June 30, 2020

STATEMENT ON PRINCIPLES AND PREREQUISITES FOR THE DEVELOPMENT, EVALUATION AND USE OF UNBIASED FACIAL RECOGNITION TECHNOLOGIES

The ACM U.S. Technology Policy Committee (USTPC) has assessed the present state of facial recognition (FR) technology as applied by government and the private sector. The Committee concludes that, when rigorously evaluated, the technology too often produces results demonstrating clear bias based on ethnic, racial, gender, and other human characteristics recognizable by computer systems. The consequences of such bias, USTPC notes, frequently can and do extend well beyond inconvenience to profound injury, particularly to the lives, livelihoods and fundamental rights of individuals in specific demographic groups, including some of the most vulnerable populations in our society.

Such bias and its effects are scientifically and socially unacceptable.

For both technical and ethical[1] reasons – pending the adoption of appropriately comprehensive law and regulation to govern its use, oversee its application, and mitigate potential harm – ***USTPC urges an immediate suspension of the current and future private and governmental use of FR technologies in all circumstances known or reasonably foreseeable to be prejudicial to established human and legal rights***.

Specifically, USTPC finds that:

- Though powerful today and likely to improve in the future, FR technology is not sufficiently mature and reliable to be safely and fairly utilized without appropriate safeguards against adversely impacting individuals, particularly those in vulnerable populations;

- Their potential to help meet significant societal needs, as well as political and marketplace forces, have driven the adoption of FR systems by government and industry ahead of the development of principles and regulations to reliably assure their consistently appropriate and non-prejudicial use;

[1] The Association for Computing Machinery (ACM) Code of Ethics and Professional Conduct counsels computing professionals to avoid harm, be cognizant of the public good, and thoroughly evaluate the impacts and risks of computing systems before deploying them. While written for ACM members and other computing professionals, these core precepts of the Code also may be employed by policy makers assessing how to effectively regulate the development and use of facial recognition technologies.

ACM U.S. Technology Policy Committee
1701 Pennsylvania Avenue, N.W. #200
Washington, D.C. 20006

+1 202.580.6555
acmpo@acm.org
www.acm.org/public-policy/ustpc

- While FR technology can be benign or beneficial in its application, its use has often compromised fundamental human and legal rights of individuals to privacy, employment, justice and personal liberty;

- Policy makers should thus immediately enjoin the use of FR technology by corporations and governments pending the creation and adoption of legal standards for its accuracy proportional to the potential harm such systems may cause to misidentified or non-identified individuals;

- Universal principles for the accurate and just use of FR technology, and for its principled regulation, must be developed without delay; and

- Relevant standards and regulations must address the accuracy, transparency, governance, risk management, and accountability of FR systems.

To these ends, USTPC offers the following guiding principles:[2]

ACCURACY[3]

- Before an FR system is used to make or support decisions that can seriously adversely affect the human and legal rights of individuals, the magnitude and effects of such system's initial and dynamic biases and inaccuracies must be fully understood.

- As the impact of each type of error is context dependent, context must be expressly identified and addressed in standards that set legally "acceptable" error rates.

- When error rates are reported, they must be disaggregated by sex, race, and other context-dependent demographic features, as appropriate.

- The accuracy of every FR system must be fully auditable over time to support third-party monitoring and robust government oversight.

---

[2] Primary contributors to this Statement and its constituent principles were USTPC Chair Jim Hendler, Vice Chair Alec Yasinsac, and Committee members Ricardo Baeza-Yates, Jeremy Epstein, Simson Garfinkel, Arnon Rosenthal, and Stuart Shapiro.

[3] The Committee also urges that practices, policies, rules and statutes governing the development and deployment of all FR technology be consistent with its *Statement on Algorithmic Transparency and Accountability* and *Statement on the Importance of Preserving Personal Privacy*. The former highlights the need to understand the consequences of software errors. (In particular, it warns of potential biases in training data and resulting potentially discriminatory harms.) The latter underscores the importance of ensuring appropriate data quality, particularly its accuracy.

TRANSPARENCY

- An FR system should be activated only after some form of meaningful advance public notice of the intention to deploy it is provided and, once activated, ongoing public notice that it is in use should be provided at the point of use or online, as practicable and contextually appropriate.

- Such notices should at minimum contain:

  - A description of the data used to develop and train the FR algorithm;

  - Sufficient detail about the algorithm's implementation so that experts can understand its performance characteristics, accuracy, and limitations;

  - A report of the algorithm's performance relative to a standardized benchmark; and

  - A clear statement of:

    - How the algorithm will be used (including particularly its role in any decisions affecting individuals and whether those decisions are to be taken automatically or by a human supported by the FR technology); and

    - The role that humans will play in application of the system.

GOVERNANCE

- No FR system should be deployed prior to establishing appropriate policies governing its use and the management of data collected by the system.

- All such policies, to the maximum extent possible, should be subject to public input, scrutiny, and oversight.

- Data retention policies and practices should be legally compliant, transparent to the public, and limit data retention to what is strictly necessary for the specific purpose for which the data was collected.

- Systems should be designed to minimize the quantity and richness of any data retained.

- FR system governance mechanisms should pay particular attention to the risks posed to, and consequently necessary protections for, vulnerable individuals and populations; the more significant the potential harm, the stricter risk management protocols should be.

RISK MANAGEMENT

- The benefits of deploying a given FR system to the deploying organization, the public, and to vulnerable subgroups should be proportional to the risks posed by its use.

- Key risks to consider must include those related to security, privacy, and safety in general, as well as to the potential for negative and discriminatory practical, legal and public policy impacts on protected and vulnerable populations.

- Organizations that use FR should empower and enable an appropriately constituted advisory board, similar to a Civilian Oversight Board or an academic Institutional Review Board, to assess whether a proposed FR system can be employed ethically before approving or deploying it for a particular proposed use.

- Such reviews should include a proactive, multi-factor impact assessment and risk analysis.

- No FR system should be made available or deployed unless its relevant material risks to vulnerable populations, or to society as a whole, can be sufficiently eliminated or remediated.

ACCOUNTABILITY

- Developers, operators, and users of any FR system must be accountable within their organization and to external stakeholders for the consequences of such systems' use and misapplication.

- When harm results from the use of such systems, the organization, institution, or agency responsible for its deployment must be fully accountable under law for all resulting external risks and harms.