U.S. Department of Homeland Security

SCIENCE AND TECHNOLOGY DIRECTORATE

2022 Biometric Technology Rally Results Webinar

Science and Technology

Arun Vemury
Lead
Biometric & Identity Technology Center

John Howard
Lead Data Scientist
The Maryland Test Facility

Yevgeniy Sirotin
Principal Investigator
The Maryland Test Facility

February, 2023

Outline

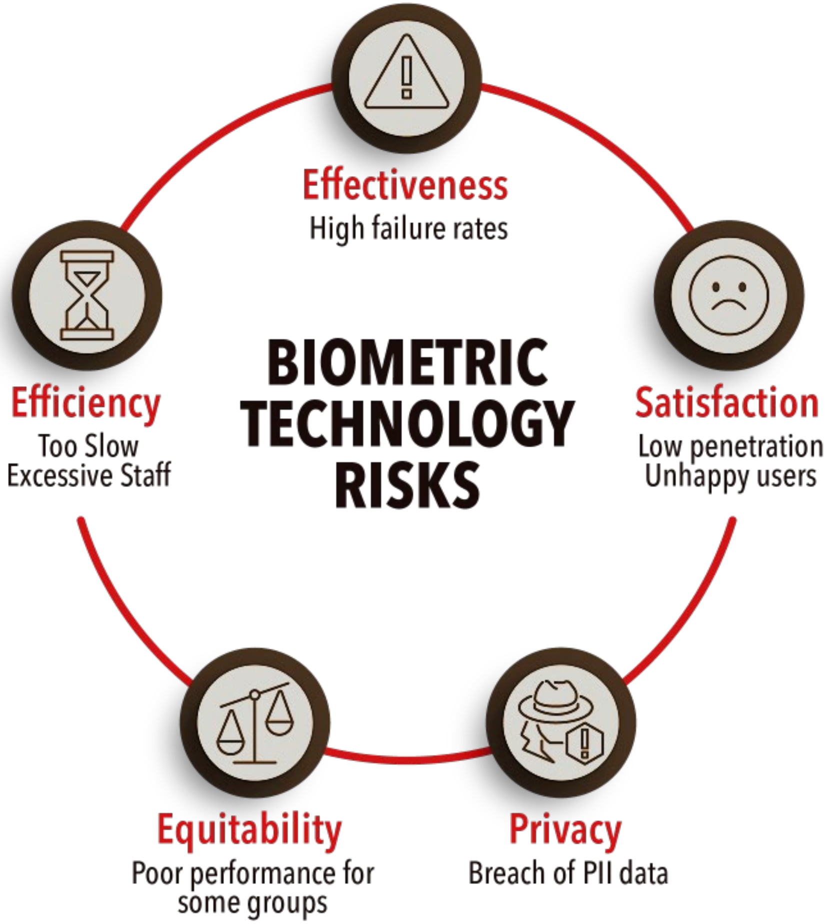
- Introduction
  - Biometric & Identity Technology Center
  - Biometric Technology Rallies
  - Past Biometric Technology Rallies
- 2022 Rally Test
  - System Functionality
  - Timeline
  - Test Approach
- 2022 Rally Results
  - Analytic Approach
  - Efficiency & Satisfaction
  - Privacy
  - Effectiveness & Equitability
- Conclusions

Biometric & Identity Technology Center

S&T conducts foundational research to ensure advancements in science and technology are harnessed for cutting-edge solutions to new and emerging operational challenges.

- Drive biometric and identity innovation at DHS through RDT&E capabilities
- Facilitate and accelerate understanding of biometrics and identity technologies for new DHS use cases
- Drive efficiencies by supporting cross cutting methods, best practices, and solutions across programs
- Deliver Subject Matter Expertise across the DHS enterprise
- Engage Industry and provide feedback
- Encourage Innovation with Industry and Academia

Biometric Technology Rallies



- Rally Goals
  - **Focus industry on a specific use-case challenge** (e.g., high throughput, small groups)
  - **Identify and mitigate risks** associated with new biometric technology
  - **Create an independent assessment** of current industry offerings
  - **Collaborate and guide** promising technologies, using Cooperative Research and Development Agreements
- Annual Rallies
  - High throughput unattended use case
  - Scenario testing with naïve test volunteers
  - Dozens of commercial devices
  - More information: https://mdtf.org

Past Biometric Technology Rallies

2018 Rally assessed acquisition systems

2019 Rally assessed acquisition systems *and* matching systems

2020 Rally assessed acquisition *and* matching systems *with* face masks

2021 Rally assessed acquisition *and* matching systems *with* face masks *and* system equitability

- Since 2018, the Rallies have demonstrated progress in the performance and maturity of biometric acquisition and matching systems
  - Rally results provide insights into how people interact with biometric systems to improve usability
  - Rally results have been used to inform participating vendors, leading to improved performance of both acquisition and matching systems
  - **There are continuing challenges with respect to reliable image acquisition in the high throughput unattended use-case**

2022 Rally

2022 Biometric Technology Rally

- Scenario test of a **“high-throughput”** use case with new challenges for acquisition systems
  - Focus on **group processing**, requiring systems to rapidly capture biometrics from multiple users
- Results demonstrate the **efficiency, effectiveness, privacy, and user satisfaction** of current commercial biometric systems:
  - Ability of acquisition systems to capture images that work across algorithms
  - Ability of matching systems to work across acquisition systems
  - Biometric system performance reported by specific demographics based on race, gender, and skin tone
  - Biometric systems protect the privacy of bystanders and non-users to preserve anonymity
- DHS S&T will continue to collaborate and guide promising technologies and share information via Cooperative Research and Development Agreements

System Functionality

- **Acquisition system** minimum requirements:
  - Acquire face images from volunteers in groups of 2-12 individuals
  - Submit exactly one image of each volunteer using the system
  - Provide no images of anyone or anything other than system users
  - Operate in an unmanned mode, i.e., without an operator/instructor
  - Submit biometric images in under six seconds per volunteer, on average

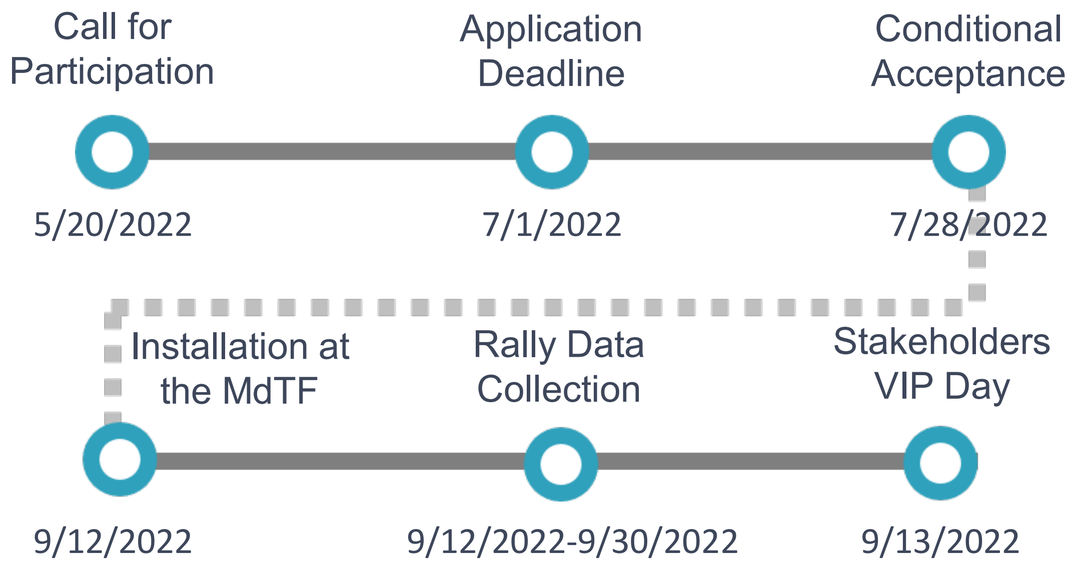
- **Matching system** minimum requirements:
  - Matching system provided as a ".tar" docker container
  - Commercially available matching algorithm
  - Less than 1.5 GB in size
  - Perform biometric operations within a time limit
  - Work without access to external networks
  - Conform to an OpenAPI 2.0 (swagger) specification

Application and Selection Process

- All Rally applications were evaluated by a panel of experts from DHS, DoD, NIST, and industry
- 16 biometric systems applied, were accepted, and participated
  - 6 face acquisition systems, 4 proceeded to evaluation
    - One vendor withdrew before installation at the MdTF
    - One vendor withdrew during installation at the MdTF
  - 10 face matching systems
- Each system was given a unique alias inspired by US rivers and mountains

2022 Biometric Technology Rally Timeline

- The 2022 Biometric Technology Rally was announced in May 2022
- Biometric technology providers had 6 weeks to submit applications
- Following conditional acceptance, technology providers had 6 weeks to build and develop their systems
  - The Maryland Test Facility (MdTF) designed and provided a cloud-hosted application programming interface (API) to allow systems to be integrated with the test prior to arrival at the MdTF



Science and Technology

Group Processing at Checkpoints (Concept):

2022 Rally Process

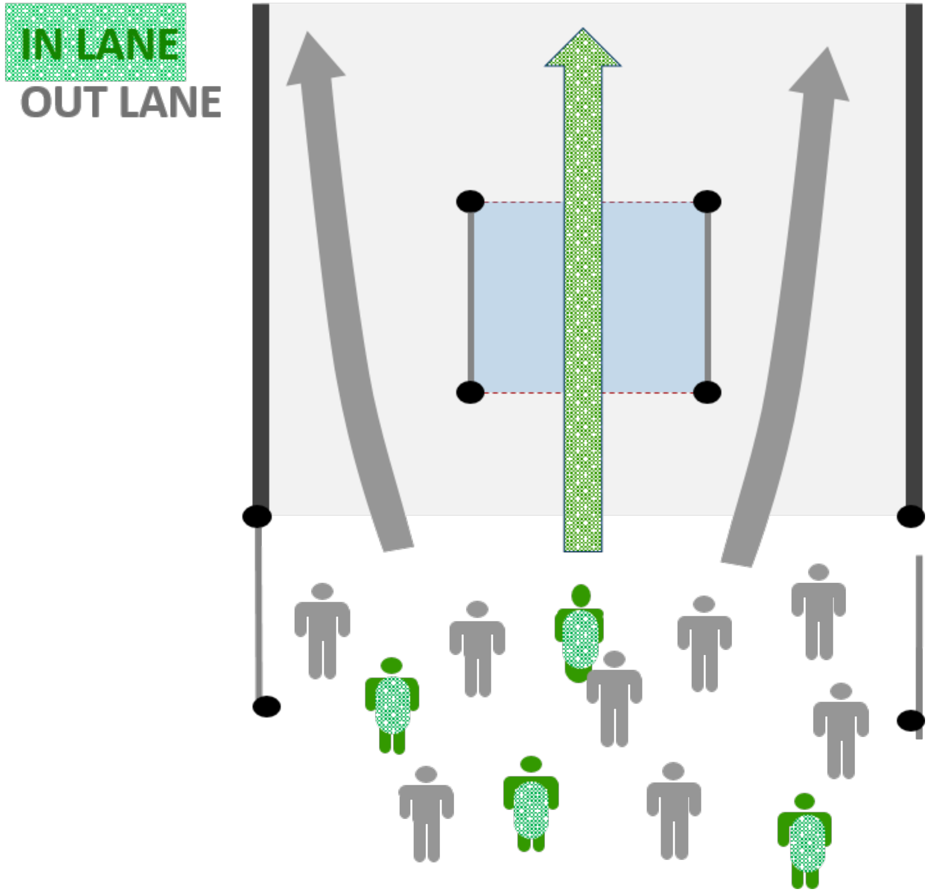
Group Processing at Checkpoints (Testing):

2022 Rally Station Configuration

Left Lane

Center Lane

Right Lane



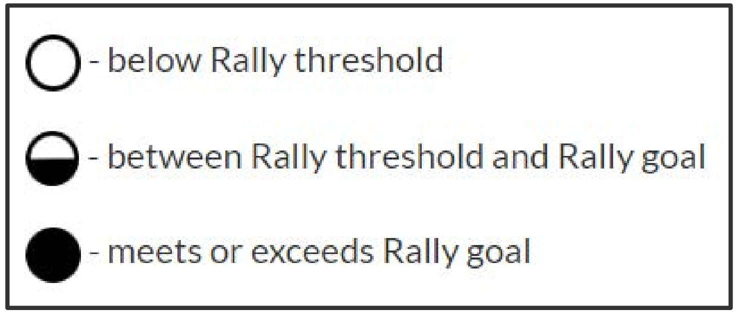
2022 Rally Results

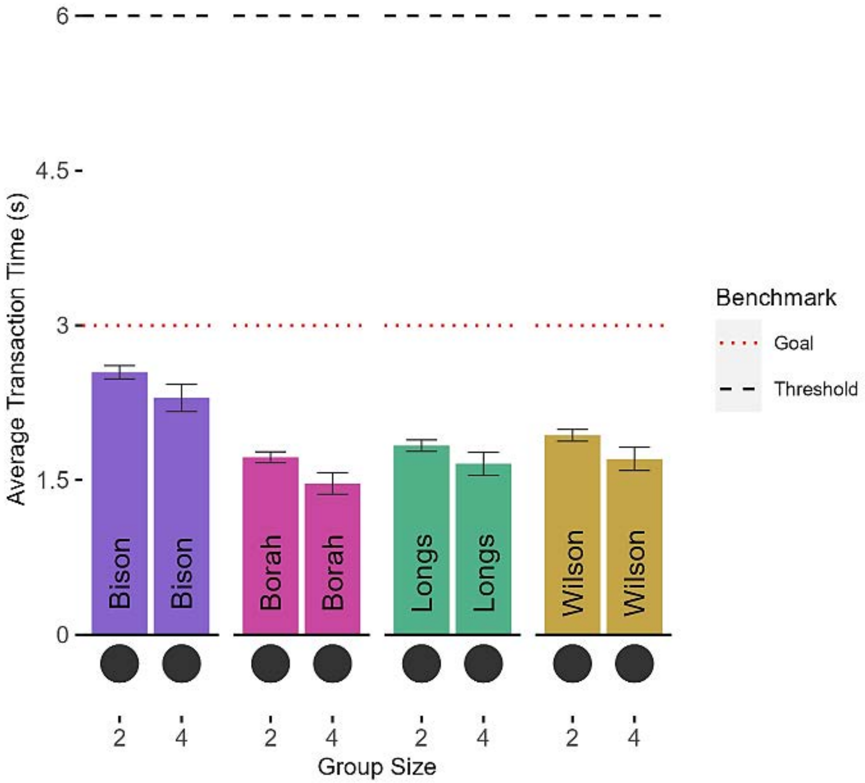
Analytic Approach

- Evaluation of biometric system combinations:
  - **Acquisition** systems only – Efficiency and Satisfaction
  - Combinations of **acquisition and matching** systems – Privacy, Effectiveness, and Equitability
  - Aggregate comparisons made against Rally **benchmarks and threshold levels**
  - Disaggregated comparisons made against **levels** and **median system performance per group**
  - No p-values or statistical checks between systems/groups or against benchmarks

Efficiency

- All acquisition systems met the goal of 3 seconds or less and had faster per person transaction times for larger groups
- Quantified as average transaction time per group size per volunteer at each Rally Station



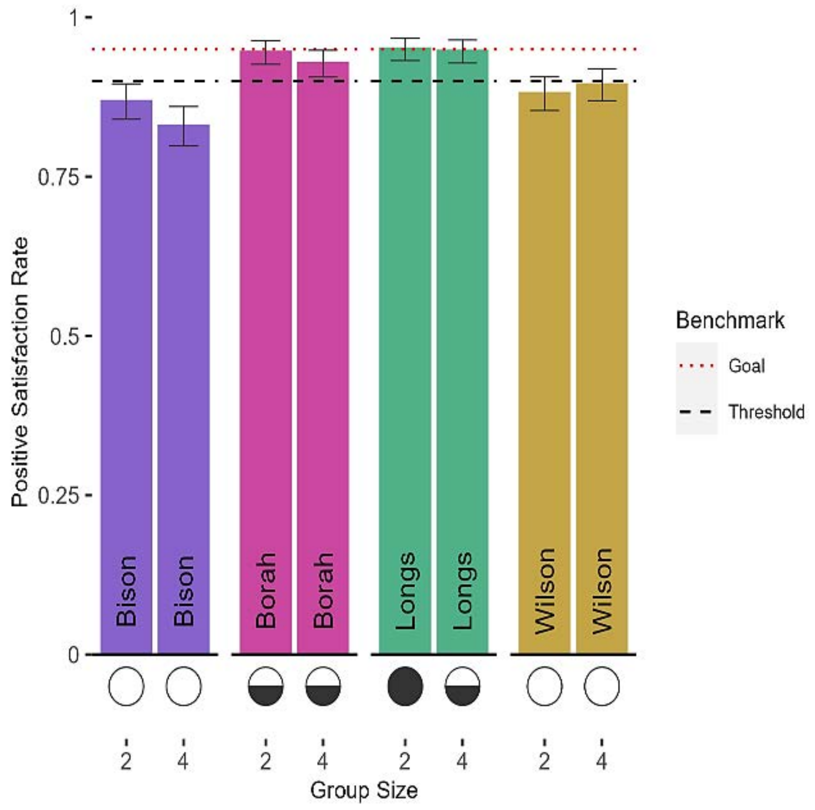


- Most efficient:
  - **Borah – 1.72 seconds per person for groups of 2, 1.47 seconds per person for groups of 4**

Science and Technology

Satisfaction

- Met the goal (Satisfaction > 95%):
  - Borah, Longs
- Test volunteers' positive attitudes toward the acquisition systems (Happy or Very Happy)



- below Rally threshold
- between Rally threshold and Rally goal
- meets or exceeds Rally goal

- Most Satisfying:
  - **Longs – 95.3% Happy or Very Happy for groups of 2,**
    95.0% Happy or Very Happy for groups of 4

Science and Technology

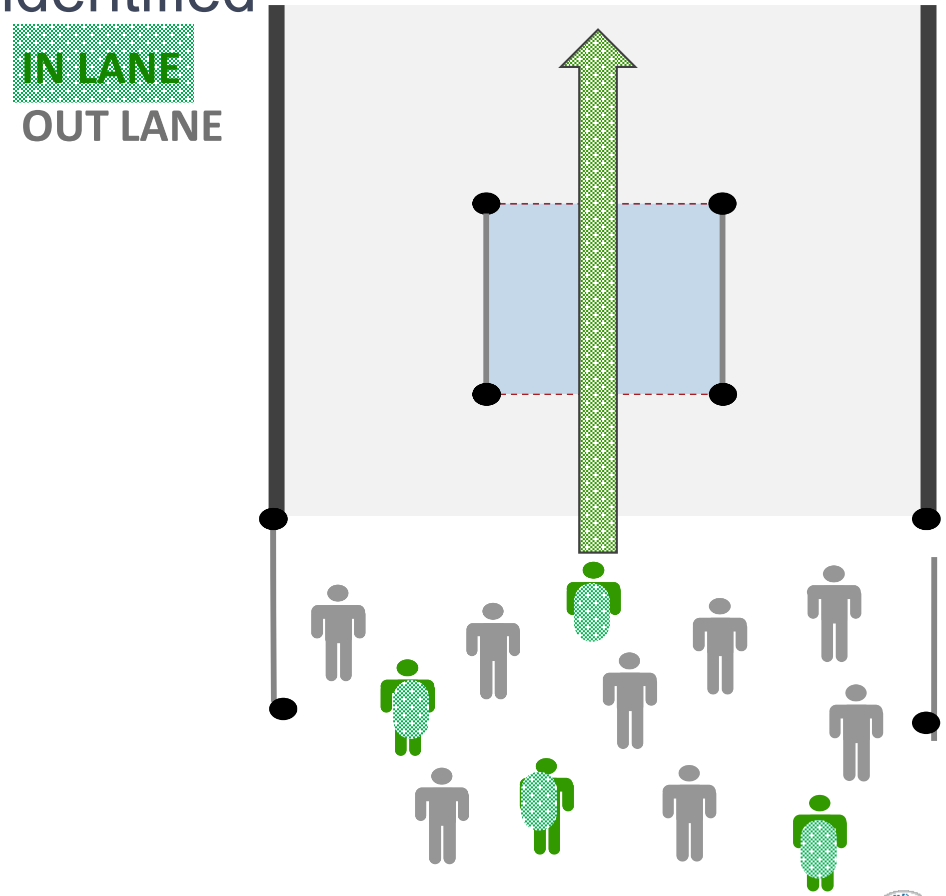
Acquisition and Matching System Combinations

Privacy – Non-User Identification Rate

- Rally Goal: < 1% of non-users acquired + identified

2022 Rally Station Configuration

Left Lane | Center Lane | Right Lane



Privacy – Non-User Identification Rate

Groups of 2

Matching System	Bison	Longs	Wilson	Borah
	Acquisition System			
Kenai	0.00 •	0.00 •	0.00 •	0.64
Miami	0.00 •	0.00 •	0.00 •	0.64
Tioga	0.00 •	0.00 •	0.00 •	1.06
Mill	0.00 •	0.00 •	0.00 •	0.85
Bronx	0.00 •	0.00 •	0.00 •	1.27
Grant	0.00 •	0.18	0.00 •	0.64
Hop	0.00 •	0.00 •	0.00 •	1.06
Entiat	0.00 •	0.00 •	0.00 •	0.85
Flag	0.00 •	0.00 •	0.00 •	0.65
Row	0.00 •	0.00 •	0.00 •	0.88

Groups of 4

Matching System	Bison	Longs	Wilson	Borah
	Acquisition System			
Kenai	0.00 •	0.00 •	0.00 •	0.22
Miami	0.00 •	0.00 •	0.00 •	0.00 •
Tioga	0.00 •	0.18	0.00 •	0.00 •
Mill	0.00 •	0.18	0.00 •	0.00 •
Bronx	0.54	0.18	0.00 •	0.22
Grant	0.00 •	0.18	0.00 •	0.00 •
Hop	0.00 •	0.18	0.00 •	0.00 •
Entiat	0.00 •	0.18	0.00 •	0.00 •
Flag	0.00 •	0.00 •	0.00 •	0.22
Row	0.00 •	0.00 •	0.00 •	0.24

- Twenty-nine (29) system combinations **met the goal** of 0 non-user identifications for both group sizes
- **Different system combinations** met the goal for groups of 2 and 4
- Overall **low NU-IR levels** across acquisition and matching system combinations

Table legend

Min •	=0%	≤1%	≤5%	>5%

Effectiveness – Operational vs Matching Focus

- During a transaction:
  - Acquisition system: expected to submit 2 or 4 images of suitable quality, one per volunteer
  - Matching system: expected to match images to a historic gallery
- **Operational** Focus – includes all sources of error
  - Failure to submit an image of an in-lane volunteer
  - Failure to submit an image of an in-lane volunteer of suitable quality
- **Matching** Focus – discounts failure to submit images & quality issues
  - When images of suitable quality were submitted, how often did they not match to a historic gallery?

Effectiveness – Operational Focus

- TIR: True Identification Rate: quantified as the **percentage of users** who were correctly identified
- (Correct Identifications / Total People)

Matching System	Acquisition System: Bison	Longs	Wilson	Borah
Kenai	97.4●	96.5	93.2	74.1
Miami	97.4●	96.5	93.2	74.1
Tioga	97.4●	96.5	93.2	73.9
Mill	97.4●	96.3	93.2	73.4
Bronx	97.0	96.3	93.0	73.6
Grant	97.4●	96.0	93.0	73.0
Hop	96.9	95.8	92.8	73.7
Entiat	96.7	95.5	92.3	73.7
Flag	97.2	93.4	93.0	72.3
Row	83.7	83.8	79.2	62.4

Groups of 2

Matching System	Acquisition System: Bison	Longs	Wilson	Borah
Kenai	97.4●	95.8	93.0	74.1
Miami	97.4●	96.0	93.0	74.1
Tioga	97.4●	96.0	93.0	74.1
Mill	97.4●	96.0	93.0	73.9
Bronx	96.8	95.7	93.0	73.7
Grant	97.2	95.1	93.0	73.7
Hop	96.8	95.7	93.0	74.1
Entiat	96.5	95.3	92.3	73.6
Flag	97.4●	94.3	92.6	72.7
Row	81.3	84.0	79.2	59.8

Groups of 4

- Seventeen (17) system combinations **met the TIR threshold of 95%** for groups of 2 and 4
- Same system combinations **across groups of 2 and 4**
- **No system** combinations met the TIR goal of 99%

Table legend

Max●	>99%	>95%	>90%	≤90%

Effectiveness – Matching Focus

- Matching -TIR: True Identification Rate, quantified as the **percentage of submitted images of sufficient quality** that were correctly identified
- (Correct Identifications / Submitted Images)

Matching System	Acquisition System: Bison	Longs	Wilson	Borah
Kenai	99.8	99.8	99.8	99.5
Miami	99.8	99.8	99.8	99.5
Tioga	99.8	99.8	99.8	99.5
Mill	99.8	99.8	99.8	98.6
Bronx	99.5	99.6	99.6	98.8
Grant	99.8	99.3	99.8	98.4
Hop	99.3	99.1	99.4	99.3
Entiat	99.1	98.7	99.1	99.1
Flag	99.8	99.8	99.8	98.8
Row	87.1	87.3	86.5	86.5

Groups of 2

Matching System	Acquisition System: Bison	Longs	Wilson	Borah
Kenai	99.8	99.5	99.8	99.8
Miami	99.8	99.6	99.8	99.8
Tioga	99.8	99.6	99.8	99.8
Mill	99.8	99.6	99.8	99.5
Bronx	99.8	99.5	99.8	99.3
Grant	99.8	98.7	99.8	99.3
Hop	99.3	99.3	99.8	99.8
Entiat	98.9	98.9	99.1	99.1
Flag	99.8	99.8	99.8	99.3
Row	86.3	87.3	86.6	85.6

Groups of 4

- Thirty-six (36) system combinations **met the 95% matching-TIR threshold** for groups of 2 and 4
- No significant impact of group size

Table legend

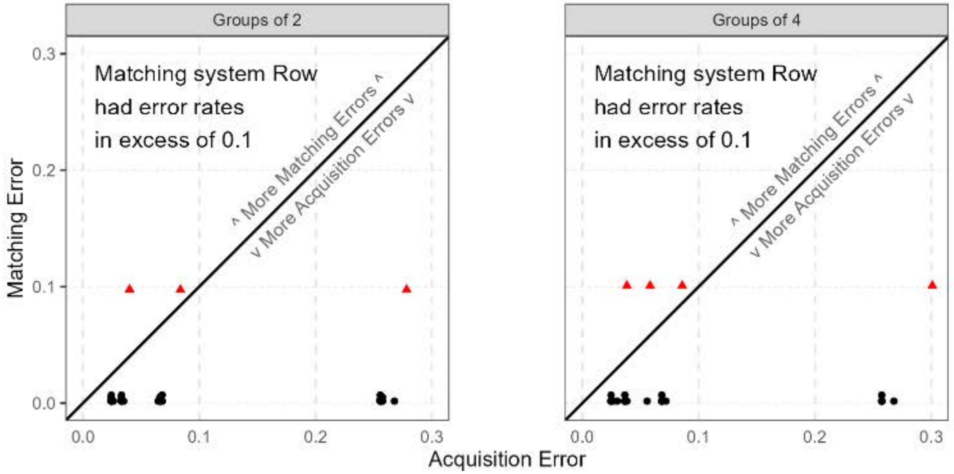
Max	>99%	>95%	>90%	≤90%

Acquisition System Errors Limit Performance

- 36 of 40 system combinations had acquisition errors in excess of matching errors

- **Recommendation:**

To reduce error rates in high-throughput unattended use-cases, biometric technology providers can focus on improving the ability of acquisition systems to capture quality images

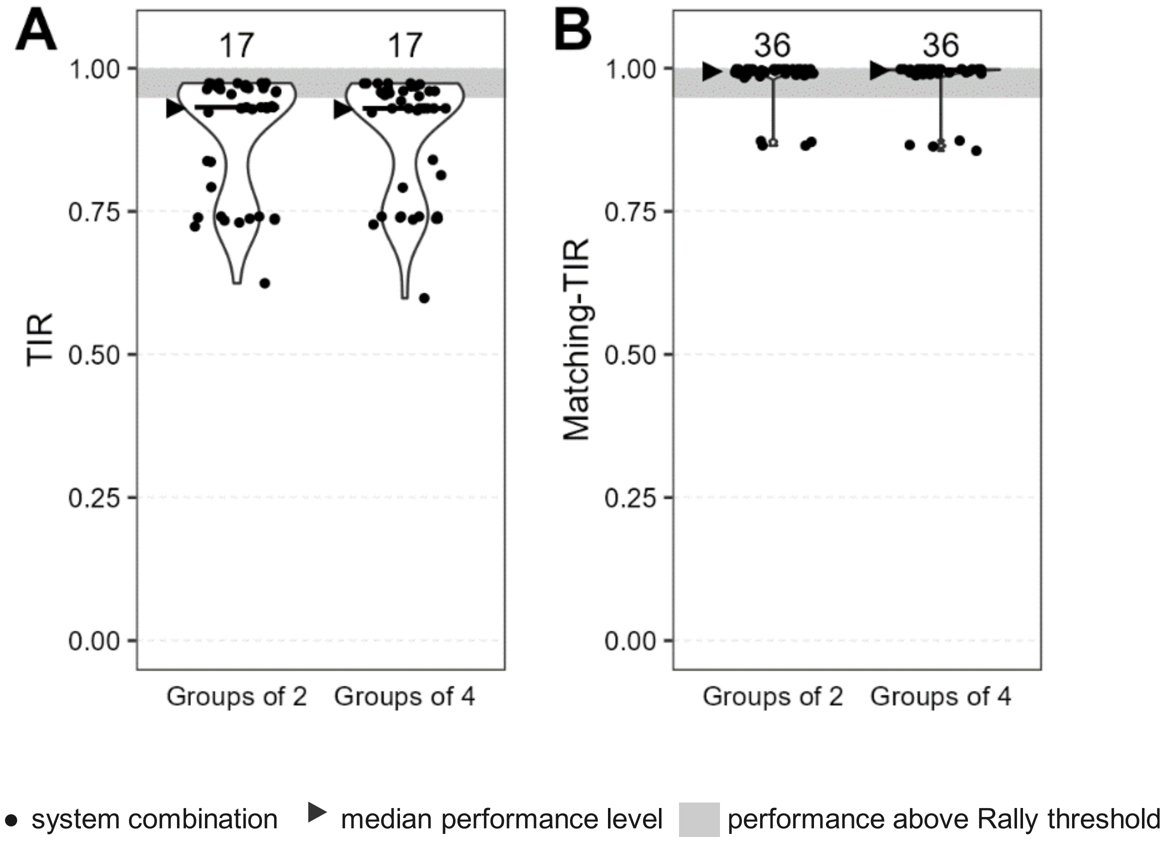


● system combinations
▲ system combinations involving matching system Row

Group Size Doesn't Affect Aggregated Performance

- Median system performance is comparable across group sizes

	Groups of 2	Groups of 4
Median TIR	93.2%	93.0%
Median Matching-TIR	99.5%	99.7%

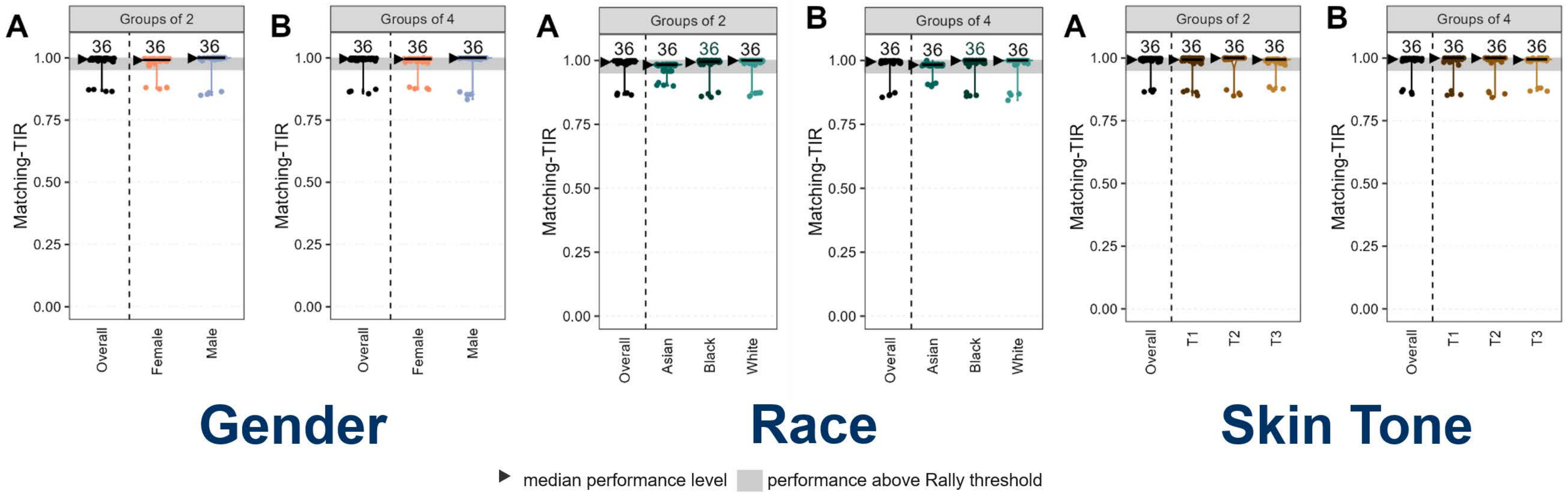


Science and Technology

Effectiveness – Demographics

- TIR performance was disaggregated into eight demographic groups
- Gender (self-reported)
  - Male, Female
- Race (self-reported)
  - Asian, Black, White
- Skin-Tone (measured)
  - Lighter, Medium, Darker

Matching Focus Demographic Differentials



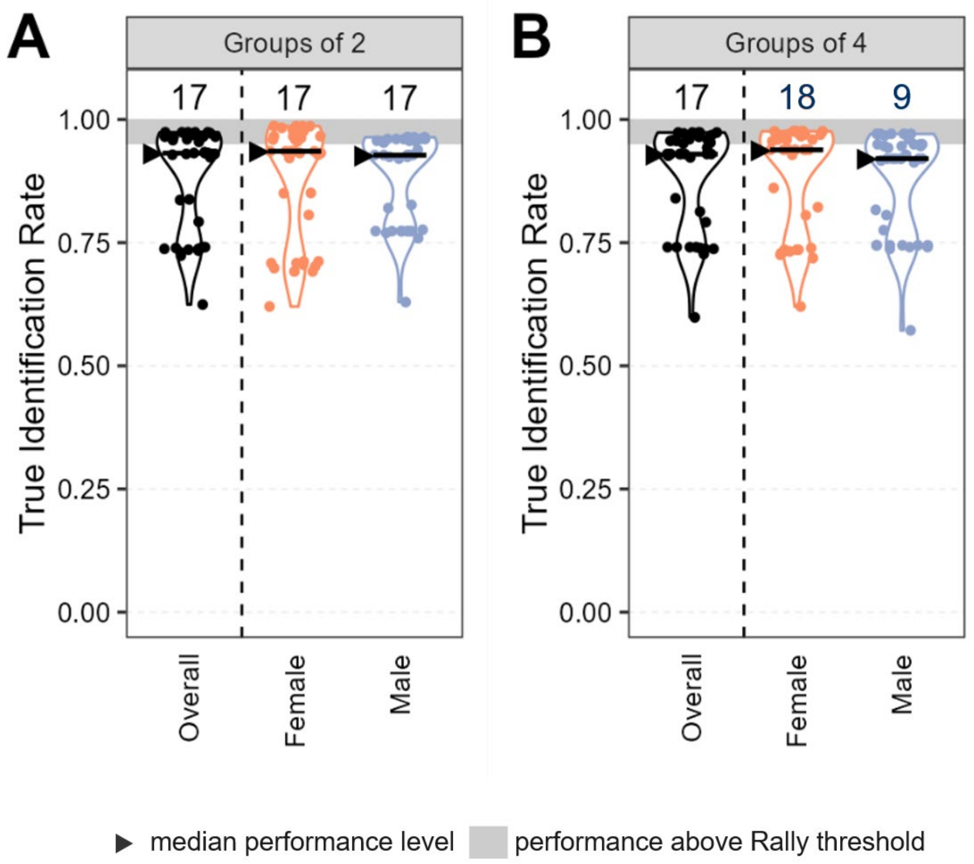
- When discounting failures to submit images of suitable quality, **most system combinations** were able to meet the 95% Rally matching-TIR threshold

Science and Technology

Operational Focus Demographic Differentials

- Some system combinations were able to meet the 95% Rally TIR threshold for all demographic groups
- However, considering acquisition some demographic differentials remained
- Median system performance was:
  - Lower for “Male” relative to “Female” volunteers **(gender differential)**

Group Size	Female		Male
2	93.5%		92.8%
4	93.9%		92.0%

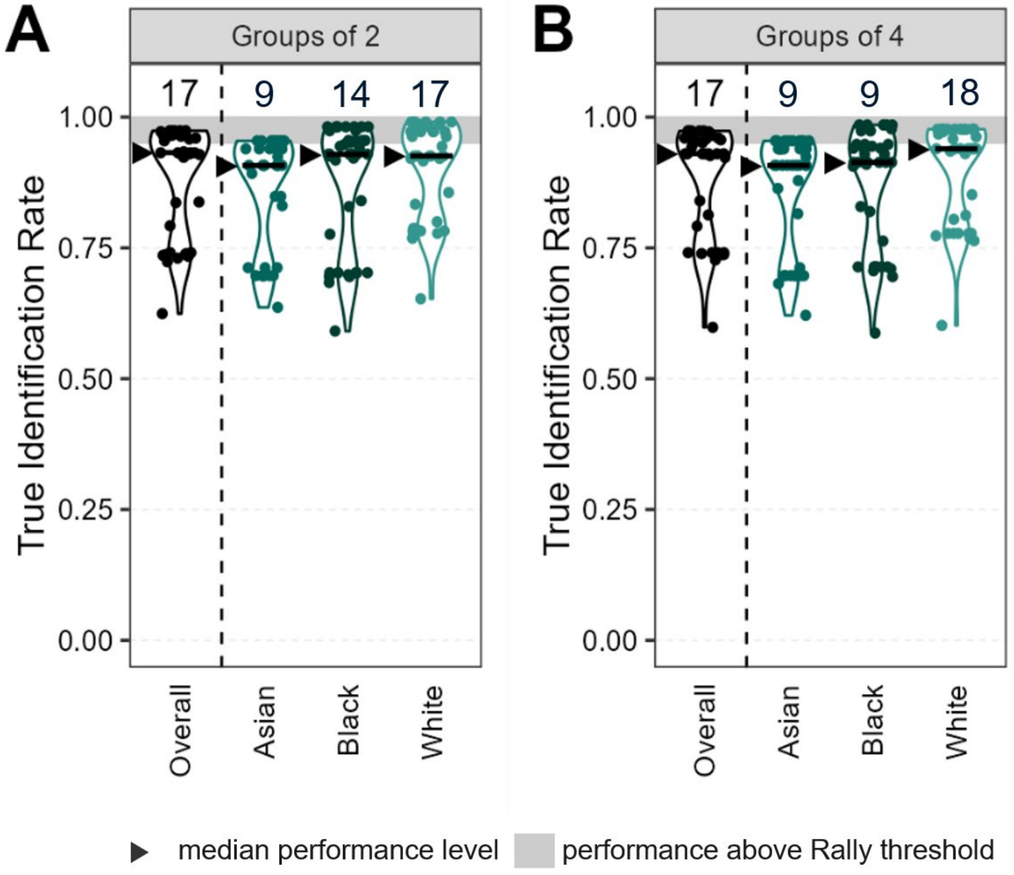


Science and Technology

Operational Focus Demographic Differentials

- Some system combinations were able to meet the 95% Rally TIR threshold for all demographic groups
- However, considering acquisition some demographic differentials remained
- Median system performance was:
  - Lower for volunteers that self-identified as "Asian" **(race differential)**

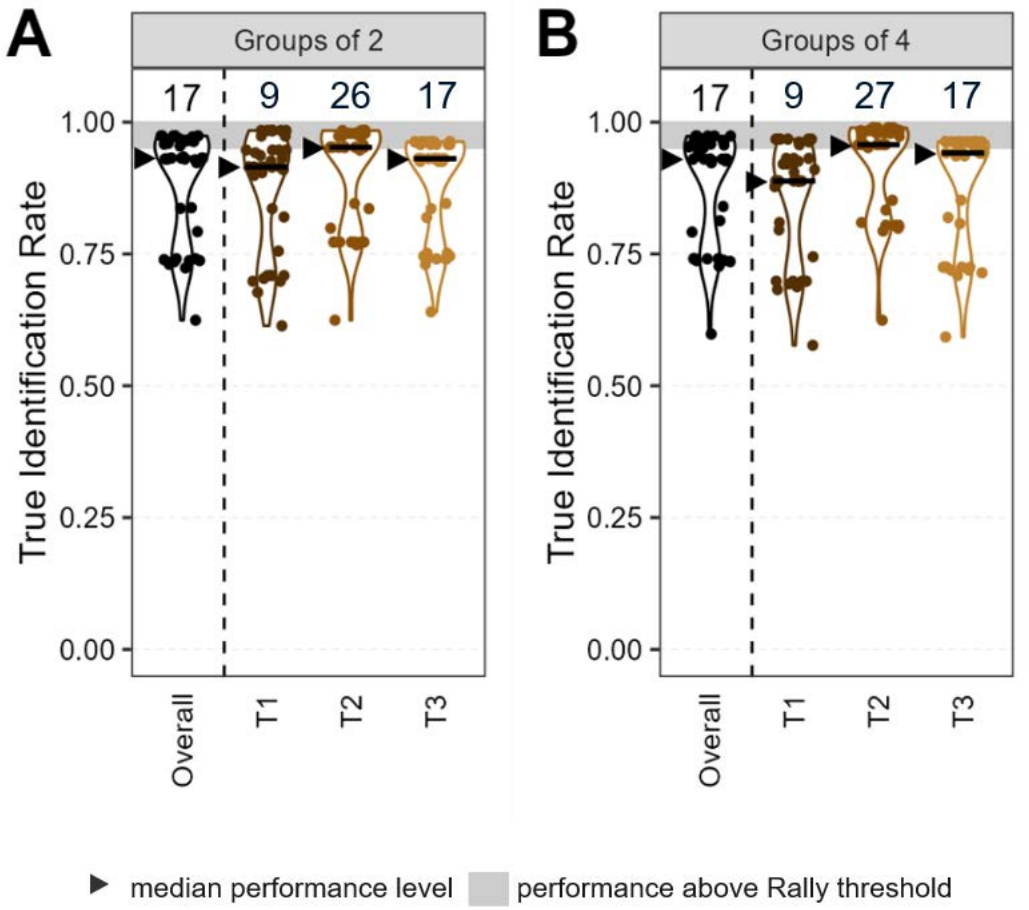
Group Size	Black	White		Asian
2	92.9%	92.5%	➘	90.8%
4	91.3%	93.9%	➘	90.8%



Operational Focus Demographic Differentials

- Some system combinations were able to meet the 95% Rally TIR threshold for all demographic groups
- However, considering acquisition some demographic differentials remained
- Median system performance was:
  - Lower for volunteers with very dark skin tone and very light skin tone **(skin tone differential)**

Group Size	Light Skin Tone		Dark Skin Tone
2	93.1%		91.4%
4	94.1%		88.8%



Demographic Summary

- When discounting failures to submit images of suitable quality, **most system combinations** were able to meet the 99% Rally match-TIR goal for all demographic groups

- Including failure to capture, **some system combinations** were able to meet the 95% Rally TIR threshold for all demographic groups

- Including failure to capture, demographic differentials in the number of systems able to achieve the 95% Rally TIR threshold were observed:
  - Lower for “Male” relative to “Female” volunteers
  - Lower for volunteers that self-identified as “Asian”
  - Lower for volunteers with darker skin tone

Conclusions

Conclusions

- Tested acquisition systems had **fast** transaction times and **high user satisfaction**
- Biometric systems can be designed to **maintain privacy** of non-users
- Effectiveness errors arose in the **acquisition** of quality images, **not matching** those images to historic samples
- **Group size** did not appear to impact effectiveness
- **Demographic differentials** remain in the **acquisition** of quality samples


CHECKPOINT
Science and
Technology
Prepared by the Identity and Data Sciences Lab

Interactive Results Available at mdtf.org

- The data presented today is available for review and exploration at https://mdtf.org
- Interactive visualization of demographically disaggregated performance
- Downloadable PDF report with detailed performance metrics for each tested system

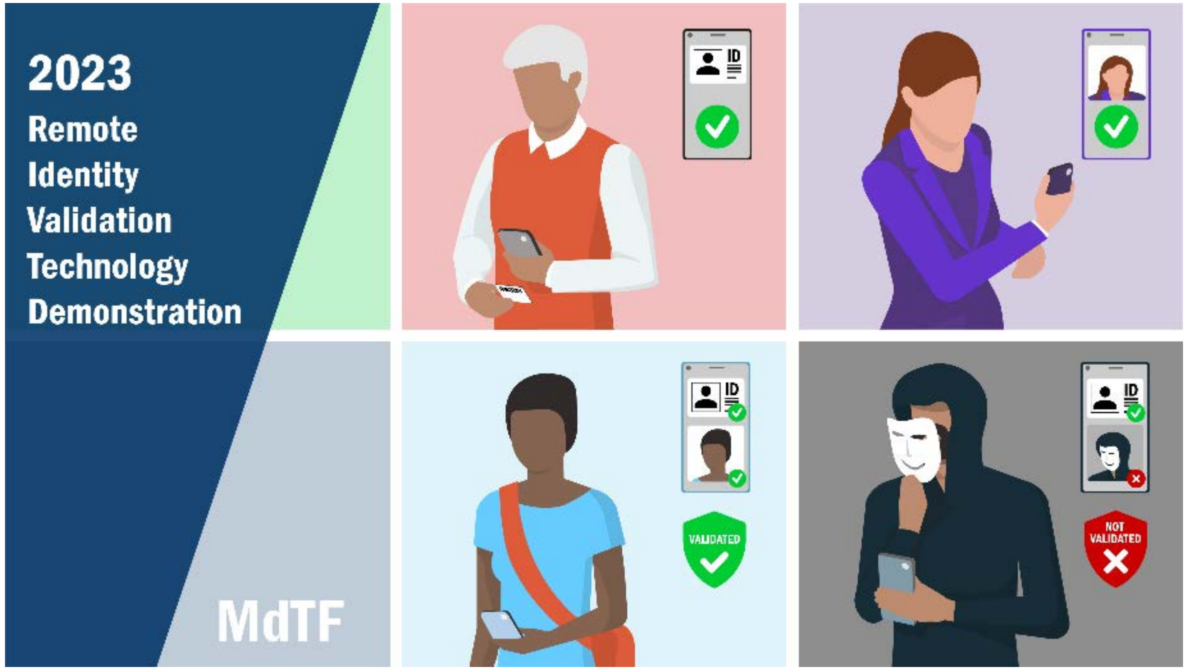
ISO/IEC 19795-10: Demographic Differentials

- DHS S&T is supporting development of standard methods of measuring demographic differentials:
  - ISO/IEC 19795-10 WD4 – Biometric performance across demographic groups
  - How to define demographic groups, including skin-tone
  - How to plan and perform an assessment of demographic differentials
  - How to calculate & report error rates across groups

Science and Technology

2023 Remote Identity Validation Technology Demonstration (RIVTD)

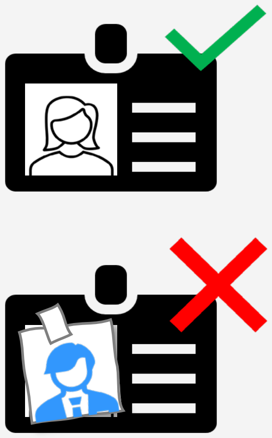
- **DHS S&T is looking for full RIV systems and/or component technologies** that are capable of:
    1. Assessing the validity of an identity document (US driver's license)
    2. Matching a "selfie" photo to the photo on the Identity Document
    3. Assessing the "liveness" of the "selfie" photograph
- **DHS S&T encourages providers of technologies that can perform any portion of the RIV process to apply to participate in this demonstration**
- This **DEMONSTRATION** will have different tracks such that each step in the RIV process will be demonstrated separately



RIVTD Tracks

Track 1: ID Validation

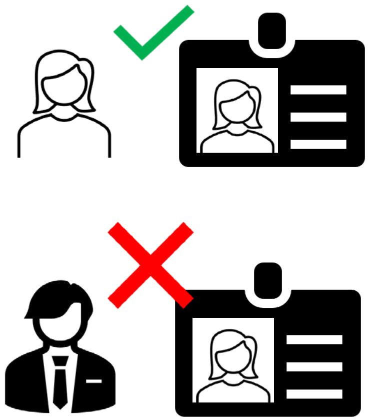
- Information Check
- Tamper Check
- Security Check



Applications Received
Feb 15, 2023

Track 2: Match to ID

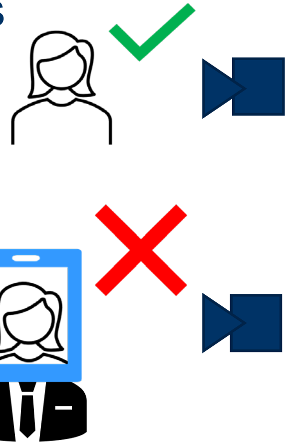
- 1:1 Verification



Applications Due
May 17, 2023

Track 3: Liveness and Presentation Attack Detection (PAD)

- Reject screens and printouts
- Reject masks and other fakes



Applications Due
Sep 13, 2023

Questions & Answers

- Contact information
  - peoplescreening@hq.dhs.gov
  - rally@mdtf.org
- Visit our websites for additional information
  - To see additional work DHS S&T supports, visit www.dhs.gov/science-and-technology
  - To view additional information about this year and prior Rallies, visit https://mdtf.org