

SUPPORTING STATEMENT FOR

Office of Biometric Identity Management Biometric Data Collection

OMB Control No.: 1601-NEW

COLLECTION INSTRUMENT(S):

DHS OBIM will not conduct any of the biometric collections itself. OBIM has agreements with performers, like JHU APL, NIST, and DHS S&T, and will be responsible for collecting relevant biometrics to answer questions about biometrics that OBIM has. These groups will use collection devices that they have access to or that DHS components use for biometric data collection. In general, the biometric devices used in the studies produce images like those taken by cameras to produce images of various biometric modalities. The studies anticipate using different biometric equipment for the collection of contact fingerprints, palm, face, and iris. Contact fingerprint and palm will be similar to ink collections, and voice collections will utilize microphone/cellular phone recording devices. There are no anticipated adverse outcomes from normal use of any of the fingerprint, face, palm, voice, and iris biometric devices.

All commercial devices used in the study will have passed the usual tests for electronic devices offered for commercial sale and should not present any generic risks above those of any other electronic device (e.g., radio, TV, camera, etc.) that might be purchased by the general public. Any contactless prototype devices will be reviewed by a device safety committee. The prototype devices will be evaluated to identify safety concerns related to their use. Safety assessments of the devices will consider the physical construction, mechanical operation, electrical system, and any other aspect with the potential to cause harm. Optical components will be evaluated to verify that the devices are eye safe. Only the prototype devices approved by the device safety committee will be added to the study.

Equipment/Device Details:

Examples of the Biometric data collection equipment are subsequently listed and include, but are not limited to, the following vendor equipment: Contact Fingerprint collection equipment (e.g., Crossmatch, Kojak, Grabba, Morpho, i3, NeoScan45, Javelin XL, BioSled, Seek, Crossmatch SEEK II, MOSAIC MCR, U.are.U.5300), Contactless Fingerprint Collection Equipment (e.g., iPhone camera, Android camera, Idemia), Face Matching Collection software devices (e.g., iPhone camera, android camera, Logitech, Axis, Iris ID, Idemia, Crossmatch SEEK II, Canon EOS, Dell Ultrasharp), Iris Collection devices (e.g., Javelin XL), Voice collection devices (e.g., computer microphone, cellular phone microphone).

B. Collection of Information Employment Statistical Methods

The agency should be prepared to justify its decision not to use statistical methods in any case where such methods might reduce burden or improve accuracy of results. When Item 17 on the Form OMB 83-1 is checked "Yes", the following documentation should be included in the Supporting Statement to the extent that it applies to the methods proposed:

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used. Data on the number of entities (e.g., establishments,

State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

Because biometrics are made up of a variety of physical, behavioral, and physiological characteristics, the breadth of types of biometric characteristics that could be collected to answer relevant OBIM biometric questions is wide. Thus, the OBIM OMB generic submission for approval to collect biometrics is broad and will include different biometrics depending on the need of DHS components and what current threats to homeland security are. To speak to the possible data analysis that can happen as part of this generic submission, OBIM Futures Identity (FI) will lay out potential groups of people and types of biometric collection that could be sampled from to answer relevant DHS questions. In OBIM FI's generic OMB submission, OBIM FI listed 1000 subjects as the average number of subjects per year OBIM plans to sample. With this number in mind, and the variety of groups and biometrics that may need to be sampled to answer relevant questions, OBIM FI has put together a table below with the total of subjects and the number of subjects within each category based upon potential projects that may be studied by performers.

Without knowing the exact study design at this current time, the numbers of subjects in total as well as in each category below are rough estimates based on the idea that at baseline. One approach to the analyses will involve a simple one-way (1X5) ANOVA such that 5 groups within one independent variable category (based on the below categories) and the average of these groups will be compared to each other based upon one dependent measure/outcome of interest. Using a very simple independent sample ANOVA power table, OBIM FI suggests that for an effect size of 0.1 with an alpha set at 0.01 and a power of 0.99 with 5 independent groups, the sample size of individuals for a biometric project would be, at minimum, 700 participants. If the complexity of the question and/or project increase, such that more groups are needed or more complex analysis is desired, the sample size would have to increase to have findings that are somewhat applicable to the public.

Participants, both staff and the general public, will be recruited. Participants will be recruited via email announcement. For staff, the email will be circulated from the Office of Human Research Protection. Interested participants will be instructed to contact a member of the research team for additional information about the study and will be given a copy of the consent form to review prior to the test date. On the test date, an overview of the study will be provided, and participants will have the opportunity to have their questions answered. Once informed consent has been obtained, participants will be asked to provide basic demographic information (age, gender, and race). Each participant will be assigned a random study ID number under which their biometric and minimal biometric information will be associated. This will not include sensitive PII information, only age, gender, and race will be stored.

Participants will be asked the following questions:

- What is your age in years?
- What is your gender?
- What is your race?
- Participants may be asked to fill out any additional questionnaires/surveys.

- o Participants are not obligated to answer any questions they do not want to answer.
- o Additional questions may involve specifics to the collection at hand. For example:
 - Voice biometric collections may require questions to be asked about how the individual is feeling and whether they feel their voice is being impacted by internal or external factors.
 - Iris collections will require questions on whether the person wears glasses or contact lenses on a regular basis.
 - Face collections will require discussions on the frequency of make-up use.
 - Fingerprint and palm collections will involve questions associated with general work field information. This can include:
 - Have you ever had any previous injuries that may have impacted your fingerprints?
 - o If yes, please identify (circle) which hand/which fingers.
 - o Do you have sensitivity to COTS hand lotion?

2. Describe the procedures for the collection of information including:

- Statistical methodology for stratification and sample selection,

Academic and national lab performers will be leading the charge in stratification and sample selection for individuals in the study as well as the appropriate data analysis that needs to be conducted to answer relevant OBIM questions. These institutions have data scientists and statisticians with advanced mathematical degrees that have years of experience understanding the appropriate sample sizes needed to answer large questions and the best methods for stratification and sample selection and ultimately data analysis that answers the question writ large. OBIM will not be directly involved in, nor will OBIM be responsible for, advising data scientists on sampling procedure or statistical methods.

- Estimation procedure

OBIM's National Lab and academic university performers will be responsible for estimating sample sizes overall and within subgroups so that project will give the best numbers of data to gather information on the best possible analysis and scalable outcomes to answer relevant questions, without engaging in unnecessary over sampling or recruiting too many participants for what is needed.

There are two commonly used methods for determining the sample size: **confidence intervals (CI)** and **statistical hypothesis tests**. CIs are estimates of the true underlying population failure rate that combine the test data with the binomial distribution to produce a range of values that are deemed to be plausible values for failure rate. Statistical hypothesis testing is a more formal method for evaluating the failure rate. Two opposite and competing hypotheses are determined prior to data collection, and the data are used to evaluate which hypothesis is more plausible. Both CIs and statistical hypothesis tests can be used for sample (trial) size determination. Table 1 shows the required elements for this determination.

Sample size determination required for a test is often determined by weighting two competing factors – science-based criteria, juxtaposed with economic realities. The amount of data that can be collected in a

specified amount of time must be weighted with the amount of precision and accuracy required for the test.

Table 1. Required Elements for Sample Size Determination

Confidence Intervals	Hypothesis Testing
• Objective failure rate p_0 • Confidence Level $1 - \alpha$ (accuracy) • Margin of error δ (precision) • Confidence interval method	• Objective failure rate p_0 • Level of significance α (type I error rate) • Threshold failure rate p_t where $\delta = p_t - p_0$ • Power to detect $1 - \beta$ • Test statistic

In statistical estimation, one typically defines a desired value for the parameter under a test (i.e., detection of a test article at an acceptable quality). The performers **define the desired parameter value as a specified objective failure rate p_0** . The **objective rate** is the goal for a new technology and in the statistical framework, the **true rate** is assumed to be p_0 until the data show otherwise. Once p_0 is chosen, one then defines an **acceptable range** (delta δ) for the value of that parameter (e.g., ± 5 percentage points, half the desired value of the parameter). Ideally, there is a specified **threshold failure rate p_t** that meets the criteria for the maximal acceptable range for the value of that parameter, such that $p_t > p_0$ and $\delta = p_t - p_0 > 0$. The threshold rate p_t represents the largest acceptable failure rate. A new technology with a true rate higher than p_t will be deemed an unsatisfactory technology. For binary outcome failure data, the underlying statistical distribution is assumed to be binomial with $X \sim \text{Bin}(n, p)$; where n is the number of trials, p is the true underlying failure rate, and X is the number of failures out of the n trials. It is also assumed that the trials are statistically independent and unique, as defined above.

- Degree of accuracy needed for the purpose described in the justification

Because OBIM's biometric repository houses hundreds of millions of unique identities, billions of biometric data points, and is involved in 400,000 identification/verification transactions a day, it is imperative that accuracy levels of collection devices or matching software are high. Ideally, OBIM would like no less than 99% accuracy when biometric systems are being utilized for comparison. Even at this 99% accuracy level that leaves 1% of 400,000 transactions a day that result in inaccuracy matching and comparison. This means that approximately 400 individuals crossing borders each day have missed identifications and it is likely that some of these mismatches may be bad actors that pose a threat to national security. Thus, high accuracy levels are a necessity and appropriate sample sizes must be collected that can scale up to the numbers of identities and transactions that OBIM is responsible for each day so DHS components accessing OBIM data can keep borders safe and bad actors out.

Based on previous studies, the performers predict the TMR (true match rate) for baseline sensors will be 99%. For a desired power of 0.70 and a Type I error rate of 0.10, we therefore need to collect a total of

317 independent samples to detect a 1% difference in TMR between contactless and baseline sensors (e.g., contactless TMR $\geq 98\%$).

To obtain this sample size, and to satisfy secondary aims to understand factors within OBIM's collection practices that may affect TMR, participants may collect multiple times on sensors under a variety of conditions (e.g., lighting, distance from camera, skin barriers). Each individual collection under separate conditions will be treated as an independent sample for the purposes of this study.

- Unusual problems requiring specialized sampling procedures

N/A.

- Any use of periodic (less frequently than annual) data collection cycles to reduce burden.

Because of yearly funds and annual project timelines and goals, there should be no need to sample individuals more than one time per fiscal year. Thus, OBIM will work with performers to understand appropriate sample sizes and what is needed for data analyses to be fruitful, scalable, and valuable without placing undue burden on the American people.

Participants will be in this study for a minimum of a single test session lasting up to 2 hours. There is a possibility that multiple test sessions will be needed to fully answer all the research questions. If the participant agrees to future contact in the consent forms, they will be contacted by the performers about participating in another test session.

Participants will be compensated after completing the full test session. Study participation is by appointment only. If staff do not agree to participate in this study, it will not affect their employment.

3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.

OBIM performers have groups of researchers and collaborations that work together to determine the appropriate sample sizes of people that must be sampled, and biometrics collected from to answer relevant questions. Actions that help to meet the appropriate sample sizes as a whole and within each stratification, including oversampling of specific, underrepresented groups that may be less responsive than other groups will be considered by performers. Other strategies include reaching out to groups and communities within certain areas (e.g., geographic regions, like recruitment companies located in the Hawaiian Islands) to ensure that the appropriate groups and numbers from these groups are represented in the overall sample size.

The performer's first step will include utilizing the existing data gathered from the study to see what answers can be obtained. The second step will be to design an initial test to answer the first round of gaps from previous data, analyze it, and identify next experiment parameters. To deal with issues of non-

response, the performer may incorporate biometric image collection and analysis from datasets obtained during previous studies they have participated in. These datasets include algorithm development/training using biometric image data. If the current study's parameters are the same or fit under the same parameters of previous studies, the datasets obtained during previous studies may be leveraged to fill in gaps.

4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.

Performers will conduct initial statistical tests to understand whether the recruitment methods and samples from individuals that are being collected provide the appropriate information from participants. If there are issue with the way recruitment occurs or problems with the biometric collection process or statistical analysis, these problems will be sussed out during the initial tests run on small samples of individuals that are initially collected. Proper changes can then be made to the procedures before full scale recruitment, collection, and testing so that undue number of individuals are not sampled, and participant time and associated funding are not wasted.

Initially, the performer will focus on the biometric sensors. Specifically, evaluating the image quality and the impact on matching algorithms for the following biometrics in studies across the three years: fingerprint (contactless and contact), face, iris, palm, and voice. Research efforts include the following examples:

- i. Evaluation of image quality of fingerprint images from “contactless” fingerprint sensors. The performer will conduct a small-scale data collection to evaluate fingerprint images from contactless Automated Fingerprint Identification Systems (AFIS). The AFIS utilize digital imaging technology to capture and analyze fingerprints. The images from the AFIS will be compared against images collected from traditional fingerprint capture systems, i.e., those that require direct contact between an individual's finger/hand and the fingerprint capture device with or without assistance from a human operator.
- ii. Evaluation of image quality of face and iris images collected simultaneous by the same sensor. The performer will conduct a small-scale data collection to evaluate the emerging technology that allows for simultaneous collection of face and iris images. One challenge previously associated with simultaneous collection of these biometrics is the differing light sources needed for face (visible) and iris (near infrared) for optimal feature collection. A second challenge included ability of cameras to reach optimal focus for both biometrics simultaneously at longer distances. The evaluation will focus on advancements in the technology to leverage balance the lighting challenge and the focusing on distance up to 2-3 ft. Many capture systems in wide use today require the irises to be held in a quite narrow range of distance from the sensor in order to capture a good-quality image.

- iii. Evaluation of image quality of palm prints and images from contact and “contactless” sensors. The performer will conduct a small-scale data collection to evaluate the emerging technology in palm print collection to include assessment of contact and contactless palm sensors; sensors that capture the full hand and segment the fingerprints and palm for analysis; and sensors that include liveness detection features for finger and palm vein collection. Palm print matching systems may use both large features of the palm such as creases and overall shape, and the small features such as friction ridges similar in nature to fingerprints.
- iv. Evaluation of audio quality of various voice sensors ranging from cellphone brand, video qualities, and microphone technologies. An initial small-scale data collection will be conducted to understand the variances in voice quality based on various collection sensors and human/environmental conditions (i.e., noise room, distance from collection sensor, sickness, loud or soft speaking). Voice collections will be captured to test the quality of the different devices to understand the speaker speaking directly into the microphone and at longer-range of capture.
- v. Evaluation of facial features from birth to adulthood leveraging a one-time collection of digital images from participants. Technology has advanced to the stage that early adults (mid-30s) have benefited from having digital photography readily available throughout their lifetime. This advancement would allow participants to provide digital imagery for each year of their life meeting key study criteria for evaluation of facial feature changes. This small-scale study will focus on variance over a demographically diverse dataset.
- vi. Follow on collections will be conducted after the initial small-scale data collections to address specific parameters identified as impactful to collection during the small scale studies and increase the collection size for operational scalability relevance.
- vii. Algorithm development/training using fingerprint, face, iris, and palm images from existing datasets. The performer aims to understand features (segments) that aid in presentation attack detection (PAD), to develop and train analysis algorithms to identify deviations from “normal” and/or “normal” versus “imposter”. One key feature of all existing biometric systems is the use of a biometric template, which is typically created right after the image post-processing steps. A biometric template is a mathematical representation of the original image. When biometric matchers perform a matching operation, they do not compare biometric images directly, but rather compare biometric templates to each other for similarity. Algorithms will be trained and developed to address any differences in image quality and matching performance between the biometric image capture methods.
- i. Matching algorithm testing to determine optimization of the quality scores and match scores of the emerging technology implemented in bullets I-V. Testing of the current operational matcher against emerging matchers that may be more advanced or tailored to the biometric use cases (i.e., contactless/contactless, contactless/match, palm, voice) to determine the best performing algorithm for future implementation in operations. Most algorithm development aims to develop techniques to make contactless images more like contact

images, so they can be handled by legacy matching algorithms. **Biometric Profile Data Collection**

The performer will compare the image matching performance of contactless fingerprint systems (AFIS) against the most used off-the-shelf (COTS) contact fingerprint capture systems. Face, iris, palm, and voice will undergo image capture for matching systems. These images will be digitized and then post-processed. The post-processing is done to make the image better suited to matching the stored biometric data.

Participants, both staff and the general public, will be recruited. Given the extensive amount of fingerprint image data collected under previous studies, the performer will prioritize recruitment of those participants provided they consented to future contact.

Participants will be recruited via email announcement. For staff, the email will be circulated from the performers' Office of Human Research Protection. For recruitment of the general public, the performer will once again seek assistance from Matthews Media Group (MMG). Interested participants will be instructed to contact a member of the research team for additional information about the study and will be given a copy of the consent form to review prior to the test date.

On the test date, an overview of the study will be provided, and participants will have the opportunity to have their questions answered. Once informed consent has been obtained, participants will be asked to provide basic demographic information (age, gender, and race) along with information on any previous injuries to their hands that may have impacted their fingerprints and if they have sensitivity to COTS hand lotion. Each participant will be assigned a random study ID number under which their fingerprint images will be stored.

The data collection is designed to determine the factors that impact image quality on several biometric sensors. The participants will visit the sensor stations to present their biometrics multiple times. Each presentation will have a different variable, e.g., low lighting, variable distance from the sensor. These parameters will be detailed to the participants in advance, and they may choose to opt out of a presentation if they are uncomfortable.

ii. Algorithm Evaluation

The performers will utilize biometric images collected under previous databases. The sponsor may provide additional data under a and/or recommend other data sources (public or private). Data Use Agreements (DUAs) will be executed as warranted.

A biometric matching system uses a matching algorithm, or matcher, to compare two biometric templates, which are derived from the biometric images (or other biometric measurements). The details of how this is done vary dramatically across modalities and can vary across matchers. For example, early face matchers identified locations on the face (i.e., eyebrows, eyes, nose, chin, corners of the mouth) and performed matching by comparing their relative positions; current face matchers almost exclusively use deep-neural-network algorithms (a machine-learning type of algorithm) to match one face image to another; iris algorithms use a handcrafted mathematical approach based on wavelet decomposition of the radial features of the iris pattern. Fingerprint algorithms use a combined approach. Fingerprint analysis for biometric identification predates computer algorithms, and the first computer algorithms simply coded the human approach of identifying minutiae points in the fingerprints and then comparing

their relative positions in an (x, y) coordinate system. The advent of deep neural networks has improved the algorithmic identification of fingerprint minutiae, but many algorithms maintain the more engineered approach to comparing the locations of minutiae.

A general feature of all biometric matchers is that they produce a similarity score; this score is a measure of how similar the specific features of interest are between the two biometric template samples being compared. It is important to note that this similarity score will not likely be linearly related to the likelihood that two images are a match, but it will be monotonically related to that likelihood.

The continued advancement of these matching systems has led to many more black box or proprietary implementations where the user does not have full visibility of methodology implemented within the matching system. Evaluations of matching systems leveraging tests that implement the same parameters and allow for 1:1 performance comparison allow for the user to have a true understanding of the variation in performance between the available matching systems.

There are two key parameters used to evaluate any matching system, the false match rate (FMR) and the false non-match rate (FNMR). The FMR is the rate at which two images which are not of the same person are falsely declared to match, and the FNMR is the rate at which two images which are of the same person are declared not to match.

- **Data Management**

Biometric image data will be stored in a separate database from the participant demographic data (i.e., contact information, PHI/PII). All data will be stored on controlled access computers, access to which will be limited to study team members and collaborators at NIST. While the biometric image data will not be shared with the Sponsors of this research, the aggregated results will be. NIST staff will not have access to participant contact information or additional PHI/PII beyond the biometric images. The people who may request, receive, or use the information include the researchers and their staff who may be a part of JHU, NIST, DHS S&T, and OBIM FI.

- **Data Sharing**

Select research staff and the National Institute of Standards and Technology (NIST) may be provided the collected biometric image data and select demographic information (i.e., age, gender, race, ethnicity, and occupation). The demographic information will be identified only by the study ID#. No other PHI/PII will be provided to these researchers, and they will not otherwise have access to that information. The data gathered may be shared with researchers at other institutions, for-profit companies, sponsors, government agencies, and other research partners. The data may also be put in government or other databases/repositories. Any results produced will be leveraged as government products and not published as academic peer review journal articles.

Participants are eligible to participate if they are:

- A U.S. citizen, between the ages of 18 and 75 able to provide proof of citizenship in the form of a valid government-issued ID, such as a state driver's license, birth certificate, or passport.
- Willing and able to provide written informed consent.

- Available to attend a test appointment lasting approximately two and a half hours at the JHU test facility in Columbia, Maryland, on a weekday.

5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency

OBIM performers will conduct official analyses via statistical teams from places like Johns Hopkins University Applied Physics Lab and the Maryland testing facility (MdTF). For OBIM analysis, OBIM Futures Identity will have individuals from our team conducting initial and in house evaluation of data for our own edification and discussion among our OBIM team writ large. For discussions on these matters, please look at the contact information below.

OBIM Futures Identity:

- 1. Kneapler, Caitlin (Technical Integrator)
 - o Caitlin.n.kneaper@hq.dhs.gov
- 2. Conde Puertas, Alexandra (Student Trainee)
 - o Alexandra.puertas@hq.dhs.gov

John Hopkins University Applied Physics Labs (JHU APL):

- 1. Engel, Marianne O.
 - o Marianne.Engel@jhuapl.edu
- 2. Mirabito, Peter V.
 - o Pete.Mirabito@jhuapl.edu

DHS Science and Technology Directorate (DHS S&T):

- 1. Vemury, Arun
 - o Arun.vemury@hq.dhs.gov

NIST

- 1. Watson, Craig I. (Fed)
 - o Craig.watson@nist.gov