PRIVACY INTERNATIONAL

Submission on Draft General Recommendation No. 36: Preventing and Combating Racial Profiling

June 2019

PRIVACY INTERNATIONAL

June 2019

Privacy International's submission on Draft General recommendation No. 36: Preventing and Combating Racial Profiling

Privacy International[1] welcomes the decision of the UN Committee on the Elimination of Racial Discrimination ('the Committee') to dedicate a General recommendation on preventing and combating racial profiling. This submission provides a few suggestions that may help clarifying the terminology used in the Draft General Recommendation 36 and also to inform the Committee's future contributions.

Artificial Intelligence (AI) is part of our daily lives. This technology has the potential to revolutionise societies in positive ways. However, as with any scientific or technological advancement, there is a real risk that the use of new tools by states will have a negative impact on human rights. It also transforms how individuals and groups can be tracked and identified, and dramatically alters what kinds of information can be gleaned about people from their data.[2] In such moments, documents, such as the draft General recommendation 36, clarifying the rules applicable are essential.

1. On 'racial profiling' for law enforcement and beyond

The draft General recommendation 36 focuses on 'racial profiling' by law enforcement. Indeed, the phenomenon of 'racial profiling' has been primarily used to describe discriminatory behaviour and activities of law enforcement authorities – in police, customs, immigration and national security agencies.[3]

[1] Privacy International (PI) PI was established in 1990 as a non-profit, non-governmental organisation based in London, working with partners around the globe, at the intersection of modern technologies and rights. It envisions a world in which the right to privacy is protected, respected, and fulfilled. PI believes that privacy is essential to the protection of autonomy and human dignity, serving as the foundation upon which other human rights are built. In order for individuals to fully participate in the modern world, developments in law and technologies must strengthen and not undermine the ability to freely enjoy this right. We are building the global movement because people must have access to privacy protection without regard to citizenship, race and ethnicity, economic status, gender, age, or education. https://privacyinternational.org/

[2] Privacy International, Submission to the UN Special Rapporteur on extreme poverty and human rights, Philip Alston, on digital technology, social protection and human rights, May 2019, available at: https://privacyinternational.org/sites/default/files/2019-05/PI%20submissions%20to%20UNSR%20Extreme%20Poverty_May%202019.pdf; Privacy International and ARTICLE 19, Privacy and Freedom of Expression In the Age of Artificial Intelligence, April 2018, available at: https://privacyinternational.org/sites/default/files/2018-04/Privacy%20and%20Freedom%20of%20Expression%20%20In%20the%20Age%20of%20Artificial%20Intelligence.pdf ; Privacy International, Submission of evidence to the House of Lords Select Committee on Artificial Intelligence, 6 September 2017, available at: https://privacyinternational.org/sites/default/files/2017-12/Submission%20of%20evidence%20to%20the%20House%20of%20Lords%20Select%20Committee%20on%20Artificial%20Intelligence%20-%20Privacy%20International.pdf.

[3] See CERD, Draft Genera recommendation 36 (CERD/C/GC/36, 14 May 2109, paras. 4-15) referring to the definition adopted during the Durban Declaration and Programme of Action, the 2015 report of the Former Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance, the 2019 publication of the High Commissioner for Human Rights.

First, Privacy International welcomes the inclusion of 'racial profiling' by customs and immigration law enforcement authorities to this recommendation. However, it would be also useful to consider the use of AI technologies by the different authorities separately, as their use and implications may differ considerably depending on the function of the law enforcement authority and may trigger different obligations. Therefore, we invite the Committee to address the separate implications of the use of AI technologies for different law enforcement purposes.

Second, racial profiling can occur in numerous contexts beyond law enforcement, particularly with the integration of new technologies in other parts of government, such as social services, health care and education. The advancement in technology has increased the powers to collect, process and gather intelligence. Recent social protection programmes include the deployment of biometric systems to access food,[4] national health insurance programmes, smart card for recipients of welfare[5] and biometric national ID systems.[6]

Newly established or reformed social protection programmes have gradually become founded and reliant on the collection and processing of vast amounts of personal data, often access and management is tied to the provision of unique identifier, and increasingly the models for decision-making include data exploitation and components of automated-decision making and profiling.[7]

These developments have increased the situations where racial profiling can be integrated. For example, new technologies are used to detect fraudulent behaviour in social welfare. The use of the basis of generalizations and stereotypes related to presumed race, colour, descent, nationality, place of birth, or national or ethnic origin – rather than objective evidence or individual behaviour – as a basis for erroneous suspicion that people with such characteristics are prone to provide fraudulent information to receive benefits can also be understood as racial profiling.[8]

Therefore, we call the Committee to consider in the future the racial biases associated with artificial intelligence in other government sectors, such as social protection.

2. On the distinction between artificial intelligence, automated decision making and algorithms in the Draft General Recommendation 36

We would like to further alert the Commission with respect to the distinct nature of each of the enumerated notions, including but not limited to artificial intelligence, automated decision making

[4] For example, in Venezuela, see: López, V., *Venezuela to introduce new biometric card in bid to target food smuggling*, The Guardian, 21 August 2014, available at: https://www.theguardian.com/world/2014/aug/21/biometric-venezuela-food-shortages-smuggling-fingerprints, and in Botswana, see: SmartSwitch, *The Electronic Food Coupon System – an African Success Story*, available at: http://www.smartswitch.co.bw/fcSuccess.

[5] Mastercard Press Release: *Ten Million SASSA MasterCard Cards Issued to South African Social Grant Beneficiaries*, available at: http://newsroom.mastercard.com/press-releases/ten-million-sassa-mastercard-cards-issued-to-south-african-social-grant/.

[6] These are mandatory to access social protection programmes in India, Peru, Pakistan, for example.

[7] Privacy International, Submission to the UN Special Rapporteur on extreme poverty and human rights, Philip Alston, on digital technology, social protection and human rights, May 2019, available at: https://privacyinternational.org/sites/default/files/2019-05/PI%20submissions%20to%20UNSR%20Extreme%20Poverty_May%202019.pdf.

[8] Private companies can also target people on the basis of race or inferred characteristics of race to direct advertising, shape content, etc. See Sweeney, Latanya, 'Discrimination in Online Ad Delivery', SSRN Scholarly Paper, Rochester, 28 January 2013, available at: https://papers.ssrn.com/abstract=2208240.

and algorithms. As a result of their distinct nature, their impact and the consequences following their use may differ.

As PI has underlined in the past:

The term 'Artificial Intelligence (AI)' is used to refer to a diverse range of applications and techniques, at different levels of complexity, autonomy and abstraction. This broad usage encompasses machine learning (which makes inferences, predictions and decisions about individuals), domain-specific AI algorithms, fully autonomous and connected objects and even the futuristic idea of an AI 'singularity'. This lack of definitional clarity is a challenge: different types of AI systems raise specific ethical and regulatory issues.[9]

From a conceptual point of view, it is important to consider the following key concepts in AI related debates[10]:

- **Artificial narrow intelligence** is the ability of machines to resemble human capabilities in narrow domains, with different degrees of technical sophistication and autonomy.[11]
- **Artificial general intelligence** is the overarching, and as yet unachieved, goal of a system that displays intelligence across multiple domains, with the ability to learn new skills, and which mimic or even surpass human intelligence. It is theorised that the creation of artificial general intelligence could lead to the 'singularity', or a period of runaway technological growth that profoundly changes human civilisation. This is still, at the very least, decades away, if not entirely implausible.
- **Algorithm** can refer to any instruction, such as computer code, that carries out a set of commands: this is essential to the way computers process data.
- **Automated decision-making** 'generally involves large-scale collection of data by various sensors, data processing by algorithms and subsequently, automatic performance.' It is an efficient means of managing, organising, and analysing large amounts of data, and structuring decision-making accordingly. It may or may not rely on AI, with varying degrees of human involvement. It can make decisions, or generate knowledge or information, that significantly shapes or influences the exercise of human rights.
- **Machine learning** is a popular technique in the field of AI which has gained prominence in recent years. It often uses algorithms trained with vast amounts of data to improve a system's performance at a task over time. Tasks tend to involve making decisions or recognising patterns, with many possible outputs across a range of domains and applications. Many of the technologies commonly referred to as AI today are, strictly speaking, machine learning systems.
- Machine learning is usually classified into these three types: **Supervised, unsupervised, and reinforced learning**. Supervised machine learning forms the majority of AI application today. It seeks to teach the computer to predict an output, assuming that the input data is labelled correctly. Supervised machine learning can either be used to predict a continuous valued output through regression, or a discrete valued output through classification. Unsupervised learning, on the other hand, depends on the computer program to find structure within data, based on particular features. Reinforced learning is the third type, wherein the program is

[9] Privacy International and ARTICLE 19, Report, *op. cit.*, pp. 6-7.

[10] For further information see Privacy International and ARTICLE 19, Report, *op. cit.*, pp. 6-7.

[11] Examples include: chatbots that assist by answering specific questions; Deep Blue, a chess-playing computer developed by IBM which famously beat world chess champion Garry Kasparov in May 1997; or the computer system which defeated the reigning master of the Chinese board game Go in May 2017. Privacy International and ARTICLE 19, Report, *op. cit.*, p. 6.

placed in an environment and must learn how to behave successfully within that environment, based on feedback of successes and failures.

We encourage the Committee to consider revising Section VI on 'Racial biases associated with artificial intelligence' to reflect these distinctions to allow for a better understanding on the potential risks of racial bias into AI technologies deployed by law enforcement authorities.

3. On predictive policing and facial recognition in the Draft General Recommendation 36

All the above terms describe general technologies and methods of treating data and making decisions. Depending on their use and the purpose for which they are employed they may have positive or negative impact on individuals' human rights. On the other hand, 'predictive data analysis' and use of algorithms to predict and combat crime (referred to also as predictive policing), which are listed in the Draft General Recommendation together with AI and automated decision making, refer to a new highly problematic policing method that is developed using one or more of the above-mentioned AI technologies.

There are two ways that profiling can be used by law enforcement. Traditionally, profiling can be used to identify individuals based on specific intelligence. Intelligence-led policing is reactive to a crime that has been committed or an alert that has been issued on a specific person. Data-driven and human processes are usually combined.[12] Predictive policing though refers to proactive, mainly data-driven analysis ('risks analysis') with no crime has been committed, or no alert has been issued on a specific person.[13] The risk of racial discrimination resulting from predictive policing is very high, particularly when the dataset upon which predictive policing is applied (e.g. the crime rate in a particular neighbourhood) is racially biased. We call the Committee to clearly state that any predictive data analysis should never be deployed if it violates the principle of non-discrimination.

Facial recognition technology uses cameras loaded with software to match live footage of people to identify or verify people.[14] Facial recognition technology also may use different techniques. Depending on the technology used – real time facial recognition or other – it may have different impact. As it was already underlined by the Committee facial recognition applications suffer from being predominantly based on white, male datasets.[15] The error margin may significantly differ depending on the type of facial recognition used.[16] In the UK, the Biometrics and Forensics Ethics Group warned that UK police's use of live facial recognition technology has the "potential for biased outputs and biased decision-making on the part of system operators".[17] We call the Committee to clearly state that no policing technology should be deployed if their use violates the principle of non-discrimination.

---

[12] EU FRA, *Preventing unlawful profiling today and in the future: a guide*, Handbook, 2018, p. 18.

[13] *idib*., p. 18.

[14] Privacy International and Liberty, Facial Recognition Explainer, available at: https://privacyinternational.org/feature/2726/police-are-increasingly-using-facial-recognition-cameras-public-spy-us.

[15] Draft General recommendation 36, para. 23.

[16] See for example, Virkam Dodd, "Met police to use facial recognition software at Notting Hill carnival", *The Guardian*, 5 August 2017, available at: https://www.theguardian.com/uk-news/2017/aug/05/met-police-facial-recognition-software-notting-hill-carnival.

[17] Biometrics and Forensics Ethics Group, Interim report, February 2019. Also, the US House Committee on Oversight and Government Reform found that the FBI facial recognition database contains photos of half of US adults without consent, and the algorithm is not only wrong nearly 15% of time, but is also more likely to misidentify black people. Olivia Solon, "Facial recognition database used by FBI is out of control, House committee hears", The Guardian, 27 March 2017, available at: https://www.theguardian.com/technology/2017/mar/27/us-facial-recognition-database-fbi-drivers-licenses-passports.

4. On the use of AI in the judicial system

On the biases in the use of AI in the judicial system, that the Committee briefly refers to in paragraph 24, Privacy International would like to bring to the attention of the committee In the United States, risk assessment software purporting to predict the likelihood of reoffending has been used to aid sentencing decisions since the early 2000s. A 2016 study by the non-profit news organisation ProPublica revealed this software's bias against African-Americans, who are more likely to be given a higher risk score compared with white offenders charged with similar crimes.[18]

Conclusion: Summary of recommendations

AI-driven identification, profiling and automated decision-making may lead to unfair, discriminatory, or biased outcomes. Individuals can be misclassified, misidentified, or judged negatively, and such errors or biases may disproportionately affect certain groups of people. Accurate predictions may reveal sensitive attributes that could be used to discriminate. On the other hand, inaccurate or systematically biased data can feed into profiles, which may lead to biased or discriminatory outcomes.[19]

The Committee concludes in paragraph 23 that 'Algorithms reproduce the inequalities of the real world.' Privacy International invites the Committee to develop further this statement. The risk with AI technologies – preferred term instead of algorithms in this sentence – and engrained biases is that not only they reproduce inequalities but actually replicate and amplify discriminatory impact. For instance, having one police officer expressing racial bias may lead to a certain number of discriminatory cases. Introducing AI technology with a racial bias risks amplifying isolated instances to an enormous scale leading to further exclusion and marginalisation of social groups that have been historically discriminated against, such as LGBTI communities, people of African descent, indigenous peoples, Roma or Jewish communities and others.

Privacy International recommends the Committee on Elimination of Racial Discrimination to consider the following, with regard to the Draft General Recommendation 36:

- To address the separate implications of the use of AI technologies for different law enforcement purposes;
- To consider examining the impact of racial profiling beyond law enforcement;
- To clarify the use of different terms invoked in the general recommendation and distinguish between general technologies, methods of treating data and making decisions, and specific technologies;
- To further reflect on the significant risks of racial discrimination in the use of predictive policing and facial recognition technologies;
- To further explore the specific consequences that racial biases associated with the use of artificial intelligence may have.

[18] Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, "Machine Bias", *ProPublica*, 2016, available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.
[19] Solon Barocas & Andrew D. Selbst, "Big Data's Disparate Impact", 104 *Cal. L. Rev.* 671 (2016).

Privacy International
62 Britton Street, London EC1M 5UY
United Kingdom

Phone +44 (0)20 3422 4321
www.privacyinternational.org
Twitter @privacyint
Instagram @privacyinternational

UK Registered Charity No. 1147471