Joy Boulamwini Testimony

Sunday, October 4, 2020 9:50 PM

QEA HW 4 Reading 1

From: https://docs.house.gov/meetings/GO/GO00/20190522/109521/HHRG-116-GO00-Wstate-BuolamwiniJ-20190522.pdf

This testimony is also available in video form: Facial Recognition Technology (Part 1): Its Impact on our Civil Rights and Liberties

United States House Committee on Oversight and Government Reform

May 22, 2019

Hearing on

Facial Recognition Technology (Part 1):
Its Impact on our Civil Rights and Liberties

Written Testimony of

Joy Buolamwini
Founder, Algorithmic Justice League

Masters in Media Arts and Sciences, 2017, Massachusetts Institute of Technology
MSc Education (Learning & Technology), 2014, Distinction, University of Oxford
BS Computer Science, 2012, Highest Honors, Georgia Institute of Technology

PhD *Pending*, MIT Media Lab

For additional information, please contact Joy Buolamwini at joy@ajlunited.org

Dear Chairman Cummings, Ranking Member Jordan, and Members of the Committee,
Thank you for the opportunity to testify on the limitations of facial recognition technology. My name is Joy Buolamwini, and I am the founder of the Algorithmic Justice League (AJL), based in Cambridge, Massachusetts. I established AJL to create a world with more ethical and inclusive technology after experiencing facial analysis software failing to detect my dark-skinned face until I put on a white mask. I've shared this experience of algorithmic bias in op-eds for Time Magazine and the New York Times as well as a TED featured talk with over 1 million views.[1] My MIT thesis and subsequent research studies uncovered large skin type and gender bias in AI services from companies like Microsoft, IBM, and Amazon.[2] This research has been covered in over 40 countries and has been featured in the mainstream media including FOX News, MSNBC, CNN, PBS, Bloomberg, Fortune, BBC, and even the Daily Show with Trevor Noah.[3]

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%

Figure 1. Intersectional Skin Type and Gender Classification Accuracy Disparities.
www.gendershades.org

Today, I speak to you as both a researcher and someone who has personally experienced erasure and bias from machines that attempt to analyze human faces.

I wish to make five main points in my testimony today:

---

[1] The Hidden Dangers of Facial Analysis, New York Times print run June 22, 2018, Page A25, online https://www.nytimes.com/2018/06/21/opinion/facial-analysis-technology-bias.html; Artificial Intelligence Has a Problem With Gender and Racial Bias. Here's How to Solve It, Time Magazine Optimist Edition http://time.com/5520558/artificial-intelligence-racial-gender-bias/; How I am Fighting Bias in Algorithms, https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms

[2] Joy Buolamwini, Timnit Gebru, Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification (February 2018), http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf; Inioluwa Raji, Joy Buolamwini, Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products (January 2019), http://www.aies-conference.com/wp-content/uploads/2019/01/AIES-19_paper_223.pdf

[3] See references of notable press mentions at www.poetofcode.com/press

- First, facial recognition technology is expanding rapidly, with little to no formal oversight.
- Second, this is occurring even though the threat of face surveillance puts civil liberties at risk, in particular endangering traditionally marginalized and vulnerable populations.
- Third, failures of a broad set of facial analysis technologies including facial recognition technology have real and dire consequences for people's lives, including in critical areas such as law enforcement, housing, employment, and access to government services.
- Fourth, the development and evaluation of these technologies raise an additional set of privacy and fairness concerns.
- Fifth and finally, given what we already know about the critical flaws of facial analysis technology, along with its rapid advancement and adoption across the country, Congress should enact a moratorium that halts law enforcement adoption of this technology unless and until appropriate regulatory mechanisms are put in place.

I. Facial Recognition Technology is Expanding Rapidly, with Little to No Formal Oversight

Facial recognition technology (FRT) which aims to analyze video, photos, thermal captures, or other imaging inputs to identify or verify a unique individual is increasingly infiltrating our lives. Facial recognition systems are being provided to airports, schools, hospitals, stadiums, shops, and can readily be applied to existing cameras systems installed in public and private spaces.[4] Companies like Facebook and Ever use FRT to identify faces in photos uploaded to their platforms with little transparency and consent practices that do not disclose the full extent to which sensitive biometric data is being used.[5] Mobile device manufacturers like Apple and Samsung enable face-based authentication to secure phones and tablets, though the systems can be fooled.[6]

Additionally, there are already documented cases of the use of FRT by government entities that breach the civil liberties of civilians through invasive surveillance and targeting. Facial recognition systems can power mass face surveillance for the government – and already there are documented excesses, such as explicit minority profiling in China[7] and undue police

---

[4] We Built an 'Unbelievable' (but Legal) Facial Recognition Machine https://www.nytimes.com/interactive/2019/04/16/opinion/facial-recognition-new-york-city.html

[5] James Vincent, "A photo storage app used customers' private snaps to train facial recognition AI" in The Verge (May 2019) https://www.theverge.com/2019/5/10/18564043/photo-storage-app-ever-facial-recognition-secretly-trained-ai. See Jennifer Lynch's 2012 Senate Testimony for a deeper dive on how Facebook failed to obtain consent in employing the photos of use to develop its facial recognition capabilities: .https://www.judiciary.senate.gov/imo/media/doc/12-7-18LynchTestimony.pdf

[6] Guy Birchall and Tom Michael, "Is the iPhone Racist? Chinese users claim iPhoneX face recognition can't tell them apart," in The Sun UK (December 2017) https://www.thesun.co.uk/news/5182512/chinese-users-claim-iphonex-face-recognition-cant-tell-them-apart/

[7] Paul Mozur, "One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority" in New York Times (April 2019) https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html

harassment in the UK.[8] Although, in the United States, the performance metrics of facial recognition systems used by the police are not mandated to be public, recent alarming individual cases of faulty facial recognition resulting in false accusations[9] and even arrests[10] add to already identified systemic causes for concern.[11]

Beyond the commercial applications, companies like Amazon and IBM put civilians at further risk by providing facial recognition systems to law enforcement and government agencies with no required oversight. On the military side, Microsoft recently signed a $480 million deal to equip the U.S. Department of Defense with HoloLens to facilitate the training and combat of active military personnel.[12] The HoloLens project could employ Microsoft's existing facial recognition capabilities in an aim to increase lethality. Similarly, reports of IBM's sale of its facial recognition technology to the Philippine government[13] and Amazon's association with the CIA and alleged sale of facial recognition technology to the FBI[14] create concern for civilian risks accrued by the widespread implementation of this technology in the absence of adequate checks and balances.

II. Face Surveillance Presents Risks, in Particular Endangering Traditionally Marginalized and Vulnerable Populations

Facial analysis technology, and face recognition in particular, raises a number of important risks, especially when adopted by law enforcement agencies. Indeed, while making investments to realize economic gains from computer vision technology, Microsoft acknowledges some of the risks posed by the face-based applications of the technology, explicitly stating:

---

[8] Silkie Carlo, "We've got to stop the Met Police's dangerously authoritarian facial recognition surveillance" in Metro UK (July 2018) https://metro.co.uk/2018/07/06/weve-got-to-stop-the-met-polices-dangerously-authoritarian-facial-recognition-surveillance-7687833/ ; Big Brother Watch, "Face Off: The Lawless Growth of Facial Recognition in UK Policing" (May 2018), https://bigbrotherwatch.org.uk/wp-content/uploads/2018/05/Face-Off-final-digital-1.pdf

[9] Jeremy C. Fox, "Brown University student mistakenly identified as Sri Lanka bombing suspect." in Boston Globe (April 2019), https://www.bostonglobe.com/metro/2019/04/28/brown-student-mistaken-identified-sri-lanka-bombings-suspect/0hP2YwyYi4qrCEdxKZCpZM/story.htm I

[10] Bah v. Apple Inc., 19-cv-03539, U.S. District Court, Southern District of New York, official lawsuit of 18-year old African American teenage boy misidentified in Apple Stores and suing for 1 Billion in damages (April 2019),

[11] Clare Garvie et al., "The Perpetual Line-up: Unregulated Police Face Recognition In America." (October 2016) https://www.perpetuallineup.org/

[12] Makena Kelly, "Microsoft secures $480 million HoloLens contract from US Army" in The Verge (November 2018) https://www.theverge.com/2018/11/28/18116939/microsoft-army-hololens-480-million-contract-magic-leap

[13] George Joseph "Inside the Surveillance Program IBM Built for Rodrigo Duterte" in The Intercept (March 2019) https://theintercept.com/2019/03/20/rodrigo-duterte-ibm-surveillance/

[14] Frank Konkel, "The Details About the CIA's Deal With Amazon" in The Atlantic (July 2014) https://www.theatlantic.com/technology/archive/2014/07/the-details-about-the-cias-deal-with-amazon/374632/ "The FBI is Trying Amazon's Facial-Recognition Software" in Next Gov https://www.nextgov.com/emerging-tech/2019/01/fbi-trying-amazons-facial-recognition-software/153888/

First, especially in its current state of a development, certain uses of facial recognition technology increase the risk of decisions and, more generally, outcomes that are biased and, in some cases, in violation of laws prohibiting discrimination. *Second*, the widespread use of this technology can lead to new intrusions into people's privacy. *And third*, the use of facial recognition technology by a government for mass surveillance can encroach on democratic freedoms.[15]
- Brad Smith, President, Microsoft - December 2018

Figure 2. Screen Capture of Microsoft Azure Computer Vision API describes Michelle Obama as a "young man," showing successful face detection accompanied by erroneous image captioning and serving as a reminder that commercial AI systems are fallible.[16]

These risks fall disproportionately on already marginalized people. This rapidly expanding technology can amplify inequalities, and poses unprecedented privacy risks as the face is an immutable high visibility identifier. The threat of face surveillance puts civil liberties at risk, in particular endangering vulnerable populations. Despite these harms and more, the technology is being readily adopted in consumer, business, law enforcement, and military contexts in the absence of regulation, safeguards, transparency and accountability. The lack of regulation persists, even in the face of resistance from concerned civilians and expressed concern from civil rights groups, including the Congressional Black Caucus[17]; opposition from tech workers

[15] Brad Smith, "Facial recognition: It's time for action" in Microsoft on the Issues (December 2018) https://blogs.microsoft.com/on-the-issues/2018/12/06/facial-recognition-its-time-for-action/
[16] See full demonstration at https://medium.com/@Joy.Buolamwini/when-ai-fails-on-oprah-serena-williams-and-michelle-obama-its-time-to-face-truth-bf7c2c8a4119
[17] Letter to Amazon about Facial Recognition Technology, Congressional Black Caucus: https://cbc.house.gov/news/documentsingle.aspx?DocumentID=896

and shareholders[18]; caution from AI experts[19], and the recent San Francisco ban[20] which all show the growing public awareness of the dangers posed by unregulated, unproven, and at times unwanted facial recognition technology. The federal government must act now to protect the public interest by putting in place limitations on facial recognition technology and measures for consent, privacy protections, meaningful transparency, and continuous oversight.

Facial Recognition is part of a wider Family of Facial Analysis Technologies that Require Oversight

Critically, we must remember that facial recognition is but a subset of a family of facial analysis tasks that can be developed to not only recognize an individual's unique biometric signature but can also learn soft biometrics like age, gender, and race or attempt to make nonbiometric inferences about emotions or neurological state even if there isn't underlying scientific evidence to support the inference. Facial analysis technology that can somewhat accurately determine demographic or phenotypic attributes like skin type can be used to profile individuals, leaving certain groups more vulnerable for unjustified stops. An Intercept investigation reported that IBM used secret surveillance footage from NYPD and equipped the law enforcement agency with tools to search for people in video by hair color, skin tone, and facial hair.[21] Such capabilities raise concerns about the automation of racial profiling by police in the United States. Deliberations to regulate facial ***recognition*** technology need to contend with a broader set of facial ***analysis*** technology capabilities that go beyond identifying unique individuals.

Letter from Nationwide Coalition to Amazon CEO Jeff Bezos regarding Rekognition, Civil Rights Group Coalition: https://www.aclu.org/letter-nationwide-coalition-amazon-ceo-jeff-bezos-regarding-rekognition, Public Petition to Stop Amazon from Selling Facial Recognition Technology: https://action.aclu.org/petition/amazon-stop-selling-surveillance

[18] Letter from Shareholders to Amazon CEO Jeff Bezos regarding Rekognition, Amazon Shareholders: https://www.aclu.org/letter/letter-shareholders-amazon-ceo-jeff-bezos-regarding-rekognition; Alexa Lardieri, "Amazon Employees Protesting Sale of Facial Recognition Software" in US News (October 2018) https://www.usnews.com/news/politics/articles/2018-10-18/amazon-employees-protesting-sale-of-facial-recognition-software

[19] Dina Bass, "Amazon Schooled on AI Facial Technology By Turing Award Winner" in Bloomberg (April 2019), https://www.bloomberg.com/news/articles/2019-04-03/amazon-schooled-on-ai-facial-technology-by-turing-award-winner:; Concerned Researchers author "On Recent Research Auditing Commercial Facial Analysis Technology" (March 2019), https://medium.com/@bu64dcjrytwitb8/on-recent-research-auditing-commercial-facial-analysis-technology-19148bda1832
Open Letter to Amazon against Police and Government use of Rekognition from Researchers: https://www.icrac.net/open-letter-to-amazon-against-police-and-government-use-of-rekognition/

[20] Kate Conger, Richard Fausset and Serge F. Kovaleski "San Francisco Bans Facial Recognition Technology" in New York Times (May 2019) https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html

[21] "IBM Used NYPD Footage to Develop Skin Color Video Search" in The Intercept (September 2018) https://theintercept.com/2018/09/06/nypd-surveillance-camera-skin-tone-search/

Biometric Status	Task Types	Facial Analysis	Body Analysis
Biometric (Uniquely Identifying)	Recognition	Face, Iris, Ear	Person, Hand, Finger
"Soft" Biometric (Somewhat Identifying)	Detection Classification	Demographic and Phenotypic characteristics *Gender / Age / Ethnicity / Race / Height / Skin Color / etc.*	
Non-biometric	Detection Segmentation	Face	Person, Pedestrian
	Detection Segmentation Classification	Face attributes *Smile / Glasses / Beard / etc.*	Body attributes *Gesture / Body part / Pose / Clothing / etc.*
	Detection Classification	Emotion, Neurological Status	

Figure 3. Example human-centric vision tasks for single images.[22]

Facial analysis technology that fall outside the traditional technical definitions of facial recognition, including those that enable the collection of demographic information like age, gender or race, the inference of health condition from the analysis of skin exposed on a face, or the assignment of behavioral traits, can propagate harms by enabling unfair differential pricing of goods, sharing health status without consent to third-parties, or systematizing discriminatory access. For example Facebook has a patent application for video technology to enable retailers to access information about an individual entering a store using the vast collection of sensitive face biometrics data they have stored on the over 2 billion users of the social media platform. One use case presented in the patent involves determining the trustworthiness of an individual based on social media activity and determining access to high value products based on the trustworthiness score.[23]

social credit

Instead of limiting legislation on facial analysis technology to facial recognition capabilities, lawmakers should make sure to consider adjacent use cases that can propagate harms or lead

[22] Dina Bass, "Amazon Schooled on AI Facial Technology By Turing Award Winner" in Bloomberg (April 2019), https://www.bloomberg.com/news/articles/2019-04-03/amazon-schooled-on-ai-facial-technology-by-turing-award-winner:; Concerned Researchers author "On Recent Research Auditing Commercial Facial Analysis Technology" (March 2019), https://medium.com/@bu64dcjrytwitb8/on-recent-research-auditing-commercial-facial-analysis-technology-19148bda1832

[23] Natasha Singer, "Facebook's Push for Facial Recognition Prompts Privacy Alarms" in New York Times (July 2018) https://www.nytimes.com/2018/07/09/technology/facebook-facial-recognition-privacy.html

to illegal forms of discrimination using information assessed from a face that is not necessarily uniquely identifying. Figure 3 provides a sampling of the different kinds of computer vision tasks that can incorporate human heads and faces.

The following section will explore failed applications of face recognition and analysis technology that are in urgent need of government oversight and complement timely reports from the Georgetown Law Center on Privacy & Technology that provide detailed explorations of face surveillance in the United States, unlawful airport use of face scans, unregulated police use of facial recognition technology, and flawed data practices employed by law enforcement officers misusing facial recognition technology.[24]

III. Failures of Face-Based Technologies Can Have Substantial Negative Consequences

Though the use of facial recognition and analysis systems is increasing, there are notable age, gender, race and phenotypic accuracy disparities that heighten the disparate impact risks of using these systems and other face-based tools in sensitive domains such as law enforcement, housing, and employment. Regardless of accuracy, face-based tools can be abused in the hands of authoritarian governments, unfettered advertisers, or personal adversaries; and, as it stands, peer-reviewed research studies and real-world failure cases remind us that the technology is susceptible to consequential bias and misuse.

Law Enforcement Misidentifications

Inaccuracies in facial recognition technology aimed at locating a unique person can result in an innocent person being misidentified as a criminal suspect and subjected to undue police scrutiny.

This is not a hypothetical situation. In April of 2019, a Brown University senior and Muslim activist, Amara K. Majeed, was misidentified by facial recognition technology as a terrorist suspect in the Sri Lanka Easter bombings.[25] As a woman of color under the age of 25, she fit demographic groups (women and youth) that former FBI facial recognition expert and colleagues recorded to be most susceptible to inaccuracies from facial recognition systems they tested in a seminal study on the the impact of demographics on the accuracy facial recognition technology. The police department later issued a statement correcting the error, yet the damage had already been done. According to the Boston Globe, Ms. Majeed received death threats as a result of the mistake, her family members in Sri Lanka were exposed to greater police scrutiny, and as a student studying for finals at the time of misidentification, her academic performance was also put at risk.

---

[24]All Georgetown Law Center on Privacy & Technology reports can be found here: https://www.law.georgetown.edu/privacy-technology-center/publications/

[25] Jeremy C. Fox, "Brown University student mistakenly identified as Sri Lanka bombing suspect." in Boston Globe (April 2019), https://www.bostonglobe.com/metro/2019/04/28/brown-student-mistaken-identified-sri-lanka-bombings-suspect/0hP2YwyYi4qrCEdxKZCpZM/story.html

On April 22, 2019, Ousmane Bah sued Apple for one billion dollars for misidentifying him as a thief.[26] Mr. Bah is an 18-year old African American teenager and, like Ms. Majeed, fits two demographic groups that FRT has been shown to struggle with, namely African-Americans and youth (18-30).[27] While it could be tempting to dismiss these recent cases as exceptional examples, we have to keep in mind that the existing real-world performance metrics available on police use of FRT indicate more misidentifications, not fewer, are likely if deployed more widely. Between May 2017 and March 2018, Big Brother Watch UK reported that the faces of over 2,400 misidentified innocent people were stored by the South Wales police department without their consent. *The department reported a false-positive facial identification rate of 91 percent.*[28]

Unlike in the UK, police departments in the United States do not report the real-world performance metrics of their facial recognition systems. Furthermore, there are no mandated accuracy requirements or real-world performance reporting mechanisms to provide critical information about whether these tools have reached technical maturity. When facial recognition technology is under consideration for deployment in sensitive contexts like policing, it is irresponsible to use these systems without legislative mandate and oversight or to employ systems that have not been proven fit for use on the intended population.

Resistance to Residential Use

Housing is another sensitive area where facial recognition technology is being introduced in the face of justifiable opposition that demonstrates the need for safeguards and governance. For example, in May 2019 the Brooklyn Legal Services Tenant Rights Coalition filed an opposition on behalf of 134 tenants in two rent-stabilized apartments to block the installation of a face recognition entry system provided by StoneLock, Inc. The vast majority of tenants belong to one or more of the groups that have among the highest failures in US government sponsored studies that examine the accuracy of facial recognition technology.[29] According to the opposition, not only did the landlord fail to obtain the expressed consent of the tenants, but despite having residents who are over 90% people of color, predominantly female identifying, and include minors alongside the elderly, the landlord also failed to request demographics performance metrics on the system in consideration.

---

[26] Bah v. Apple Inc., 19-cv-03539, U.S. District Court, Southern District of New York, official lawsuit of 18-year old African American teenage boy misidentified in Apple Stores and suing for 1 Billion in damages (April 2019), https://www.scribd.com/document/407291893/Bah-v-Apple-Inc-19-cv-03539-U-S-District-Court-Southern-District-of-New-York

[27] Brendan Klare et al., "Face Recognition Performance: Role of Demographic Information," in *IEEE Transactions on Information Forensics and Security* (December 2012) https://ieeexplore.ieee.org/document/6327355

[28] Big Brother Watch, "Face Off: The lawless growth of facial recognition in UK policing," (May 2018) https://bigbrotherwatch.org.uk/wp-content/uploads/2018/05/Face-Off-final-digital-1.pdf

[29] Cynthia M. Cook et al., Demographic Effects in Facial Recognition and Their Dependence on Image Acquisition: An Evaluation of Eleven Commercial Systems," in *IEEE Transactions on Biometrics, Behavior, and Identity Science* (February 2019), https://ieeexplore.ieee.org/document/8636231; Klare et al.Note Brendan Klare is a Former FBI Facial Recognition Expert.

The accuracy of a facial recognition system cannot be assumed to hold constant across demographic groups. In an April 2019 report conducted by the National Institute for Standards and Technology, audited facial recognition models on average had decreased accuracy for matching the faces of individuals over 71 and under 17 as compared to other age groups. In regards to the faces of minors,[30] the report states, "Younger subjects give considerably higher FNMR. This is likely due to rapid growth and change in facial appearance."[31] Furthermore, companies like StoneLock who have deployed their facial recognition technology in one context cannot assume that in a different context with a different population acceptance or performance will be comparable. In a document sent to tenants, StoneLock attempted to inspire confidence claiming to have successfully deployed its technology to 40% of fortune 100 companies. However the demographic composition of corporate employees who have used the system does not necessarily match the composition of the predominantly black and brown tenants.

FLOO being very diverse

The Brooklyn housing example shows how the current lack of regulations and oversight permits marginalized communities to be further exposed to facial recognition technology that instead of proving beneficial can become an imposition on the rights, privacy, and security of civilians. No person should be required to submit face biometrics to law enforcement or immigration officials in exchange for a roof over their head, but in the current unregulated climate, there is a significant risk the information could leak from the landowners' hands into the government's. The absence of regulation and privacy law in this space cautions against the use of facial surveillance technologies in sensitive locations like apartment buildings.

Employment Complications

Alongside government agencies, companies are looking to use face-based verification for gate keeping and fraud detection. Companies are also using facial analysis technology to make inferences about potential and current workers.

Uber reportedly deactivated the accounts of transgender drivers[32], erroneously denying economic opportunity and highlighting how gender-minorities face additional harms from these face-based tools.[33] Facial analysis technology that incorporates emotion recognition[34] is being

---

[30] Face Recognition Vendor Test (FRVT) from National Institute of Standards and Technology (NIST), https://www.nist.gov/sites/default/files/documents/2019/04/04/frvt_report_2019_04_04.pdf

[31] FNMR is an acronym for False Negative Match Rate. FNMR is used to determine how likely a genuine subject is to be rejected by the matching algorithm. For the case of entrance to a residential property, this metric can be interpreted as the estimated likelihood a tenant is blocked from entry.

[32] Jaden Urbi, "Some transgender drivers are being kicked off Uber's app" in CNBC (August 2018) https://www.cnbc.com/2018/08/08/transgender-uber-driver-suspended-tech-oversight-facial-recognition.html

[33] See more about the harms transgender and other gender minorities face from automated decision making systems: Sasha Costanza-Chock, "Design Justice, A.I., and Escape from the Matrix of Domination" in Journal of Design and Science (July 2018), https://jods.mitpress.mit.edu/pub/costanza-chock . For the limitations and harms of binary gender classification see: Os Keyes. 2018. "The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition" https://ironholds.org/resources/papers/agr_paper.pdf

integrated into hiring tools. Hiring intelligence company HireVue, allows employers to interview potential job candidates on camera, using artificial intelligence to rate videos of each application according to verbal and nonverbal cues.[35] The system is reportedly trained on the current top performers of a company.[36] Should those employees be largely homogenous there is a risk that the data-centric AI system learns not to discriminate on features for applicant ability but discriminate on identity based features like gender. Amazon learned this lesson when an internal AI hiring tool developed to increase efficiency was reported to have harmful gender bias after the system was training on 10 years of hiring data. If the word "women's" and certain women's colleges appeared in a candidates' resumes, they were ranked lower.[37]

The Amazon's internal tool did not use video input which introduces new risks. As I wrote in my New York Times op-ed, "Given how susceptible facial analysis technology can be to gender and racial bias, companies using HireVue, if they hope to increase fairness, should check their systems to make sure it is not amplifying the biases that informed previous hiring decisions. It's possible companies using HireVue could someday face lawsuits charging that the program had a negative disparate impact on women and minority applicants, a violation of Title VII of the Civil Rights Act." Beyond having companies implement internal bias mitigation processes, there needs to be external testing and validation to assess the use of face-based tools in employment contexts.

Phenotypic Failures Impeding Access To Government Services

In addition to documented accuracy disparities in facial recognition technology based on demographic attributes like age, gender, and race, phenotypic traits can have an impact on the accuracy of facial recognition systems.

A February 2019 government-backed study revealed that for 11 commercial facial recognition systems examined in the recent U.S. Department of Homeland Security, Science and Technology Directorate biometric technology rally, skin reflectance had the most impact on the accuracy of the systems. In addition to being less accurate on individuals with darker complexions, these facial recognition systems took longer to process the faces of darker-skinned individuals as compared to lighter-skinned individuals.[38]

---

[34] Emotion recognition is also referred to as affect recognition. See the AI Now 2018 report which explores how the use of facial analysis technology to link external expressions to inferences of internal states can be akin to digital phrenology. https://ainowinstitute.org/AI_Now_2018_Report.pdf

[35] Corporate Financial Institute, "HireVue Interview Guide: How to prepare for a HireVue interview," accessed on 20 May 2019 https://corporatefinanceinstitute.com/resources/careers/interviews/about-hirevue-interview/

[36] https://www.businessinsider.com/hirevue-ai-powered-job-interview-platform-2017-8

[37] Jeffrey Dastin (October 2018) https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

[38] Demographic Effects in Facial Recognition and Their Dependence on Image Acquisition: An Evaluation of Eleven Commercial Systems," in IEEE Transactions on Biometrics, Behavior, and Identity Science (February 2019), thttps://ieeexplore.ieee.org/document/8636231

Reuters reported a case where a New Zealand man of Asian descent had his photo rejected by an online passport photo checker run by New Zealand's department of internal affairs. The facial recognition systems registered his eyes as being closed by mistake.[39] When government agencies attempt to integrate facial recognition into verification processes phenotypic and demographic bias can lead to a denial of services that the government has an obligation to make accessible to all constituents. Should it be deemed suitable and attending privacy risks are addressed, any US government agency considering using facial recognition technology for access to services needs to make sure other means of verification exist. The agencies need to also assess the phenotypic and demographic performance of the system on the intended population that will be using the system.

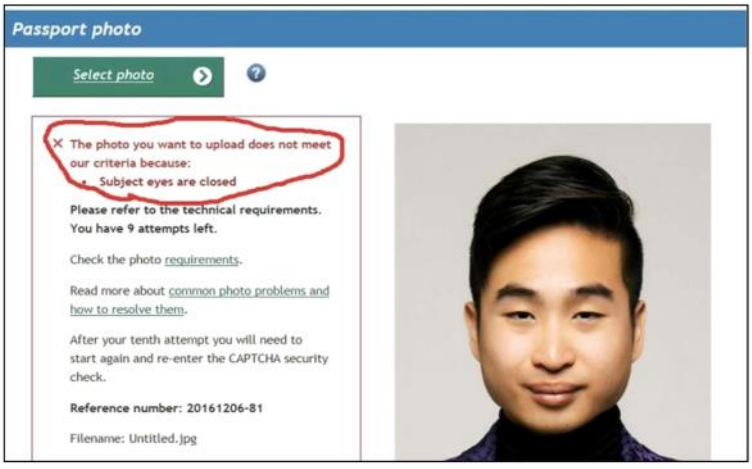

Figure 4. A screenshot of New Zealand man Richard Lee's passport photo rejection notice stating "Subjects eyes are closed", supplied to Reuters.

IV. Key Challenges with the Development and Evaluations of Facial Analysis Technology

Alongside the need for regulation and oversight of facial recognition and related technologies is the need to reexamine the existing processes for the development and evaluation of these technologies. As these technologies heavily rely on large-scale biometric data collection for development and evaluation, they immediately present data privacy risks. Furthermore, the act of labeling human faces with demographic information like race and gender risks[40] reifying

[39] James Regan, "New Zealand passport robot tells applicant of Asian descent to open eyes," in Reuters (December 2016) https://www.reuters.com/article/us-newzealand-passport-error/new-zealand-passport-robot-tells-applicant-of-asian-descent-to-open-eyes-idUSKBN13W0RL

[40] See scholarship from Os Keyes and Morgan Klaus Scheuerman about the potential harms of automatic gender recognition for transgender communities, as the use of automated gender recognition can range from the humiliating (showing a public advertisement which implicitly misgenders the individual) to the violent (rejecting identification or entry to a public restroom).

social constructs and can rob individuals the agency of self identification and disclosure. In order to assess the performance of facial analysis technology to assess accuracy on different subpopulations thoughtful classification in addition to the the acknowledgement of the limitations and politics of classification is necessary. We must also set limits on what kinds of inferences are made based on a face as controversial research studies that claim to assess sexuality and criminality from face remind us of vestiges of physiognomy that reinforce stereotypes often at the expense of marginalized communities.[41]

Skewed Training Data

Currently, state-of-the art facial analysis systems rely on an approach to artificial intelligence known as machine learning. Machine learning models are trained on vast quantities of 2D or 3D inputs to learn how to analyze patterns of human faces. The inputs can use visible light, near-infrared light, or other imaging approaches to capture human face data. Regardless of the imaging techniques used, a machine learning face-based biometric identification system has risks for bias and accuracy disparities. In an IBM Research report exploring existing limitations in face datasets, the authors note, "Face recognition systems that are trained within only a narrow context of a specific data set will inevitably acquire bias that skews learning towards the specific characteristics of the dataset."[42] The chart below provides information about notable imbalances by age, gender, and/or skin type for seven prominent face datasets:

	Age Group							Binary Gender[43]		Skin Color / Type	
Dataset	**0-3**	**4-12**	**13-19**	**20-30**	**31-45**	**46-60**	**>60**	**Female**	**Male**	**Darker**	**Lighter**
LFW	1.0%	10.6%	25.4%		29.6%		33.4%	22.5%	77.4%	18.8%	81.2%
IJB-C*	**0.0%**	**0.0%**	**0.5%**	**16.2%**	**35.5%**	**35.1%**	**12.7%**	**37.4%**	**62.7%**	**18.0%**	**82.0%**
Pub fig	1.0%	10.8%	55.5%		21.0%		11.7%	50.8%	49.2%	18.0%	82.0%
CelebA	77.8%					22.1%		58.1%	42.0%	14.2%	85.8%
UTKface	8.8%	6.5%	5.0%	33.6%	22.6%	13.4%	10.1%	47.8%	52.2%	35.6%	64.4%
AgeDB	0.1%	0.52%	2.7%	17.5%	31.8%	24.5%	22.9%	40.6%	59.5%	5.4%	94.6%
IMDB-Face	0.9%	3.5%	33.2%	36.5%	18.8%	5.4%	1.7%	45.0%	55.0%	12.0%	88.0%

Table 1. Age, Binary Gender, and Skin Color/Type Distribution of 7 Prominent Face Datasets
Data reproduced from IBM Research Diversity in Faces Report: https://arxiv.org/pdf/1901.10436.pdf
*IJB-C is a US Government Face Dataset Produced by the National Institute for Standards and Technology

[41] Wang et al.. 2018. "Deep Neural Networks Are More Accurate than Humans at Detecting Sexual Orientation from Facial Images." Journal of Personality and Social Psychology 114 (2): 246–57; Wu, Xiaolin, Xi Zhang. 2016. "Automated Inference on Criminality Using Face Images." CoRR abs/1611.04135.
[42] IBM Research, Diversity in Faces (April 2019), https://arxiv.org/pdf/1901.10436.pdf
[43] No systematic information is yet available about face based biometric identification system failure rates for gender nonconforming, nonbinary gender, agender, and/or transgender people, specifically.

While private companies may have datasets with different population distributions, without disclosure of the characteristics of the training data they use, we simply do not know and cannot assume the datasets are more diverse than publicly available datasets. Though research has shown better balanced datasets can lead to more accurate models, we also cannot assumed that more inclusive training datasets by themselves will completely address accuracy disparities. In the 2019 Algorithmic Justice League Actionable Auditing study, my colleagues and I found that even when target companies improved binary-gender classification performance publicly attributed to improved training data, they still performed better on lighter-skinned than darker-skinned faces, performed better on male-identified faces than female-identified faces, and performed worst on women of color. Even if accuracy disparities are within a few percentage points, differential accuracy on millions or hundreds of millions of people will impact substantial quantities of individuals.

Limited Single-Axis and Demographic Performance Evaluation

In addition to having transparency about the demographic and phenotypic composition of face training data of models, models that are aimed at sensitive use cases like law enforcement, housing, or employment must be externally evaluated to assess suitability of use on intended populations should there be legislative approval for deployment.

What is intended population? Messy

Such tests cannot rely on a single aggregate metric for accuracy and must be constructed to disaggregate differences between subpopulations, which can be substantial.

GENDER CLASSIFIER	TYPE I	TYPE II	TYPE III	TYPE IV	TYPE V	TYPE VI
Microsoft	1.7%	1.1%	3.3%	0%	23.2%	25.0%
Megvii (Face++)	11.9%	9.7%	8.2%	13.9%	32.4%	46.5%
IBM	5.1%	7.4%	8.2%	8.3%	33.3%	46.8%

Table 2. From 2018 Gender Shades Study: Binary-Gender Classification Error Rates on Women by Fitzpatrick Skin Type

For example, when evaluating error rates for the the facial analysis task of binary-gender classification (which does not account for gender nonconforming people, nonbinary people, agender people, and/or transgender people), our 2018 Gender Shades audit showed women with skin types associated with blackness had error rates as high as 47%. In the same study for men with skin-types perceived as white, error rates were no more than .08% in aggregate. The 47% error rate is of note because binary-gender classification has a 50/50 chance of success based on a random guess. Facial recognition technology that attempts to match a face against

thousands or millions of potential face matches must deal with much larger odds. Furthermore, though there is research that shows skin properties as opposed to self-identified racial subgroups provide more insights into accuracy disparities,

To date as it relates to face-based technologies used in sensitive contexts intersectional demographic and phenotypic error rate distribution information is not required to inform procurement decisions, nor are there standards that provide guidelines on when the use of models that shows problematic bias.

No Guidelines Around Harvesting of Face Data

The development and evaluation of facial analysis technology is dependent on the large-scale collection of sensitive face data. Private sector, government and academic collection face data often happens in the absence of consent or knowledge of the individuals whose face data is harvested.

In the United States there are no restrictions that prevent private companies from collecting face data without consent. In 2011, Facebook automatically enrolled all of its users in its face recognition program without obtaining consent.[44] At times, face data collected expressly collected for one use is repurposed for an entirely different use, without any notice to the people whose face data is collected. For example, the photo storage app Ever, which began as a cloud storage app but shifted to face recognition four years after its founding, did not explicitly inform its customers they had begun using their faces to train face recognition technology.[45]

On behalf of the United States government, the National Institute of Standards and Technology(NIST) maintains a dataset of faces that companies can test against in order to evaluate their accuracy when developing facial analysis technology. However, the dataset NIST maintains for private companies to practice on contains photos certainly not given with consent, including photos of children exploited for child pornography, immigrant visa application photos, and mugshots.[46] NIST even makes public a mugshot dataset of deceased persons.[47]The inclusion of mugshot photos in particular may be especially pernicious, as inclusion in a

---

[44] Jennifer Lynch's 2012 testimony: https://www.judiciary.senate.gov/imo/media/doc/12-7-18LynchTestimony.pdf

[45] James Vincent, "A photo storage app used customers' private snaps to train facial recognition AI," in The Verge (May 2019), https://www.theverge.com/2019/5/10/18564043/photo-storage-app-ever-facial-recognition-secretly-trained-ai; for the original report see: Olivia Solon and Cyrus Farivar, "Millions of people uploaded photos to the Ever app. Then the company used them to develop facial recognition tools," in NBC News (May 2019), https://www.nbcnews.com/tech/security/millions-people-uploaded-photos-ever-app-then-company-used-them-n1003371

[46] Os Keyes, Nikki Stevens, and Jacqueline Wernimont, "The Government Is Using the Most Vulnerable People to Test Facial Recognition Software," in Slate (March 2019) https://slate.com/technology/2019/03/facial-recognition-nist-verification-testing-data-sets-children-immigrants-consent.html

[47] https://www.nist.gov/itl/iad/image-group/special-database-32-multiple-encounter-dataset-meds

mugshot dataset does not indicate criminality, and will disproportionately impact black and brown people.

Unregulated face data harvesting also extends into academic research.[48] For the Algorithmic Justice League's Pilot Parliaments Benchmark developed in 2017, we deliberately choose to use images of parliamentarians since they are public figures with known identities and photos available under non-restrictive licenses posted on government websites. Given the ease of collecting mass face datasets and the risks to privacy, guidelines around the selection of subjects and images should apply to researchers and other entities developing public or private datasets.

V. Recommendations

The Algorithmic Justice League (AJL) urges Congress to consider adopting a moratorium prohibiting law enforcement use of face recognition or other facial analysis technologies unless and until regulations are adopted—in partnership with communities—that ensure these powerful tools will be used in ways that are responsible, accountable, and transparent. These tools are too powerful, and the potential for grave shortcomings, including extreme demographic and phenotypic bias is clear. We cannot afford to allow law enforcement agencies to adopt these tools and begin making decisions based on their outputs today, and figure out later how to rein in misuses and abuses.

This recommendation stems in part from the fact that law enforcement agencies themselves and vendors of facial analysis technologies have demonstrated unwillingness to engage in meaningful self-regulation. AJL made recommendations to companies in the past on what we believe to be the most basic measures that companies should take in order to demonstrate their commitment to developing facial analysis and recognition technology that aligns with the ethical treatment of the public. In partnership with the Georgetown Center on Privacy & Technology, we released the "Safe Face Pledge" project (https://www.safefacepledge.org/), an opportunity for organizations to make public commitments towards mitigating the abuse of facial analysis technology.

The Safe Face Pledge calls on organizations to declare the following priorities:

- **S**how Value for Human Life, Dignity, and Rights
- **A**ddress Harmful Bias
- **F**acilitate Transparency
- **E**mbed Commitments into Business Practices

[48] Olivia Solon, "Facial recognition's 'dirty little secret': Millions of online photos scraped without consent" in NBC News (March 2019), https://www.nbcnews.com/tech/internet/facial-recognition-s-dirty-little-secret-millions-online-photos-scraped-n981921

But although these priorities are fundamental and should be uncontroversial, to date very few vendors of facial analysis have signed on. And many high-visibility vendors of face recognition technology were approached directly regarding the Pledge, but ultimately declined to sign.

Again, given what we know today about how law enforcement agencies use and abuse facial analysis technology, as well as the susceptibility of these tools to unreliability and bias, the responsible step forward would be to adopt a moratorium that halts adoption of these tools until appropriate protections can be adopted. After that, much more is needed, but the basic priorities outlined in the Safe Face Pledge can inform the framing of regulatory recommendations for the government to take action to protect its citizens from the harmful misuse of these technologies. We thus present our official recommendations within this framework to highlight minimal requirements to mitigate the harms associated with facial analysis technology.

Detailed recommendations are thus outlined below:

Commitment One: Show Value for Human Life, Dignity, and Rights

- **Institute Moratorium on Face Surveillance:** As there is an absence of regulations and privacy laws that address face-based biometric use and substantial risks, especially to marginalized and vulnerable communities, use of this technology should be halted unless and until there are are legal limitations and protections.

- **Mandate Affirmative Consent:** Facial analysis technology should not be used in consumer products or online platforms without explicit opt-in consent. Some companies complying with the GDPR in Europe already have adopted an opt-in approach to facial recognition.[49] People in the United States deserve the same assurance that they will not be subjected to facial analysis technology unless and until they have been asked for, and have provided, their expressed permission.

Commitment Two: Address Harmful Bias

- **Require Vendors of Facial Analysis Technology to:**
    - Implement internal bias evaluation, mitigation, and reporting procedures;
    - Regularly report performance on available national benchmarks; and
    - Support independent evaluation from research community and third-party testing.

- **Decriminalize Beneficial Research:** Researchers investigating bias in commercial systems may need to employ methods that arguably fall under the Computer Fraud and Abuse Act. This threat looms particularly as some vendors of facial analysis technology push back against outside criticism by making it more, rather than less, difficult for researchers to use and manipulate their products for the purpose of evaluating performance and bias.

---

[49] Thuy Ong, "Facebook announces new European privacy controls, for the world" in The Verge (Apr. 18, 2018) https://www.theverge.com/2018/4/18/17250840/facebook-privacy-protections-europe-world-gdpr

- **Require National Institute of Standards & Technology to:**
  - Make public the demographic and phenotypic composition of its benchmark datasets;
  - Report accessible intersectional performance metrics as part of its ongoing Face Recognition Vendor Test;[50] and
  - Address data ethics concerns with its collection,use, and/or dissemination of vulnerable children's images, deceased persons' images, mugshots, and other unconsented datasets.

Commitment Three: Facilitate Transparency

- **Institute Transparency Requirements:** Facial analysis technology should not be used to track, monitor, and analyze human subjects in secret, incapacitating our ability to object or even to decline to use products that violate our preferences. Customers and users need to know when and how the is being used in consumer products and services and have a choice in whether or not their face data is captured, stored, sold, and/or used to enhance the technical capabilities of the vendor or third parties.[51]

- **Mandate Disclosure of Employment, Education, Housing, or Health Use—Private or Public:** Facial analysis technology that has not been evaluated for bias and demonstrated fit for purpose for intended populations should not be used in areas with material consequences for the lives of civilians. Should any cases be proven fit for purpose, consent must be obtained from individuals to use the technology and disclosure of the potential to use the technology must be mandated.

Commitment Four: Embed Commitments into Practices

- **Regulate Process Not Just Products :** Vendors should ensure that the priorities outlined above are embedded and considered at every level of their operations. Enforcement agencies, including the Federal Trade Commission and State Attorneys General, should evaluate failures with respect to these priorities under existing frameworks prohibiting unfair and deceptive trade practices.

- **Require Agencies Overseeing Critical Processes to Check Unproven Uses:** Agencies should restrict use of facial analysis technology to inform decisions regarding employment, credit, housing, healthcare, and other critical services unless and until entities produce evidence that the technology does not introduce bias in violation of anti-discrimination laws and has sound scientific basis. For example, facial analysis technology should not be permitted in hiring decisions where it has not been clearly established that use of that technology would not result in violations of Title VII.

---

[50] National Institute of Standards and Technology, "Face Recognition Vendor Test" https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt

[51] As discussed above, some vendors have used photos collected for one purpose to train facial analysis products without further consent from the creators and subjects of those photos. This should not be allowed.

Conclusion

Facial recognition and other facial analysis technologies can amplify inequalities while breaching civil rights and liberties, They pose unprecedented privacy risks as the face is an immutable high visibility identifier. Peer-reviewed academic studies show face-based technology can be susceptible to age, gender, race and phenotypic accuracy disparities that heighten the disparate impact risks of using these systems in domains such as law enforcement, housing, employment, and access to government services. Real-world failures and problematic deployments including mass state surveillance, false arrests, and the denial of working opportunities remind us of what is at stake in the absence of oversight and regulation. Congress must act now to protect the public interest.

Thank you,
Joy Buolamwini

Biography

Joy Buolamwini is a graduate researcher at the Massachusetts Institute of Technology who researches algorithmic bias in computer vision systems. She founded the Algorithmic Justice League to create a world with more ethical and inclusive technology. Her TED Featured Talk on algorithmic bias has over 1 million views. Her MIT thesis methodology uncovered large racial and gender bias in AI services from companies like Microsoft, IBM, and Amazon. Her research has been covered in over 40 countries, and as a renowned international speaker she has championed the need for algorithmic justice at the World Economic Forum and the United Nations.She serves on the Global Tech Panel convened by the vice president of European Commission to advise world leaders and technology executives on ways to reduce the harms of A.I. In late 2018 in partnership with the Georgetown Law Center on Privacy and Technology, Joy launched the Safe Face Pledge, the first agreement of its kind that prohibits the lethal application of facial analysis and recognition technology.

As a creative science communicator, she has written op-eds on the impact of artificial intelligence for publications like TIME Magazine and New York Times. In her quest to tell stories that make daughters of diasporas dream and sons of privilege pause, her spoken word visual audit "AI, Ain't I A Woman?" which shows AI failures on the faces of iconic women like Oprah Winfrey, Michelle Obama, and Serena Williams as well as the Coded Gaze short have been part of exhibitions ranging from the Museum of Fine Arts, Boston to the Barbican Centre, UK. A Rhodes Scholar and Fulbright Fellow, Joy has been named to notable lists including the Bloomberg 50, Tech Review 35 under 35, BBC 100 Women, Forbes Top 50 Women in Tech (youngest), and Forbes 30 under 30. Fortune magazine named her "the conscience of the AI revolution". She holds two masters degrees from Oxford University and MIT; and a bachelor's degree in Computer Science from the Georgia Institute of Technology. Her final degree will be a PhD from MIT. Learn more at www.poetofcode.com