Biometrics: how unique are you?

Richard Harvey
IT Livery Company Professor of Information Technology, Gresham College

Professor of Computer Science, School of Computing Sciences,
University of East Anglia

#richardwharvey

The need for digital identity

- Imagine your landlord has thrown you out of your house
- In a rage he has destroyed all your possessions
- Which identity card do you want to get first?

Bank card?

Driving licence?

Passport?

Bank card	Driving licence	Passport
Varies but typically	Passport or	Birth certificate + parents' birth certificates
Passport or	Birth certificate +	2 passport photos (countersigned)
Driving licence	NI card or letter from DWP or P45 or Payslip or original benefits claim letter or marriage cert or university card	
	Countersigned photos	Interview
	£34	£85

What is a biometric?

Not … statistics in biology

rather …

Distinct, measurable characteristics of people

an alternative to …

Tokens.

Or.

Humans.

How accurate are humans?

- Shoppers were issued with photo-id credit cards
- Sales assistants were paid £50 to process the transactions “quickly and accurately” with a £25 bonus payable for good performance

More than 50% of fraudulent cards were accepted

When Seeing should not be Believing: Photographs, Credit Cards and Fraud., R Kemp, N Towell and G Pike, Applied Cognitive Psychology. Jun1997, Vol. 11 Issue 3, p211-222

How accurate are humans?

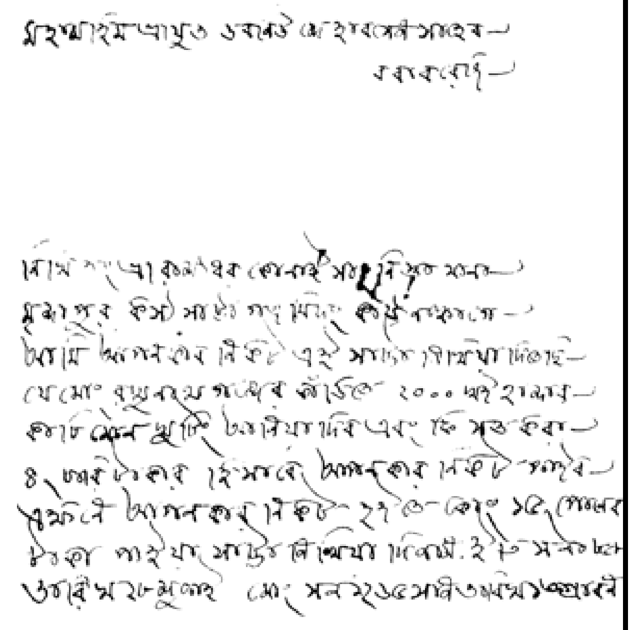
- A tourist asks for directions
- There is a switcheroo
- Can people spot the switch?

https://www.youtube.com/watch?v=FWSxSQsspiQ

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin and Review, 5,* 644–649.

Contract “signed” between W J Hershel and Rājhadyar Kōnāi in Janigipur in 1858

As described in *The origin of finger-printing* by William J Hershel 1916



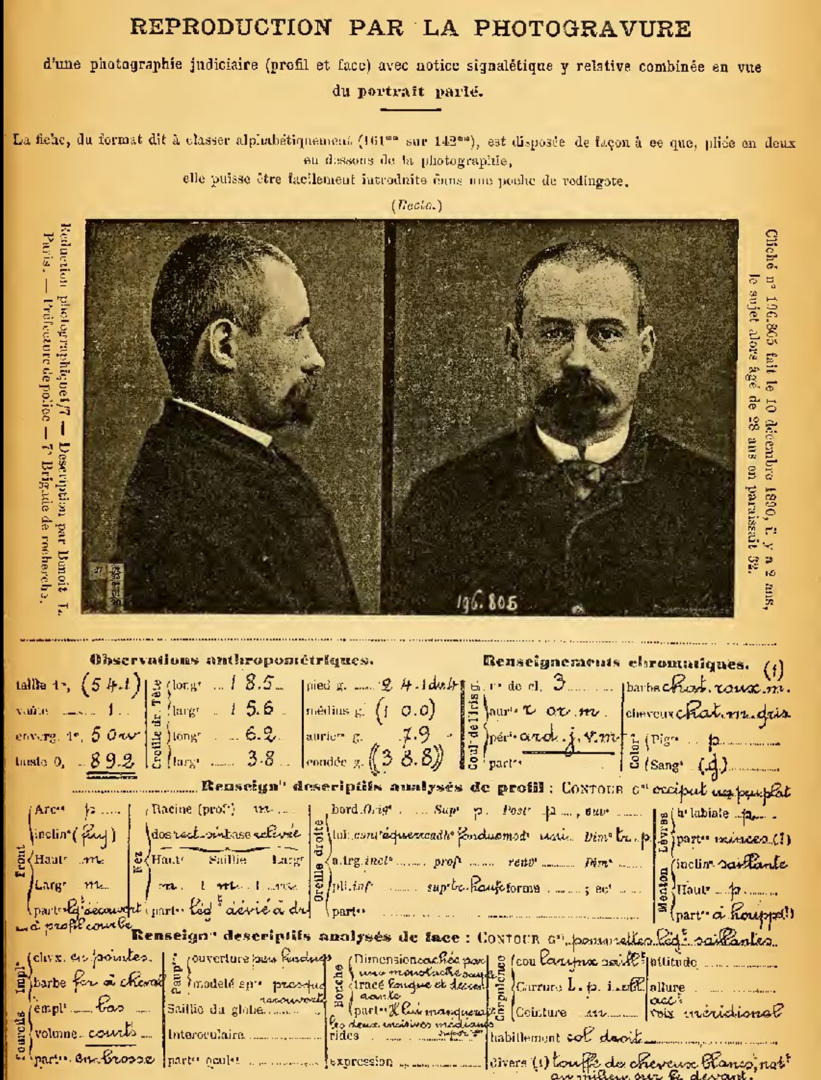
Contract for 2,000 maunds of road-metalling, between W. J. Herschel and Rajyadhar Konai, in Konai's handwriting

KONAI'S HAND

Bengal 1858

Identification anthropométrique: instructions signalétiques

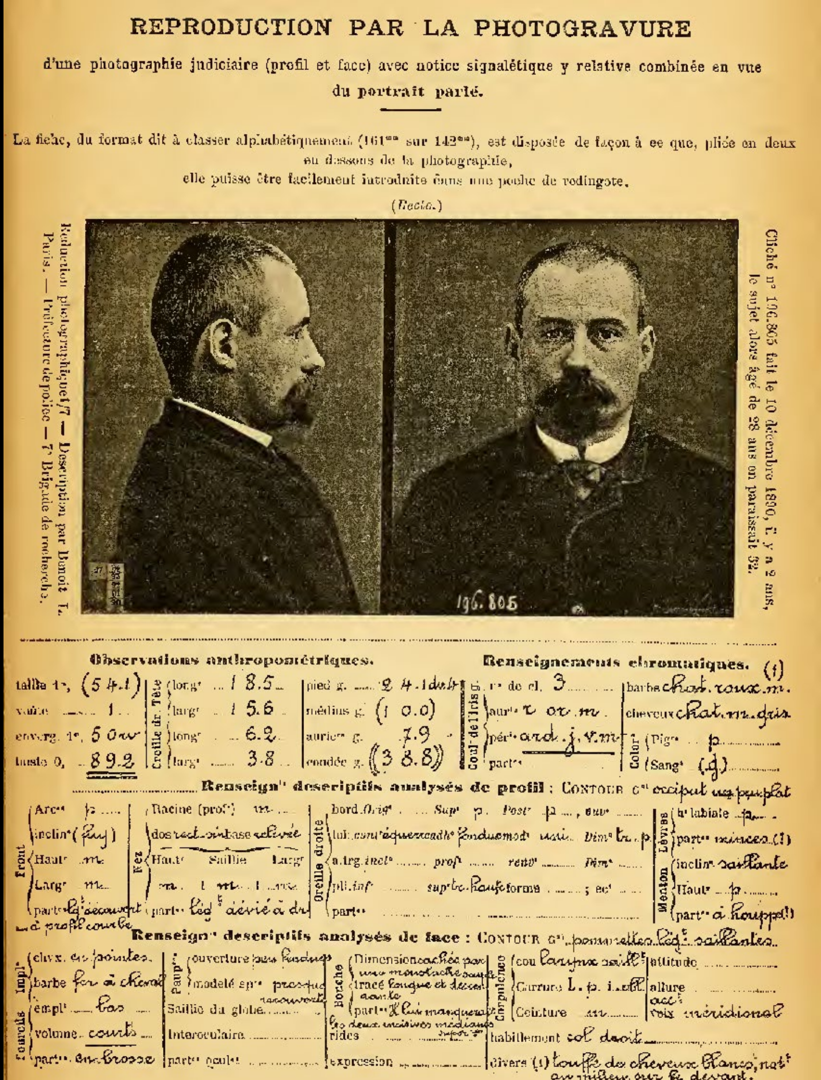
Alphonse Bertillon, New Edition 1893, First edition 1883.



"I came to you, Mr. Holmes, because … I am suddenly confronted with a most serious and extraordinary problem. Recognizing, as I do, that you are the second highest expert in Europe—"

"Indeed, sir! May I inquire who has the honour to be the first?" asked Holmes with some asperity.

"To the man of precisely scientific mind the work of Monsieur Bertillon must always appeal strongly.

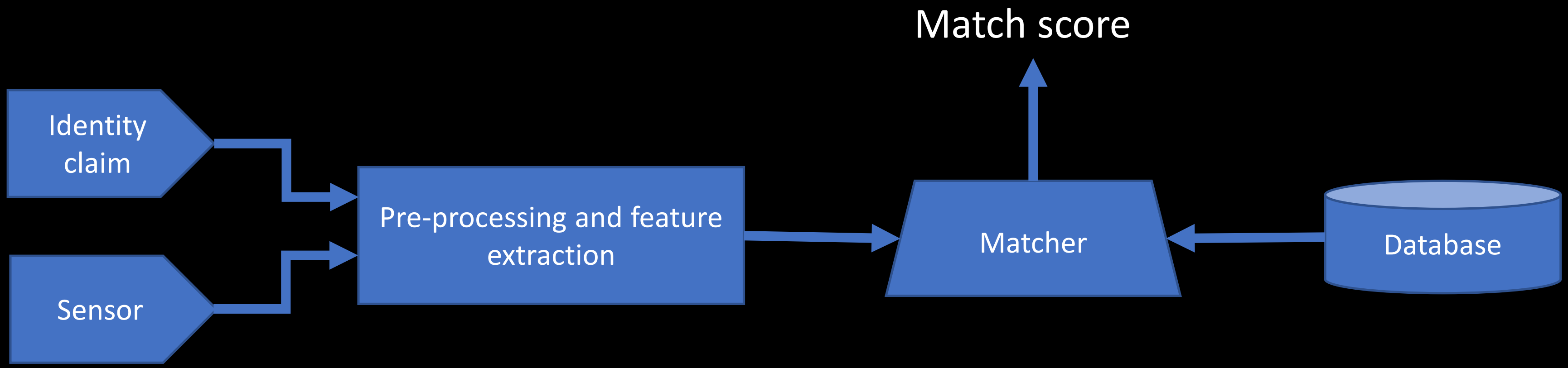


Biometrics and crime fighting

Identification of baddies

Identification of baddies

Biometric system

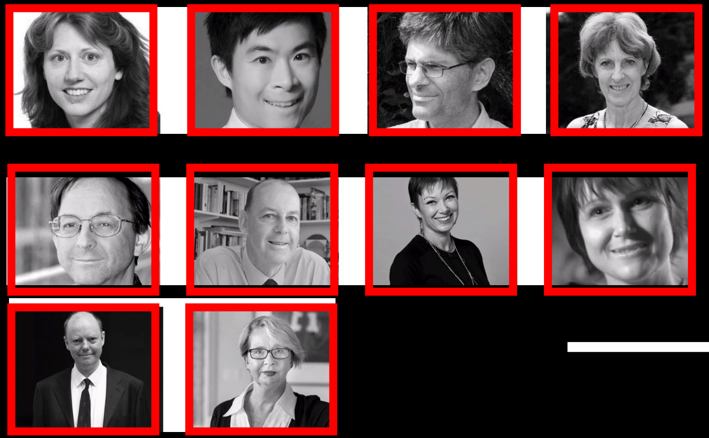


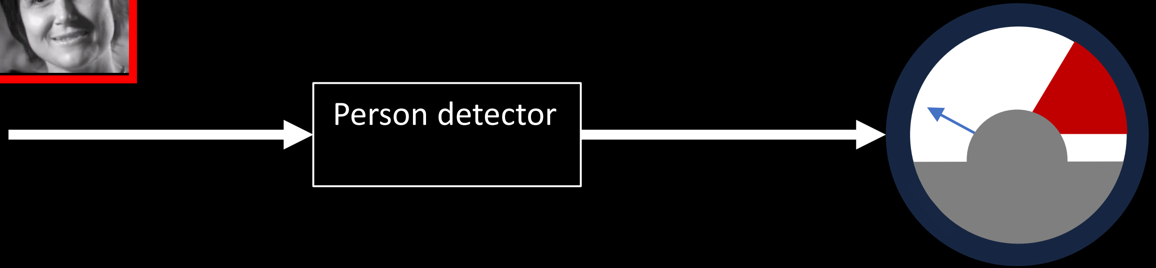
Ratha, Nalini & Connell, Jonathan & Bolle, Ruud. (2001). *Enhancing Security and Privacy in Biometrics-Based Authentication Systems.* IBM Systems Journal. 40. 614-634. 10.1147/sj.403.0614.

What makes a good biometric?

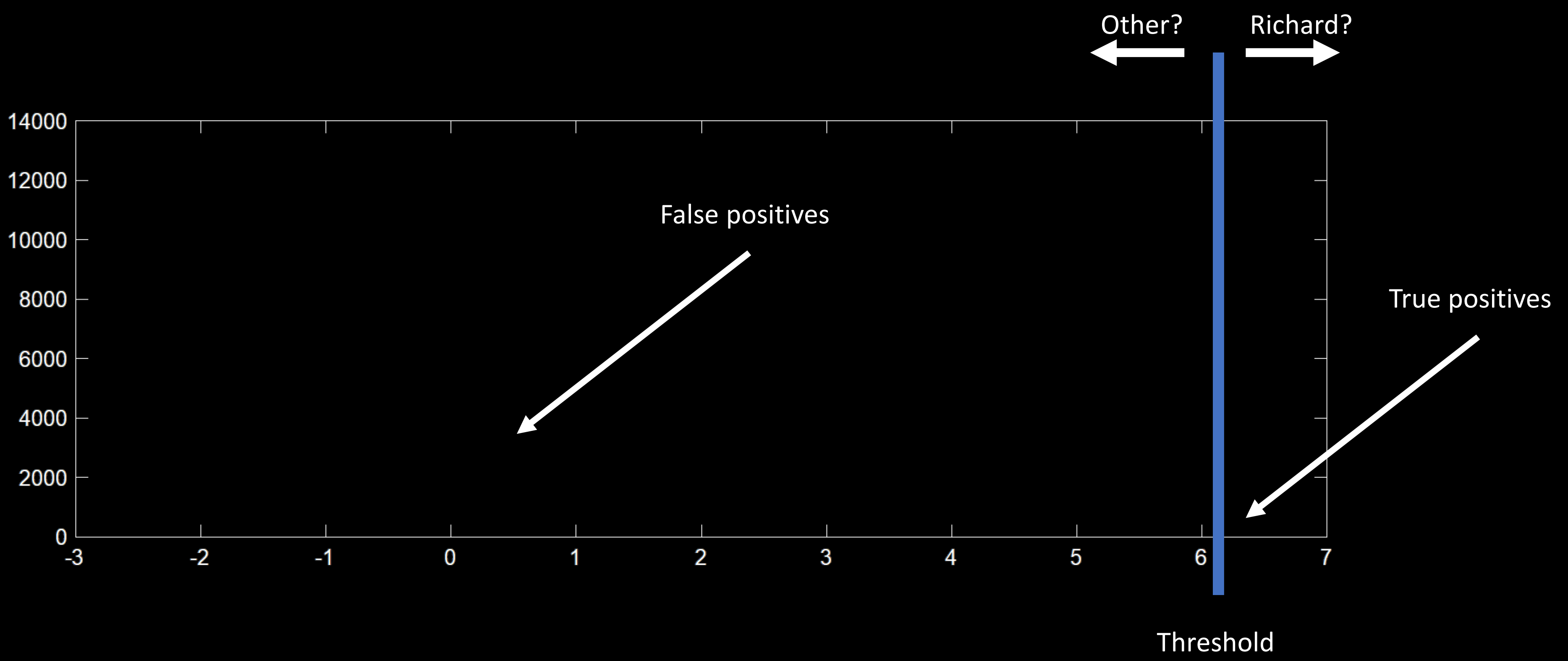
	Universal?	Unique?	Permanent?	Collectible?	Acceptible?
DNA	Yes	Yes	Yes	Tricky	Poor
Iris	Yes	Yes	Yes	Yes	Marginal
Fingerprint	Yes	Yes	Yes	Yes	Fair
Face	Yes	Yes?	No	Yes	Good
Voice	No	Yes	No	Yes	Good
Gait	No	No	No	Yes	Good
Keystrokes	No	No?	No	Yes	Good?

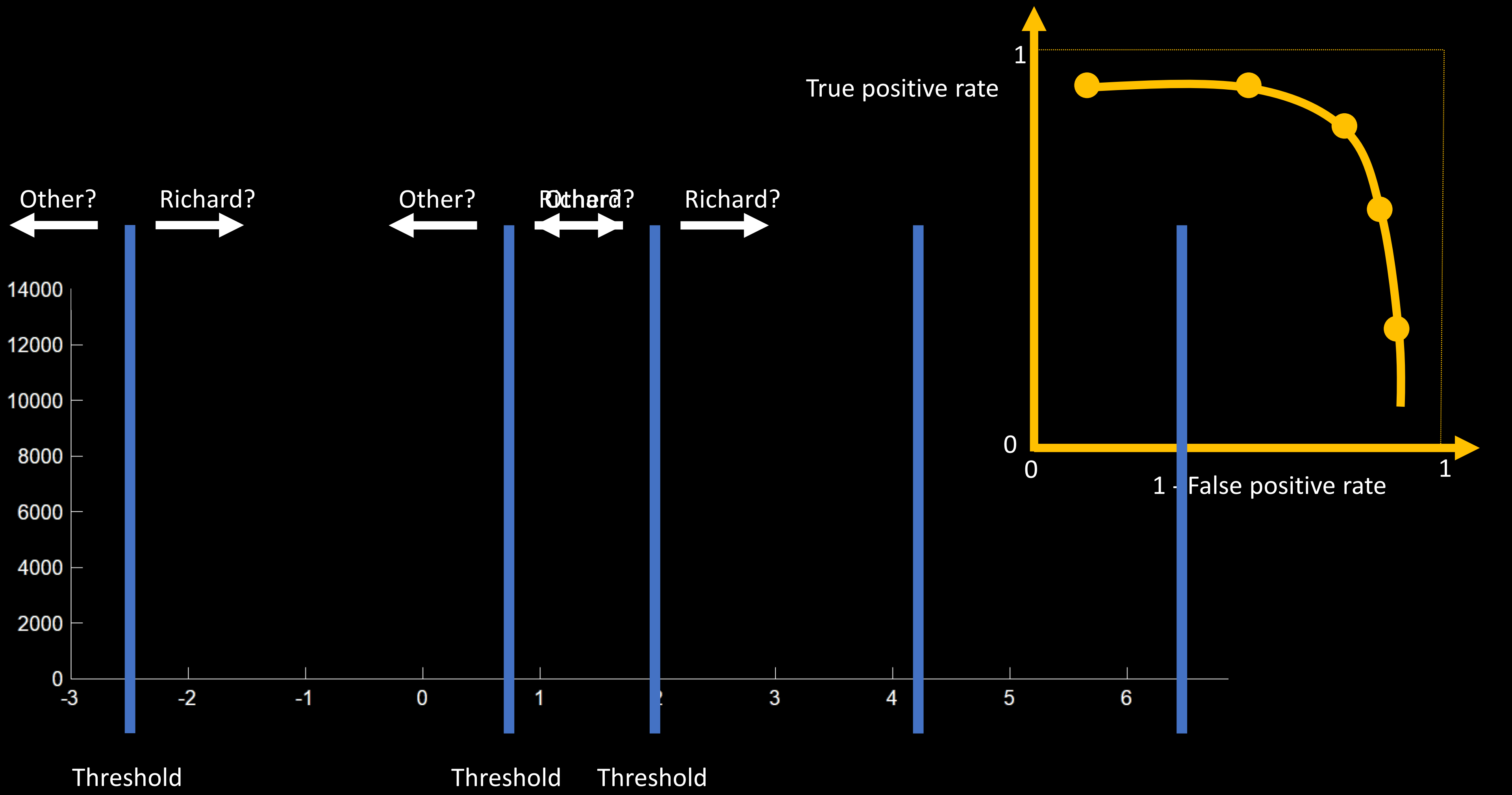
Evaluating a detector





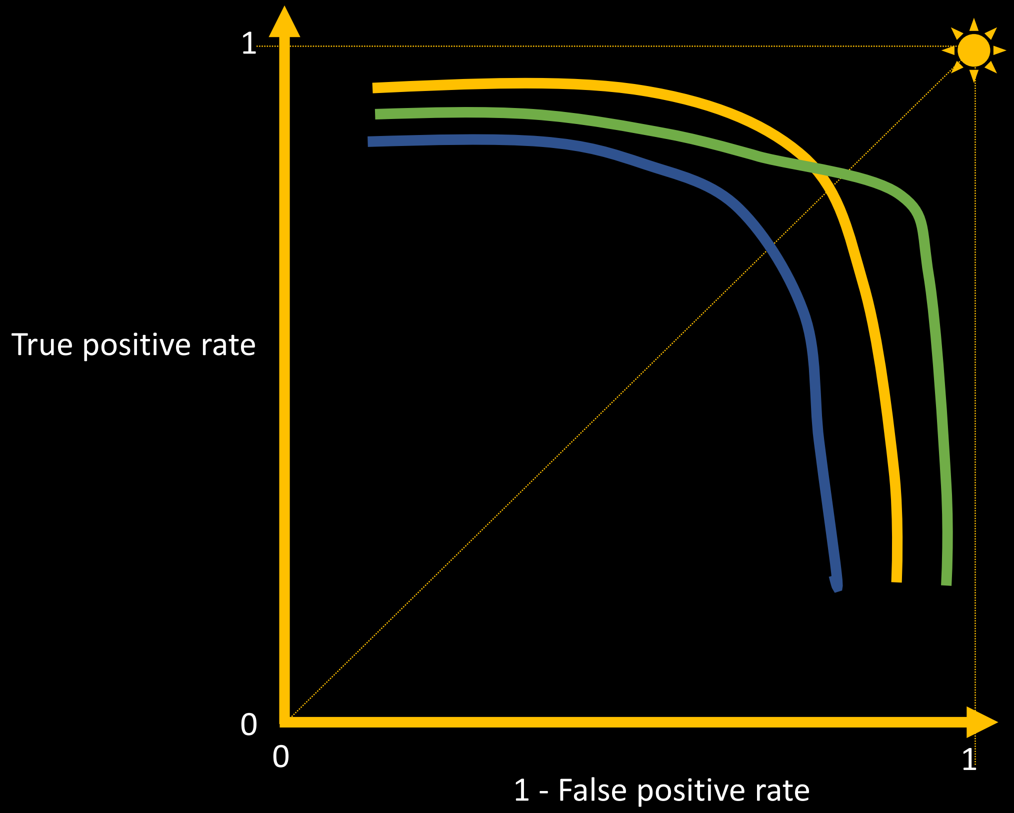
Evaluating a detector





Other?
Richard?
Other?
Richard?
14000
12000
10000
8000
6000
4000
2000
0
-3
-2
-1
0
1
2
3
4
5
6
Threshold
Threshold
Threshold
True positive rate
1
0
0
1
1 - False positive rate

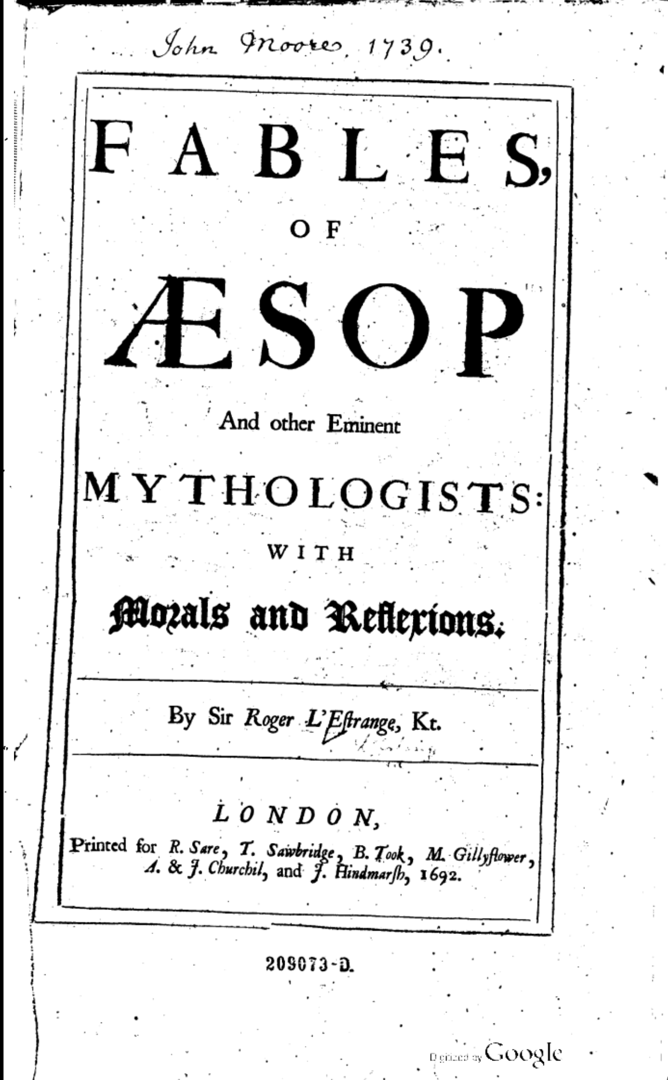
Receiver Operating Characteristic (ROC)



Types of error

False positive = Type 1 error
= false alarm = crying wolf

False negative = Type II error
= missed detection



Operating point

- Assign costs to the bad event

  (and the good ones)

- Choose an operating point that makes sense for the application

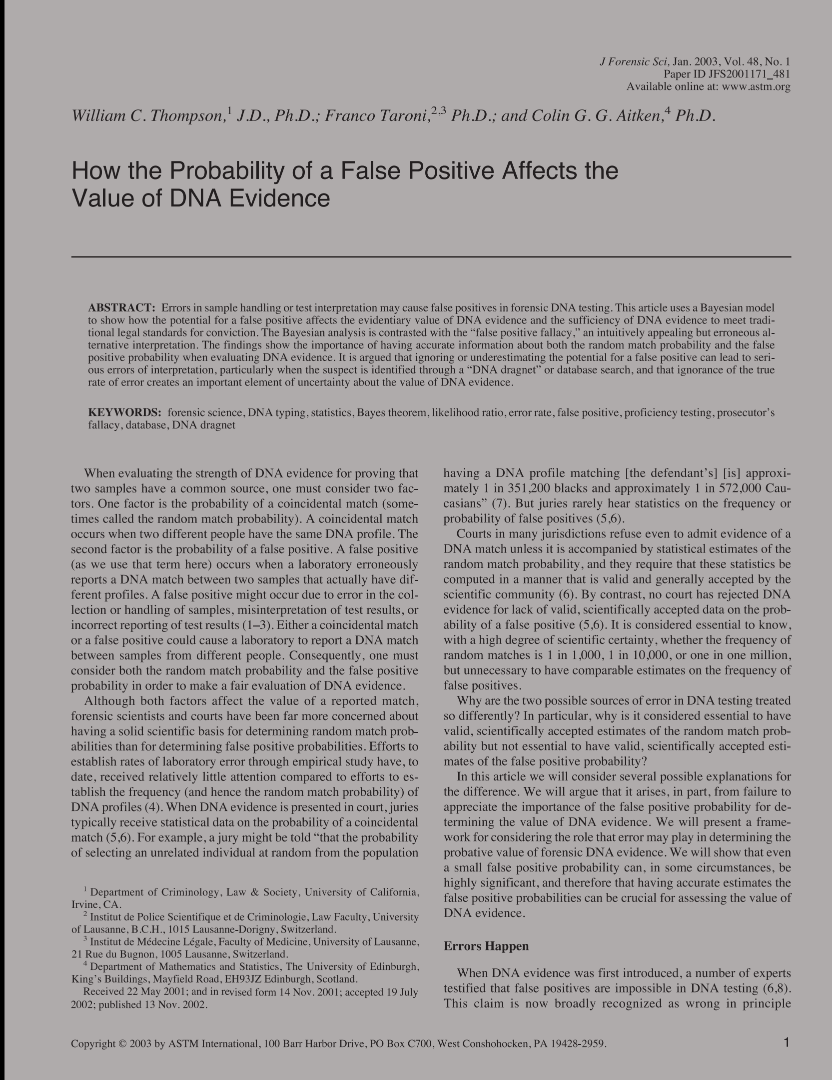
- Idea is to choose a point to minimise total cost or *risk*.

Stanislav Petrov photographed in 2016. Taken from https://en.wikipedia.org/wiki/Stanislav_Petrov

DNA

- False positive rate of zero?
- Timothy Durham convicted of rape in 1993
- 3000 year sentence

- Lab error mixed-up victim's DNA with Daltons
- Durham released in 1997



J Forensic Sci, Jan. 2003, Vol. 48, No. 1
Paper ID JFS2001171_481
Available online at: www.astm.org

William C. Thompson,[1] J.D., Ph.D.; Franco Taroni,[2,3] Ph.D.; and Colin G. G. Aitken,[4] Ph.D.

How the Probability of a False Positive Affects the Value of DNA Evidence

ABSTRACT: Errors in sample handling or test interpretation may cause false positives in forensic DNA testing. This article uses a Bayesian model to show how the potential for a false positive affects the evidentiary value of DNA evidence and the sufficiency of DNA evidence to meet traditional legal standards for conviction. The Bayesian analysis is contrasted with the "false positive fallacy," an intuitively appealing but erroneous alternative interpretation. The findings show the importance of having accurate information about both the random match probability and the false positive probability when evaluating DNA evidence. It is argued that ignoring or underestimating the potential for a false positive can lead to serious errors of interpretation, particularly when the suspect is identified through a "DNA dragnet" or database search, and that ignorance of the true rate of error creates an important element of uncertainty about the value of DNA evidence.

KEYWORDS: forensic science, DNA typing, statistics, Bayes theorem, likelihood ratio, error rate, false positive, proficiency testing, prosecutor's fallacy, database, DNA dragnet

When evaluating the strength of DNA evidence for proving that two samples have a common source, one must consider two factors. One factor is the probability of a coincidental match (sometimes called the random match probability). A coincidental match occurs when two different people have the same DNA profile. The second factor is the probability of a false positive. A false positive (as we use that term here) occurs when a laboratory erroneously reports a DNA match between two samples that actually have different profiles. A false positive might occur due to error in the collection or handling of samples, misinterpretation of test results, or incorrect reporting of test results (1–3). Either a coincidental match or a false positive could cause a laboratory to report a DNA match between samples from different people. Consequently, one must consider both the random match probability and the false positive probability in order to make a fair evaluation of DNA evidence.

Although both factors affect the value of a reported match, forensic scientists and courts have been far more concerned about having a solid scientific basis for determining random match probabilities than for determining false positive probabilities. Efforts to establish rates of laboratory error through empirical study have, to date, received relatively little attention compared to efforts to establish the frequency (and hence the random match probability) of DNA profiles (4). When DNA evidence is presented in court, juries typically receive statistical data on the probability of a coincidental match (5,6). For example, a jury might be told "that the probability of selecting an unrelated individual at random from the population having a DNA profile matching [the defendant's] [is] approximately 1 in 351,200 blacks and approximately 1 in 572,000 Caucasians" (7). But juries rarely hear statistics on the frequency or probability of false positives (5,6).

Courts in many jurisdictions refuse even to admit evidence of a DNA match unless it is accompanied by statistical estimates of the random match probability, and they require that these statistics be computed in a manner that is valid and generally accepted by the scientific community (6). By contrast, no court has rejected DNA evidence for lack of valid, scientifically accepted data on the probability of a false positive (5,6). It is considered essential to know, with a high degree of scientific certainty, whether the frequency of random matches is 1 in 1,000, 1 in 10,000, or one in one million, but unnecessary to have comparable estimates on the frequency of false positives.

Why are the two possible sources of error in DNA testing treated so differently? In particular, why is it considered essential to have valid, scientifically accepted estimates of the random match probability but not essential to have valid, scientifically accepted estimates of the false positive probability?

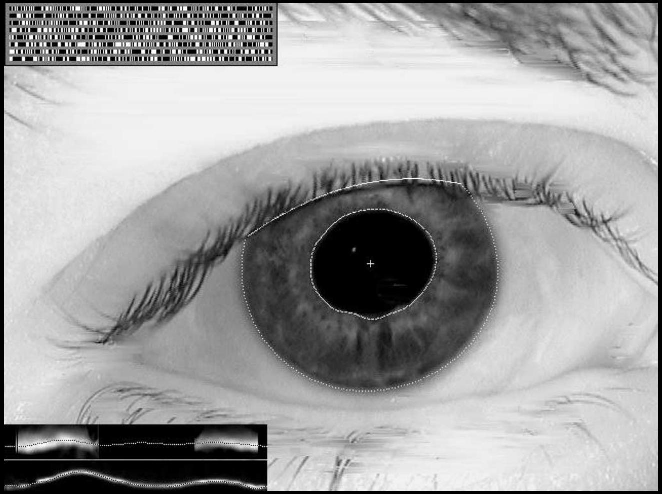
In this article we will consider several possible explanations for the difference. We will argue that it arises, in part, from failure to appreciate the importance of the false positive probability for determining the value of DNA evidence. We will present a framework for considering the role that error may play in determining the probative value of forensic DNA evidence. We will show that even a small false positive probability can, in some circumstances, be highly significant, and therefore that having accurate estimates the false positive probabilities can be crucial for assessing the value of DNA evidence.

Errors Happen

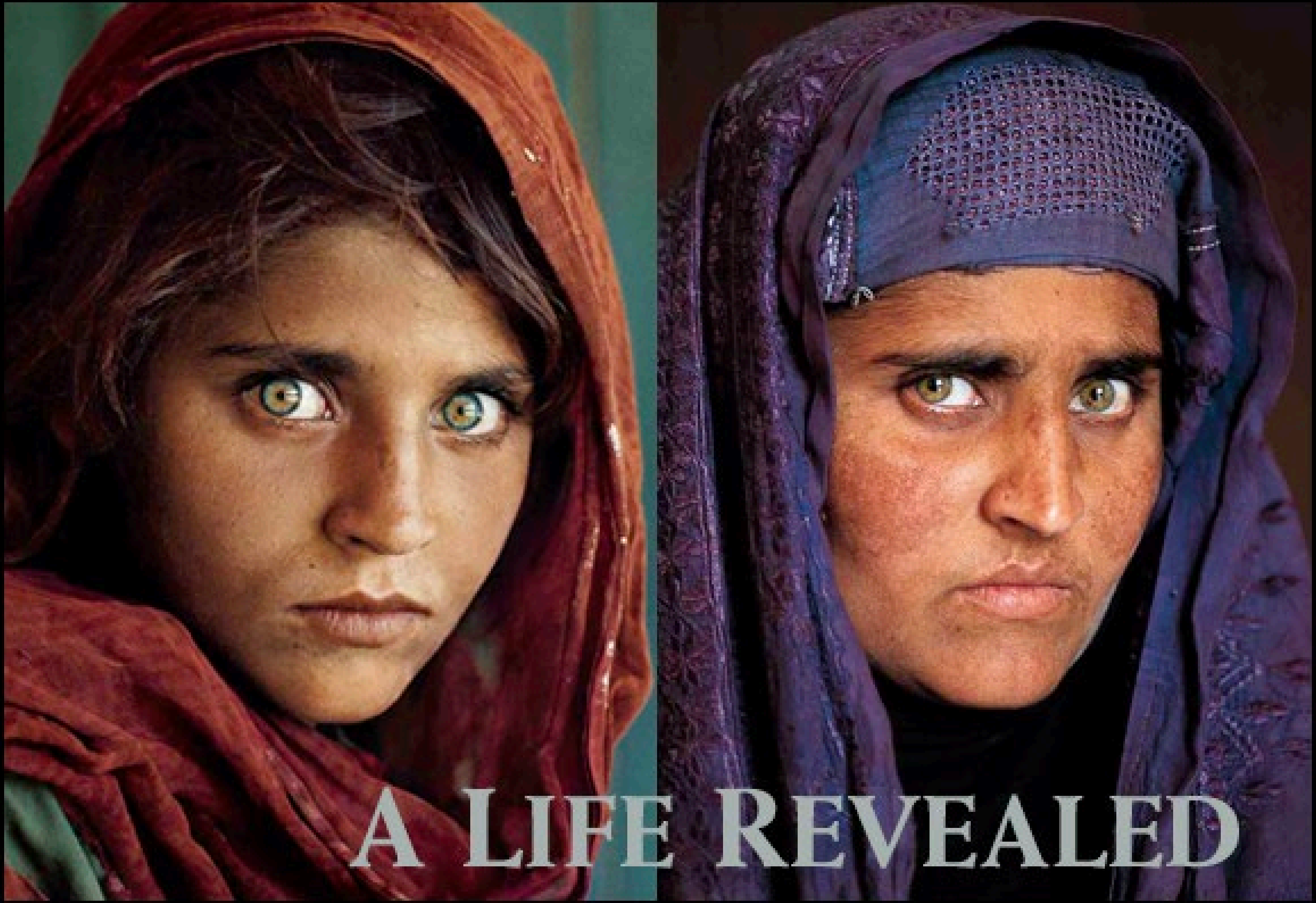
When DNA evidence was first introduced, a number of experts testified that false positives are impossible in DNA testing (6,8). This claim is now broadly recognized as wrong in principle

[1] Department of Criminology, Law & Society, University of California, Irvine, CA.
[2] Institut de Police Scientifique et de Criminologie, Law Faculty, University of Lausanne, B.C.H., 1015 Lausanne-Dorigny, Switzerland.
[3] Institut de Médecine Légale, Faculty of Medicine, University of Lausanne, 21 Rue du Bugnon, 1005 Lausanne, Switzerland.
[4] Department of Mathematics and Statistics, The University of Edinburgh, King's Buildings, Mayfield Road, EH93JZ Edinburgh, Scotland.
Received 22 May 2001; and in revised form 14 Nov. 2001; accepted 19 July 2002; published 13 Nov. 2002.

Iris scan



New methods in Iris Recognition, John Daugman, IEE Trans, Systems, Man and Cybernetics – Part B, Vol 37, No 5, Oct 2007, pp 1167--1175


A LIFE REVEALED

Aadhaar number

- 1.2 Bn people enrolled
- Around 40M authentications per day
- Links biometric information with identity
- Uses Iris scans

"the most sophisticated ID programme in the world"

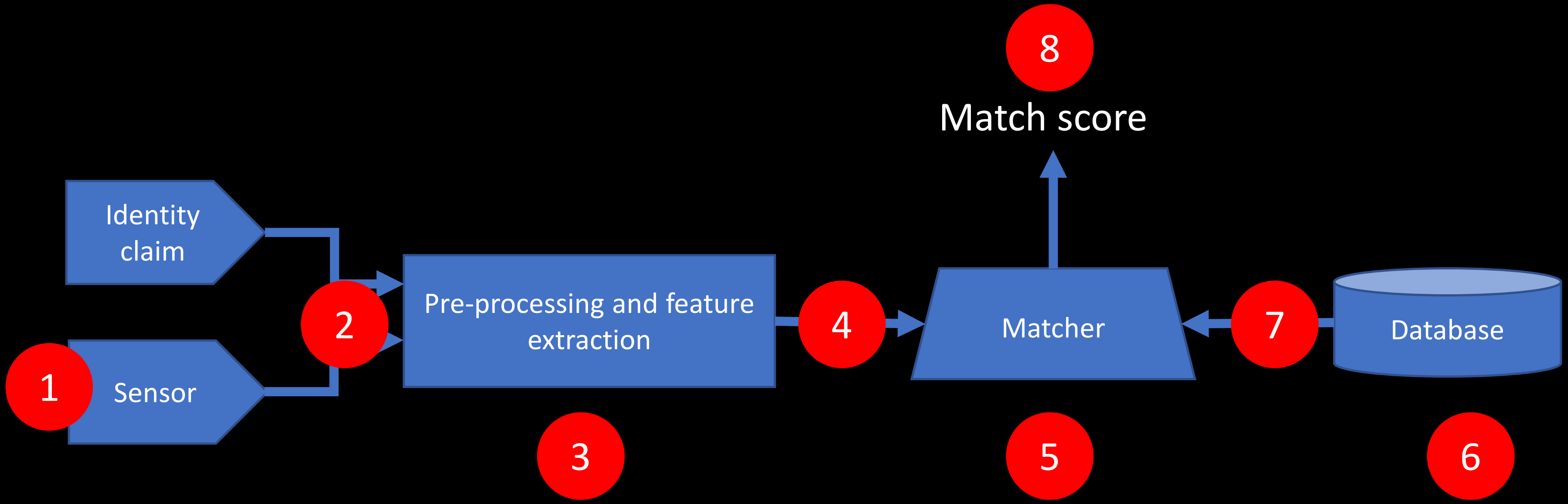
Paul Rohmer, Chief Economist, World Bank

Face recognition

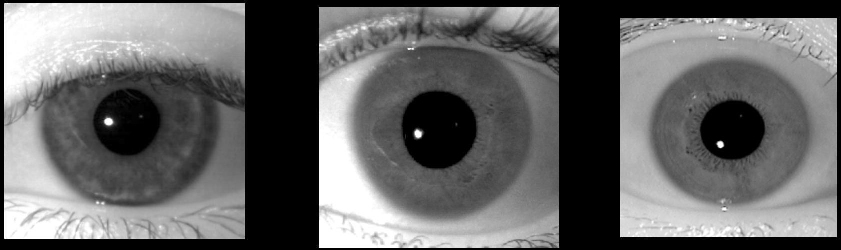
- Much less reliable
- Some ethnicity issues
- Various public CCTV trials

Cvdazzle (cvdazzle.com),
Adam Harvey

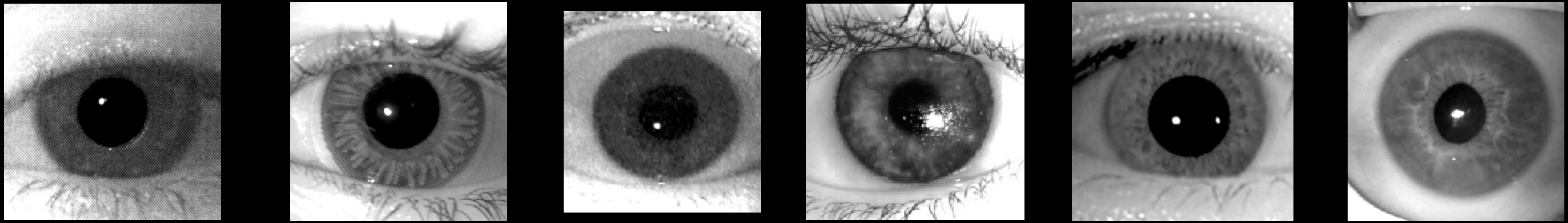
Biometric attacks



Presentation attack



Adam Czajka and Kevin W. Bowyer. 2018. *Presentation Attack Detection for Iris Recognition: An Assessment of the State-of-the-Art*. ACM Comput. Surv. 51, 4, Article 86 (July 2018), https://doi.org/10.1145/3232849



print | contact lens | kindle | glass | syn image | cadaver

GDPR 2018

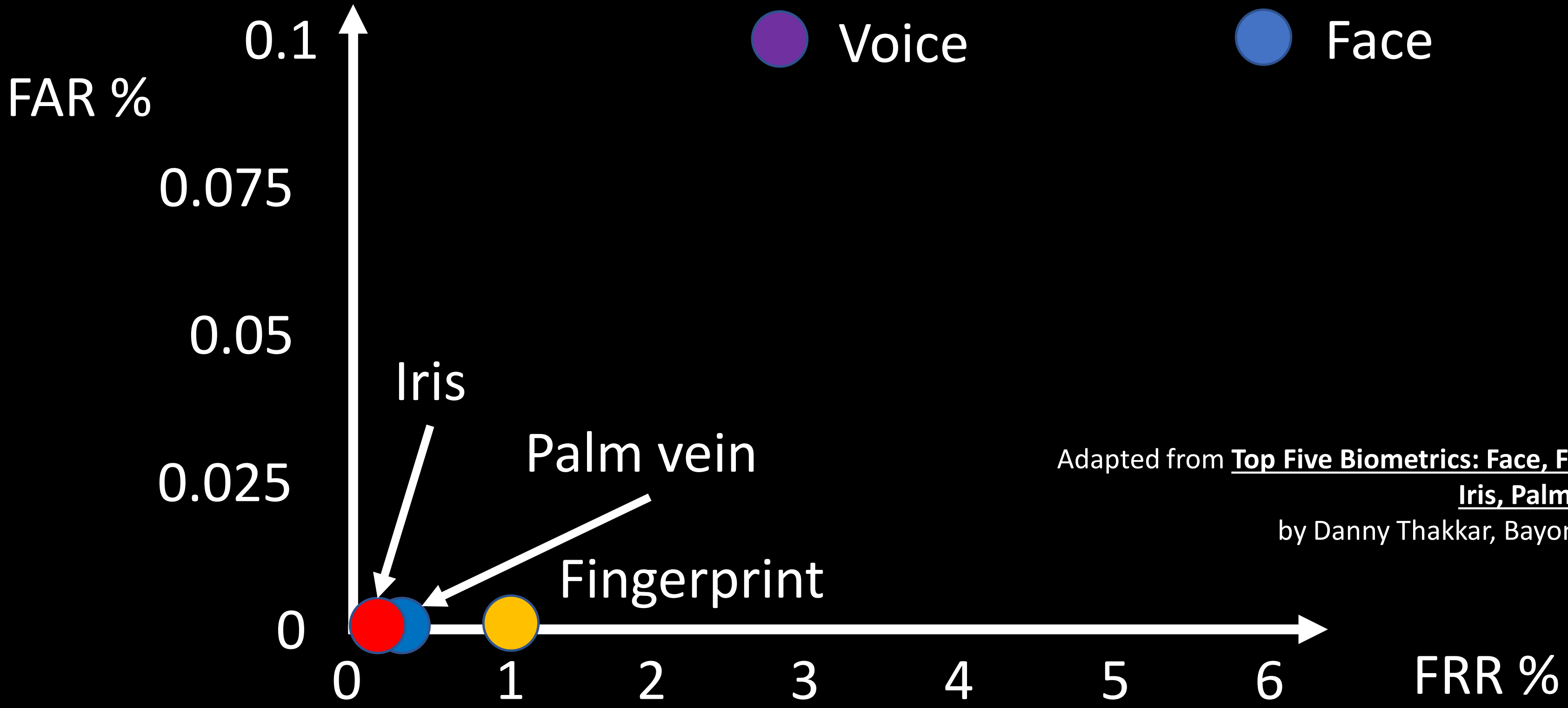
1. Right to transparency
2. Right to access their data
3. Right to rectification and erasure
4. Right to avoid automatic processing

Biometric data is regarded as “sensitive category of personal data” so requires special care

CCPA 2020

1. Right to transparency
2. Right to know whether personal data is sold or disclosed
3. Right to block the sale of personal data
4. Right to access their personal data
5. Right to erasure
6. Right to not be discriminated against for exercising their privacy rights.

Biometric performance



Adapted from **Top Five Biometrics: Face, Fingerprint, Iris, Palm and Voice** by Danny Thakkar, Bayometric.com

Next lecture:

"Taming the trolls of social media"
Tuesday Nov 6th 2019 at 18:00 (6pm) London Time

www.gresham.ac.uk