Request for Information (RFI) on Public and Private Sector Uses of Biometric Technologies: Responses

DISCLAIMER: Please note that the RFI public responses received and posted do not represent the views or opinions of the U.S. Government, the Office of Science and Technology Policy (OSTP), or any other Federal agencies or government entities. We bear no responsibility for the accuracy, legality, or content of these responses and the external links included in this document. Additionally, OSTP requested that submissions be limited to 10 pages or less. For submissions that exceeded that length, the posted responses include the components of the response that began before the 10-page limit.
January 14, 2022

Via Electronic Submission to

Office of Science and Technology Policy (OSTP)
Executive Office of the President
Eisenhower Executive Office Building
1650 Pennsylvania Avenue
Washington, D.C. 20504

Re: RFI Response: Biometric Technologies

Dear Members of the Office of Science and Technology Policy:

I, Orissa Rose, graduate student at University of California, Berkeley am pleased to submit the following comments on Public and Private Sector Uses of Biometric Technologies. As a candidate for a Masters of Information Management and Systems from Berkeley’s School of Information, my work focuses on expanding access to data and improving its usability, reliability, and credibility while preserving human rights, security and privacy. My research on information policy investigates the role public institutions like you play in mitigating algorithmic harms and establishing data standards that protect citizens, especially the most marginalized by rapid technology developments. These comments do not address the entirety of RFI sections available, but rather, focus on Section 6 of the RFI: Governance programs, practices or procedures.

I welcome any opportunity to continue this conversation, provide future resources or answer questions. Although the undersigned is a member of UC Berkeley, I am signing as an individual and do not represent this organization.

In appreciation,

Orissa Rose
Masters Student and Technology Consultant
**Section reference:**
General comments on data collection ethics.

**Recommendation:**
Prioritize investigation into disparate impacts on the most vulnerable communities throughout your investigations of biometric data use. These communities are most likely to be negatively impacted by discriminatory algorithms and data treatments, and least likely to have their interests advocated for and funded.

**Comment:**
A common mistake of our times is the assumption that when we hand activities off from human actors onto technical systems, things will go smoothly so long as the components we swap are modular in their functional scope, and we have accounted for errors with math.¹ The growing popularity of using AI built on biometric data to identify people, surveil activity, or assess human emotion and health has gained momentum on the notion that employing technical systems to do this work will generally leave ethical and political dimensions intact. We rationalize that these tools and uses of data don’t cause more harm than good, and we believe these narratives because the system developers claim to have attended to bias in their models, or because we are aware of error rates in human computing, or because we do not yet have evidence to counter the claims.

Recent literature on the failings of pretrial risk assessment tools like COMPAS² can be a useful and cautionary resource to your team about the pitfalls of reliance on subjective datasets to make serious determinations about a person's future. Legal experts and technology scholars have consistently demonstrated how risk assessment tools -- which purport to leverage data to predict a criminal’s pretrial custody options, appropriate length of parole, and more -- are more truly tools to predict and map trends in policing behavior. The use of biometric technologies to monitor behavior and predict emotional state or physical health is not too dissimilar. Just as risk recidivism tools use partial datasets to advance prejudice and racialized decisions, biometric technologies that are not built with the explicit intention to mitigate known and unknown harms could result in worsening occurrences of undue prosecution, misidentification of innocent people, and entrenchment of medical racism and elderly abuse.

The FTC’s 2016 report on data collection abetted by IoT devices details the disparate impact these technologies have on disadvantaged communities. Inaccurate or biased analyses of biometric data collected by a range of IoT devices can lead to consumers being denied opportunities for education, employment, healthcare or credit.³ The imperative of our federal agencies to protect equal opportunity and fair treatment compel OSTP to prioritize these most vulnerable communities in your investigations of biometric data collection and standards.

---

¹ (Nissenbaum & Mulligan, 2020)
³ (Federal Trade Commission's Bureau of Consumer Protection and Office of Policy Planning, 2016)
**Section Reference:**
6a) Stakeholder engagement practices for systems design, procurement, ethical deliberations, approval of use, human or civil rights frameworks, assessments, or strategies, to mitigate the potential harm or risk of biometric technologies.

**Recommendation:**
Resist the urge to convey the use of biometric data for healthcare and social welfare programs as a panacea to industry bias and inefficiency. Advocate for biometric data collection standards that are based on human rights principles. Additionally, evaluating proposed biometric technologies against the five “abstraction traps” could help OSTP and your partners prioritize technologies that place human rights over profit incentives.

**Comment:**
The language in OSTP’s RFI demonstrates your optimism that biometric technologies can solve many societal and institutional shortcomings. While I acknowledge and celebrate the important ways technological developments have contributed to incredible achievements in medicine, education and social connection, I am also compelled to warn OSTP against the five common traps that indiscriminate support of AI can create. Machine learning scholars from Microsoft, Princeton, and the Data and Society Research Institute identified five “Abstraction Traps’ that result from failing to properly account for or understand the interactions between technical systems and social worlds”. This framing can support policymakers’ and technology professionals’ essential considerations before proceeding with adoption and implementation of new machine learning systems”.  

1. Framing Trap - Failure to model the entire system over which a social criterion, such as fairness, will be enforced.
2. Portability Trap - Failure to understand how repurposing algorithmic solutions designed for one social context may be misleading, inaccurate, or otherwise do harm when applied to a different context.
3. Formalism Trap - Failure to account for the full meaning of social concepts such as fairness, which can be procedural, contextual, and contestable, and cannot be resolved through mathematical formalisms.
4. Ripple Effect Trap - Failure to understand how the insertion of technology into an existing social system changes the behaviors and embedded values of the pre-existing system.
5. Solutionism Trap - Failure to recognize the possibility that the best solution to a problem may not involve technology.

Analyses of marketing trends amongst leading biometric technology companies found that the arguments they put forward skillfully position the deregulation of biometric data as essential to the advancement of medical science. Many of these parties “assumed that patients want ‘their’ data to be used and prioritized long-term use of patient data above actual patient expectations or consent.” In practice, and as evidenced by the lawsuits, recent media firestorms, and “levels of indignation [from newly informed citizens] about perfectly lawful data practices”, we know that

---

4 (Selbst, Boyd, Friedler, Venkatasubramanian, & Vertesi, 2019)  
5 (Felt & Starkbaum, 2019)
that status quo needs to change. A recent investigation into facial recognition programs at large tech companies revealed how the lack of public knowledge about biometric data collection and use has “allowed technology companies to create a climate favorable to their use of customers' biometric data, largely without their knowledge or consent.” To support data collection standards that prioritize the data subjects and are especially sensitive to vulnerable communities, these principles of designing for social justice will be useful to incorporate in your investigation:

- Design for Transformation - A long-term approach focusing on the role of structural inequalities; designing for the evolving, emergent social, economic, and political relations that produce inequalities and perpetuate social injustices.
- Design for recognition - Focuses on identifying unjust practices, policies, laws, and other phenomena that create data realities, as well as identifying those people who are most negatively impacted by such phenomenon.
- Design for Reciprocity - Focuses on relationships and the ways in which those relationships may need to change to become more equitable for all stakeholders; the relationship between those who are owed justice and what needs to occur for the obligations of justice to be fulfilled.
- Design for Enablement - a multi-level focus on developing opportunities for change, including scaffolding individual behavior change as well as the practices and policies related to fostering structural change.
- Design for Ethics - Contending with the individual's immediate needs, and attempts to address the larger context and imbalanced power relations that produce the potential for and actual oppression.

Section Reference
6c) Practices regarding data collection, review, management, storage, and monitoring practices.

Recommendation:
Leverage the maturity and effectiveness of existing data collection and storage standards from GDPR, CCPA and existing state biometric legislation (Illinois, Oregon, Texas, Washington, City of New York) to advance proven practices whose path to compliance is already known.

Comment:
The benefit of OSTP supporting data collection articles similar to the GDPR, CCPA and other existing biometric policies is that many large companies and international researchers have already had to build out their systems to achieve compliance with those standards, and they are proven effective. The GDPR articles most apt for consideration in OSTP’s biometric data collection investigations are:

6 (Nissenbaum, 2015)
7 (Pope, 2018 discussing April Glaser’s investigation into Facebook’s collection of the facial data of millions of its users and is now the biggest lobbying force against biometric data privacy laws).
8 Including: Illinois' Biometric Information Privacy Act (BIPA); City of New York Administrative Code, Title 22, Chapter 12; Portland City Code, Title 34- Digital Justice, Washington Rev. Code §§ 19.375.010 et seq
- **Responsibility of the controller** ([Article 24](#)): Controller (data collector) shall implement appropriate technical and organizational measures to ensure and to be able to demonstrate that data processing and protection is performed in accordance with all regulations.

- **Data minimization** ([Article 25](#)): Defaults data collection to only the data “which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility”.

- **Data protection** ([Article 25](#)): Data controllers must employ practices like pseudonymization (replacing identifiable fields within a data record with pseudonyms or dummy values), to protect personal data, (and sensitive data categories) and must certify compliance with this requirement.

**Section Reference**
6c) Practices regarding data collection, review, management, storage, and monitoring practices.

**Recommendation:**
Standardize compliance with conscientious data collection tools to help biometric data companies examine their positionality and the ethical implications of their methods and findings. Develop sensitivity ratings for different types of biometric data. In an effort to minimize paperwork, focus dataset evaluations on biometric metadata that is most sensitive and influential to care, treatment or sentencing decisions. Consider an online survey or form like HEVAC[^9] that requires participation in 70% of recommended activities when handling these highly sensitive data categories to receive a passing mark. Use this as a gatekeeping mechanism for any program seeking federal funding, partnership or approvals.

**Comment:**
Reflexivity is a practice of critically assessing “the significance of environment, power, and context as well as subjectivity in the delineation and construction of knowledge”.[^10] Without reflexivity, biometric data researchers, private companies and public agencies may unwittingly impose the values, prejudices and practices of their specific culture onto the data and people their tools are working on. Prominent technology scholars and data ethicists have developed tools that data collectors, engineers, and cross-disciplinary teams can use to get clear on the history, context, limitations and ethical implications of the datasets they’re working with.

- **Datasheets for Datasets:** Require development of datasheets for any public agency or private partner working with biometric data. In 2018, researchers from Microsoft, Georgia Institute of Technology, Cornell University, University of Maryland, and the Now Institute

[^9]: Developed by the Higher Education Information Security Council, the HECVAT is a questionnaire specifically designed to measure vendor risk. Before institutions partner with third-parties, the partner must complete a HECVAT tool to confirm that information, data, and cybersecurity policies are in place to protect sensitive information and constituents’ PII.

[^10]: (Ryan & Walsh, 2018)
introduced the idea of a narrative “datasheet” to accompany any dataset. Recognizing that “the risk of unintentional misuse of datasets increases when developers are not domain experts”, or are using data collected by third parties, and hoping to “initiate a broader conversation about data provenance, ethics, privacy, and documentation”, the team developed questions that prompt users to describe dataset’s creation, their strengths, and their limitations. Sections of the Datasheet tool prompt reflection on: motivation for dataset creation; dataset composition; data collection processes; data preprocessing; dataset distribution; dataset maintenance; and legal and ethical considerations. The outcome of datasheet creation is better and more appropriate use of data, and the intentional consideration of the implications and logic models the dataset embodies. Datasheet questions that are particularly important to the ethical collection and use of biometric data, plus my commentary, are as follows:

○ What preprocessing/cleaning was done? (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values, etc.) Judicious and non-discriminatory data cleaning is an essential step in making sure models working off these datasets do not compound discriminatory or sensitive characterizations of people.

○ Were there any ethical review applications/reviews/approvals? (e.g. Institutional Review Board applications) Beyond protecting citizens, obtaining ethical review approvals helps mitigate legal fillings.

○ Were data subjects told what the dataset would be used for and did they consent? What community norms exist for data collected from human communications? Understanding and respecting community norms protects against skewed data resulting from collection under duress or discomfort.

○ Could this dataset expose people to harm or legal action? (e.g., financial, social or otherwise)? What was done to mitigate or reduce the potential for harm?

○ Does it unfairly advantage or disadvantage a particular social group? In what ways? How was this mitigated? Explore what funding resources could be made available to subsidize harm mitigation efforts for qualifying partners.

● **Data User Guides** are another useful tool whose utilization OSTP would be wise to promote across public and private biometric technology organizations.\(^\text{11}\) Similar to Datasheets for Datasets, a Data User Guide is a narrative document prepared by data controllers and engineers that addresses key issues anyone working with the data should be aware of.

---
\(^{11}\) (Data User Guides, 2015)
Bibliography of References & Recommended Reading


Data user guides. (2015). *Data-User-Guides/README.md*


