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.
Office of Science and Technology Policy

In response to Document Number: 2021-21975

Request for Information for Public and Private Sector Uses of Biometric Technologies

January 12, 2022

Submitted to:
Office of Science and Technology Policy
BiometricRFI@ostp.eop.gov

Submitted by:
Tracy Hulver | Senior Director, Product Management
Aware, Inc.
40 Middlesex Turnpike
Bedford, MA 01730-1404
www.aware.com
Public and Private Sector Uses of Biometric Technologies

**Scope:** OSTP invites input from any interested stakeholders, including industry and industry association groups; civil society and advocacy groups; state, local, and tribal governments; academic researchers; technical practitioners specializing in AI and biometrics; and the general public. In particular, OSTP is especially interested in input from parties developing biometric technologies, parties acquiring and using such technologies, and communities impacted by their use. Input is welcome from stakeholders, including members of the public, representing all backgrounds and perspectives.

**Information Requested:** Respondents may provide information for one or as many topics below as they choose. Through this RFI, OSTP seeks information on the use of biometric technologies in the public and private sectors, including on the following topics:

Please see Aware, Inc.’s responses below.

1. **Descriptions of use of biometric information for recognition and inference:** Information about planned, developed, or deployed uses of biometric information, including where possible any relevant dimensions of the context in which the information is being used or may be used, any stated goals of use, the nature and source of the data used, the deployment status (e.g., past, current, or planned deployment) and, if applicable, the impacted communities.

Aware allows for the capture, matching, and verification of biometric data. This includes fingerprint, facial, iris and voice. Aware customers typically want to utilize biometric data for two primary purposes: 1) capturing a biometric for matching; and 2) capturing a biometric for verification.

Our primary markets are law and policy enforcement organizations and entities that want to use biometric capture for end-user verification applications such as multifactor authentication (MFA). Biometrics are a desirable form for MFA due to the ease-of-use, speed, fidelity, and accuracy of biometric capture and matching. Aware believes that giving the user control of their biometric data is a preferred method of deploying biometric solutions. MFA is a proven way of increasing the security of an ecosystem, such as access to financial information, healthcare data, and other sensitive information. Giving the user control of how and where their biometric information is being stored and used increases privacy of the user as well as providing a strong method of securing transactions.
2. **Procedures for and results of data-driven and scientific validation of biometric technologies:** Information about planned or in-use validation procedures and resulting validation outcomes for biometric technologies designed to ensure that the system outcomes are scientifically valid, including specific measures of validity and accuracy, resulting error rates, and descriptions of the specific measurement setup and data used for validation. Information on user experience research, impact assessment, or other evaluation of the efficacy of biometric technologies when deployed in a specific societal context is also welcome.

There are two primary settings in which validation procedures and performance metrics are relevant: 1) in R&D, and 2) in deployment. In R&D, of critical importance are the following:

1. Devising a test set that is independent of the data used to train and/or develop biometric algorithms, in order to ensure that algorithms generalize to real world operational settings
2. Persisting that test set to provide a way to benchmark algorithm improvements
3. Measuring performance metrics that can accurately assess the relevant operational characteristics of a particular biometric technology

Performance metrics can vary depending on the type of biometric algorithm.

The following table details the primary performance metrics used by R&D for various biometric algorithms.

<table>
<thead>
<tr>
<th>Biometric Algorithm</th>
<th>Relevant Performance Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liveness</td>
<td>Attack Presentation Classif. (Spoof Detection) Error Rate [APCER] vs. Bonfide Presentation Classif. (Live) Error Rate [BPCER]</td>
</tr>
<tr>
<td></td>
<td>APCER vs. BPCER as a function of spoof type</td>
</tr>
<tr>
<td></td>
<td>BPCER at fixed APCER as a function of race (racial bias)</td>
</tr>
<tr>
<td></td>
<td>BPCER at fixed APCER as a function of gender (gender bias)</td>
</tr>
<tr>
<td>Verification (all biometrics)</td>
<td>False Non-Match Rate (FNMR) vs. False Match Rate (FMR)</td>
</tr>
<tr>
<td></td>
<td>Failure to Enroll Rate (frequency with which the algorithm cannot generate a template for verification)</td>
</tr>
<tr>
<td></td>
<td>FNMR vs. FMR as a function of image quality</td>
</tr>
<tr>
<td></td>
<td>FNMR at fixed FMR as a function of race (racial bias)</td>
</tr>
<tr>
<td></td>
<td>FNMR at fixed FMR as a function of gender (gender bias)</td>
</tr>
<tr>
<td>Identification (all biometrics)</td>
<td>False Negative Identification Rate (FNIR) vs. False Positive Identification Rate (FPIR)</td>
</tr>
<tr>
<td></td>
<td>Failure to Enroll Rate (frequency with which the algorithm cannot generate a template for identification)</td>
</tr>
<tr>
<td></td>
<td>FNIR vs. FPIR as a function of image quality</td>
</tr>
<tr>
<td></td>
<td>FNIR at fixed FPIR as a function of race (racial bias)</td>
</tr>
<tr>
<td></td>
<td>FNIR at fixed FPIR as a function of gender (gender bias)</td>
</tr>
<tr>
<td>Biometric Sample Quality Assessment</td>
<td>Regression/Classification accuracy based on groundtruthed samples</td>
</tr>
</tbody>
</table>

*Table 1: Performance metrics for biometric technologies*
Clearly, the same performance metrics are relevant in both R&D and deployment settings, but vendor access to operational metrics is severely curtailed after deployment, primarily due to privacy restrictions and client security. In deployment, metrics are often indirect measurements of biometric algorithm performance that are computed by the client, and typically reflect the impact they have on their business use case. For example, usability (failure to enroll) is often an important operational consideration at operating points that satisfy minimum client security requirements.

A supporting infrastructure that enables collection of performance data is vital to providing clients with the capability to monitor operational system performance. As an example, Aware provides a reporting mechanism on backend servers that enables clients to observe incoming data streams and to review the results of algorithm measurements. This can be used to adjust thresholds or operational configurations to optimize system performance. In addition, data streams can be anonymized and provided to Aware to further improve and enhance algorithm performance in cases where a collaborative relationship exists with a deploying partner.

3. **Security considerations associated with a particular biometric technology.** Information about validation of the security of a biometric technology, or known vulnerabilities (such as spoofing or access breaches). Information on exhibited or potential leaks of personally identifying information via the exploitation of the biometric technology, its vulnerabilities, or changes to the context in which it is used. Information on security safeguards that have been proven to be efficacious for stakeholders including industry, researchers, end users, and impacted communities.

Although this question specifically refers to security, we see security and privacy as two intimately connected, but separate and critically important factors in promoting the widespread use and adoption of biometrics. For the purpose of this discussion, we define the two as follows. Security in biometrics can be considered from two perspectives: security of a particular biometric in transactions as it relates to error rates, and the security of the biometric in transit or at rest, as it relates to its possible compromise. Privacy in biometrics relates to concealing the connection between an individual’s identity and their biometrics, regardless of the success or failure of any security measures in place. A solution that provides complete privacy for a biometric obviates (or at least reduces) the need for its security.

Security in biometrics as it relates to error rates, focuses on false match rate (FMR) in the case of verification, false positive identification rate (FPIR) in the case of identification, and attack presentation classification error rate (APCER) in the case of liveness. Willful circumvention of biometric technologies is becoming increasingly difficult as matching and liveness technologies continue to improve, enabling systems to operate at lower and lower error rates. However, these must be implemented in smart workflows that lowers the statistical risk of a single malicious user exploiting potential vulnerabilities. Biometric solutions are highly robust, but they are ultimately reliant on the statistical improbability of a single user...
being able to randomly match a lookalike (in the case of verification or identification) or to spoof a liveness solution with multiple, creative attempts. For example, limiting the number of allowable failed attempts to some small number (e.g., 1 or 2) limits the ability of a user to learn and take potential advantage of a system’s vulnerabilities.

There are also multiple points of potential failure in the security of a biometric system: at the point of biometric sample creation, in storage and during transmission. As an example, Aware’s liveness solution provides several mechanisms during mobile image acquisition to bind the image capture to the real-time transaction, to eliminate the possibility of emulator injection attacks during image acquisition (biometric sample creation). Sophisticated watermarking techniques then guarantee the integrity of the biometric sample storage to prevent replay attacks. Finally, end-to-end encryption protects against tampering or substitution during transmission. All three points of failure must be addressed to provide secure and tamper proof execution.

Regarding privacy, recent advances in crypto-biometric authentication approaches are enabling the use of biometrics as unique keys in zero-knowledge proofs for authentication. A user uses their own biometric (which must be proven to be live) to unlock a secret known only to the interacting institution to prove their identity without giving up their own biometric. Privacy is enhanced because no biometric ever needs to leave the device of the user and no central repository of biometric templates is required to authenticate. Obviating the need for a central database also enhances the security of the authentication system because there is no biometric template to steal.

4. **Exhibited and potential harms of a particular biometric technology:** Consider harms including but not limited to: Harms due to questions about the validity of the science used in the system to generate the biometric data or due to questions about the inference process; harms due to disparities in effectiveness of the system for different demographic groups; harms due to limiting access to equal opportunity, as a pretext for selective profiling, or as a form of harassment; harms due to the technology being built for use in a specific context and then deployed in another context or used contrary to product specifications; or harms due to a lack of privacy and the surveillance infrastructure associated with the use of the system. Information on evidence of harm (in the case of an exhibited harm) or projections, research, or relevant historical evidence (in the case of potential harms) is also welcome.

Like any technology, biometrics can be used for ill or good, depending on the application. From our (vendor) perspective, the harms we consider are related to harms we can mitigate through the development and control of our own technology. For example, gender and race bias have recently been the focus of concerns regarding the use of face recognition algorithms. Aware has made a great effort and taken significant strides to develop training and scoring methodologies for our matchers that minimize race and gender bias. At a time when training data is becoming scarcer due to privacy concerns, it is more important now...
than ever to develop training methodologies that are not at the mercy of the data being trained on. Blindly relying on data without a true understanding of the effect of gender and race on facial recognition and liveness is a mistake that is often made by fledgling biometric companies.

5. **Exhibited and potential benefits of a particular biometric technology:** Consider benefits including, but not limited to: Benefits arising from use in a specific domain (absolute benefit); benefits arising from using a specific modality of biometric technology combination thereof) compared to other modalities in a specific domain (relative benefit); and/or benefits arising from cost, consistency, and reliability improvements. Information on evidence of benefit (in the case of an exhibited benefit) or projections, research or relevant historical evidence (in the case of potential benefit) is also welcome.

The value of biometric fusion is often underestimated or discounted due to the cost of implementation. Multiple biometric data collection requires multiple sensors, and for that reason is often rationalized away. However, the fusion of face and voice is a powerful combination that is frequently overlooked. The technologies are easily integrated on mobile phones and can provide an added level of security not possible by either biometric alone. Further, they allow the straightforward implementation of both authentication and liveness. Score level fusion allows increased accuracy for both verification and liveness, with redundancy in biometrics if one modality is unusable due to a sub-optimal capture environment. Again, a good understanding of fusion and score mapping in particular is required to extract full advantage from such an integration. Aware’s historical expertise in biometric fusion has enabled it to do this, as well as to apply that expertise to its other multi-biometric offerings.

6. **Governance programs, practices or procedures applicable to the context, scope, and data use of a specific use case:** Information regarding:

a. 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;

As a publicly traded company, Aware is bound to certain ethical and business requirements that we not only follow, but embrace. In addition, because we are a global organization with customers all over the world across many different vertical markets, we abide by numerous business, security, and privacy best practices to meet both governmental and customer-required compliance and regulatory guidelines and mandates. Some of these include best practices on privacy (i.e., GDPR, NIST Privacy Framework, California Consumer Privacy Act, etc.) and security (i.e., NIST 800-53, SOC II, FedRAMP, etc.).
b. Best practices or insights regarding the design and execution of pilots or trials to inform further policy developments;

It is recommended that prior to an organization deploying a biometric verification system they clearly identify the biometrics and explain how the data will be captured, stored, and used. The production of a Privacy Impact Analysis is also created and made available to all parties that will be utilizing the biometric system. Next, the organization should plan on having 3 phases during the deployment 1) Proof-of-Concept (POC), 2) limited pilot, and 3) full production.

POC: During the POC, a limited number of users will be asked to interact with the biometric system and test the usability and accuracy of the biometric capture and matching. The size of the POC will vary depending on the final size and demographic makeup of the deployment but typically, limiting the POC to 5-100 users is desirable. The length of the POC will also vary depending on the availability of the users testing the system but 30-60 days should be a reasonable target. The POC should be limited to a user base that is internal to the organization of to external users who understand that they are using a non-production system. A clear set of agreed upon use cases need to be identified and what success criteria need to be met in order to move into the pilot phase. Items that require testing should not only focus on the workflow of the biometric system but also the various environmental variables that need to be tested such as, for face capture and verification different lighting environments (i.e., bright light, low light, reflected sunlight, direct sunlight, etc.) For voice capture and verification different background noise levels need to be analyzed (i.e., no background noise, windy conditions, background voice/music, talking through a mask, etc.). Adjustments to the sensitivity of the matching and liveness scores should be available for fine tuning of the biometric system to achieve the proper balance of ease-of-use and security.

Pilot: Moving into the pilot phase will expand the user base access to the system as well as expanding the overall user cases identified. Typically, a pilot will include external users accessing production systems but will limit the user community to a specific user case or cases. An example of this would be targeting a specific demographic group or geolocation for the pilot. The size of the pilot should include a representative sample of the final user community and therefore will vary greatly as to the number of users, but a rule of thumb is to target a pilot size of 10%-15% of the final deployment size. Obviously for large deployments a pilot may be accomplished in stages to ensure a manageable size of users during the project roll-out. The length of the pilot phase will vary depending on the size of the user community being included as well as the smoothness of the pilot deployment, but a timeframe of 3-9 months should be planned.

Production: Finally, after confidence in the system is achieved, full deployment can proceed.
d. Safeguards or limitations regarding approved use (including policy and technical safeguards), and mechanisms for preventing unapproved use;

A detailed analysis of the requirements for data collection, storage, and use will be needed before safeguards can be implemented. The requirements will vary greatly depending upon the use cases. An example of this is that use of biometric data within a closed ecosystem such as employment enrollment and verification will be much more controllable by the organization than implementing a biometric authentication and verification system for citizens and consumers. One variable that will not change is the requirement for the ecosystem to follow best practices on privacy (i.e., GDPR, NIST Privacy Framework, California Consumer Privacy Act, etc.) and security (i.e., NIST 800-53, SOC II, FedRAMP, etc.)

These security and privacy controls will help establish best practices that the vendors as well as the organizations deploying the solution can implement and follow. Of course, local and federal laws will also have to followed where applicable.

e. Performance auditing and post-deployment impact assessment (including benefits relative to current benchmarks and harms);

A supporting infrastructure that enables collection of performance data is vital to providing clients with the capability to monitor operational system performance. As an example, Aware provides a reporting mechanism on backend servers that enables clients to observe incoming data streams and to review the results of algorithm measurements. This can be used to adjust thresholds or operational configurations to optimize system performance. In addition, data streams can be anonymized and provided to Aware to further improve and enhance algorithm performance in cases where a collaborative relationship exists with a deploying partner.