

# #EqualAI

*Bulletin on Bias & Artificial Intelligence:  
Issue 5, September 2020*

## Spotlight of the Week: Kathy Baxter



*“Ethics must be at the foundation of AI. For all the good that AI can bring, responsible tech companies must recognize, prepare for, and mitigate the potential unintended or harmful effects of AI, while being inclusive of everyone it impacts.”*

*- Kathy Baxter, Architect, Ethical AI at Salesforce*

## Hot Topics in Business and Technology

- **Microsoft Granted License to OpenAI’s Language Model:** Microsoft announced plans to license OpenAI’s highly sophisticated GPT-3 technology, an autoregressive language model that outputs the most human-like text available through AI to date (9/22).
- **Amazon Expands AI Device Offerings:** Amazon launched over a dozen new AI-powered accessories, including a car security device to record police stops, new Alexa software that can interact with multiple people at once, and the ‘Ring Always Home’ camera drone for home security (9/24):

- *"We do not want customers to have to think they are choosing between innovation and privacy...They get to have both." - Tom Taylor, Alexa head, Amazon*
- **Racial Bias in Twitter Image Previews:** Users discovered that tweets containing photos of both Black and white faces showed only the photo of the white face on the image preview, which Twitter is now investigating (9/21).
- **Facebook Addresses Disinformation:** Facebook announced the detection and shut down of three disinformation networks operating on their platform. Two of the networks were linked to Russian intelligence services; the third with a 'troll farm' that interfered in the 2016 U.S. presidential election (9/24).
- **TikTok Update:** U.S. District Judge Carl Nichols announced that the Trump administration must either file a response to a request by TikTok for a preliminary injunction or delay the Commerce Department's order by 2:30 p.m. EDT on Friday, September 25. If the order is not delayed, TikTok will be banned from Google and Apple app stores on Sunday, September 27 (9/24).
- **Snowflake:** Hitting a market value of \$70 billion, the AI-powered cloud data platform company was valued as the biggest tech IPO of the year, a broader trend of an increasingly AI-driven financial market (9/16).
- **AI-Driven Anti-Corruption Practices:** DOJ is indicating an interest in decreasing penalties for companies that incorporate data analytics or AI-driven monitoring tools into their compliance practices (9/22).
- **Self-Driving Cars:** Zoox, a self-driving car startup acquired by Amazon, is now permitted by the state of California to test its vehicles on certain streets without previously required 'back-up' human passengers (9/18).
- **AI to Combat Anti-Semitism:** An international team of scientists have commenced project 'Decoding Anti-Semitism' by teaming up discourse analysts, computational linguists and historians who will develop a "highly complex, AI-driven approach to identifying online anti-Semitism," to develop AI tools to combat escalating radicalization online (9/22).
- **AI-Powered Mayflower Ship:** 400 years after the wooden ship set sail across the Atlantic Ocean, the U.S. ambassador to Britain launched a self-driving Mayflower ship developed by IBM and Promare. The ship will reach areas of the ocean deemed unsafe for human researchers (9/16):
  - *"...assessing the impact of climate change, measuring micro-plastic pollution and studying populations of whales and dolphins. Along the way, its AI captain will have to make complex decisions in response to wind, waves, vessels and unknown surprises."*
- **AI Insurtech:** Agile startups developing AI- driven risk models are booming, as they meet growing demands from the current climate of economic and environmental uncertainty (9/21).

- **Tech Giants Scrutinized for Stifling Competition:**
  - **Google:** In anticipation of a lawsuit against Google, the DOJ is expected to announce an antitrust investigation into how the company uses AI to suppress their competitors (9/21).
  - **Amazon:** From limiting some vendors from buying advertisements to blocking others from selling their products altogether, Amazon uses its platform to drive demand for their own products. Amazon officials denied wrongdoing, arguing that their practices are legal and common for e-commerce sites (9/22).
- **Boycotting Advertisers Struck Deal with Facebook, YouTube, Twitter:** After months of advertiser boycotts due to perceived failure of platforms to act upon harmful content, the World Federation of Advertisers announced the platforms will adopt shared universal standards for reporting harmful content. Some practices will now be reviewed by external auditors, and advertisers will have more control of content displayed alongside ads (9/23).
- **AI & Gender Bias on the Supreme Court:** University of Washington researchers and lawyers are training AI models to analyze gender bias in the Court's oral arguments (9/22):
  - *"We find that as the first person in an exchange, female speakers and attorneys are spoken to more competitively than are male speakers and justices. We also find that female speakers and attorneys speak more cooperatively as the second person in an exchange than do male speakers and justices." - Haley Lepp and Gina-Anne Levow*
- **CSET Released AI Policy for 2020 President:** Georgetown University's Center for Security and Emerging Technology published policy guidelines to encourage promotion of U.S. leadership in international AI development while defending against global security threats. Key recommendations include: imposing export controls on AI equipment and supporting domestic research to attract global talent (9/22).
- **Site Launched at Art Exhibit Results in Removal of Biased Images:** ImageNet, a popular online database of images, will remove 600,000 pictures from its system after an art project debuted this week, *ImageNet Roulette*, revealed racist and sexist outcomes based on crowdsourced tags (9/24):
  - *"[It] inspired a long-overdue public conversation about the politics of training data, and we hope it acts as a call to action for the AI community to contend with the potential harms of classifying people."*
- **IBM Announced Tech Initiative for HBCUs:** The IBM-Historically Black Colleges and Universities (HBCU) Quantum Center will foster collaboration between the technology company and HBCUs on academic, education, and

community outreach programs. The initiative includes an \$100M in technology assets and skill development at HBCUs (9/17).

- **Using AI to Study Loneliness:** Researchers at UCSD are harnessing AI analytics to eliminate human bias from research around how people experience loneliness and react to isolation (9/24):
  - *“Most studies use either a direct question of ‘how often do you feel lonely,’ which can lead to biased responses due to stigma associated with loneliness or the UCLA Loneliness Scale which does not explicitly use the word ‘lonely,’... For this project, we used natural language processing or NLP, an unbiased quantitative assessment of expressed emotion and sentiment, in concert with the usual loneliness measurement tools. - Ellen Lee, MD, UC San Diego School of Medicine*
- **AI Robot Beats Curling Athletes:** Curly, a robot-powered by deep reinforcement learning, beat top-ranked athletes from the South Korean curling team, demonstrating the capacity of AI to function in the physical world rather than just in computer- simulated environments (9/23).

## US Federal and State Government News

- **CIA Labs:** In an effort to drive technical innovation in national security, the CIA announced the formation of CIA Labs. This initiative intends to foster partnerships between the CIA and academia, and will permit CIA employees to patent and reap partial proceeds from their inventions (9/23).
- **AI Ethics & the Energy Department, Trump Administration:** U.S. Chief Technology Officer Michael Kratsios and Energy Secretary Dan Brouillette opined on how the Administration is taking ethics into account as it pushes for AI developments (9/22):
  - *“The same cost-benefit analysis that is required by statute for any other regulation should also be done in the context of AI.” - U.S. Chief Technology Officer Michael Kratsios*
- **Facebook Biometric Data Lawsuit:** The social media giant is facing new accusations of obtaining and profiting off biometric information of more than 100 million Instagram users without their consent. In July, the company paid \$650 million to settle a lawsuit over the illegal collection of users' biometric data through a photo-tagging tool (9/17).
- **Trump Administration Focuses on AI Bias Against Conservatives:** The administration called for investigations into how social media companies censor conservative content. The Administration also launched a legislative proposal to encourage accountability around content censorship (9/23).



## EqualAI Partners in the News

- **LivePerson, Inc.**, a global leader in Conversational AI announced this week the launch of 'Performance Optimizer,' a new, AI-powered tool for brands to measure and manage the health of their conversational operations (9/23):
  - *"Just as we at LivePerson have analyzed our decades of conversational data to build and optimize our AI, brands that understand and use their conversational data to the fullest will be best positioned to deliver outstanding customer experiences,"* Robert LoCascio, founder and CEO of LivePerson, Founder and Board Member of EqualAI.
- **Pymetrics'** co-founder and CEO, Frida Polli, talks with Harvard Business School Professors Bill Kerr and Joe Fuller and presents ways to overcome inherently biased human judgement and to bring empirical rigor to the task of matching talent to fast-changing job categories (9/23):
  - *"That is the problem with System 1 thinking. That is the ultimate design fault, the human brain. And again, it's not because we're bad people, it's because the System 1 thinking has developed to be fast... I think that's the promise of audited technology is that we can get to that place of parity."*



## AI Around the World

- **Russia:** Russian facial recognition company NtechLab is developing global facial recognition technology to detect "aggression." This announcement gives rise to concerns about the tool's accuracy and bias, as well as fear of increased surveillance similar to the backlash against NtechLab's FindFace app (9/22).
- **Singapore:** Singapore's government is encouraging academia to prepare students for an increasingly digital labor economy through initiatives including career-transition programs to teach technical skills and expanded remote learning programs (9/22).
- **Brunei:** Minister at the Prime Minister's Office and Minister of Finance and Economy called for embracing an AI-ready future workforce (9/20).
- **India:** Tamil Nadu became the first state in India to come out with AI, blockchain, and cybersecurity policies (9/19).

## Op-Eds: Top Picks of the Week 🤖

- **AI Gets Closer to Decoding our Emotions:** A team at MIT Technology Review [explores](#) what emotion AI is, where it is, and what it means on this podcast episode. If we want to build trustworthy machines, we need machines that can understand our emotions. Understanding personality and gait and behavior “has become increasingly the frontier of AI” (9/24).
- **Minecraft AI Holds the Future to City Design:** MIT Technology Review examines the potential for [Minecraft](#) machines to design cities autonomously. The annual Generative Design in Minecraft competition challenges participants to build an AI for urban planning (9/22).
- **Synthetic Data Combats Bias in AI:** Increasingly, industry is looking to solve the issue of insufficient data sources to build AI with the use of synthetic data. It can be cheaper and faster but carries the risk of exacerbating harmful biases. Academic researchers are [developing](#) synthetic data with an eye toward ‘fairness’ instead of solely on ‘accuracy’ to train algorithms, hoping these ‘fake’ datasets can be used to combat harmful data-driven biases (9/17).

## Upcoming Events 🎉

- Politico’s AI Summit, ‘Can Europe keep up with its digital legacy and be a global AI-rule maker?’: [September 30- October 1](#).
- AI in Healthcare Summit, ‘Business and Academia for Innovations in Healthcare,’ co-hosted by MIT CSAIL and STEMM Global: [October 1-2](#).
- Security and Emerging Technology Seminar Series, ‘Building Trust Through Testing,’ Hosted by CSET: [October 6](#).
- HAI Summit, ‘Triangulating Intelligence: Melding Neuroscience, Psychology, and AI,’ hosted by Stanford University: [October 7](#).
- 2020 Ethics of Big Data Symposium on ‘Race, Representation & Justice,’ featuring keynote speaker Ruha Benjamin: [October 7](#).
- ‘How Will Quantum Computing Shape the Future of AI?’ hosted by The Center for Data Innovation: [October 14](#).