AnnaMaria Andriotis, Freddie Mac Tests Underwriting Software That Could Boost Mortgage Approvals, Wall Street Journal, September 24, 2019, https://www.wsj.com/articles/freddie-mac-tests-underwriting-software-that-could-boost-mortgage-approvals-11569333848

[28] Marcel Deer, AI Adoption in the Financial Sector, Marcel Deer, Medium, February 7, 2019 https://medium.com/towards-artificial-intelligence/ai-adoption-in-the-financial-sector-77c6bb81cfd3

and identify bias, correct for it, then ensure future processing outputs are fairer. Unfortunately, there is no one best or proven way to do this evaluation for every case. Research is progressing by academics and industry research and development to find ways to accomplish this analysis. These questions are an example of a broad framework for ways to check for systematic bias:[29]

1. Are any identifiable groups suffering from systematic data error?
2. Has any group been ignored, or underrepresented?
3. Are groups represented proportionally, particularly along protected class categories?
4. Are there enough features to sufficiently include minority groups?
5. Is the model using or creating factors that are proxies?
6. Are there stereotyping features?
7. Is the model appropriate for the underlying use case?
8. Is the output accuracy similar for all groups? (Are predictions skewed any identifiable subsets or groups?
9. Is the model optimizing all required metrics?

There are other ways to design or describe useful frameworks, with similar considerations for the analysis and reviews. This set of recommendations is another way to consider what the model impact is.[30] By creating alternative groups to simulate protected classes, and reviewing factors to ensure these groups have equal predictive values and equality across false positive and false negative rates, it is possible to detect and potentially measure bias in your AI.

1. Ensure all data groups have an equal probability of being assigned to the favorable outcome for a protected/sensitive class.
2. Ensure all groups of a protected/sensitive class have equal positive predictive value.
3. Ensure all groups of a protected/sensitive class have predictive equality for false positive and false negative rates.
4. Maintain an equalized odds ratio, opportunity ratio and treatment equality.
5. Minimize the average odds difference and error rate difference.

---

[29] Aarwal, Fair AI: How to Detect and Remove Bias from Financial Services AI Models https://www.finextra.com/blogposting/17864/fair-ai-how-to-detect-and-remove-bias-from-financial-services-ai-models
[30] Id.

Explainability of these systems in a way that is sufficient to satisfy everyone from regulators, to consumers, and industry experts will remain challenging.[31]

There are other tools and solutions that can be applied to the datasets as well. Statistical analysis tools like aggregation, masking records or fields, injecting "noise" into datasets, blurring and perturbations are all ways to manipulate the data to both provide protection for individual data, and to improve the evaluative accuracy of the dataset as a whole.

Differential privacy and synthetic data are also options. Synthetic data, while still in its early stages, shows promise for many of the challenges for correcting historically biased data. Synthetic data is a generated dataset of fake individual records that sufficiently represent the scale and scope of actual data to be useful for many of the analysis functions that do not reflect upon specific individuals or impact individual accounts. The synthetic datasets can be can optimized for accuracy, to mirror as closely as possible the details of actual data, but they can also be optimized for less bias. There are always tradeoffs for these types of optimization, in this case, a likely loss of some accuracy or functionality. However, the balance of accuracy and fairness can be managed to ensure that the resulting dataset is sufficient for internal sharing, access management, research, and modeling – this keeps risk lower with minimal numbers of individuals having access to "real" customer data. This type of artificial data might be sufficient for designing user interfaces or testing for accessibility from third party platforms.

V. What Actions Could be Taken in the Legal and Regulatory Environment?

There are times when discussing legal and regulatory standards for AI and ML-based systems when the concerns and arguments expressed imply we are starting from some sort of

[31] P. Hall, et al, Proposed Guidelines for the Responsible Use of Explainable Machine Learning, November 2019, https://arxiv.org/pdf/1906.03533.pdf

blank slate, and that when the challenges of bias in these systems become apparent, we must immediately take targeted action to prevent harm. But in fact, AI systems operate in the same regulated world that exists for other technology platforms.

Since the civil rights movement of the 1960s, there have been claims against the financial services industry that institutions treated some individuals less favorably than others. Once the civil rights laws established "protected classes" for particular oversight, the focus was on discrimination affecting individuals in those classes. In 1971, the term "disparate impact" was first used in the Supreme Court case *Griggs v. Duke Power Company*. The Court ruled that it was illegal for a company to rely on factors which were shown to unfairly favor white applicants to make hiring or promotion decisions, whether or not the discrimination was intentional.[32] This lack of intent is still applicable – and any AI systems that yield recommendations that demonstrate a disparate impact on protected classes would still be illegal.

In addition to intent, more recent cases have made disparate impact claims that focus on the effect, instead of the intention, of lending policies. The Supreme Court ruling in *Texas Department of Housing and Community Affairs v. Inclusive Communities Project* affirmed the use of the disparate impact theory based on outcomes. In this case a statistical analysis of housing patterns showed that a tax credit program resulted in effective segregation by race.[33] Affirming disparate impact should be a flag for technology and compliance managers in financial services. An algorithm that inadvertently disadvantages a protected class continues to be unacceptable under existing laws.

---

[32] Davis, Wright, Tremaine, Discrimination and Algorithms in Financial Services: Unintended Consequences of AI, March 6, 2018, https://www.dwt.com/blogs/payment-law-advisor/2018/03/discrimination-and-algorithms-in-financial-service

[33] Id.

Other current laws and regulations still apply as well, including the general laws against unfair or misleading trade practices, labor and employment laws, applicable privacy laws, as well as the entire regulatory structure around financial services in particular. Therefore, taking new action to legislate AI specifically should be approached with caution.

As discussed in earlier sections, there are developing best practices for overall AI fairness implications. These emergin AI governance practices and standards should be the baseline of any further guidance. However, AI risk-benefit comparisons are vastly different depending on context and application, and it is impossible to consider that any one rule could successfully address bias concerns across the entire range of use cases. It is possible that some level of legislative guidance would be appropriate in the new digital environment of automated decision making using these complex systems, but if so, the most effective would likely be based on protecting the underlying values and principles[34] at issue rather than seeking to set detailed technical standards or create performance rules that could easily be avoided or outdated in a short time.

VI. Conclusion:

Financial services organizations have the responsibility, both legally and ethically, to treat their customers, whether other businesses or individuals, fairly and equally. As more players in this industry employ AI systems in more use cases, it is incumbent on them to ensure that their algorithms are fair and explainable.

Similar challenges regarding new technology applications have been faced before. From wiretapping phones, to accessing the contents of emails, consumer protection laws have had to address the issues around particular technology platforms and determine how best to provide

[34] D. Mulligan, et al, This Thing Called Fairness, September 2019, https://arxiv.org/pdf/1909.11869.pdf

appropriate levels of privacy and security for individuals, protect their interests as consumers, and also facilitate business models that provide useful features and services.[35] These historical examples reflect the ongoing need to determine the appropriate balance of technological, legal, and policy standards and protections, along with the underlying threshold question of whether some applications, or some use cases, are simply too high risk to implement regardless of perceived benefits.

AI systems offer many potential benefits, including the opportunity to improve on biased human systems, and to increase fairness and equality at scale, but to do so there must be appropriate accountability across developers and users for their impacts, and clear evaluations of how these models are applied or used in ways that affect individuals. How we face these challenges will determine how we move further into the conveniences of a digital world, while continuing to embrace our fundamental ideals of personal liberty and freedom.

[35] J. Black, A. Murray, Regulating AI and Machine Learning: Setting the Regulatory Agenda, 2019, http://ejlt.org/article/view/722/980