

What Constitutes Biometric Data Quality?

**Austin Hicklin
Mitretek Systems**

8 March 2006

Recent interest in data quality

Department of Justice,
Bureau of Investigation.

GEN/LMR

✓ June 25, 1924.

TO ALL CHIEFS OF POLICE, MARSHALS, SHERIFFS, AND OTHER OFFICERS ENGAGED IN THE ENFORCEMENT OF LAW.

This is to advise you of the establishment of an Identification Division in the Bureau of Investigation of the United States Department of Justice at Washington, D.C., effective July 1, 1924, under authority of an Act of Congress which authorizes the exchange of identification records with officers of the cities, counties and states. The purpose of this work is to supply you with a service to fill a long felt need in the proper enforcement of law. It is the culmination of the plans that have been under consideration for a considerable period of time with a view of bringing about a more efficient enforcement of the law and greater expedition in the apprehension, conviction and confinement of violators of the law. In order that this service may be a success, it is necessary to have the closest cooperation of all law enforcement officers and it is for this reason that you are being requested to cooperate to the fullest extent in making the operation of the Identification Division a success by contributions in the form of information and data upon criminals arrested by you. The necessary forms upon which such information and data should be submitted will be sent to you under separate cover.

In view of the accumulation of work incident to the establishment of the Identification Division, there will, no doubt, at first be a number of minor delays in answering inquiries which you may have occasion to make. This is due to a number of difficulties beyond the control of this Department which have been in existence for several months. There is now available, however, an appropriation and a trained personnel to operate the Identification Division and, commencing on July 1st, every effort will be made to render the proper service to the law enforcement officers of the United States.

There is being printed and will be forwarded to you as soon as it is off the press a pamphlet giving complete and illustrated instructions on the preparation of good prints. I cannot impress upon you too forcibly the necessity for submitting good prints.

-2-

There will be forwarded to you also a liberal supply of the printed forms for the prints and photographs, where you use photographs. Where a photograph is used please see that it is pasted upon the finger-print form instead of being mounted on a stiff cardboard and attached to the finger-prints with a paper clip. I remind you that a perfect print is useless unless we are sure to have the post office address of the place where taken. This is more valuable than the perfection of the print.

Where inquiries are made by telegraph the subscriber, of course, will bear the charge.

All communications and correspondence without exception intended for the Identification Division or connected with finger-prints should be addressed impersonally:

"Director, Bureau of Investigation,
Department of Justice,
Washington, D. C.

Identification Division."

I want you to feel free at all times to let me have, without hesitation, any expression which you feel will result in the improvement of the service and I assure you that if you have any complaint upon any matter connected with the work which reasonably can be corrected, immediate steps will be taken towards correction. I feel that the identification work for the first time is being placed upon a positive operating basis and merely ask your indulgence and consideration of the enormity of the task confronting us in bringing up to date the classification and filing in order that we may render you more efficient service in the months to come.

Very truly yours,

J. E. Hoover,

Acting Director.

~~Recent~~ interest in data quality

Sample Quality

I cannot impress upon you too forcibly the necessity for submitting good prints.

Standards

There will be forwarded to you also a liberal supply of the printed forms for the prints and photographs, where you have photographs. Where a photograph is used please see that it is pasted upon the finger-print form instead of being mounted on a stiff cardboard and attached to the finger-prints with a paper clip.

Best Practices

There is being printed and will be forwarded to you as soon as it is off the press a pamphlet giving complete and illustrated instructions on the preparation of good prints.

Data Integrity

I remind you that a perfect print is useless unless we are sure to have the post office address of the place where taken. This is more valuable than the perfection of the print.

What does it mean to talk about biometric data quality?

- It depends on who you are. We can't forget that a biometric system exists in the overlap of four domains:
 - Pattern recognition
 - Domain expertise
 - System engineering
 - Database

How is quality defined?

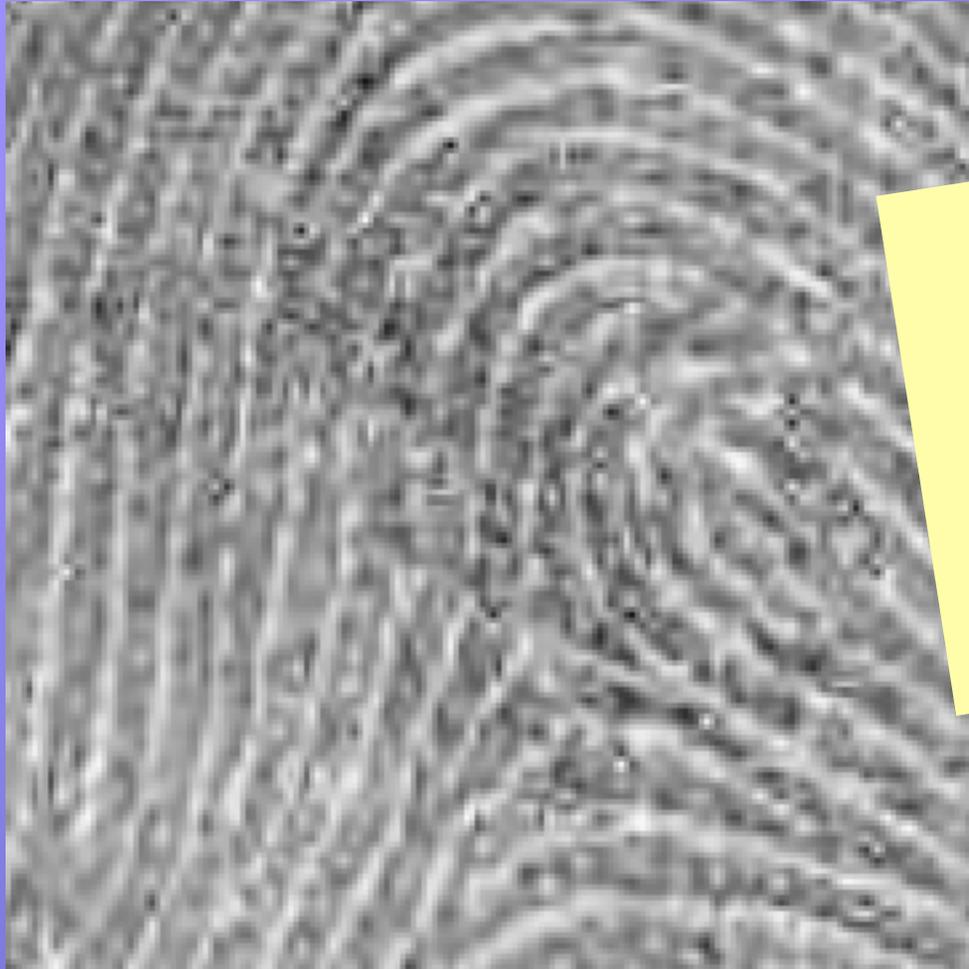
According to M1 DRAFT Biometric Sample Quality Standard:

- **Character**
 - Intrinsic physical data content

- **Fidelity**
 - The accuracy with which physical characteristics are represented in a sample (image)
 - *This is the only definition that everyone seems to agree with*

- **Utility**
 - The value of a sample for a given purpose
 - *Arguably the only way of quantifying how effective a quality metric is*

What is poor quality?



Overcompression:

Poor **Fidelity**

What is poor quality?



Interlacing:
Poor **Fidelity**

What is poor quality?



Distortion:
Poor **Fidelity**

What is poor quality?



Limited data content:

Poor **Character**

What is poor quality?

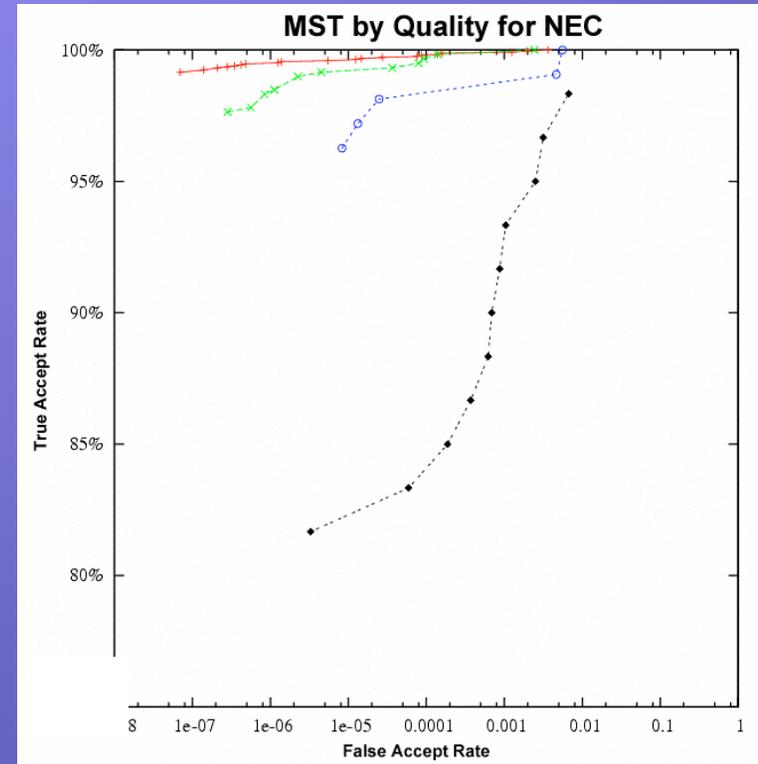
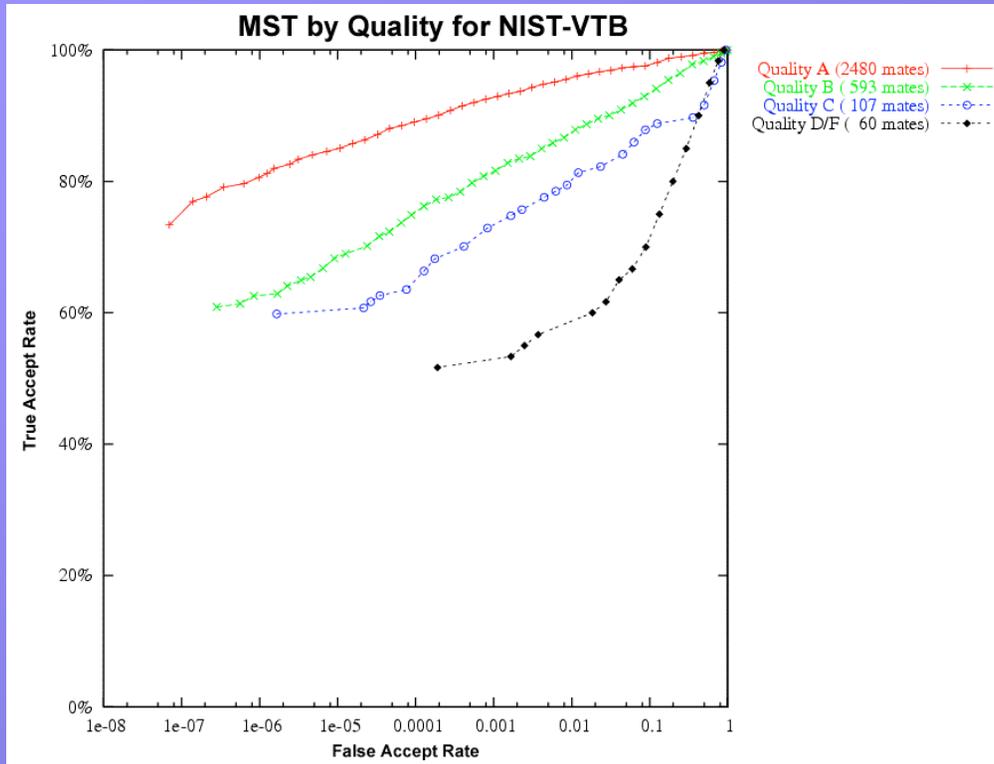


Dissociated ridges:

Poor **Character**

(hard for most matchers,
but especially distinctive
to a fingerprint examiner)

Defining quality in terms of utility



... it depends on what you use it for.

Defining quality in terms of utility

- Matching accuracy is affected by
 - Probe sample (image) quality
 - *Character & fidelity*
 - *Much but not all of this is matcher-independent*
 - Gallery sample (image) quality
 - *Character & fidelity*
 - *Much but not all of this is matcher-independent*
 - Variability/correspondence between samples
 - *Overlap and relative distortion between probe & gallery sample*
 - *Some argue that these are really just immeasurable cases of quality factors*
 - Metadata quality / Data integrity
 - *Biographic, cross-reference, or indexing errors*
 - *Data aging*

Please see our report to US-VISIT:
Hicklin, Khanna; “The Role of Data Quality in Biometric
Systems”; February 2006.
http://www.mitretek.org/Role_of_Data_Quality_Final.pdf

Austin Hicklin
Mitretek Systems
hicklin @ mitretek.org