DATA TALK · GDI FORUM

Facial biometric data at the crossroads of privacy and security

Nuno Gonçalves
DEEC · Faculty of Sciences and Technology, University of Coimbra | VisTeam · Institute of Systems and Robotics – Coimbra

nunogon@deec.uc.pt · visteam.isr.uc.pt

THE CROSSROADS

One technology, two simultaneous demands

Security

- Border control and travel documents (ICAO)
- Online services, banking, public administration
- Prevention of identity fraud
- **Deep learning models require exceptionally large volumes of facial training images**

Privacy

- Biometric data: a special category of personal data (GDPR Art. 9)
- Irrevocable: no one changes their face the way they change a password
- Risk of re-identification and purpose creep
- **Citizens' trust is a precondition for the research itself**

More data → better models. Less data → more privacy. ***Biometrics research lives at this crossroads.***

BY THE NUMBERS

The scale of the data — and of the consent gap

260.9M

faces in WebFace260M

4.0M identities · all web-scraped

30 bn

faces scraped by Clearview AI

≈ 5 images per person on Earth

10M

MS-Celeb-1M images — deleted

~100K people · retracted 2019

1,016

people in our DFIC — who said yes

58,633 images · consented

MegaFace: 3,311,471 Flickr photos

100% required attribution — none was given; 69% barred commercial use. Decommissioned 2020.

The fairness cost of imbalance

NIST FRVT: 10×–100× more false positives across demographic groups. Gender Shades: 34.7% vs 0.8% error.

WHY STREAMLINE DATA ACCESS

The race against the forgers

Attacks in constant evolution

Face morphing, deepfakes, presentation attacks (masks, screens, printouts), and document fraud.

Generation tools within reach of any attacker

Forgers do not ask for consent, do not wait for ethics boards, and do not publish their data.

Defense depends on real, recent data

Data-access cycles that are too slow leave detectors trained for yesterday's attacks.

Streamlining is not loosening: we need fast, responsible data-access pathways with safeguards — not fewer safeguards.

STANDARDIZATION

Standards as the infrastructure of data

ISO/IEC 39794-5

Facial image data format for travel documents; ICAO transition underway from ISO/IEC 19794-5.

ISO/IEC 29794-5:2025

Facial image quality (FIQA): common metrics, published in 2025.

ISO/IEC 30107-3

Presentation attack detection (PAD): test methodology and metrics.

ICAO Doc 9303 · CT226

Requirements for travel documents; in Portugal, CT226 (Biometrics) mirrors ISO/IEC JTC 1/SC 37.

Common formats, metrics, and protocols are the "I" of FAIR in action: they make biometric data interoperable, comparable, and reusable across labs, industry, and States.

LEGISLATION

The legal framework shapes how data is designed

GDPR

- Art. 9 — biometric data as a special category: processing prohibited by default
- Exceptions: explicit consent; scientific research with safeguards (Art. 89)
- Art. 17 — right to erasure ("right to be forgotten")

EU AI Act

- Biometric systems generally classified as high-risk
- Data governance requirements: quality, representativeness, and bias examination
- Documentation, traceability, and human oversight

Practical consequence: informed consent, minimization, impact assessments, and ethics are part of any biometric dataset's design — from day one, not as a final patch.

FAIR DATA

FAIR when the data are people's faces

F

Findable

Persistent identifiers, rich metadata, publication, and indexing of the datasets.

A

Accessible

Access mediated by license and consent — controlled and auditable, never simply open.

I

Interoperable

Standardized formats and metrics (ISO/IEC, ICAO); shared evaluation protocols.

R

Reusable

Clear licenses and purposes, documented provenance and consent, benchmark partitions.

"As open as possible, as closed as necessary" — in biometrics, the necessary is considerable.

CASE 1 · DATASET

DFIC — Diverse Face Images Coimbra

A facial dataset designed for ICAO compliance

- ICAO compliance annotations with graded severity
- Eight-region facial segmentation masks
- Balanced demographic partitions for fairness analysis
- Paired high- and low-quality captures + video with head rotation
- **GDPR consent and ethics approval from the design stage**

Status
In finalization and submission; derived works already accepted for publication.

FAIR in practice
Rich metadata, defined evaluation protocols, license-mediated access.

People first
Demographic diversity as a requirement, not an accident.

DFIC-HC — head coverings and respect for religion

The problem: automatic ICAO compliance checkers tend to reject photos with religious head coverings — a practical discrimination in access to identity documents.

The response: the DFIC-HC dataset, with religious and non-religious head coverings, and a verification method trained to tell them apart from non-admissible occlusions.

≈ 39.7%

State-of-the-art HTER (ICAONet) on religious coverings

≈ 3.5%

HTER of our method, trained on DFIC-HC (accepted · EURASIP J. Image and Video Processing)

We can respect people's religion without sacrificing automatic verification quality — provided the data allow it.

CASE 2 · IJCB 2025

VOIDFace — privacy by construction in training

The risk

Face recognition models are vulnerable to model inversion and training-data replication: faces can be recovered.

The idea

Each image is split into facial patches, encrypted via visual secret sharing and distributed: no training node stores or reconstructs the full face.

Enforceable right to erasure

Deleting the authentication share removes the subject from future training (GDPR Art. 17); VOIDFace-X (2026) adds verifiable model unlearning.

Data management as architecture: protection is not in a contract — it is in the way the training data themselves are stored and used.

CASE 3 · HORIZON EUROPE

ACHILLES — legally compliant, auditable AI

16	€9.6 M	EU
European partners	funding	Horizon Europe

VisTeam's contributions to fairness in facial biometrics:

- Measurement and mitigation of bias in face recognition systems.
- Skin-tone detection and estimation — an objective basis for anti-discrimination policy.
- Mitigation of bias in automatic descriptions of human faces by LLMs/VLMs.
- Integration of VOIDFace: privacy and fairness handled together, not in competition.

CASE 3 · IWBF 2026

Identity unlearning — the right to be forgotten, measured

The requirement: GDPR grants a right to erasure — but a face that shaped a trained model does not disappear when its images are deleted. Its influence stays encoded in the weights.

Forget one identity, keep the rest

Fine-tune a trained model as if it had never seen the target person — not by destroying accuracy, but by making that identity statistically indistinguishable from identities the model never encountered.

A metric for "how well did we forget?"

Our USDTO metric (Unlearning Similarity Distribution Trade-Off) scores the balance between erasing the target and preserving the feature space for everyone else — computed from templates, no class labels needed.

Fast, and enough at step one

Two strategies compared (classification- and contrast-confusion); sufficient unlearning is reached after a single update — efficient enough for real deployment (ACHILLES).

Medvedev, Muhammed & Gonçalves, "Feature-based Identity Unlearning Evaluation for Biometric Applications," IWBF 2026 · ISR-Coimbra · funded by ACHILLES (Horizon Europe, GA 101189689).

SYNTHESIS

We do FAIR — even without calling it FAIR (work in progress)

Our practice	Principle
Publishing datasets with identifiers and rich metadata (papers, repositories, arXiv)	**F — Findable**
Access mediated by license/EULA, with documented consent and purpose	**A — Accessible (controlled)**
ISO/IEC and ICAO formats and metrics; evaluation protocols shared by the community	**I — Interoperable**
Defined benchmark partitions, recorded provenance, explicit reuse conditions	**R — Reusable**

The biometrics community has practiced FAIR principles for years, at least partially — out of scientific and legal necessity.
What we mostly lack is systematizing it.

Synthetic faces — the way out, and its new risks

The promise: faces of people who never existed sidestep the consent problem entirely — no subject to notify, no right to be forgotten to enforce. This is where the field is heading.

Collision with real people

- As synthetic faces multiply, some inevitably resemble real individuals — risking false identification
- Generators may leak or memorize real training identities
- "Fake" and "real" faces start to blur — a problem for both recognition and forensics

Not yet fully representative

- Generators inherit the demographic gaps of the data they learned from
- Human diversity — age, ethnicity, conditions — is not yet faithfully reproduced
- Training on synthetic data can amplify existing bias rather than remove it

Synthetic data is necessary, not sufficient: *it must be validated against real, consented, diverse benchmarks — exactly the role datasets like DFIC play.*

VULNERABLE GROUPS

The people hardest to represent — and to consent

Older adults
Ageing changes faces; often under-sampled

Children
Fast-changing faces; consent must come from guardians

Patients & disabled
Medical conditions rarely represented in datasets

Under-served regions
Whole populations thinly covered by public data

Under-representation
The groups most likely to be misidentified are the ones least present in training data — and often those most affected by errors (border control, healthcare, access to services).

Consent is genuinely hard
Minors, cognitively impaired or ill people cannot always give informed consent; guardianship, capacity and withdrawal make lawful, ethical collection far more complex.

Fairness and consent pull in opposite directions here: *the data we most need for equity is the data hardest to gather ethically.*

CONCLUSION

Four messages to take away

1. Privacy and security are not opposites: they are simultaneous design requirements for both the data and the models.

2. Standards, laws and norms are essential infrastructure — not obstacles — that make biometric data interoperable, fair and reusable.

3. **But process is now the bottleneck: bureaucracy is measurably slowing our ability to protect systems and improve their performance. Accelerating responsible data access is urgent.**

4. FAIR in biometrics means as open as possible, as closed as necessary — with the "necessary" well documented, and the pathway to access fast.

Thank you.

Questions & Answers

Nuno Gonçalves
DEEC · FCTUC | VisTeam · Institute of Systems and Robotics – Coimbra
nunogon@deec.uc.pt

visteam.isr.uc.pt
datasets · papers · team

Scan for our datasets, papers and team

BACKUP · FOR Q&A

Data behind the numbers

Dataset sizes · consent & retractions · demographic bias · deepfake/morphing sets · our datasets

Face-recognition training datasets

Dataset	Year	Images	Identities	Note
WebFace260M	2021	260,890,076	4,008,130	Largest public; web-scraped
WebFace42M (clean)	2021	42,474,558	2,059,906	Largest cleaned set
Glint360K	2020	17,091,657	360,232	Largest non-retracted
MS-Celeb-1M	2016	~10,000,000	~100,000	Retracted 2019
MegaFace	2016	4,753,320	672,057	Decommissioned 2020
VGGFace2	2018	3,310,000	9,131	Withdrawn — privacy
CASIA-WebFace	2014	494,414	10,575	Early workhorse set
LFW	2007	13,233	5,749	Historic small benchmark
DFIC (ours)	2026	58,633	1,016	Consented · GDPR

Private sets, for scale: *Facebook DeepFace ~4.4M / ~4,030 · Google FaceNet ~200M / ~8M (reported in original papers; not publicly released).*

Consent, retractions and licensing

MS-Celeb-1M — retracted (2019)
~10M images of ~100K people. Pulled after a Financial Times investigation and the Exposing.ai analysis showed journalists, artists and academics were included without consent.

MegaFace — licensing violations
3,311,471 Flickr photos: 100% required attribution (none given), 69% barred commercial use. Included photos of children (potential BIPA issues). Decommissioned 2020.

Clearview AI — 30 billion faces
Confirmed by the Dutch DPA (2024), which called facial recognition "highly intrusive"; €30.5M fine. No consent, no notification, no usable opt-out.

Deletion is not enough
A Princeton study (2021) found retracted sets such as MS-Celeb-1M and Duke MTMC persist in derivative datasets and hundreds of citations.

Demographic bias and benchmark scale

NIST FRVT Part 3 (2019)

- Test scale: 18.27M images of 8.49M people; ~189 algorithms evaluated
- 10×–100× higher false positives for Asian and African-American faces (1:1)
- Highest false positives for African-American women (1:N) — the most consequential case

Gender Shades (2018)

- Error 34.7% (darker-skinned women) vs 0.8% (lighter-skinned men)
- Benchmarks audited were 79.6% (IJB-A) / 86.2% (Adience) lighter-skinned
- FairFace (2021): 108,501 images balanced across 7 race groups — a corrective

The upstream cause is the training data: imbalance in what the models see becomes disparity in what they do. This is what DFIC's balanced partitions and ACHILLES directly target.

Deepfake and manipulation datasets

DFDC (Facebook, 2020)

128,154 clips from 3,426 paid, consenting actors — the largest public face-swap video set, built on explicit consent.

FaceForensics++ (2019)

1,000 real videos + 4,000 manipulated (DeepFakes, Face2Face, FaceSwap, NeuralTextures).

Celeb-DF v2 / DeeperForensics

Celeb-DF: 590 real + 5,639 synthesized. DeeperForensics-1.0: 60,000 videos / 17.6M frames.

Proliferation (Sensity, 2019)

Deepfake videos online rose from 7,964 to 14,678 in nine months (~84%); 96% non-consensual pornography.

Note the contrast: *DFDC solicited consent from every actor — the ethical inverse of the scraped recognition sets.*

Our datasets and projects — the sources

DFIC — Diverse Face Images Coimbra

58,633 images + 2,706 videos, 1,016 subjects; >2,000,000 annotations across 26 ICAO categories. Balanced 50/50 gender; ~34% Asian / 35% White / 31% African. Ethics CEIUC_11/R-9; GDPR-compliant. arXiv:2602.10985 (2026).

DFIC head coverings

Guerra, Marcos & Gonçalves, "Automatic validation of ICAO compliance regarding head coverings," BIOSIG 2023 (IEEE). Head coverings = ICAO requirement #8, highest non-compliant share (21.54%).

VOIDFace

Muhammed, Medvedev & Gonçalves, arXiv:2508.07960 (IJCB 2025). Visual secret sharing to prevent training-data replication and enforce GDPR Right-To-Be-Forgotten; evaluated on VGGFace2.

Identity Unlearning

Medvedev, Muhammed & Gonçalves, "Feature-based Identity Unlearning Evaluation for Biometric Applications," IWBF 2026. Introduces the USDTO metric; classification- vs contrast-confusion strategies on MS-Celeb-1M / RFW.

ACHILLES

Horizon Europe, GA 101189689. 16-partner consortium; privacy-by-design, explainable, fair, demographically balanced models with identified data sources — auditability under the EU AI Act.