

Briefing for the General Scheme of the Garda Síochána (Recording Devices) (Amendment) Bill

18 January, 2024

Dr Abeba Birhane
Dr Elizabeth Farries

We recognise that this Draft General Scheme reflects the Department of Justice's purported and ongoing plans to enable An Garda Síochána's use of Facial Recognition Technology systems (collectively Policing FRT).¹ We will therefore refer to Policing FRT for the purpose of this submission.

In reference to the author - Dr Birhane's - research, research from the Centre for Technology and Democracy ([Radiya-Dixit & Nuff 2023](#)), and leading recent publications, we submit evidence outlining Policing FRT's disproportionate impact on vulnerabilised communities, the importance of identifying less intrusive methods through democratic consultation, and minimum safeguard thresholds that must be established *prior* to the deployment of policing FRT.

1. Disproportionate impacts on vulnerabilised communities

A person's facial biometric data is sensitive and personal. Processing this data for policing purposes is highly intrusive in that it represents a serious interference with rights including privacy, data protection, expression, assembly, and equality and non-discrimination.² Policing FRT disproportionately limits the rights of us all, but particularly those with racial and gender vulnerabilisations, through misidentification, dehumanisation, and over surveillance.³

a. Invasive and ineffective

Although often presented as a cost and resource effective aid to policing, FRT has proven to be the least effective and most intrusive technology. In a recent survey by [Big Brother Watch](#) reviewing police use of FRT across Wales where over 508,542 faces were scanned, over 3,000 people were wrongfully identified, over 88% inaccuracy recorded in the period of 2016-2023, and only 3 arrests made.⁴ FRT, therefore, is extremely invasive and a technology that expands and normalises surveillance state while also largely failing to aid effective policing.

¹ See a Minister's statement that the Department of Justice seeks to legalise Policing FRT by amending Garda Síochána (Recording Devices) Bill (now the Garda Síochána (Recording Devices) Act 2023) <https://www.irishtimes.com/crime-law/2023/04/06/oireachtas-committee-wants-to-scrutinise-use-of-facial-recognition-technology-by-gardai/>

² Farries, E. and Cronin, O. (4 June 2022) Submission to inform the report by the United Nations High Commissioner for Human Rights on the right to privacy in the digital age at its 51th session in 2022, Human Rights Council adopted resolution 48/4 [https://files.inclo.net/content/pdf/72/FINAL_%20Right%20to%20privacy%20in%20the%20digital%20age.%20HRC%2048_4%20\(1\).pdf](https://files.inclo.net/content/pdf/72/FINAL_%20Right%20to%20privacy%20in%20the%20digital%20age.%20HRC%2048_4%20(1).pdf)

³ Ibid.

⁴ Big Brother Watch (n.d.) Stop Facial Recognition <https://bigbrotherwatch.org.uk/campaigns/stop-facial-recognition/>

b. Misidentification

There is a current, peer reviewed, and established body of evidence that women, men and people who have darker skin are more likely to be misidentified by FRT.⁵ Numerous, rigorous, evidence based studies speak to these findings, not least the author's - Dr Birhane's - own [2023 publication](#) demonstrating how black men and black women have the highest rate of being classified as 'criminals' and 'suspicious person' by computer vision models, the basis for FRT.⁶ The larger the data set the higher the misclassification rate. Misclassification is proven to occur at the expense of darker skinned people.

c. Dehumanisation

The author's findings follow the findings of a [2021, Radford et al Open AI](#) audit demonstrating how image classification models tend to mislabel and misclassify images of people that have darker skin colours.⁷ This study that evaluated computer vision models using the Fairface dataset, showed that darker skinned men and women are classified as non-human animals, including chimpanzees, gorillas, and orangutans, and suspicious person, criminal and thief at a higher rate compared to other less vulnerable races and genders. The reduction through technology of people to non-human animals and suspicious characters is dehumanising.

d. Over surveillance of us all and vulnerabilised communities especially

Error prone FRT systems create over surveillance as a problematic norm: anyone's facial image captured by this technology is subject to rights implicating surveillance. However, given the historic and discriminatory over surveillance of vulnerabilised communities by law enforcement, including those with darker skin,⁸ a particular consequence and risk attached policing FRT is detainment or incarceration without cause of people from vulnerabilised communities. In the US alone, six known cases of law enforcement have been documented, all falsely incarcerating people on the basis of

⁵ See for example Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, 81, 1-15. <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>; NIST. (2019, December 19). NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software. <https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software>; UK's Metropolitan Police FRT found to have an error rate of 81 per cent, see 81% of 'suspects' flagged by Met's police facial recognition technology innocent, independent report says, Sky News, July 2019, <https://news.sky.com/story/met-polices-facial-recognition-tech-has-81-error-rate-independent-report-says-11755941> ; MIT and Stanford University tested three different commercial FRT systems; less than 1% errors for light skinned men, 20% of the cases related to faces of dark-skinned women, see Study finds gender and skin-type bias in commercial artificial-intelligence systems, MIT News, February 2018, <https://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212>

⁶ Birhane, A., Prabhu, V., Han, S., & Boddeti, V. N. (2023). On hate scaling laws for data-swamps. Ithaca: Cornell University Library, arXiv.org. <https://doi.org/10.48550/arxiv.2306.13141>

⁷ Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning transferable visual models from natural language supervision. Ithaca: Cornell University Library, arXiv.org. <https://doi.org/10.48550/arxiv.2103.00020>

⁸ See for example "Sus Laws" (stop and search laws) that were critiqued for their discriminatory application amongst black and Irish people in the 1970s. See McCluskey, S. (2016). The crime of being suspicious: British counter-terrorism legislation and the history of discriminatory preventative laws in the United Kingdom. Rutgers Race & the Law Review, 17(1), 131-165.

their incorrect FRT classifications.⁹ Every single person was black. It is important to note that these are only reported cases. Many people don't have the resources to know that they were jailed because of a policing FRT error or the means to contest it. Or if they know, they don't have the means to contest or communicate this to wider society or policy makers.

We therefore submit that this invasive and ineffective trend of misidentification, dehumanisation, and over surveillance attached to policing FRT presents the risk of undue interference into our rights that is experienced disproportionately according to race and gender.

2. Identifying less intrusive measures through adequate consultation

In addition to disproportionate rights limits of oversurveilled communities, we submit that the Department of Justice should find less intrusive measures, given the disproportionate negative impact and ineffectiveness of FRT. We highlight here the importance of proactive and direct consultation, prior to the legislative stage, with community representatives, policing and technology experts, and civil society organisations. This would include per ([Radiya-Dixit & Nuff 2023](#)):

- Via direct consultation, proactively considering views of the public, especially marginalised communities, on the particular type of Policing FRT and justified a disregard of the views if relevant;
- Conducting transparent, proactive consultations with civil society and independent experts on the particular type of Policing FRT;
- Establishing that it has considered the advice from consultations and transparently explained the outcomes, including providing a justification if the advice is not followed;
- Outlining clear, proactive processes for the public, especially marginalised communities, to influence if and how Policing FRT is implemented;
- Outlining clear, proactive processes for the public, especially vulnerabilised communities, to contest and challenge decisions from FRT, given the ample evidence showing failures of FRT; and
- Approaching the consultation process with materials accessible to people with disabilities and provided in immigrant languages.¹⁰

We see no evidence that the Department of Justice has engaged proactive consultation or established that policing FRT is less intrusive in comparison to other measures. Indeed we note the government has not adequately engaged with the proactive outreach from a consortium of experts from all **7 universities in Ireland and 13 NGOs**, including experts penning our:

- Open letter in the [Irish Times](#) (June 2022)
- [Letter to Oireachtas Cabinet members](#) (June 2022)
- [Follow up letter to the Minister of Justice](#) (November 2022)
- [Op-Ed in the Irish Times](#) (April 2023)
- [Expert Briefing Note to Oireachtas Members](#) (May 2023)

⁹ Swarns, Christina. (19 Sept 2023) When Artificial Intelligence Gets It Wrong. Unregulated and untested AI technologies have put innocent people at risk of being wrongly convicted. The Innocence Project. <https://innocenceproject.org/when-artificial-intelligence-gets-it-wrong/>

¹⁰ Radiya-Dixit, E., & Neff, G. (2023). A sociotechnical audit: Assessing police use of facial recognition. Paper presented at FAcCT '23. 1334-1346. <https://doi.org/10.1145/3593013.3594084>

3. Pre-establish minimum thresholds to safeguard Policing FRT deployment

Given 1 and 2, the Department of Justice should pre-establish minimum thresholds before deploying Policing FRT systems. Supporting Radiya-Dixit & Nuff (2023), we submit that the Department should:

- a. Carry out proactive expert, community and public consultations, according to criteria itemised in point 2 to determine the least intrusive measures for policing.
- b. Established through independent (i.e. a designated conflict-free expert auditor) auditing obligate minimum operational thresholds of selected Policing FRT systems. These thresholds include but are not limited to precision, false positive rate, true positive rate, etc.
- c. Establish safeguards precluding the use of Policing FRT with an unsuitable low-quality probe or image.
- d. Carry out and publish a data protection impact assessment and appropriate policy document for sensitive data processing
- e. Create clear, objective, and limited criteria concerning third-party access to the data collected or retained, including with regard to what data can be shared, with whom it can be shared, and for what specific purpose it can be shared.
- f. Establish and deploy accessible communication protocols across ability and languages to inform potential data subjects and most people impacted jurisdiction in advance about when, where, why, and how Policing FRT will be (or is currently being)
- g. Establish clear measures to ensure data subjects can exercise their individual rights including the rights to rectification, erasure, and object with clear justifications if exemptions apply
Deploy accessible communication protocols as above to be used and how data subjects can exercise their individual rights.¹¹

If these criteria are not met and, particularly, if Policing FRT Systems are proprietary and/or independent auditors cannot access the training data sets or models to audit them, then the Policing FRT systems should not be used. We invite the Department of Justice to identify less intrusive measures through the democratic consultation process and to include these measures in any forthcoming legislation.

Abeba Birhane is a cognitive scientist, currently a Senior Advisor in AI Accountability at Mozilla Foundation and an Adjunct Assistant Professor at the School of Computer Science and Statistics at Trinity College Dublin, Ireland. She researches human behaviour, social systems, and responsible and ethical AI – work for which she was recently featured in [Wired UK](#) and TIME on the [TIME100 Most Influential People in AI](#) list. Birhane also serves on the United Nations Secretary-General's [AI Advisory Body](#) and the newly-convened AI Advisory Council in Ireland.

Dr Elizabeth Farries is the Director of the UCD Centre for Digital Policy and a Senior Fellow with the International Network of Civil Liberties Organizations.

¹¹ *Ibid.*