

ANALYSIS:

Police technologies: Gender dimension bias and the impact upon criminal justice

3rd December 2021 | Andrew Staniforth | Policing Insight



Picture © [metamorworks](#) / iStock

Artificial intelligence and machine learning have the potential to revolutionise identification processes for law enforcement; but if they are to be reliable and trusted, those commissioning their use must ensure concerns around gender dimension bias have been addressed from the outset, says Policing Insight’s Andrew Staniforth.

Senior police leaders must acknowledge that [gender dimension](#) goes far beyond gender equality, in so much as the desire to have equal numbers of identified sexes as active participants across the rank structures and the diverse operating landscape of policing.

“For the developers of algorithms that form the basis of automated data processes and artificial intelligence, it’s imperative that due diligence is afforded to the growing concern of gender dimension bias.”

Gender equality is a fundamental human right, and for social justice and security for civil society requires that everyone – regardless of their age, race, class, religion, disability, sexual orientation or gender – has equal rights and opportunities in society in general (including economic resources and decision-making), as reflected in efforts to improve the gender balance of the police services which serves to keep society safe and feeling safe.

Beyond the gender balance, it is vital to recognise the criticality of gender dimension in contemporary criminal justice, especially matters directly relating to the integrity of new technologies and identity related matters.

For the developers of algorithms that form the basis of automated data processes and artificial intelligence (AI) which underpin biometrics and facial recognition tools, techniques and technologies for security, it's imperative that due diligence is afforded to the growing concern of gender dimension bias.

Gender-by-design

All tech developers must use a 'gender-by-design-approach', created to address aspects of gender dimension which may, if not identified, examined and explored at the very outset of concept design, negatively impact upon technology outputs being used in policing.

Senior police leaders, especially those charged with the [digital transformation](#), procurement of technology products and leading on ethics, must recognise the potential of gender imbalance which is inherent in some technology design.

“They have the potential to undermine the robustness and reliability of ML operations that supplement law enforcement forensics and human investigations, which form the basis of sharing and constructing evidence that should be able to be trusted, tested and verified.”

For example, to actively address this challenge within the development and implementation of technological solutions envisaged by machine learning (ML) using biometrics, developers must start from the [ethical principle of beneficence](#), by minimising harm to citizens and potential harms to commercialised outputs.

In practice, this means that a 'gender-by-design approach' starts from acknowledging that algorithms are biased, and as such their capacity for harm is in-built, replicable and in the case of [blockchain](#) could multiply and embed biases (and hence harms).

As a result they have the potential to undermine the robustness and reliability of ML operations that supplement law enforcement forensics and human investigations, which form the basis of sharing and constructing evidence that should be able to be trusted, tested and verified.

From this general principle of inbuilt bias, if any ML biometric decision-making biases are replicated in identity (ID) management, including widely available facial recognition

systems and [combi-multimodal biometrics](#), this issue can be compounded with unintended consequences. These systems can discriminate based on characteristics such as race, age and gender, and their intersections.

Ballooning error rates

For example, in a [study exploring gender shades](#) from researchers at [Massachusetts Institute of Technology \(MIT\)](#) and [Stanford University](#) presented at the [Conference on Fairness, Accountability, and Transparency](#) in February 2018, researchers examined three commercially released facial-analysis programs from major technology companies that demonstrated both skin-type and gender biases.

[Joy Buolamwini](#), a researcher at the [MIT Media Lab](#), discovered that the dataset populated with men and lighter-skinned individuals misidentified darker-skinned females more often. In the researchers' experiments, the three programs' error rates in determining the gender of light-skinned men were never worse than 0.8%. For darker-skinned women, however, the error rates ballooned – to more than 20% in one case and more than 34% in the other two.

So while facial recognition technology is improving by leaps and bounds, and some commercial software can now tell the gender of a person in a photograph, according to the systems studied by from MIT and Stanford University, when the person in the photo is a white man, the [software is right 99% of the time](#) – but when it's a Black female it could be as low as 65%.

The findings of the research provided examples of intersectional bias, in which different types of discrimination amplify negative effects on an individual or group.

More problematically, it also highlighted instances of the reinforcement of ML stereotypes that compromise the reliability of any sifting and electronic ID recognition of authentic versus manipulated or faked IDs, even when adjusted for societal segmentation.

Fact or fiction

Three years on from the gender shades study, the findings continue to raise questions about how today's neural networks, which learn to perform computational tasks by looking for patterns in huge data sets, are both trained and evaluated.

“All in authority need to seek reassurance from tech developers that they started the design concept of their commercial product from the ethical principle of minimising harm.”

For instance, according to the paper published following their gender shades research, Joy Buolamwini and [Timnit Gebru](#), former AI Chief at Google, claimed an accuracy rate of more than 97% for a face-recognition system they had designed. But the data set used to assess its performance was more than 77% male and more than 83% white.

These disparate results, show how some of the biases in the real world can seep into AI, the computer systems that inform facial recognition.

All in authority must recognise the potential impact that gender imbalance may have on the development and implementation of technological solutions envisaged and developed by ML using biometrics.

They also need to seek reassurance from tech developers that they started the design concept of their commercial product from the ethical principle of minimising harm; anything less will lead to further mis-identifications, flawed evidence produced in court, and potentially serious miscarriages of justice.

Link to online article at Policing Insight: <https://policinginsight.com/features/opinion/police-technologies-gender-dimension-bias-and-the-impact-upon-criminal-justice/>

Policinginsight



Andrew Staniforth is Director of Innovation at SAHER (Europe), a security research, training and consultancy operating at a global level, participating in the NOTIONES (*iNteracting netwOrk of iTelligence and securIty practitiOners with iNdustry and acadEmia actorS*) network, funded by the European Commission Horizon 2020 Programme (No. 101021853). Andrew is also an active Researcher of Cyber Threats at the Hillary Rodham Clinton School of Law and Criminology, University of Swansea. As a former Counter-Terrorism Detective, he has worked across the world and supported missions of the United Nations Terrorism Prevention Branch. **Contact:** Andy@saher-eu.com
To learn more about NOTIONES visit: <https://www.notiones.eu>
