Monsters of the DiD

Jon Crowcroft

2/3/26

Identity friction

UX – not a number

Re-Decentralised

“We live in two worlds…

the world into which we were born, and the otherworld that was born within us.

Both may be a blessing or a curse. We choose.” – Druid saying.

NO2ID

Stop the Database State

Very Big Id Systems of the World

e.g.

AADHAAR

मेरा आधार
मेरी पहचान

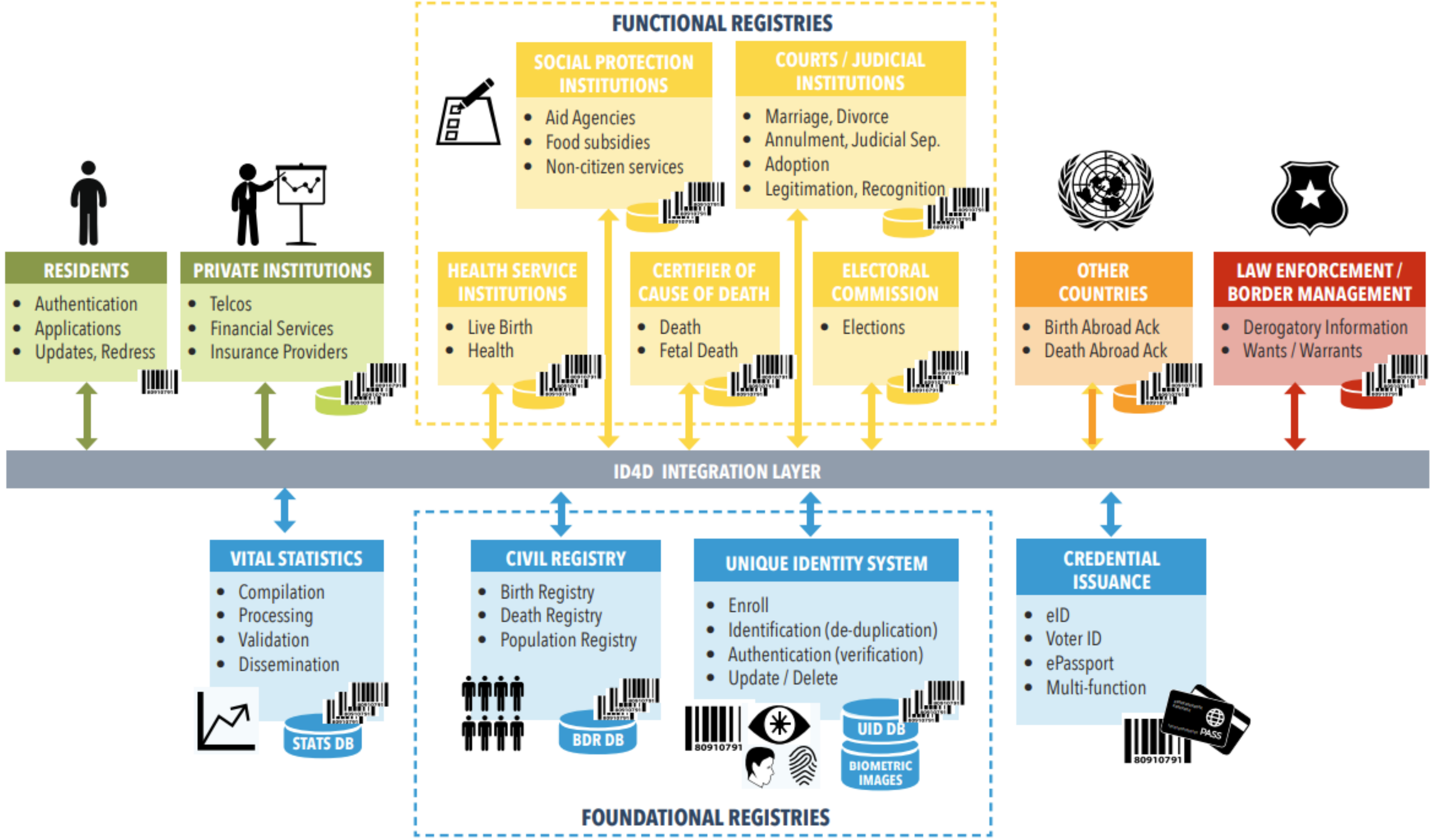
What makes you you? What is identity

- = or == ?
    - The Ship of Theseus
    - Trigger’s Broom
- Heraclitus – objective context
    - You can’t step in the same river twice
- Locke – cognitive context
    - Memories (solipsistic)
- ***You*** are who ***they*** say you are – personal context
    - E.g. you are always your parents child (etc) (at least til you take over care/attorney!)
    - Petal Kelly
- Crises
    - Cultural context
- Economicrises
    - metrics
- A2D
    - Analog – Mind/Body (Sajnani) ->
    - How does being digital change anything?
    - Mind/Body/Bits/Body/Mind journey

Identity & Self

“The whole of this doctrine leads us to a conclusion, which is of great importance in the present affair, viz. that all the nice and subtile questions concerning personal identity can never possibly be decided, and are to be regarded rather as grammatical than as philosophical difficulties. Identity depends on the relations of ideas; and these relations produce identity, by means of that easy transition they occasion. But as the relations, and the easiness of the transition may diminish by insensible degrees, we have no just standard, by which we can decide any dispute concerning the time, when they acquire or lose a title to the name of identity. ” David Hume

DiD Basics:: Functional vs Foundational



FUNCTIONAL REGISTRIES

SOCIAL PROTECTION INSTITUTIONS
- Aid Agencies
- Food subsidies
- Non-citizen services

COURTS / JUDICIAL INSTITUTIONS
- Marriage, Divorce
- Annulment, Judicial Sep.
- Adoption
- Legitimation, Recognition

RESIDENTS
- Authentication
- Applications
- Updates, Redress

PRIVATE INSTITUTIONS
- Telcos
- Financial Services
- Insurance Providers

HEALTH SERVICE INSTITUTIONS
- Live Birth
- Health

CERTIFIER OF CAUSE OF DEATH
- Death
- Fetal Death

ELECTORAL COMMISSION
- Elections

OTHER COUNTRIES
- Birth Abroad Ack
- Death Abroad Ack

LAW ENFORCEMENT / BORDER MANAGEMENT
- Derogatory Information
- Wants / Warrants

ID4D INTEGRATION LAYER

VITAL STATISTICS
- Compilation
- Processing
- Validation
- Dissemination

STATS DB

CIVIL REGISTRY
- Birth Registry
- Death Registry
- Population Registry

BDR DB

UNIQUE IDENTITY SYSTEM
- Enroll
- Identification (de-duplication)
- Authentication (verification)
- Update / Delete

UID DB

BIOMETRIC IMAGES

CREDENTIAL ISSUANCE
- eID
- Voter ID
- ePassport
- Multi-function

FOUNDATIONAL REGISTRIES

The Unique ID system and Civil Registry are integrated through the ID4D Integration layer. This architecture preserves the modular nature of the model versus customized direct interactions.

https://www.biometricupdate.com/wp-content/uploads/2015/09/ID4D-report-World-Bank-Accenture.pdf

Uniqueness – Is that needed?

	Uniqueness Matters	Uniqueness is Less Critical
	Foundational ID systems (e.g. national ID, civil registry)	**Loyalty programs / memberships**
Functional ID Systems	**Voter registration & elections**	**Transportation cards / event passes**
	Social protection/welfare programs	**E-commerce or social media accounts**
	Subsidy distribution (e.g. food, fuel, cash transfers)	**Education records (basic)**
	Immigration/border control	**Digital login/auth systems** (e.g. OAuth, SSO)
	Pension/employment tracking	**Short-term or anonymous surveys/census**
	Criminal justice	
	Refugee registration / humanit	
	National health insurance s	

"Uniqueness of identity is essential when a system confers rights, benefits, or responsibilities to individuals — where one person must not claim entitlements meant for another.

In contrast, when systems support preferences, pseudonymity, or general usage tracking without legal or social consequences, strict uniqueness becomes less critical"

Non Biometric Models

– For Avoiding Double Spend (so often about money!)

Mechanism	Description
Cryptographic Tokens	Unique, signed tokens (e.g. QR, NFC) issued per user or claim, validated on use
Verifiable Credentials (VCs)	Issued by trusted parties; users prove eligibility or uniqueness cryptographically
Zero-Knowledge Proofs (ZKPs)	Users prove they haven’t claimed before, without revealing identity
Fuzzy Demographic Deduplication	Matching individuals based on partial or noisy personal data (name, DOB, etc.)

Pros and Cons

Mechanism	Pros	Cons
Cryptographic Tokens	- Easy to deploy offline- Low-cost- Strong one-time use control	- Requires secure issuance and tracking- Token loss = benefit loss
Verifiable Credentials	- Privacy-preserving- Reusable & portable- Enables selective disclosure	- Requires digital wallets & trusted issuers- Not trivial to deploy at scale
Zero-Knowledge Proofs	- Maximum privacy- Strong uniqueness without identity disclosure	- Technically complex- Harder to implement in low-tech environments
Fuzzy Demographic Matching	- No biometric or advanced tech needed- Scalable for existing databases	- Prone to errors (false matches or misses)- Lower trust in low-data-quality contexts

Identity Collision

- Capacity of biometrics crucial for deduplication.
- Analogous to “birthday problem” – how many people must be assembled before it is likely two share the same birthday.
- Theoretical capacity of IrisCode (245-bits, Daugman 2024) and false match rate of 10^{-20} enough to prevent identity collision of 12 billion persons.
- Biometric entropy of other modes is lower e.g. 82-bits for fingerprints (Young et al. IJCSET 2013), and 40-bits for face (Adler et al. 2006)

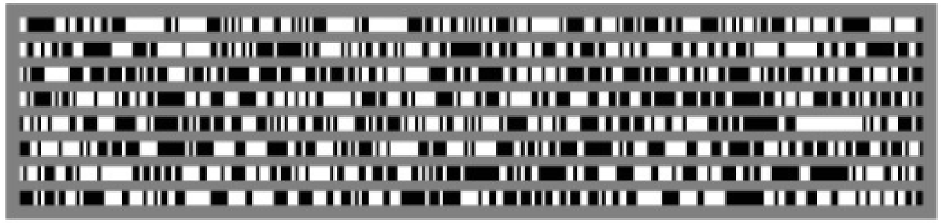
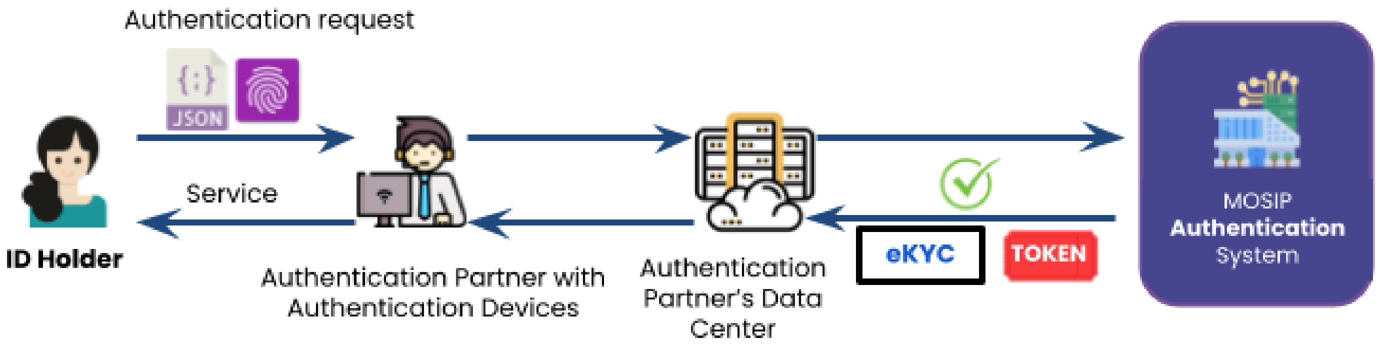


Figure 1: Representation of the IrisCodes [9], produced by four different eyes. The eight rows within each can be regarded as eight concentric rings, each encoding a $[0, 2\pi]$ traversal around the iris. (Eyelid masking is not shown.)

https://oajaiml.com/uploads/archivepdf/802142123.pdf

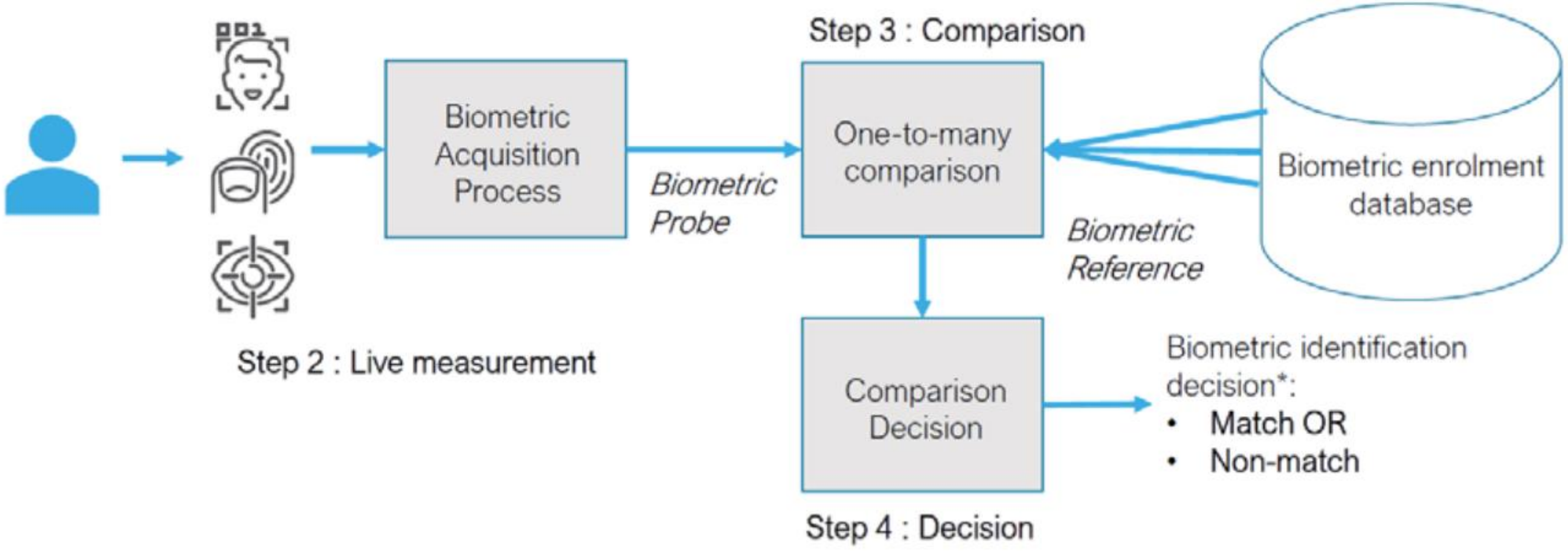
Identity Verification – typical workflow

Authentication Process


Authentication request
JSON
ID Holder
Service
Authentication Partner with Authentication Devices
Authentication Partner's Data Center
eKYC
TOKEN
MOSIP Authentication System

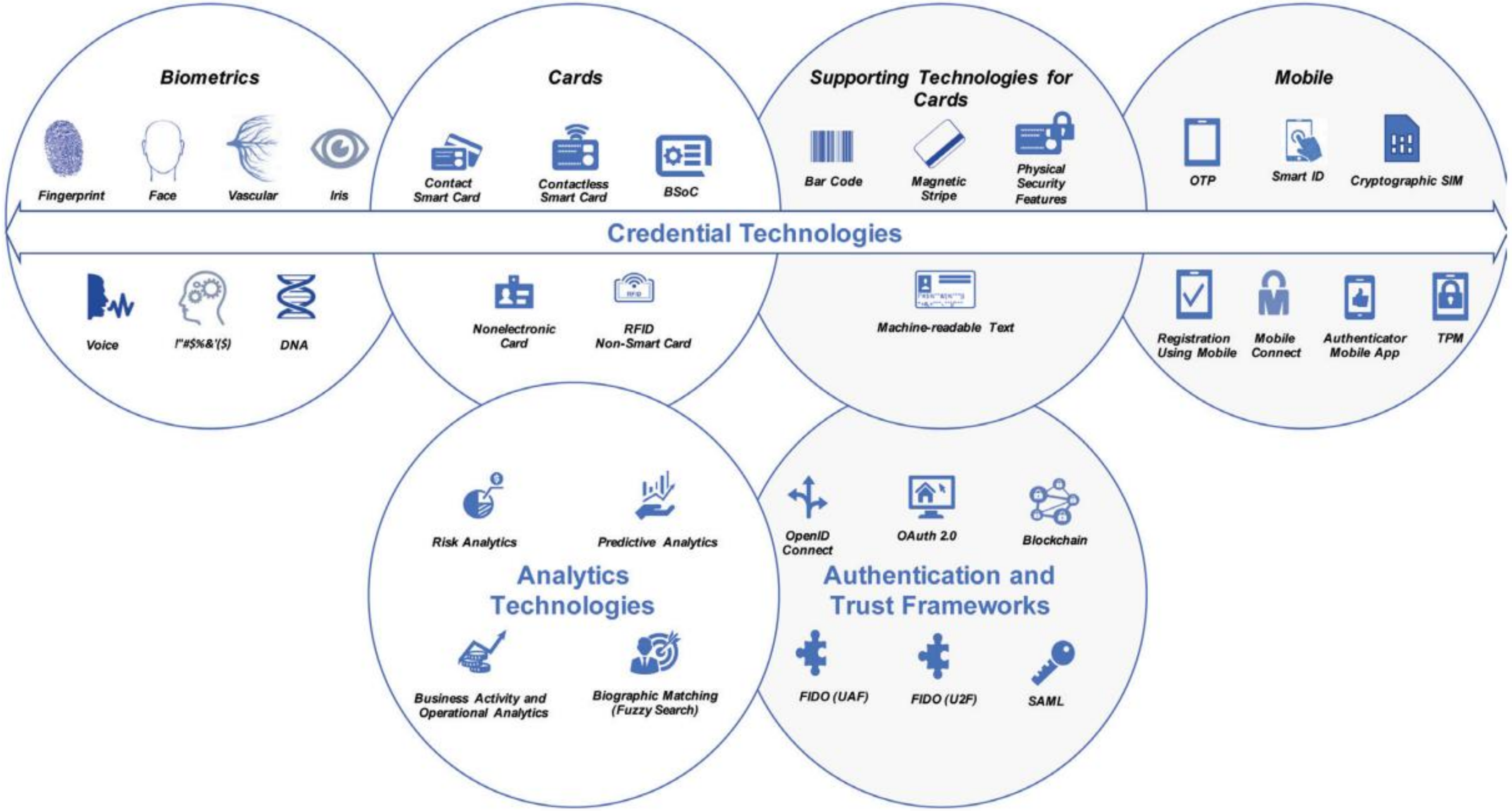
https://docs.mosip.io/1.2.0/id-lifecycle-management/identity-verification/id-authentication

Deduplication/ Uniqueness



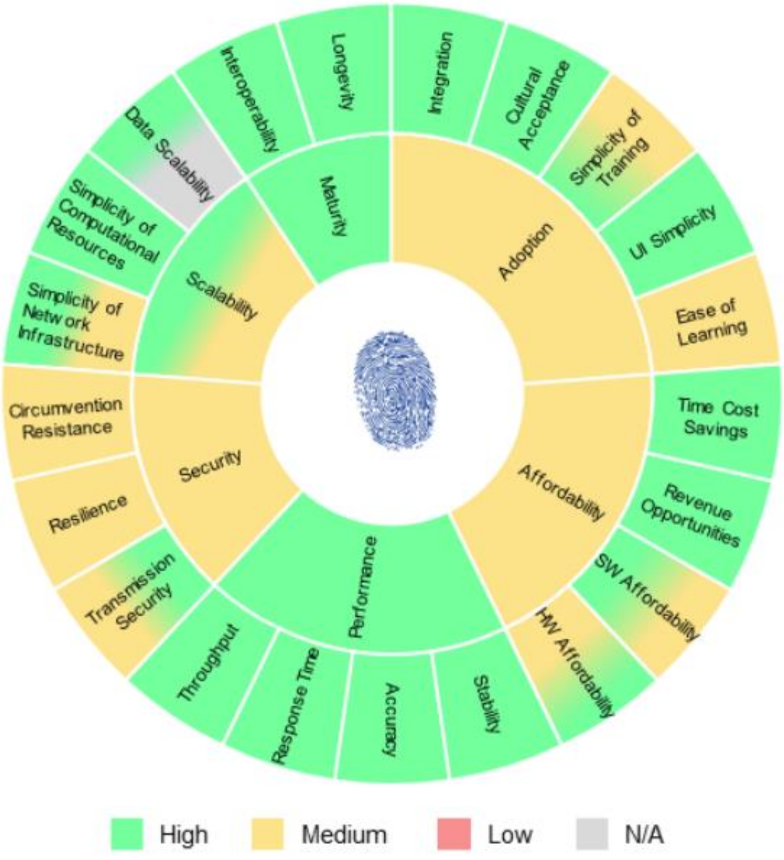
https://documents1.worldbank.org/curated/en/199411519691370495/Technology-Landscape-for-Digital-Identification.pdf

Technology Landscape

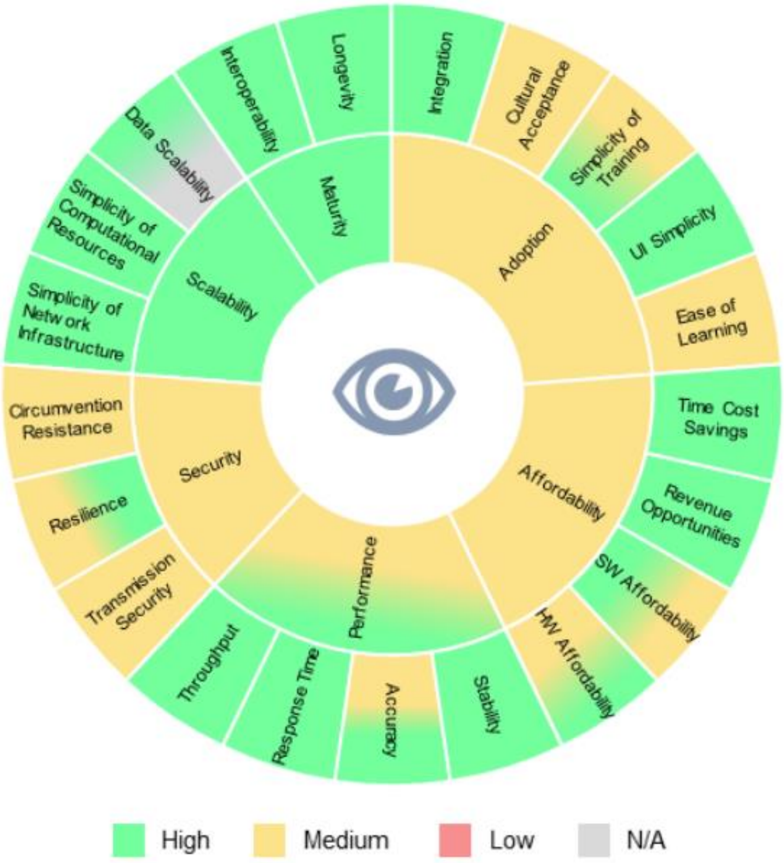


https://documents1.worldbank.org/curated/en/199411519691370495/Technology-Landscape-for-Digital-Identification.pdf

Fingerprint & Iris



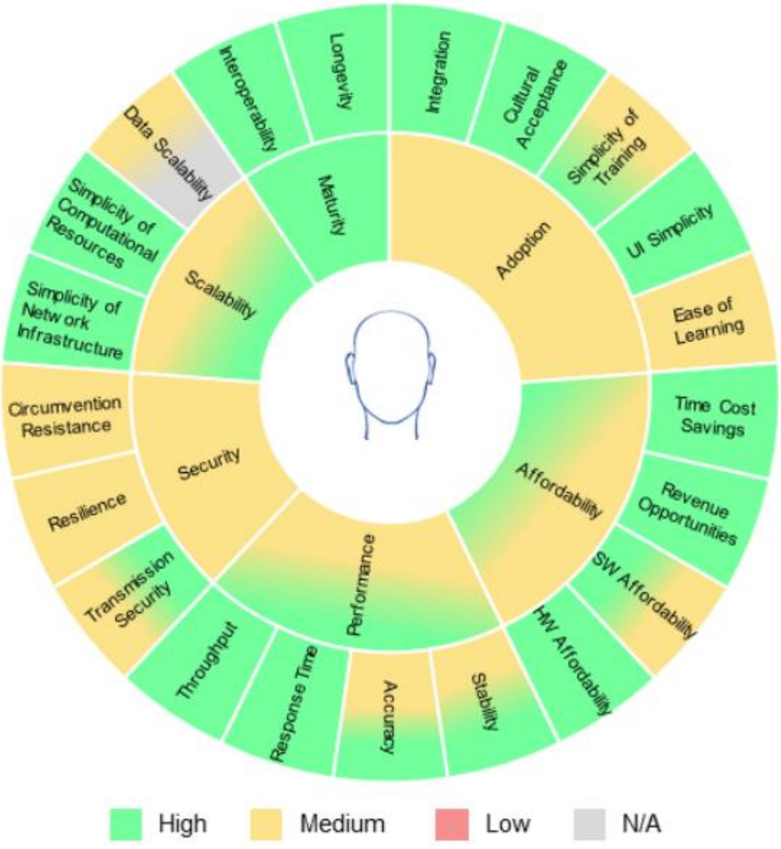
Fingerprint Capture and Matching



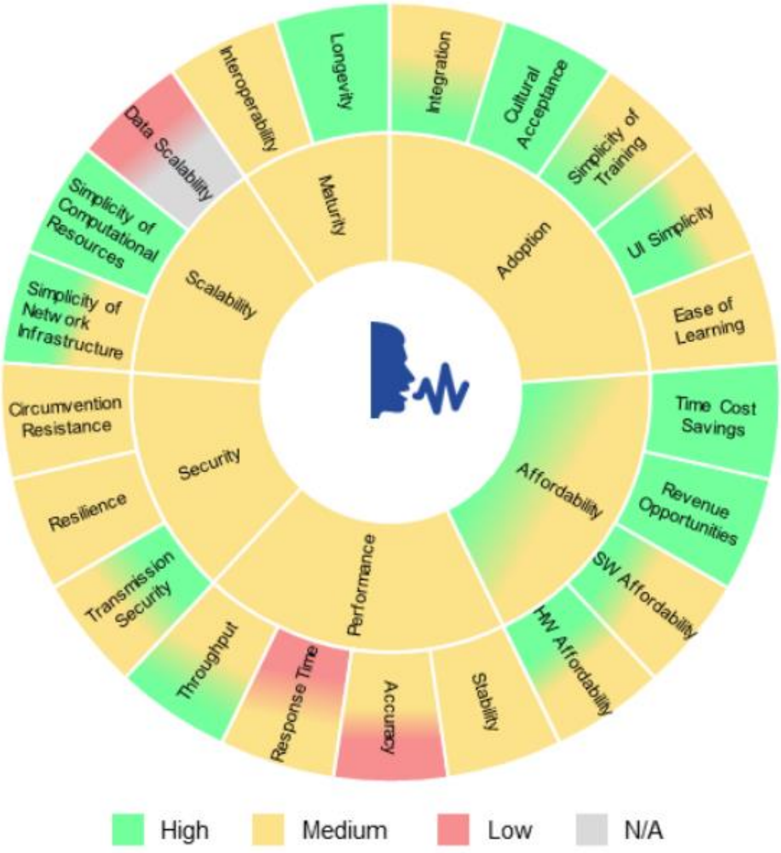
Iris Capture and Matching

https://documents1.worldbank.org/curated/en/199411519691370495/Technology-Landscape-for-Digital-Identification.pdf

Face & Voice



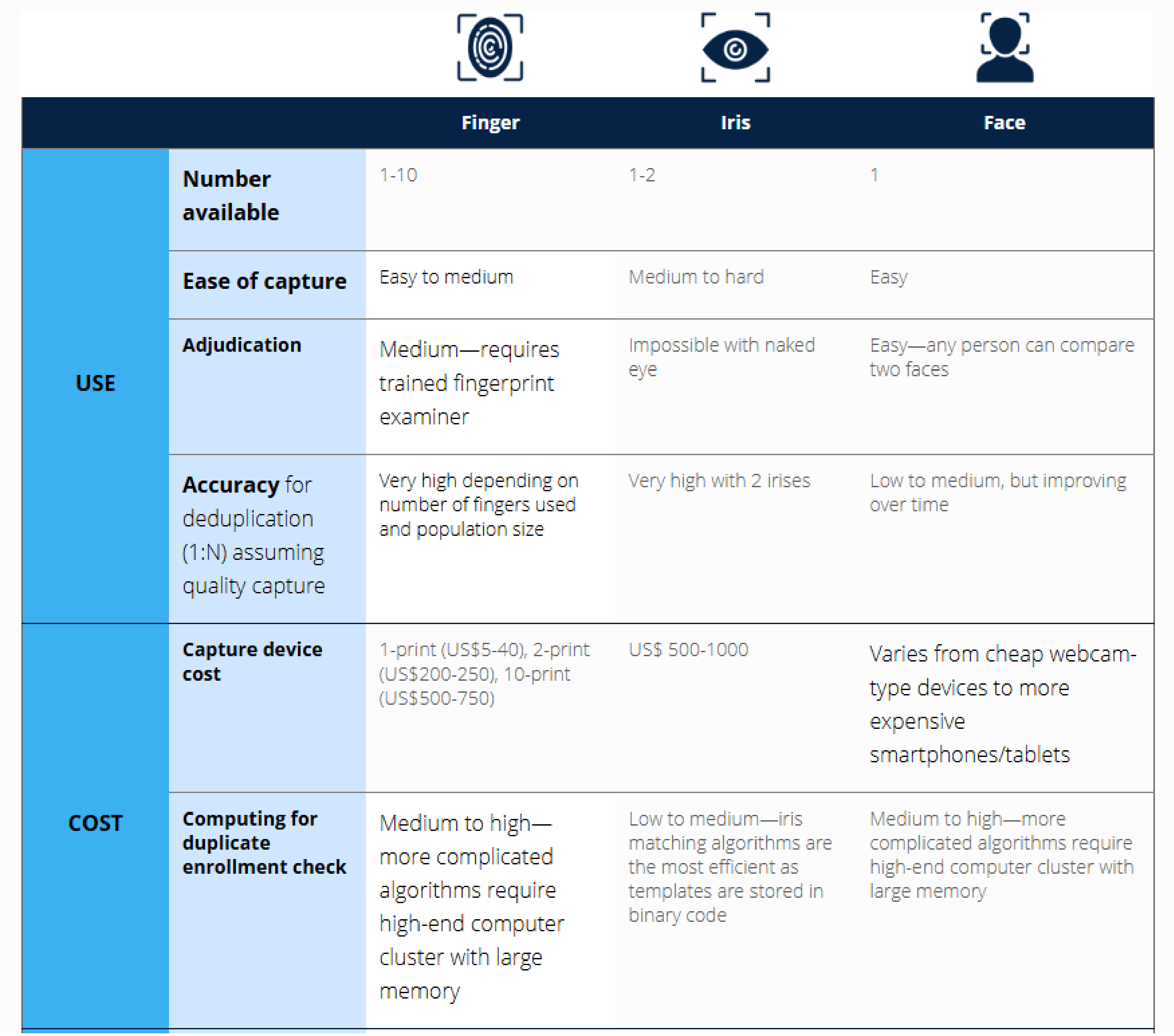
Face Capture and Matching



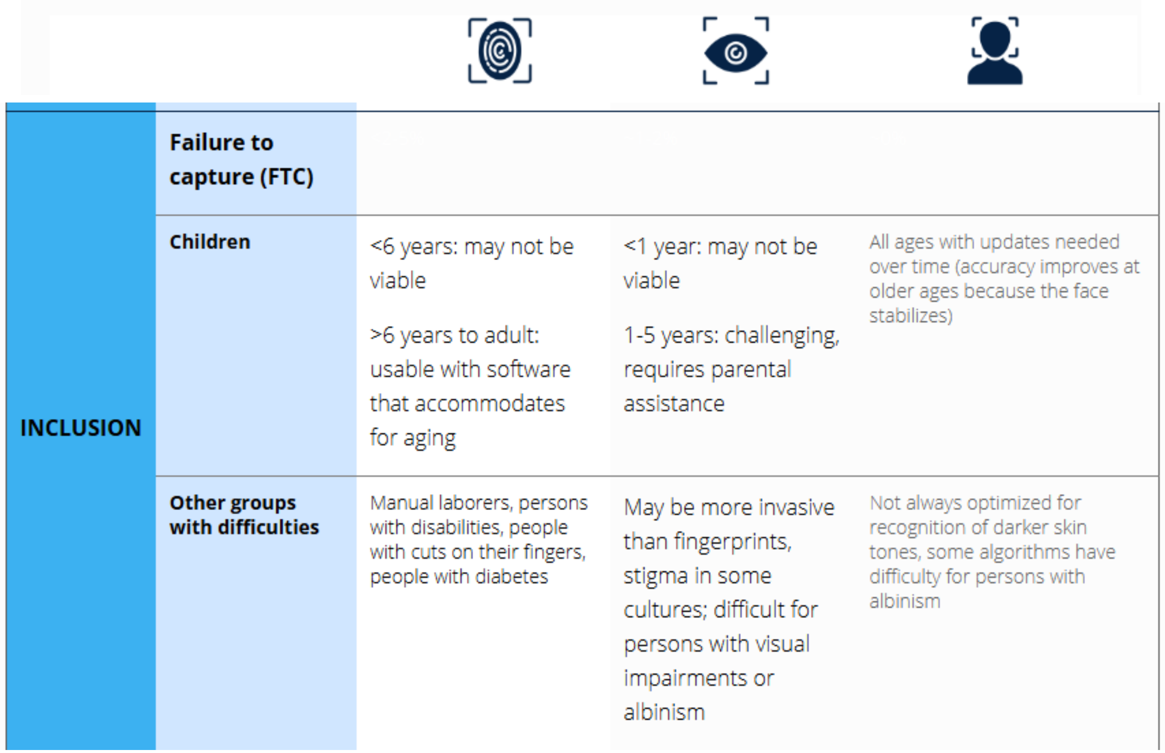
Voice Capture and Matching

https://documents1.worldbank.org/curated/en/199411519691370495/Technology-Landscape-for-Digital-Identification.pdf

Tech Comparison



		Finger	Iris	Face
USE	Number available	1-10	1-2	1
	Ease of capture	Easy to medium	Medium to hard	Easy
	Adjudication	Medium—requires trained fingerprint examiner	Impossible with naked eye	Easy—any person can compare two faces
	Accuracy for deduplication (1:N) assuming quality capture	Very high depending on number of fingers used and population size	Very high with 2 irises	Low to medium, but improving over time
COST	Capture device cost	1-print (US$5-40), 2-print (US$200-250), 10-print (US$500-750)	US$ 500-1000	Varies from cheap webcam-type devices to more expensive smartphones/tablets
	Computing for duplicate enrollment check	Medium to high—more complicated algorithms require high-end computer cluster with large memory	Low to medium—iris matching algorithms are the most efficient as templates are stored in binary code	Medium to high—more complicated algorithms require high-end computer cluster with large memory



INCLUSION	Failure to capture (FTC)			
	Children	<6 years: may not be viable >6 years to adult: usable with software that accommodates for aging	<1 year: may not be viable 1-5 years: challenging, requires parental assistance	All ages with updates needed over time (accuracy improves at older ages because the face stabilizes)
	Other groups with difficulties	Manual laborers, persons with disabilities, people with cuts on their fingers, people with diabetes	May be more invasive than fingerprints, stigma in some cultures; difficult for persons with visual impairments or albinism	Not always optimized for recognition of darker skin tones, some algorithms have difficulty for persons with albinism

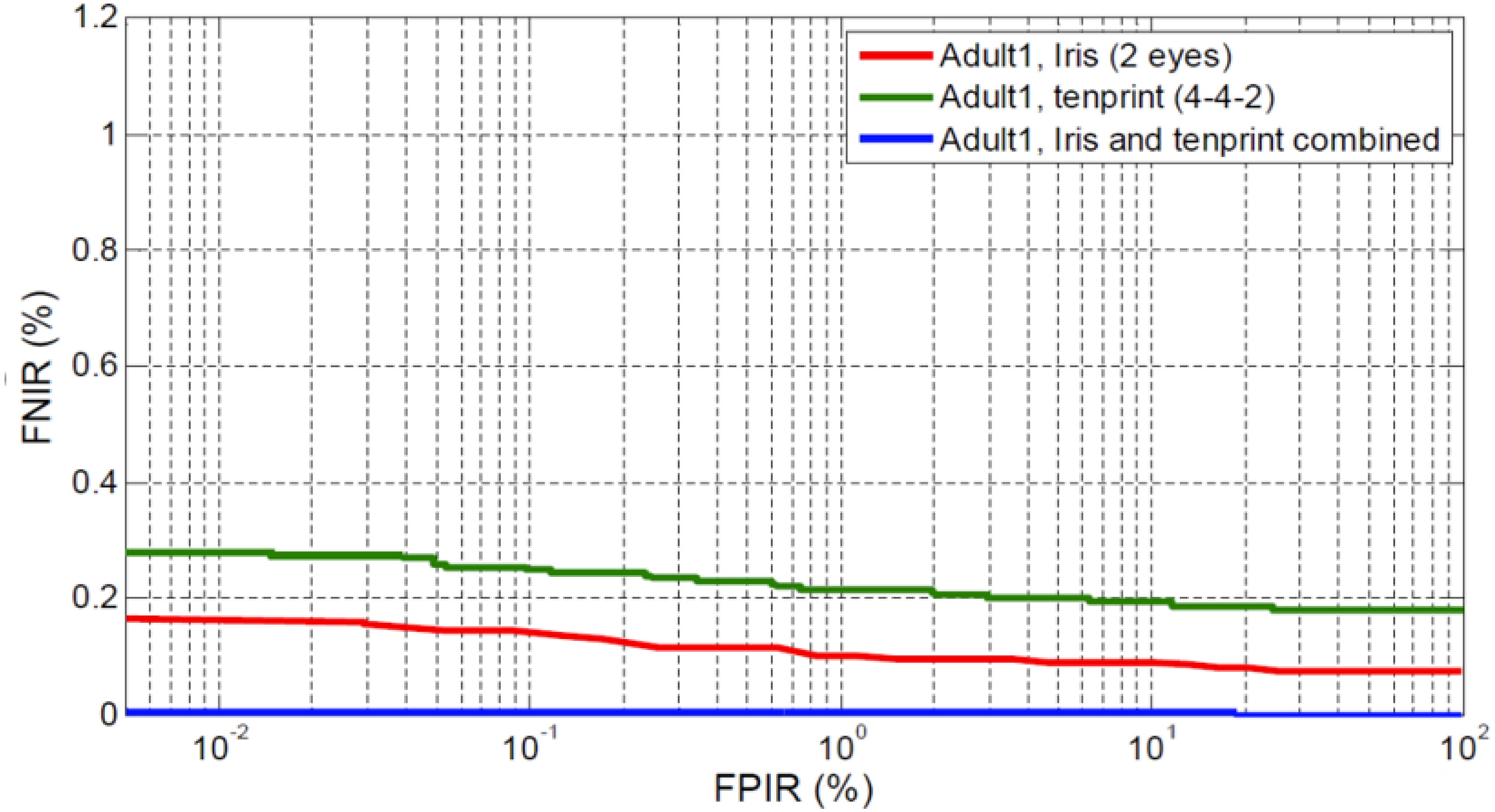
https://documents1.worldbank.org/curated/en/199411519691370495/Technology-Landscape-for-Digital-Identification.pdf

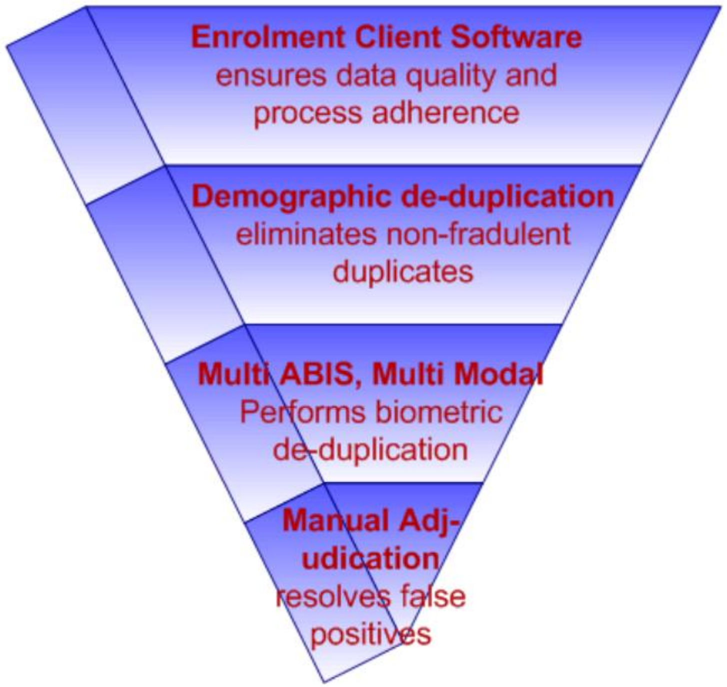
Single vs Multi Modal vs Population

Population Size	Recommended Modalities	Can One Modality (Face) Be Enough?	Evidence / Notes
Small* (< 5 million)	Face **or** Fingerprint	**Yes**, if conditions are controlled	ID4D finds face-only viable in **low-scale, low-risk** cases when capture quality is high and manual review is available
Medium* (5–50 million)	Fingerprint **+** Face	**Possible**, but risk of false duplicates increases	ID4D classifies face deduplication as **low–medium accuracy**, with fingerprint adding stronger identity uniqueness .
Large* (50+ million)	Fingerprint **+ Iris + Face**	**No**, face-only is not reliable	NIST FRVT shows false-positive identification rates (FPIR) for face grow nearly linearly with database size (FPIR ≈ N × FMR) . Best algorithms still struggle with scale. |

Non standard definition

Multimodal Deduplication Accuracy





Enrolment Client Software
ensures data quality and process adherence

Demographic de-duplication
eliminates non-fradulent duplicates

Multi ABIS, Multi Modal
Performs biometric de-duplication

Manual Adj-udication
resolves false positives

UIDAI POC for Enrolment (135000 records)
(2008, but still relevant)

Accuracy Levels – NIST Evaluations

Mechanism	Modality	Accuracy
1:1 Comparison (Authentication) **FAR* @ 0.001%**	Fingerprint	TAR* = 99.56% (Verifinger V12.3)
	Iris	TAR = 99.43% (NIST IREX IX)
	Face	TAR = 99.83 % (NIST FRVT 2022)

1:N Comparison (Identification) **FPIR* = 0.001**	Fingerprint (10 Fingers) Fingerprint (1 Finger)	FNIR* = 0.001 (5M Gallery) FNIR = 0.019 (100K Gallery)
	Iris (Both Eyes)	FNIR = 0.0035 (500K Gallery)
	Face	**FNIR = 0.03 (12M Gallery)**

*

FAR = False Acceptance Rate
TAR = True Acceptance Rate

**

FPIR = False Positive Identification Rate
FNIR = False Negative Identification Rate

- NIST FRVT 1:N Identification: https://pages.nist.gov/frvt/html/frvt1N.html;
- NIST FpVTE: https://nvlpubs.nist.gov/nistpubs/ir/2014/NIST.IR.8034.pdf
- NIST IREX 10 Identification Track: https://pages.nist.gov/IREX10/
- https://biometrics.cse.msu.edu/Presentations/Israeli%20School%20on%20Biometrics%20April%2021-2025-FINAL.pdf

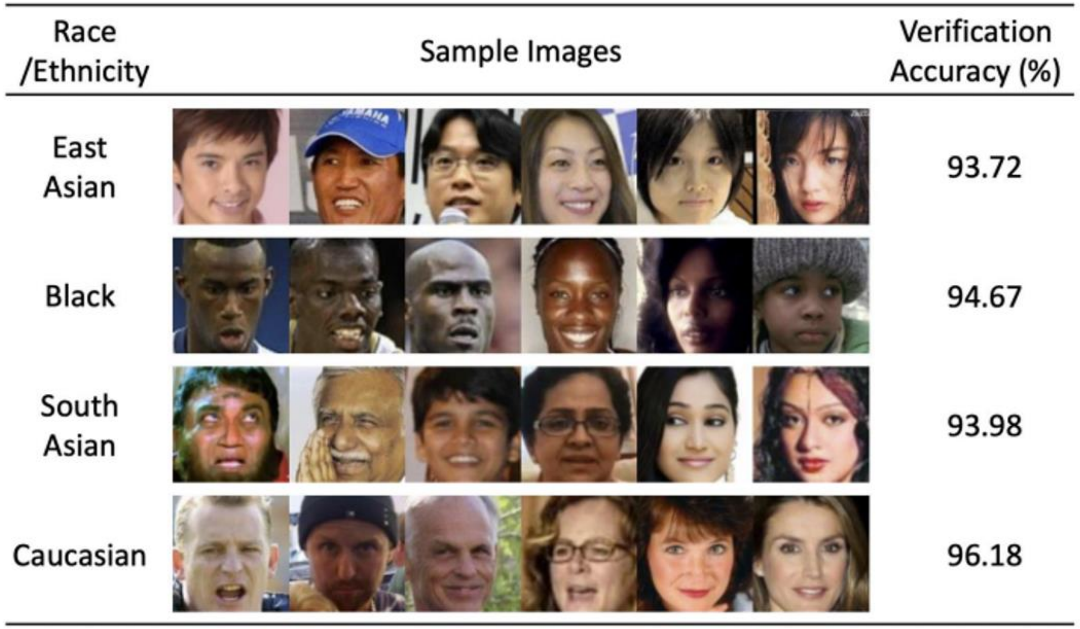
Face – Why not for Large Population?

Particular	Details
Scaling Effect	According to NIST, in a 1:N search, the FPIR approximately equals FMR × N. More manual deduplication requirement.
NIST FRVT evidence	On the 1.6 million-record Visa–Border dataset, NIST fixed FPIR at 0.3%. At that level, top algorithms showed FNIRs of ~0.16% to ~0.72%, missing 1 to 7 duplicates per 1,000.
Demographic biases	Rate varies by age, gender, race:

https://pages.nist.gov/frvt/reports/demographics/implications_for_1N.pdf
https://www.paravision.ai/whitepaper-enterprise-grade-1n-face-recognition/
https://neurotechnology.com/awards-frvt-1-n.html
https://nvlpubs.nist.gov/nistpubs/ir/2019/nist.ir.8280.pdf

Demographic Bias & Age Impact

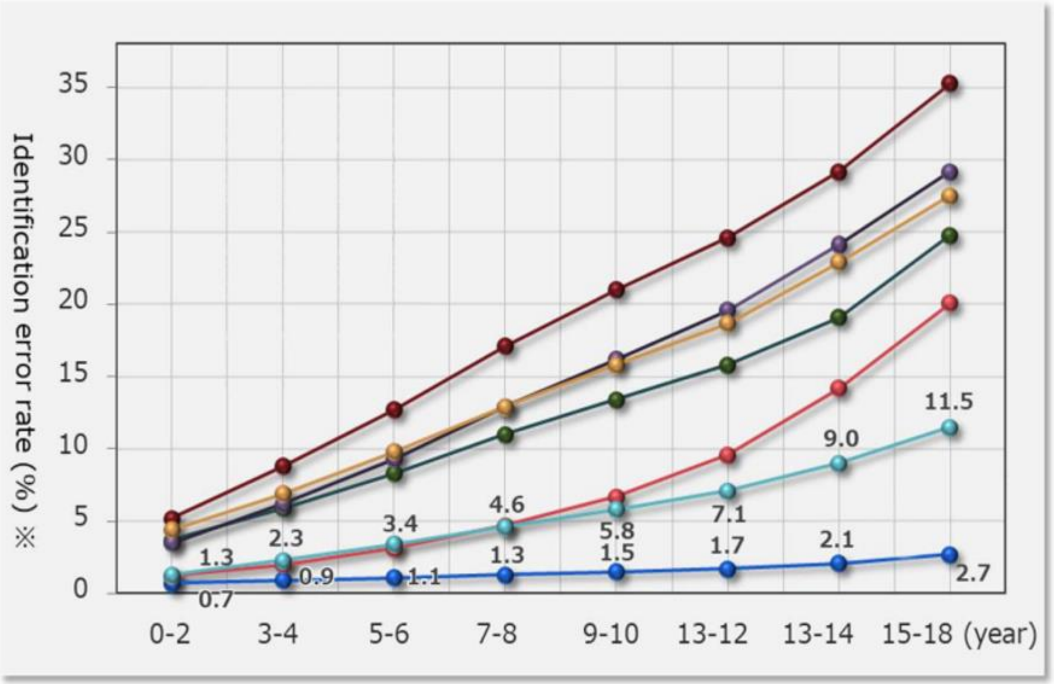
Demographic Bias



Race /Ethnicity	Sample Images	Verification Accuracy (%)
East Asian		93.72
Black		94.67
South Asian		93.98
Caucasian		96.18

Face Verification performance by ArcFace [1] on each race/ethnicity cohort in RFW dataset [2]

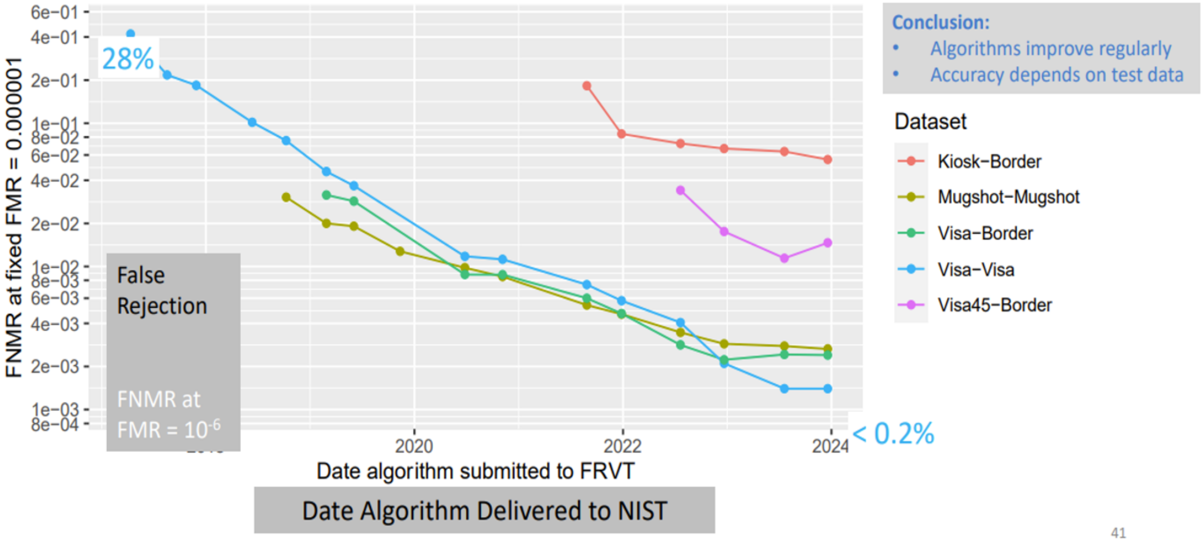
Algorithmm Accuracies over Aging Photos



※ False-rejection discrimination rate at a false acceptance discrimination rate of 0.1% at the time of registration of 31000 people

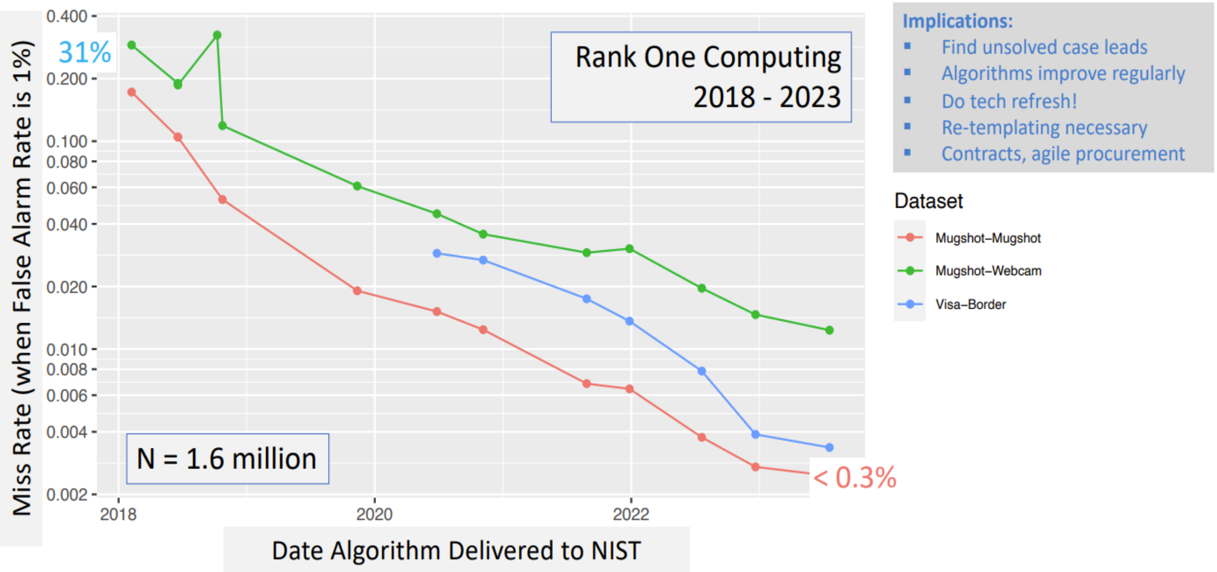
https://biometrics.cse.msu.edu/Presentations/Biometric_Summer_School_2021_Final.pdf
https://biometrics.cse.msu.edu/Presentations/MBZUAI_Sept_1_2020.pdf

Gain Continues for Face



Conclusion:

- Algorithms improve regularly
- Accuracy depends on test data



Implications:

- Find unsolved case leads
- Algorithms improve regularly
- Do tech refresh!
- Re-templating necessary
- Contracts, agile procurement

1:1

1:N

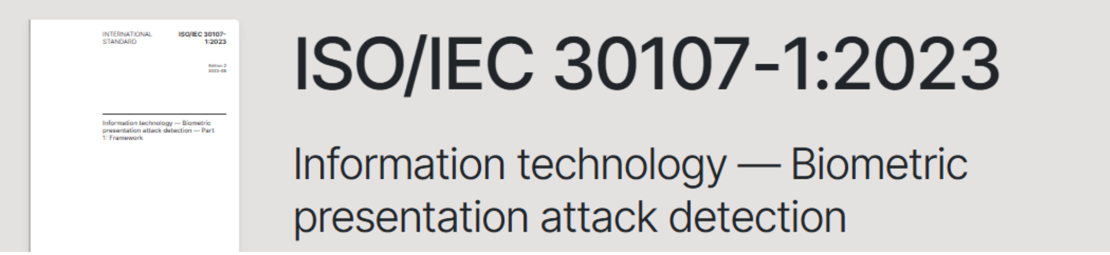
NIST FRVT,
MSU

Other Areas

Liveness

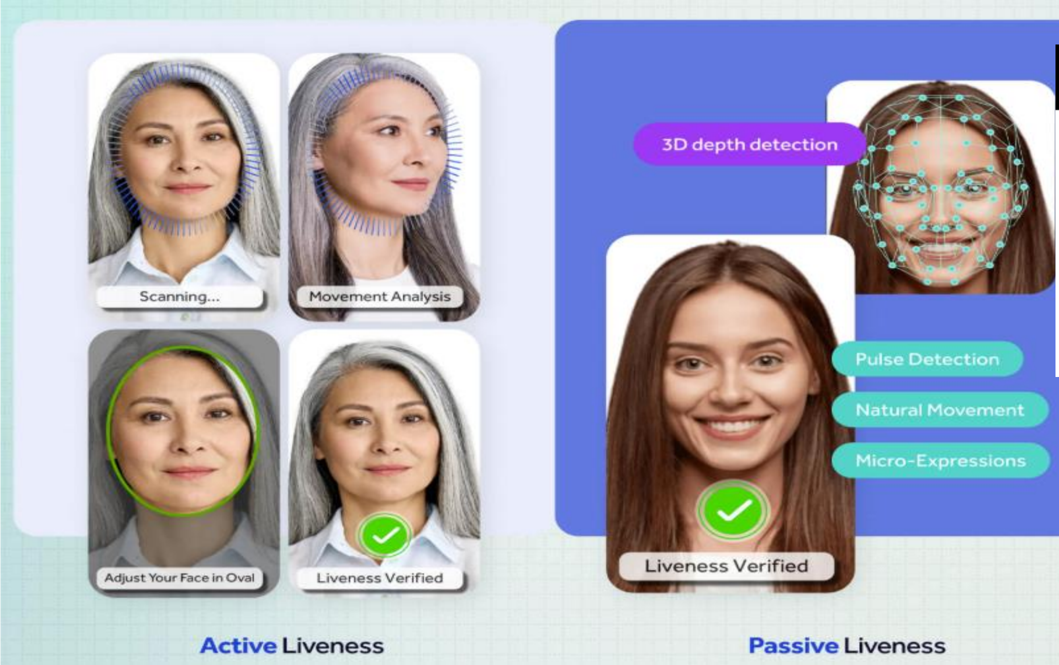
Privacy

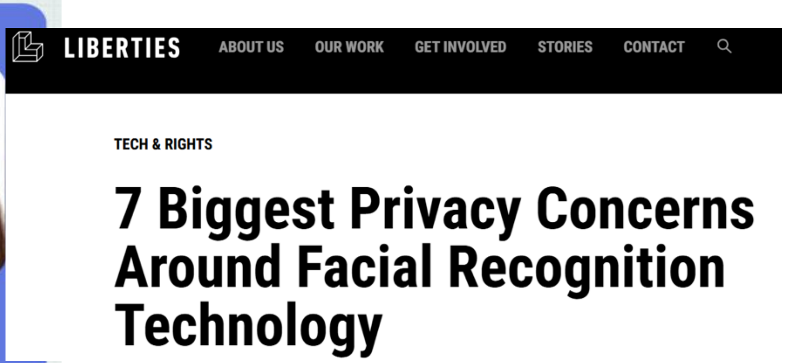
Consent



ISO/IEC 30107-1:2023

Information technology — Biometric presentation attack detection



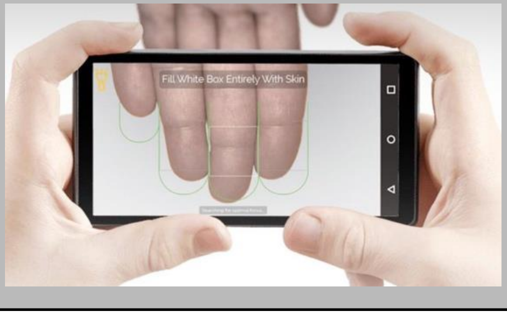


TECH & RIGHTS

7 Biggest Privacy Concerns Around Facial Recognition Technology

https://facia.ai/blog/active-liveness-vs-passive-liveness-key-differences-and-how-they-work/
https://www.liberties.eu/en/stories/facial-recognition-privacy-concerns/44518

Contactless Fingerprinting

Advantages	Challenges
• Hygienic (no contact) • Portable & potentially low-cost • Software-driven (mobile capture possible) • Smudge-free, less hardware • Larger capture area & multi finger possible 	• No global standards, vendor lock-in risks • Lower accuracy due to 3D-to-2D conversion artifacts • Poor interoperability with legacy ABIS • Capture issues: motion blur, lighting, background noise • Device variability (camera optics, OS) • Longer capture time, higher failure rates • Weak liveness detection vs. traditional sensors

- NIST 2015: Programme: Contactless Fingerprinting Capture.
- NIST (2020). Study Measures the Performance of Contactless Fingerprinting Technologies
- World Bank ID4D. (2021). *ID and COVID-19: Overview of Country Examples in Safety Protocols and Practices*
- NIST (2022). NIST Special Publication 500-336 *Specification for Interoperability Testing of Contactless Fingerprint Acquisition Devices, v1.0*
- NIST (2023) Specification for Certification Testing of Contactless Fingerprint Acquisition Devices, v1.0
- NIST (2024). *Contactless Fingerprint Capture and Data Interchange Best Practice Recommendation*

Contactless Fingerprinting

Key Evaluations:

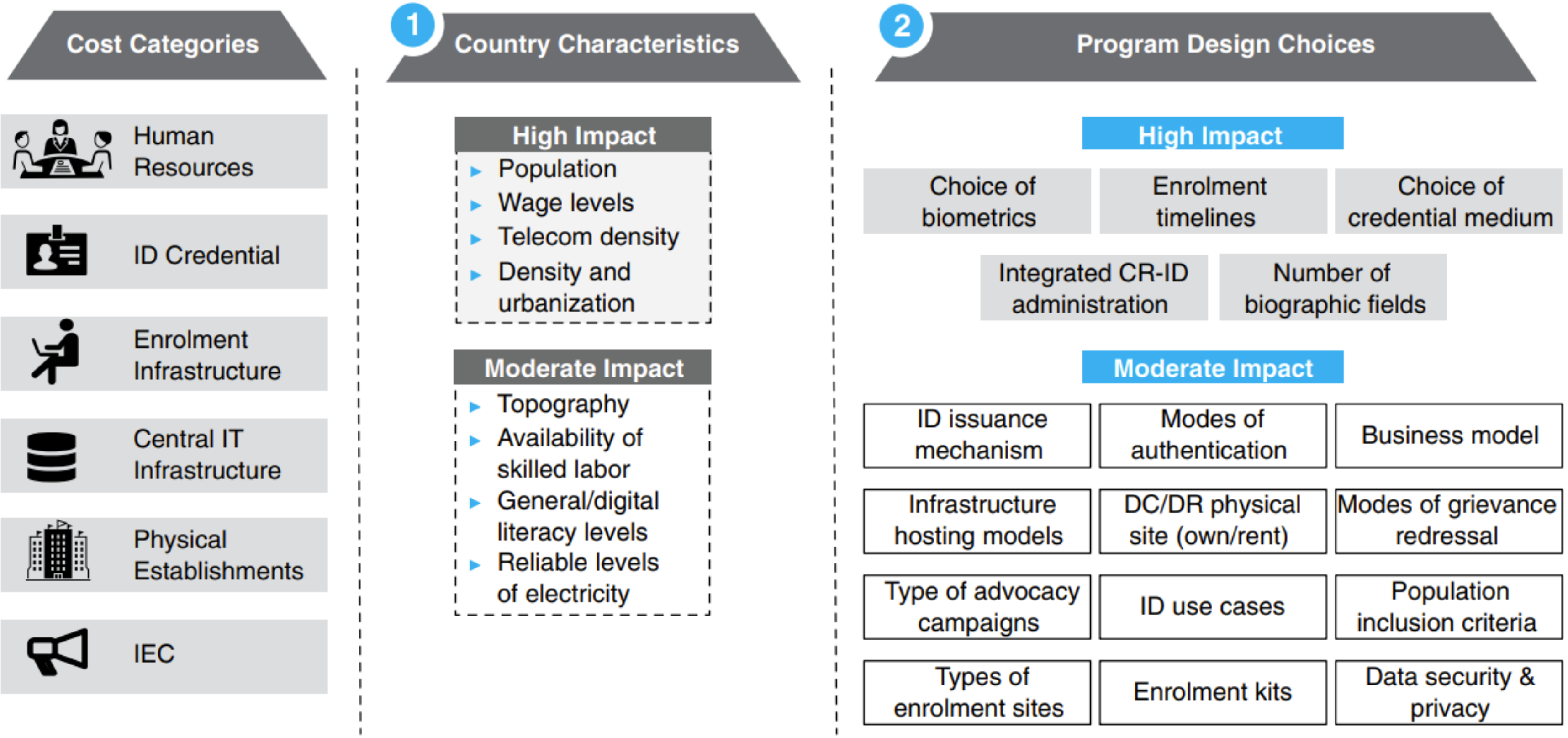
- NIST (2020): Wide performance range (20–80%); only 1 hardware based solution >90% (multi-finger)
- NIST (2022–24): Warns against reusing contact-based quality algorithms; no cert like FBI Appendix F yet; legacy impact must be evaluated

Conclusion:

While promising for self-service or low-assurance use cases, contactless fingerprinting is *not yet reliable enough for deduplication* or secure foundational ID enrollment without further standardization and validation.

- NIST 2015: Programme: Contactless Fingerprinting Capture.
- NIST (2020). Study Measures the Performance of Contactless Fingerprinting Technologies
- World Bank ID4D. (2021). *ID and COVID-19: Overview of Country Examples in Safety Protocols and Practices*
- NIST (2022). NIST Special Publication 500-336 *Specification for Interoperability Testing of Contactless Fingerprint Acquisition Devices, v1.0*
- NIST (2023) Specification for Certification Testing of Contactless Fingerprint Acquisition Devices, v1.0
- NIST (2024). *Contactless Fingerprint Capture and Data Interchange Best Practice Recommendation*

Cost Categories – ID System



Cost Categories

- Human Resources
- ID Credential
- Enrolment Infrastructure
- Central IT Infrastructure
- Physical Establishments
- IEC

1 Country Characteristics

High Impact

- Population
- Wage levels
- Telecom density
- Density and urbanization

Moderate Impact

- Topography
- Availability of skilled labor
- General/digital literacy levels
- Reliable levels of electricity

2 Program Design Choices

High Impact

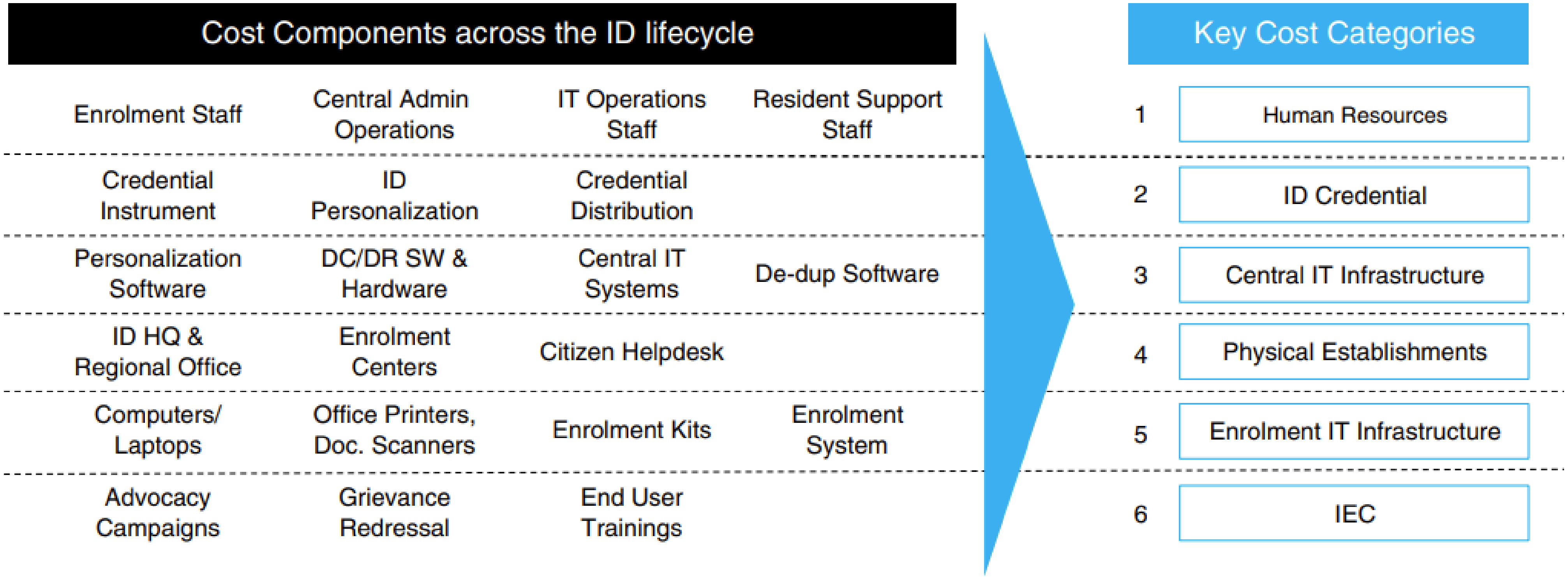
- Choice of biometrics
- Enrolment timelines
- Choice of credential medium
- Integrated CR-ID administration
- Number of biographic fields

Moderate Impact

- ID issuance mechanism
- Modes of authentication
- Business model
- Infrastructure hosting models
- DC/DR physical site (own/rent)
- Modes of grievance redressal
- Type of advocacy campaigns
- ID use cases
- Population inclusion criteria
- Types of enrolment sites
- Enrolment kits
- Data security & privacy

https://documents1.worldbank.org/curated/en/248371559325561562/pdf/ID4D-Practitioner-s-Guide.pdf

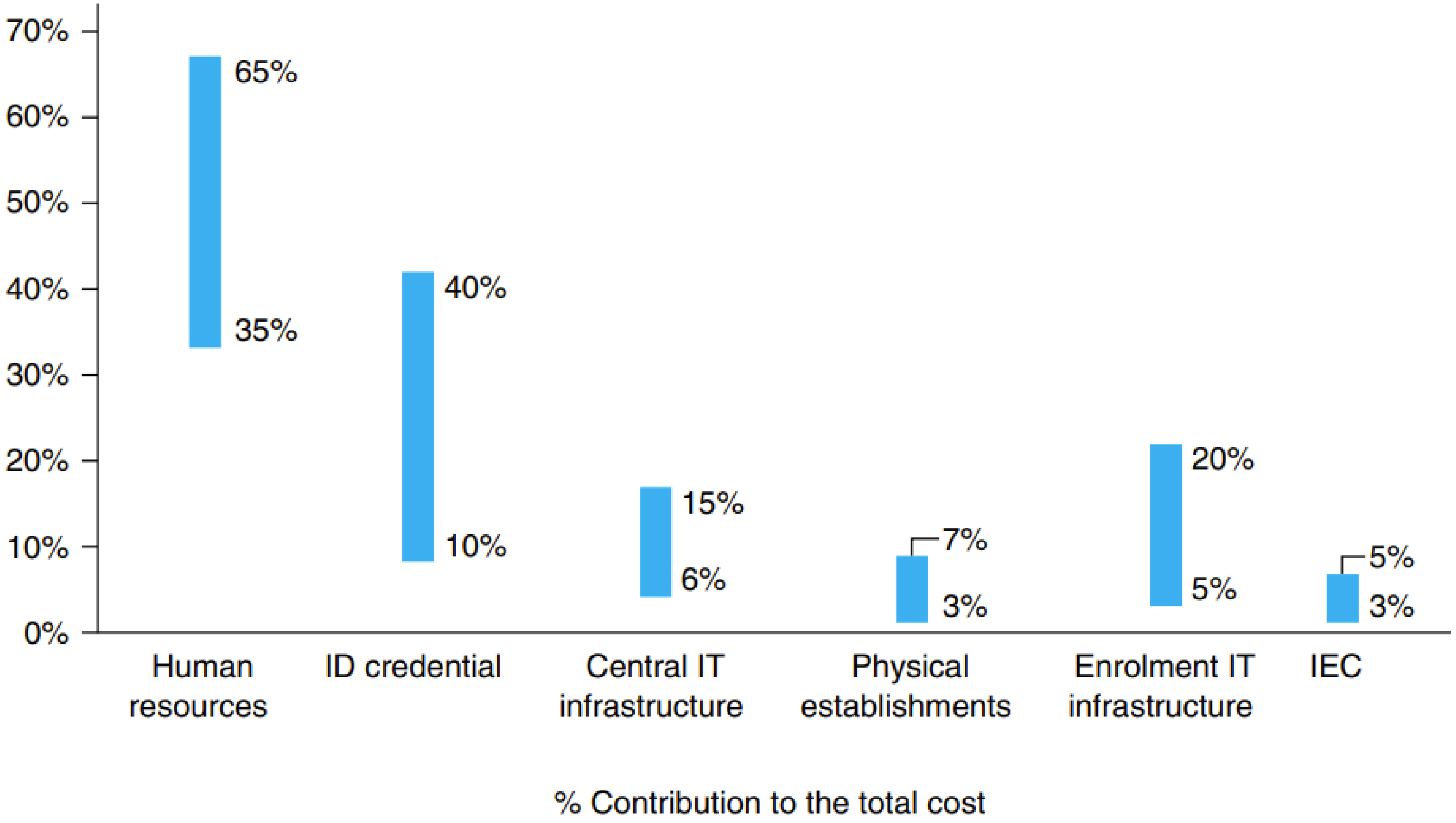
Key Cost Categories



Cost Components across the ID lifecycle					Key Cost Categories
Enrolment Staff	Central Admin Operations	IT Operations Staff	Resident Support Staff	1	Human Resources
Credential Instrument	ID Personalization	Credential Distribution		2	ID Credential
Personalization Software	DC/DR SW & Hardware	Central IT Systems	De-dup Software	3	Central IT Infrastructure
ID HQ & Regional Office	Enrolment Centers	Citizen Helpdesk		4	Physical Establishments
Computers/ Laptops	Office Printers, Doc. Scanners	Enrolment Kits	Enrolment System	5	Enrolment IT Infrastructure
Advocacy Campaigns	Grievance Redressal	End User Trainings		6	IEC

https://documents1.worldbank.org/curated/en/702641544730830097/pdf/Understanding-Cost-Drivers-of-Identification-Systems.pdf

Cost Contribution



https://documents1.worldbank.org/curated/en/702641544730830097/pdf/Understanding-Cost-Drivers-of-Identification-Systems.pdf

Cost – Additional Biometric Modality

“Each additional biometric modality increases accuracy and inclusion, and is estimated to increase enrolment costs by only about 5–10%.”

“Although multimodal biometrics may represent an added cost (compared with single mode biometrics, e.g., fingerprints), their use can—depending on population size and other characteristics— reduce overall costs because it will reduce the rate of manual adjudications during deduplication, as well as improve the accuracy and flexibility of authentication”

(study covers a group of 15 countries across Europe, South America, Africa, and Asia)

https://documents1.worldbank.org/curated/en/702641544730830097/pdf/Understanding-Cost-Drivers-of-Identification-Systems.pdf (2018)

Cost – Biometric Devices

Name of the item / Service	Unit cost (US $)
Computer / Laptop	**$ 1,200.00**
Mobile / tablet	$ 200.00
Camera / webcam	**$ 100.00**
Multifunction Printer	$ 200.00
Power back up	$ 100.00
Finger print scanner - One finger	**$ 200.00**
Finger print scanner - Slab scanner	**$ 1,000.00**
Iris scanner - one	**$ 300.00**
Iris scanner - two	**$ 1,000.00**
Signature pad	$ 150.00
GPS dongle	$ 100.00
Additional Screen	$ 150.00
Voice recording device	$ 150.00
Case for kit	$ 200.00

Tri Modal Enrolment Kit ~ 3K -4K USD

Depends on Volume, Geography, Businesss Risk etc.

https://documents1.worldbank.org/curated/en/702641544730830097/pdf/Understanding-Cost-Drivers-of-Identification-Systems.pdf (2018)

Cost – Biometric Deduplication

Cost per deduplication (For first biometric)	**$ 0.25**
Add on dedup cost for additional biometric	**$ 0.10**

→ **ID4D**

Vendor 1

Records	Cost	Cost/ID
10K	20K	**2**
50K	35K	**0.7**
500K	100K	**0.2**
1M	170K	**0.17**
10M	600K	**0.06**
100M	3.4M	**0.034**

Vendor 3

Records	Cost/ID	
1M	**1**	With Hardware
5M	**0.9**	
10M	**0.8**	
1M	**0.5**	Only software
5M	**0.45**	
10M	**0.4**	

Matcher SDK could be 10 % of ABIS costing

10-12% AMC

40% for DR

30-40% AMC

Vendor 2

Records	Cost	Cost/ID
1 M	500K	0.5
5M	1.4 M	0.28
10M	1.6M	0.16

Vendor 4

Records	Cost	Cost/ID
1M	200K	**0.2**
5M	800K	**0.16**
10M	1.5M	**0.15**

Average

1 M = 0.35
5 M = 0.29
10 M = 0.19

https://documents1.worldbank.org/curated/en/702641544730830097/pdf/Understanding-Cost-Drivers-of-Identification-Systems.pdf (2018)

TriModal Data from Vendor Interviews; software cost by default

Identity friction – say no…or?

- UX – I am not a number

versus

- Re-Decentralised Viz Estonia…

NO2ID
Stop the Database State

Recall John Perry Barlow

- “Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather.” … + …

- “We will create a civilization of the Mind in Cyberspace. May it be more humane and fair than the world your governments have made before.”

And Dave Clark

- "We reject: kings, presidents, and voting. We believe in: rough consensus and running code."

- And Larry Lessig: "Code is Law" (in Code and Other Laws of Cyberspace)

- See also (1997) https://web.archive.org/web/20020420162518/http://www.telegraph.co.uk/connected/main.jhtml?xml=/connected/1997/04/22/ecdip22.xml

What went wrong?

- Centralisation (rich get richer) +
  - Moore’s Law + Metcalfe’s “Law”
  - i.e. growth 100%/18months, and super-linear value proposition
  - Hyperscale (scale out) is a natural consequence of these net fx
- Imagine if Musk didn’t just own Starlink (and supply via SpaceX)
  - But also, say, GPS. Or a nation’s power grid?
- The forces of centralization are remarkably strong

Apple (hyperscalars) market cap >> UK GDP

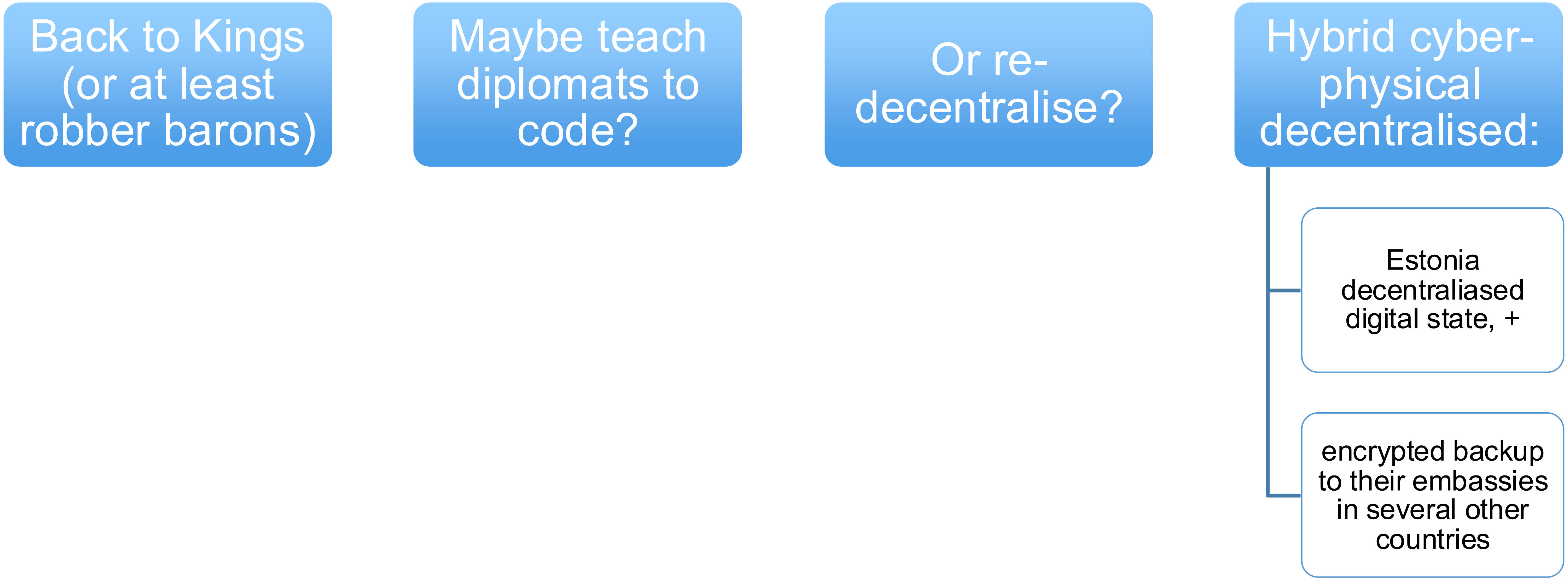
Lots of other examples

AI startup valuations >> GDP of small nations

Yet energy consumption to have MVP also >> small nations

Few people (even US e.g. FBI v. Apple) can co-erce compliance

Impact on Nations



Back to Kings (or at least robber barons)

Maybe teach diplomats to code?

Or re-decentralise?

Hybrid cyber-physical decentralised:

- Estonia decentraliased digital state, +
- encrypted backup to their embassies in several other countries

Now for some SF
What if….we could mod our biometrics?

- Link between bio- and digital no longer immutable (or unique)