



Testimony of
Miriam Vogel
President & CEO, EqualAI

Hearing entitled: “Beyond I, Robot: Ethics, Artificial Intelligence, and the Digital Age”

Before the Task Force on Artificial Intelligence
United States House Committee on Financial Services
October 13, 2021

Chairman Foster, Ranking Member Gonzalez, and distinguished members of the Task Force on Artificial Intelligence (AI), thank you for conducting this critical hearing and for the opportunity to submit this testimony. The work you have done in [past hearings](#) has been important to clarify and understand the issues and challenges surrounding AI development and use, and I commend you for delving into this next question of the ethical implications of AI.

My name is Miriam Vogel, and I am the President and CEO of [EqualAI](#), a nonprofit organization that was founded with the express purpose of reducing unconscious bias in AI systems. At EqualAI, we are AI net-positive. We believe that AI is and will be a powerful tool to advance our lives, our economy, and our opportunities to thrive. But only if we are careful to ensure that the AI we use does not perpetuate and mass produce historical, and new unanticipated forms, of biases and discrimination.

I was asked to lead this organization with a background that is currently unconventional in this space but hopefully will become more common as we invite collaborative, multi-stakeholder efforts. I worked at the intersection of technology, policy and the law as a lawyer in private practice and in government. I previously had the opportunity to address the problem of bias in more traditional contexts, such as leading President Obama’s Equal Pay Task Force and the effort to create implicit bias training for federal law enforcement at the Department of Justice under the direction of Deputy Attorney General Sally Yates. Given that orientation, our focus at EqualAI is on driving multi-staker efforts to ensure our technology platforms are equitable and inclusive. We perceive implicit bias in AI as an age-old issue that is surfacing in a new medium, but now at scale and with graver potential impacts.

We believe we are at a critical juncture because AI is becoming an increasingly important part of our daily lives - while decades of progress made and lives lost to promote equal opportunity can essentially be unwritten in a few lines of code. And the perpetrators of this disparity may not even realize the harm they are causing. Our country’s long history of housing discrimination is

being [replicated](#) at scale in mortgage approval algorithms that determine credit worthiness using proxies for race and class. An exciting innovation in AI deep learning language modeling, GPT-3, is also demonstrating its problematic biases, such as generating stories depicting sexual encounters involving [children](#) and exhibiting biases against people based on their [religion, race and gender](#). Our goal at EqualAI is to help avoid perpetuating these harms by offering programs, frameworks and strategies to establish responsible AI governance.

We focus our efforts at EqualAI on supporting three main stakeholders: companies, policymakers and lawyers. We work with companies to help them address and reduce the infusion of implicit bias in their AI systems. We aim to support policy makers in the essential task of establishing the appropriate guardrails that support innovation while mitigating harmful bias in AI. For instance, we look forward to supporting the important work [underway at the National Institute of Standards and Technology \(NIST\)](#) next week by moderating an AI Risk Management Framework (RMF) workshop panel. Finally, we bring lawyers into this effort. Lawyers need to help companies understand and manage the risks in the AI systems they are building, acquiring, using and deploying by ensuring they employ frameworks to reduce both harms and liabilities.

Often, the first step in our work is helping companies come to terms with the reality that they are now effectively AI companies. Two decades ago, most companies did not realize that they needed to have cohesive plans and contingencies in place to protect against the unauthorized exploitation of systems, networks and technologies. Today, however, companies widely recognize the importance of cybersecurity. Much like the trajectory of cybersecurity awareness, companies now need to adjust and understand that they use AI in one or more pivotal functions- hiring, credit lending, health care determinations, to name a few- and must, therefore, have a governance plan in place to address the potential discrimination that these systems can dispense at scale. This is of particular concern with AI given that key determinations and assessments occur behind the thick veil of the proverbial ‘black box’, where the algorithm’s inputs and operations are generally unknown to the end user. We support businesses’ efforts to establish AI governance and best practices in a variety of ways. For instance, we just launched a [badge program](#), in collaboration with the World Economic Forum, to train senior executives on how to understand and implement responsible AI governance and create a community comprised of companies and individuals committed to this effort.

Implicit Bias in AI

Your [past hearings](#) provided an important overview of both the benefits and risks of our accelerated use of AI in the financial and housing sectors. One grave concern that has been repeatedly articulated is the infusion and dissemination of implicit bias at scale through AI systems. [Implicit or unconscious bias](#) is based on a stereotype, or characterization of people of a

certain group which can be positive or negative. It is noticing patterns and making generalizations based on those assumptions. As referenced in your past [hearings](#), implicit bias embeds in AI in a variety of ways. Our operating thesis is that bias can embed in each of the human touch points throughout the lifecycle of the creation of an AI system (see *Diagram 1* below). From the ideation phase- deciding which problem you want to use AI to help solve, to the design, data collection, development and testing phases. Each stage is limited by the experience and imagination of those on that team, which is reinforced by historical and learned biases in the data. But we are also optimistic and think each touchpoint is an opportunity to identify and eliminate harmful biases. As such, risk management should occur at each stage of the AI lifecycle.

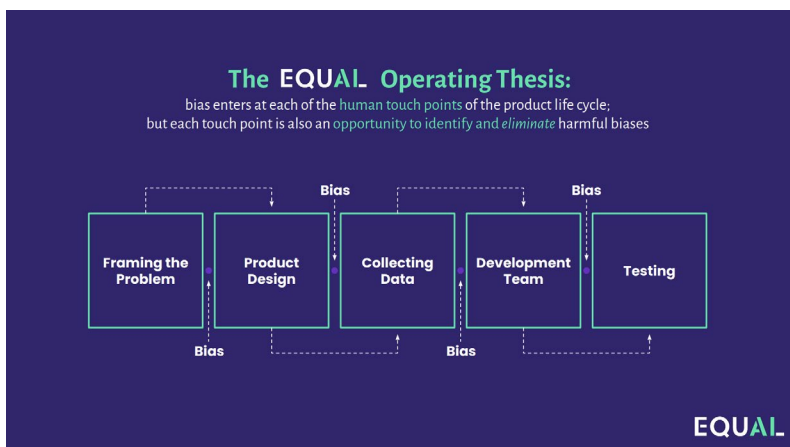


Diagram 1

Framework to Combat Implicit, Harmful Bias in AI

There are an increasing number of frameworks that provide helpful guidance on methods to identify and reduce harms from AI systems before they materialize (e.g. [GAO Framework](#); [GSA Center of Excellence Guide to AI Ethics](#), [DoD Ethical Principles for AI](#), [AI Ethics Framework for the Intelligence Community](#)). There are also efforts underway to clarify and standardize frameworks and best practices, such as the important work at NIST to support AI standards development, develop a risk management framework for trustworthy AI systems, and develop best practices for documenting and sharing data sets used to train AI systems, pursuant to the [National Defense Authorization Act for Fiscal Year 2021](#) (NDAA).

We offer an EqualAI Framework (see *Diagram 2* below) as a general guide with five “pillars” a company should consider as part of its effort to establish enterprise-wide responsible AI governance. These recommendations are particularly important in sensitive sectors such as finance and housing, given the real potential for biased AI to perpetuate discrimination against job candidates, renters, mortgage seekers, insurance applicants, and disadvantaged small businesses seeking capital. These “pillars” are as follows:

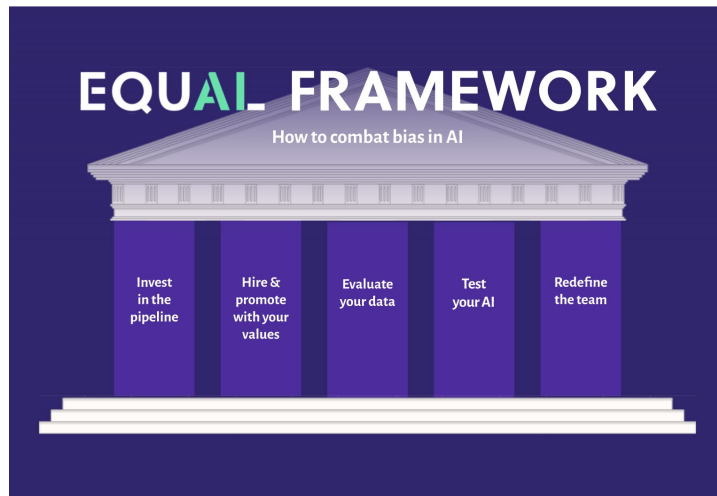


Diagram 2

1. Invest in the pipeline

Artificial intelligence needs to be created by and for a broader cross-section of our population. Research [suggests](#) that homogeneous teams – like those that comprise [many](#) of the teams coding our AI – are [more likely](#) to generate biased algorithms than [diverse teams](#). We have [also seen](#) that lack of diversity in AI creation could rise to a life or death issue given the ultimate uses for many of these technologies, such as determining who can access [ventilators](#) and [critical health care services](#) during a crisis or [deciding](#) an individual’s fate in the criminal justice system.

Many organizations are ensuring that our next generation of coders and tech executives represent a broader cross-section of our population (e.g., [AI4All](#), [Black Girls Code](#), [Girls Who Code](#), [Code.org](#), etc.). Their work is critical to ensuring that our AI benefits from broader perspectives and that more communities thrive in the AI economy.

We currently offer inadequate access to computer science and engineering courses in our classrooms. Reports indicate that only [22](#) states have K-12 standards for computer science education, and only 15 require high school computer science courses. Where these courses are offered, they are often rife with race, class, geographic and [gender imbalances](#), depriving our workforce, and resulting AI, of the full breadth of American talent.

Congress can play a key role by offering funding for teachers to learn and teach coding as part of the K-12 curriculum in our nation’s public schools. Congress can also direct additional funding to ethical AI research and development in our higher education system, including at HBCUs, minority-serving institutions, and community colleges.

2. *Hire and promote with your values*

To create and sustain a diverse workplace, and [produce better AI](#), employees and managers in particular, must be trained to recognize and address implicit bias in human resource functions (e.g., hiring, evaluation, promotion and termination decisions). AI programs used in these human resource functions should be checked to ensure they are in sync with a company's values. AI used for [hiring](#), evaluations, promotions and terminations could be infused with bias and as such, must be checked- and constantly rechecked- for harmful biases given the likelihood it will constantly learn new patterns and may offer inequitable employment decisions. A few best practices we recommend at this stage include: (1) ensuring a broad cross-section of diversity in each candidate panel, (2) keeping humans involved and 'in the loop' in decision-making processes and (3) constantly checking for biased patterns or outcomes with simulated or hypothetical personas.

Congress can help by clarifying that hiring and employment laws prohibiting discrimination are equally applicable to AI-driven or supported recommendations.

3. *Evaluate your data*

The more we know about the datasets on which AI is built and trained, the safer we are as a society. We encourage companies to identify the gaps in the data so that they can be rectified or at a minimum, clarified for end users. We offer the [EqualAI Checklist](#)© as a starting place to evaluate data sets and identify possible liabilities. There are other helpful resources that provide best practices to ensure data sources, as well as gaps, over- and under-representations, are identified, such as [Federal Trade Commission \(FTC\) guidance](#). Best practices include seeking to answer: (1) How representative is the data set? (2) Does the data model account for biases? (3) How accurate are the predictions that the AI offers? (4) [For whom](#) could this system fail?

Congress can play a key role by clarifying expectations and specifying the information that should be provided when brokering data sets or AI training data, similar to current expectations for nutrition labels. Industry, and society as a whole, would benefit from notice and clarity on what data points should be highlighted and particularly by those selling or sharing the data sets to help with predictive modeling and other AI uses. This can help ensure uniform standards and identify gaps that need to be addressed when using the data that would otherwise be unknown or unspecified, leading to potentially [harmful](#) and [inequitable](#) outcomes.

In the financial services context, intensive data evaluation can also help ensure the use of AI does not result in violations of existing statutes. Laws such as the Fair Housing Act and the Equal Credit Opportunity Act are critical, albeit imperfect, safeguards against discrimination for those seeking to build equity, access capital, and pursue better opportunities for themselves and

their families. Congressional vigilance is critical to ensuring that these safeguards are not eroded given that creditworthiness, underwriting, and other decision-making processes are increasingly automated. These hearings by the Task Force on AI are important steps in attaining this goal.

4. Test your AI

AI systems, and particularly those that are customer-facing or used in human resource functions, should be checked for bias on a routine basis. Given that AI constantly iterates and learns new patterns as it is fed new data, it will often adopt new biases. Good AI governance includes routine audits and checks to ensure recommendations are consistent with expectations and that outlier outcomes are investigated.

Responsible AI governance includes:

- ❖ identifying which values and biases will be tested routinely;
- ❖ articulating the stages of the AI lifecycle development at which the testing will be conducted (e.g., pre-design, design and development, deployment);
- ❖ establishing the cadence for testing;
- ❖ documenting relevant findings and the completion of each stage to promote consistency, accountability and transparency; and
- ❖ identifying the designated point of contact who owns this responsibility ultimately, including: coordinating incoming questions and concerns, ensuring that responses are consistent and that new challenges are elevated and addressed appropriately.

The [FTC guidance](#) and [EqualAI Checklist](#)© are two sources for additional guidance and there is a growing body of experts and algorithmic auditors to help test AI systems. Congress can help normalize the practice of algorithmic audits and by collecting information about a company's high level AI governance plan and the appropriate point person. The submission of this information could be made a common practice as part of routine filings, such as with the Departments of Housing and Labor, Office of the Comptroller of the Currency (OCC), Consumer Financial Protection Bureau (CFPB), FTC or Securities and Exchange Commission (SEC).

5. Redefine the team

AI products should be tested prior to their public release by and for those under-represented in its creation or in the data used to build and train the system. Special consideration should be given, and broader audiences should be brought in to help determine potential end users and those who could be impacted downstream who were insufficiently represented on the creation team and in the datasets used to build and train the program.

Congress can help support this important aspect of responsible AI governance by ensuring that it is following these and other best practices with its own internal AI development and procurement processes and by sharing best practices, resources, and lessons learned with industry and the general public. This Committee can help ensure that our nation’s financial regulators, and the institutions they oversee, are leaders in this regard.

Additional Proposed Solutions

In addition to frameworks, there are numerous additional ways that Congress can play an instrumental role in ensuring more effective, less harmful AI development and deployment.

The National Security Commission on Artificial Intelligence (NSCAI), like prior [cybersecurity reports](#), has [warned](#) that “America is not prepared to defend or compete in the AI era.” The report recommends the establishment of the foundations for widespread integration of AI by 2025, including digital infrastructure and developing a digitally-literate workforce. These recommendations are critical to ensuring we have the critical mass of Americans necessary to perform AI-created and supported jobs. They also will enable us to support vulnerable and underrepresented populations that otherwise could fall subject to an even wider and more dangerous income disparity gap.

1. Auditing

We support mandates for [auditable AI](#) for systems used in pivotal functions, where AI systems are queried externally with hypothetical cases that are either synthetic or real. The more transparency the better, and in particular, notice of populations who are under or over represented in underlying datasets and for whom the AI system will have different success rates should be encouraged in the form of nutritional labels, as mentioned above. However, when this is not possible, and even when it is, there should be an expectation of routine, external audits with publicly available and easily accessible results.

The necessity of algorithmic auditing is particularly evident in the financial services context given that these AI-supported recommendations directly impact people’s lives and opportunities and yet, are rooted in a part of our history, and thus trained on data, that is rife with imbalance and discrimination.

2. Incentivize investment in the Future of Work

There is a palpable concern that AI will edge humans out of the workplace, as addressed at a [congressional hearing](#) on the Future of Work in 2019. To be sure, automation, in tandem with the COVID-19 recession, is creating a [‘double-disruption’](#) scenario for workers. Like all transformative technologies, AI has and will inevitably eliminate jobs, but it will also open up

possibilities, many of which we do not yet realize. Some [estimate](#) that by 2025, 85 million jobs may be lost but 97 million new roles may emerge due to automation in the workforce.

The U.S. is [estimated](#) to spend approximately [0.1% of its GDP](#) on retraining programs, which is one-fifth of the average expenditure for countries in the Organization for Economic Cooperation and Development (OECD).

To lead in the AI revolution, safeguard our economy and support greater prosperity in more communities, we need to reskill our workforce. Some [estimate](#) that businesses could collectively reskill 45% of workers at risk of losing their jobs but, if governments join this effort, we could reskill as many as 77% of at risk workers. This would benefit government and society directly with increased tax returns and lower social costs, including reduced homelessness and food insecurity.

In addition, Congress could commission a study to better understand and articulate the type of skills and jobs that will likely emerge, enabling us to educate and upskill accordingly. One often cited [study](#) found that there are as many predictions of what the new jobs will look like as there are experts. The best way for us to plan for the future workforce is to offer clarity and evidence-based assessments of what it will look like. This study could be part of current, related efforts, such as the [National Academies AI Impact Study on Workforce](#), (per Section 5105 of the NDAA), and the [National AI Advisory Committee](#) (per Section 5104(d) (4) of the NDAA).

An additional significant contribution could be the inclusion of additional tax breaks for investments in upskilling employees, loan forgiveness for graduates with computer science degrees who spend a minimum number of hours teaching in K-12 classes and increasing opportunities for secondment both for those with technical skills to support schools and government regulators and for government employees to spend time in the private sector.

3. Bill of rights

We need to ensure that the general public is empowered to require that AI systems, and other technologies we use, respect our democratic values and right to be free from discrimination. One such solution is the new “AI bill of rights” [proposed](#) last week by [Dr. Eric Lander](#) and [Dr. Alondra Nelson](#) of the White House Office of Science and Technology. It would ensure that the public is put on notice of critical information, such as when and how AI is influencing a decision that affects our civil rights and civil liberties and when we are using AI that has not been audited for implicit biases or trained on sufficiently representative data sets. Likewise, it envisions an opportunity for meaningful recourse for individuals harmed by such algorithms.

In conclusion, as we noted at the outset, we believe that it is imperative at this critical juncture to ensure that AI is built by and for a broader cross-section of our population. It is not only the right thing to do, a strong U.S. economy and our leadership depend on it.

Thank you again for the opportunity to testify before the Task Force. I look forward to answering your questions.