

#EqualAI

*Bulletin on Bias & Artificial Intelligence:
Issue 3, September 2020*

Weekly Spotlight: Victoria Espinel



"AI can help people do great things, but it must be developed and deployed responsibly. I support the mission of EqualAI and the work being done to ensure that AI systems are inclusive by design and used in ways that enhance public trust. While the risks of AI bias are real, so too are the opportunities for AI to broaden inclusion and enhance public wellbeing. There's no one-size-fits-all solution to AI bias, so organizations that design and deploy high-impact AI must integrate risk management safeguards that extend throughout a system's life cycle – accounting for the earliest stages of data collection all the way through to the point at which a system is being decommissioned. AI can and should be used in a way that will not exacerbate discrimination, but instead actively and affirmatively broaden inclusion."

- Victoria Espinel, President and CEO, Business Software Association

Hot Topics in Business and Technology 🔥

- **With Election Day Looming, Tech Giants Develop New Policies:**
 - Google announced changes in advance of US elections, including (9/11):
 - removal of AI Autocomplete predictions from Google Search,
 - formation of an internal Intelligence Desk to monitor news,
 - new programs designed to prevent vandalism on Wikipedia from sneaking onto the Knowledge Graph panels that appear alongside search results, and
 - fact-check labels in Google Images.

- Twitter announced it will:
 - remove or add label to tweets presenting false or misleading information about laws on civic processes, and the officials and institutions overseeing them, and
 - prepare for unclear or contested election results, including plans to either remove or attach a warning label to any claims of victory prior to election results being official (9/10).
- Facebook announced plans to pay selected users to stop using their platforms at the end of September in advance of the elections to study how social media usage affects voter outcomes (9/8).
- **TikTok:**
 - In response to the Trump Administration's moves to tighten national security threats surrounding technology, China announced plans to launch its own initiative to set global standards on data security and AI. These restrictions place further challenges on plans to sell US TikTok operations as they call for governments to respect other countries' data sovereignty (9/8).
 - Skeptisim towards TikTok's AI algorithm grew as TikTok struggled to sensor a suicide video the algorithm failed to delete (9/9).
 - TikTok confirmed that the algorithm censors LGBTQIA content to comply with local laws in Russia, Bosnia, and Jordan (9/10).
 - US TikTok executives established a new 'transparency center' to showcase the workings behind TikTok's famous algorithms, despite restrictions placed on the export of data analysis technology for personalized information recommendation services by Chinese officials earlier this month (9/11).
- **AI Tools Improve Trading Outcomes:** Tickeron, an investment and trading analytical platform, conducted an internal study indicating advantages to using AI-powered tools over human intelligence, particularly regarding AI's ability to eliminate emotional bias for higher yields (9/4).
- **AINow released Biometric Surveillance Report:** The report provides insight into global biometric surveillance, analyzing the spectrum of calls for deregulation and advocacy (9/4).
- **Instagram launched in-house DEI team:** On the motivation behind launching an in-house equity team, Instagram head Adam Mosseri wrote of their goal to "*focus on better understanding and addressing bias in our product development and people's experiences on Instagram*" (9/9).
- **AI Used to Combat Hate Speech:**
 - USC's Computational Social Sciences Lab is studying the implicit bias behind hate speech to use AI for better detection and response strategy (9/10).
 - Areto Labs launched an AI bot to recognize and respond to hate tweets targeted at female election candidates (9/8).
- **Ford's AI Robot Dog:** Ford recently sought the help of Boston Dynamics' robot dog "Fluffy" to aid their logistical operations and design processes (9/8).
- **AI Knowledge Graph:** DiffBot uses image recognition and natural-language processing to constantly evaluate and consolidate information across the internet. Its unique AI turns text into three-part factoids using subject, verb, object recognition (9/4).

- **AI & Facial Recognition:** Charlette N'Guessan, CEO and co-founder of BACE, won the Royal Academy of Engineering's Africa Prize for Innovation Award, making her the first woman and Ghanaian honoree. BACE is one of the first facial recognition platforms created with and for Black communities (9/7).
- **AI-Powered Audio:** Facebook revealed prototype headphone software that uses AI to amplify distant or muffled sounds while silencing background noise (9/4).
- **AI & Music:**
 - Google's Lo-Fi Player uses AI to open the door for more people to mix music (9/4).
 - Using materials from Spotify, DataRobot trained its AI to predict genres based on album cover art (9/10).

AI & Public Health

- **AI & Veterans' Mental Health:** A new VA program utilizes customer experience AI to monitor veterans' mental health through tone and language analysis. The technology has been effective in aiding at-risk veterans (9/8).
- **GlaxoSmithKline AI Research Lab:** The healthcare giant announced plans to open a research base in King's Cross which will use AI to find new treatments for cancer and other serious diseases (9/2).
- **AI & Medical Imaging:**
 - NYU Langone Health is partnering with Facebook AI Research to explore the benefits of increasing the speed of an MRI scanning. Their research indicates that AI could improve patient experience and overall efficiency in radiology departments (9/4).
 - Intel and Samsung described their efforts to use AI to improve ultrasound imaging, which could aid doctors in determining best methods for delivery and reduce unnecessary cesarean sections (9/10).
 - Researchers from Carnegie Mellon are using AI to analyze placenta samples for signs of health risks in future pregnancies (9/8).
 - Lunit is developing an AI solution that matches the rate of accuracy of human radiologists in identifying breast cancer, potentially reducing workloads and scaling early-stage cancer detection (9/10).
- **Digital Ethics Lab at the University of Oxford:** The Lab released a report mapping out the landscape of AI ethics in healthcare (9/10).

US Federal and State Government News

- **DoD:**
 - Air Force Gen. John E. Hyten spoke at the 2020 Department of Defense AI Symposium and Exposition about the DoD's move to take the lead in global and national AI research and development (9/10).
 - *"We want peace on this planet. We want peace to be the world that our children live in. That's the world we want. That's what we need to use all of our capabilities for, and artificial intelligence is one of those" - Gen. John E. Hyten*

- DoD hosted the AI Symposium and Exposition, spotlighting their Responsible AI Subcommittee (9/10):
 - *“Our objective is to field AI and help our nation prevail on tomorrow’s battlefield...yet the way we win matters! This is precisely why the JAIC policy team is leading a far-reaching initiative that brings together thought leaders from across the Department through the Responsible AI Subcommittee to develop and standardize procedures for implementing the AI ethical principles that were adopted by Secretary Esper earlier this year.” - Hon. Dana Deasy, DoD Chief Information Officer*
- **AI & Rental Regulation:** In response to plummeting tax revenues, city governments across the US are using AI and machine learning to improve ordinance enforcement (9/4).
- **AI Facial Recognition Banned in Portland:** Without federal guidelines, Portland city government joined the growing list of states to ban AI facial recognition technology. Portland’s laws appear to be the broadest to date, banning both local government and business from using the AI technology (9/9).

EqualAI in the News

The Social Dilemma, initially released in January 2020, is now available for streaming on Netflix. The documentary reports on the behavioral psychology behind AI in social media, and spotlights EqualAI Senior Advisor, Cathy O’Neil.

AI Around the World

- **China:**
 - Hong Kong Science and Technology Parks Corporation, a public innovation fund of the HK Government, released a testing platform to promote B2B growth by validating AI technology capabilities as well as establishing industry benchmarks on AI standards (9/10).
- **India:**
 - The American Express Data Analytics, Risk and Technology Laboratory was inaugurated at IIT Madras. The lab will focus on mitigating risk at the intersection of human behavior and technology (9/4).
 - GoLearn, India's first AI-powered English Language Teaching & Learning app, was launched to address the global shortage of English teachers, growing demand for English speakers, and increasing class sizes (9/7).
 - Biometric surveillance expands across Indian railways, as AI-based thermal screening technologies track the spread of Covid-19 (9/11).
- **Australia:**
 - NSW Government AI Review Committee established an AI strategy which aims to create tech jobs as well as foster innovative services on its way to becoming the ‘digital capital of the southern hemisphere’ (9/4).

- **Pakistan:** New AI-based career counseling app MeraFuture evaluates students' personality, interests, and subject expertise to offer careers options (9/10).
- **Senegal:** Rapper Akon released plans to build a city inspired by the Wakanda city of Black Panther film on Senegalese coastline. Ethical usage of AI is at the core of the city's proposed design (9/7).
- **UAE:** 101 graduate students from 31 countries will attend the world's first AI university in Abu Dhabi, when it opens for its first academic year on January 10, 2021 (9/10).

Op-Eds: Top Picks of the Week 🤖

- **Harvard Business Review:** David De Cremer wrote about how to achieve fairness in AI (9/3):
 - *"An organization's data scientists need to know and agree with the values and moral norms leadership has established. At most organizations, a gap exists between what data scientists are building and the values and business outcomes organizational leaders want to achieve. The two groups need to work together to understand what values cannot be sacrificed in the use of algorithms. For example, if the inclusiveness of minority groups, which are usually poorly represented in available data, is important to the company, then algorithms need to be developed that include that value as an important filter and ensure that outliers, not just commonalities, are learned from."* - David De Cremer

Upcoming Events 🎉

- AI in Healthcare Summit, 'Business and Academia for Innovations in Healthcare,' co-hosted by MIT CSAIL and STEMM Global: October 1-2.
- HAI Summit, 'Triangulating Intelligence: Melding Neuroscience, Psychology, and AI,' hosted by Stanford University: October 7.