Research Brief: Facial Recognition Technology, AI Bias & Judicial Education

Evidentiary Competence, Not Advocacy: Building the Case for Equity-Threaded Curriculum

Objection, That's a Deepfake ▪ A Demonstrative Workshop for Judges ▪ Summer 2026

Prepared by Alexander Karsh, Program Manager ▪ Coordinator: the coordinating judge

Purpose of This Brief

This document supports the development of a proposed learning outcome for the AI & Digital Evidence workshop addressing the documented differential reliability of facial recognition and AI identification tools across demographic groups. The proposal is grounded in evidentiary competence: judges who evaluate expert testimony on AI-mediated identification evidence must understand that these tools have known, quantified performance gaps that vary by skin tone, age, and sex.

This is not a standalone bias module. It is equity content threaded directly into the authentication and admissibility framework that judges are already being taught.

Proposed Learning Outcome

> *Recognize the documented limitations and differential reliability of AI detection and identification tools — including facial recognition systems — across demographic groups, and consider these factors when evaluating expert testimony and authentication arguments.*

Why This Matters for Judges

The 2025 Judicial Conference session on AI and the Law (the presenting judge) established a conceptual foundation: judges must become gatekeepers for AI-generated evidence, navigate authentication challenges, and understand the evidentiary "Catch-22" of unacknowledged deepfakes.

This July workshop moves from that awareness into hands-on application. One critical dimension the conference did not develop in depth: the AI tools judges are being asked to evaluate do not perform equally across all populations. Facial recognition systems, in particular, have been extensively documented to exhibit higher false positive rates for darker-skinned individuals, women, and younger people.

A judge who does not know this cannot properly:

- Evaluate expert testimony about a facial recognition match used as an investigative lead
- Assess the probative value vs. prejudicial effect of AI-mediated identification evidence
- Ask informed questions during authentication challenges under MR 5-901 or the proposed Fed. R. Evid. 901(b)(9)
- Instruct juries appropriately on the limitations of such evidence

Evidence Base: Key Sources

Federal Testing Data

NIST Face Recognition Vendor Test (FRVT) — Demographic Effects (NISTIR 8280, with updates through 2025). The most comprehensive federal testing program for facial recognition algorithms. Documents differential false positive rates across demographic groups. Maintained as a living resource with updated summary tables.

Foundational Research

Buolamwini & Gebru, "Gender Shades" (2018) — MIT Media Lab. Demonstrated that commercial facial analysis systems had dramatically higher error rates for darker-skinned women compared to lighter-skinned men. Foundational study that catalyzed industry and policy responses.

Legal-Policy Analysis

Georgetown Law Center on Privacy & Technology — "The Perpetual Line-Up" (scope of law enforcement facial recognition networks) and "A Forensic Without the Science" (courtroom-facing reliability critique, including compounding human + machine error).

Documented Wrongful Identification Cases

The following cases represent documented instances where facial recognition technology contributed to wrongful arrests. Each presents a distinct teaching opportunity for judicial education:

Case	Jurisdiction	What Happened	Teaching Hook
Robert Williams	Detroit, MI	Arrested based on a faulty FR match; held for 30 hours. Charges dismissed. ACLU lawsuit and settlement.	What safeguards should exist between an FR "match" and an arrest? How does a judge evaluate the investigative chain?
Porcha Woodruff	Detroit, MI (2023)	8 months pregnant; arrested on robbery/carjacking warrant generated from FR match. Charges dropped.	When FR is the sole or primary investigative lead, what is the probative value? What corroboration should be required?
Nijeer Parks	Woodbridge, NJ	Wrongly accused of shoplifting and assault based on FR match. Spent 10 days in jail. Case dismissed.	How should courts evaluate FR evidence when the defendant challenges the match accuracy? What expert testimony is needed?

How This Threads Into the Workshop

The equity content enters through the existing program structure, not as an addition:

Morning: AI Fundamentals	A faculty member and/or a second faculty member cover how facial recognition systems work, including training data, algorithmic decision-making, and known performance variation across demographic groups. This is presented as a reliability variable, not a social commentary.
Midday: Authentication	A faculty member and a second faculty member frame authentication challenges for AI evidence, including how to evaluate expert testimony about tool reliability. The differential error rate data becomes part of the admissibility analysis.
Afternoon: Mock Hearing	The family law hypothetical can incorporate an FR-related exhibit or expert testimony where tool reliability across demographic groups is contested. Judges experience the authentication challenge in real time.
Closing: Bench Card	The bench card includes questions about tool validation, known error rates, demographic testing, and corroboration requirements — practical takeaways grounded in the session.

Framing Discipline

This content is strongest when framed as evidentiary competence. Recommended language:

Use This Language

- Reliability and validity of AI-mediated identification evidence
- Differential error rates and probative value
- Automation bias and downstream contamination
- Procedural accuracy safeguards
- Risk management in evidentiary gatekeeping

Avoid Leading With

- DEI training
- Anti-racism workshop
- Implicit bias module
- Equity lecture

These aren't wrong — they're just not necessary here. The pilot's power is that it's unavoidably relevant to competent judging.