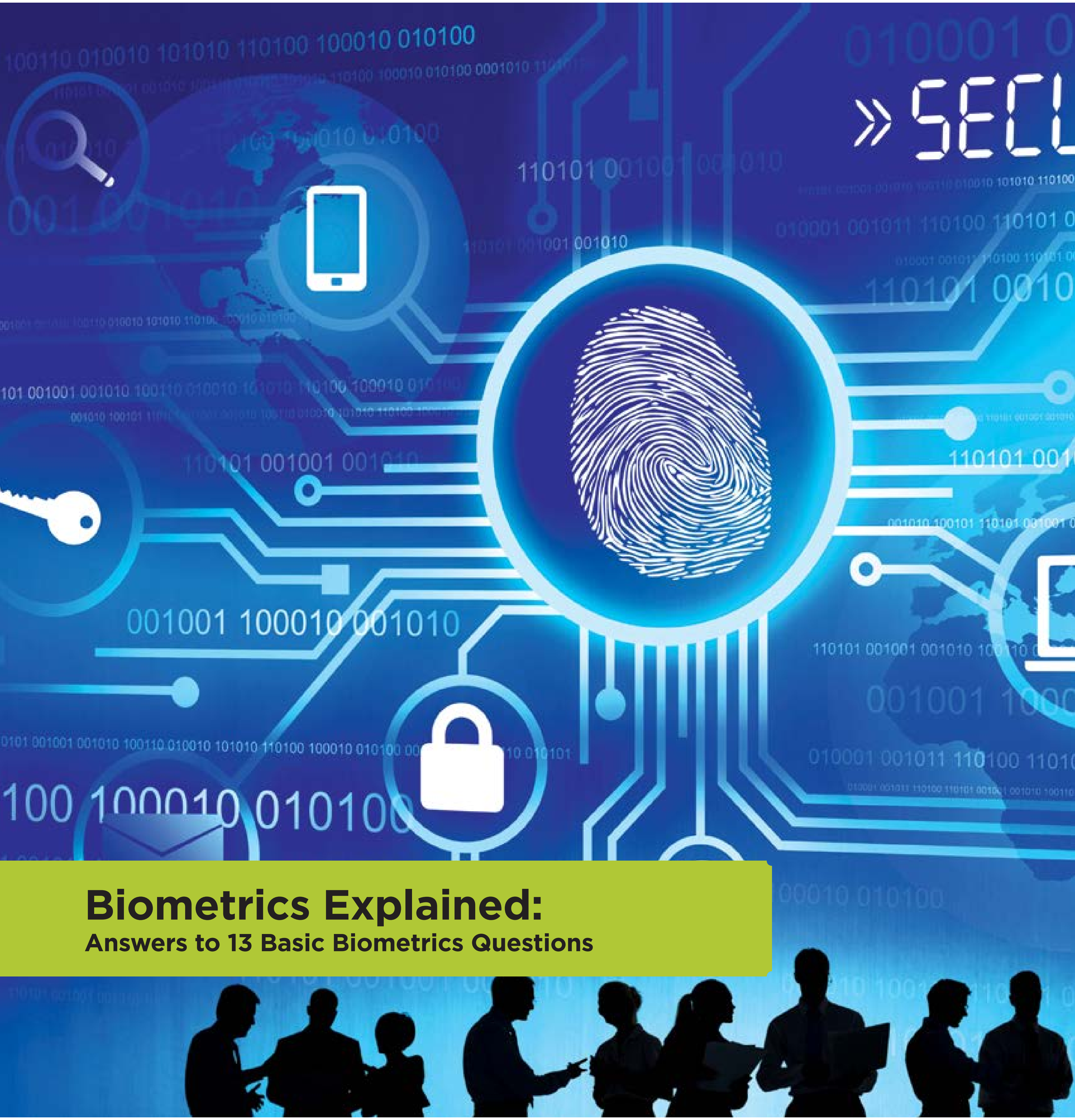

Biometrics Explained:

Answers to 13 Basic Biometrics Questions

IBIA
International
Biometrics+Identity
Association

IBIA advances the adoption and responsible use of technology-based identification solutions to enhance identity security and privacy and to facilitate convenience and productivity for government, business and consumers.

Biometrics Explained: Answers to 13 Basic Biometrics Questions

As biometric technology has become more widely adopted, it has brought with it a number of questions. In this paper, IBIA seeks to answer many of the most commonly raised questions in a forthright manner.

What are biometrics?

The National Research Council of the National Academy of Sciences, in its 2010 report "Biometric Recognition: Challenges and Opportunities," offered the definition: "Biometrics is the automated recognition of individuals based on their behavioral and biological characteristics. It is a tool for establishing confidence that one is dealing with individuals who are already known (or not known)—and consequently that they belong to a group with certain rights (or to a group to be denied certain privileges)."

Are biometrics really effective given the common complaint they are not perfect?

No system is perfect. The real question is whether the particular pros and cons of a biometrics system make it workable, and if so, whether it's better than the available alternatives. In its 2010 report, the National Research Council made the point: "Human recognition systems are inherently probabilistic, and hence inherently fallible. The chance of error can be made small but not eliminated."

For example, high-end commercial fingerprint matching systems have fully automated search reliability of 99.6% or better. Coupled with human examiners, which is typically done in government systems, the search reliability can exceed 99.99%. This is exceptionally high reliability for most purposes. Iris based matching accuracy is comparable to known source fingerprint matching. DNA based identification of unrelated persons is yet more accurate, but in comparison to fingerprints, face, or iris is much more expensive, has a far slower response time, is considerably more invasive, and has legal restrictions on its use for many purposes. Face matching systems, under ideal conditions, approach 98% reliability. High-resolution portrait-style capture under favorable lighting with no others in the frame lends itself to high accuracy searching. However, facial recognition systems very rarely are employed under ideal conditions.

The performance of nearly all existing facial recognition systems degrades rapidly with increasing yaw, roll, or pitch of the head, and common occlusions of the face like glasses, hats, or hair covering one or both eyes. As these are commonly encountered situations, facial recognition can be less effective for some use cases. However, this is true of most methods of identification, including biometrics. The best ones to use are dependent on the environment, distance between the user and the system doing the identification, time required, and the application. Behavioral trait systems are most useful in situations where humans are in continuous contact with the systems that use this modality, like smart phones or commercial or government computer applications. Speaker recognition has proven valuable for use in call centers and other situations where remote identification of telephone callers is required. Speaker recognition or speaker verification can be made most effective when used in conjunction with other identity factors, like caller-ID with a registered phone number, a concurrently presented password or ID, another biometric, an electronic token, texted "out-of-band" PIN numbers, and other factors. There are dozens of less well known biometric modalities (e.g. subdermal vein structure is widely used in Asian financial institutions) that can also be useful for particular applications.

In all systems, manual and automated, there are false negatives as well as false positives. In border crossing identification applications, a small segment of the population will be referred to a secondary process for final determination. How large that segment will be is one of the considerations in the architecture and design stage. Typically, false positives are referred to human agents for final determination and the threshold is set at a rate the available labor can comfortably support.

As an example, for a visa or passport photograph comparison, at a false match rate of 1 in 10,000, the best performing automated systems are better than 99% reliable (1 miss in 100), which compares to a roughly 20% human examiner miss rate (according to the 2014 study "Passport Officers' Errors in Face Matching" conducted by David White, Richard Kemp, Rob Jenkins, Michael Matheson, and A Mike Burton). Error rates for other human recognizers (most of the rest of us) have been reported as high as 50% or more, though training can reduce this error rate, and some super-recognizers (less than 2% of the population) perform exceptionally well with little training. Other end of the spectrum of human face recognition performance, about 2% of the population can't recognize faces at all, a condition known as prosopagnosia. Other limitations of humans that don't apply to automated systems are attention span, emotional impairment, or observational skills. This effect is often observed investigators when interviewing multiple witnesses of the same crime.

For each modality, except for DNA, there are small segments of the public for which it is unusable, such as amputees for fingerprints, those with damaged eyes for iris, and so on. Thus, every biometric system will require a secondary process to address population outliers. In some cases, the secondary process may be yet another biometric. To allow for universal usage, there always needs to be a provision for referral to human final determination.

Should I be concerned that biometrics are not secret?

Indeed, biometrics are not secret. The most reliable form of authentication is personal recognition. We recognize each other through our faces, voices, gaits, scents and so on. All these things are publicly revealed throughout our lives. Secret does not equal "protected" and publicly revealed does not mean "vulnerable." Both manual and automated identification systems can be, and have been, compromised by human error, intrigue, and technology. Biometric authentication and identification systems are simply not based on hidden information. While fingerprints and iris scans are not accessible to most people, the underlying information is revealed whenever we are in public. The lack of secrecy doesn't make biometric systems less effective. It is in fact quite difficult to fool biometric systems, particularly when human supervision is an element of the process. Nor is it so easy to fool unattended sensors as is sometimes asserted.

Why is biometrics better than name-based identification — what is different about biometrics?

The simple answer is that 1) names can be easily changed or 2) translated incorrectly or 3) entered erroneously. In contrast, biometrics, when implemented well, can be difficult to change or spoof without detection, and we all carry them through life.

Names and other biographic data can be easily changed and are susceptible to entry errors in databases. Terrorists and criminals often change their names to avoid detection.

Foreign names translated into English can prove problematic when doing a names-only match against watch lists. For example, the name "Muhammad" has at least 19 different spellings (Moohammed, Mahmad, Mehmed, Mahamed, Mohamad, Mohamed, Mohammad, Mohammed, Muhamad, Muhamed, Muhamet, Muhammed, Muhammet, Mahammud, Mehmet, Mohd, Muh,"Mohamed","Mahamid" – see https://en.wikipedia.org/wiki/Muhammad_(name)).

This also illustrates the third difficulty - the potential to mistype or misspell a name. For instance, Boston bomber Tamerlan

Tsarnaev was supposed to be pulled aside for questioning during one of his transits through JFK airport, but he slipped through undetected because someone had misspelled his name in a security database.

As another example, compare your biographic information recorded by the three major credit rating companies Experian, Equifax, and TransUnion – like most of us, you'll find discrepancies.

A number of years ago the Bureau of Justice Statistics, responding to an ever-expanding set of difficult policy issues associated with the increasing demand to conduct national criminal history background checks on individual applicants, commissioned a study of Interstate Identification Index Name Check Efficacy. Then, as now, many private and government organizations wished to avoid the delays and expense of fingerprint based background checks compared to name based (more accurately biographical information including names, ages, addresses, and so forth) background checks. "Interstate Identification Index Name Check Efficacy", Report of the National Task Force to the U.S. Attorney General, dated July 1999, NCJ-179358 was the result.

Independent subject matter experts, well versed in matters of criminal history, name based checks and fingerprint based checks, conducted the study. Name based checks resulted in 11.7% false negatives and 5.5% false positives. That is almost 12% of persons with felony level arrests or convictions were missed while approaching 6% of persons with no criminal record whatever were incorrectly reported as having a criminal record. Eighteen years later there is no reason to believe the situation is different either for public source information or for government held records.

How does biometrics prevent ID theft?

ID theft takes many forms, ranging from a minor borrowing, or fraudulently obtaining, a driver license to gain adult privileges, to a validly issued passport but based upon someone else's documentation, to filing to receive another's tax refund, to setting up a false account in someone else's name, and more. Often ID theft occurs using personal or biographic data obtained through a database hack or through social engineering. Social engineering is the use of deception to manipulate individuals into divulging confidential or personal information that may be used for fraudulent purposes. But incidents of "dumpster diving" have occurred to obtain such information. Beyond that, most biographic information is a matter of public record. At one time, obtaining those records meant an in-person trip to a courthouse or other public records repository and was generally impractical; whereas today most such information is readily available for a small fee. However, biometrics are not today public records and, even though they are publicly displayed, are not readily collected in a useable form. Requiring biometrics as one or more of the factors needed to authenticate transactions greatly enhances the security. Likewise, requiring biometrics in new account setup, then requiring the presentation of a biometric to conduct transactions, will make the account more secure.

How do you protect biometric data?

Biometric data are no different from other personal data – they need to be protected appropriately, and this includes adherence to good cyber hygiene practices. Encryption, access controls, timely security patches, and other standard security measures are still essential. Companies whose biographic databases were hacked, such as Equifax, did not follow these principles. Like a company that keeps its account usernames and passwords stored together, in clear text form, a company that stores my data improperly is subject to compromise.

However, not all compromised data is equally useful. A list of clear-text passwords can be immediately and quickly used in a multitude of problematic ways. But the template encoding typically used in the biometric match processing (the data that would be obtained in a breach) is a type of encoding that must be reverse-engineered to be useful to either access existing accounts, or set up new ones, a not insignificant task requiring less common technological resources and skills to accomplish.

Have there been breaches of biometric databases?

Generally, biometric data exposures are almost non-existent, for two reasons. The systems are difficult to compromise (as noted above) and are not attractive as targets like personal biographic data or medical records, which can more easily be monetized directly or used for identity theft.

In contrast to biographic data, which has been subject to numerous high-profile and highly damaging breaches, the only known biometric breach of any kind was the 2015 OPM fingerprint hack in which the fingerprints of 5.6 million people were reported by OPM to have been stolen. Consequences of this reported exposure are yet to be realized. If the hack was by Chinese state actors as theorized, we may never see a widespread impact comparable to that possible when the hackers are criminals, hacktivists, terrorists, or similarly malevolent players. However, a potential exposure exists for any of the 5.6 million people involved when crossing the

Chinese border. If the Chinese digitized the fingerprint images and enrolled them in their border security fingerprint check system, it is possible that undercover agents or people with high security clearances could be identified and detained or subject to coercion.

What happens if someone steals my biometrics?

The argument is sometimes advanced that biometrics as an identification or authentication mechanism presents a unique cause for concern. While it is always possible to revoke a username and password, it is not possible to revoke a biometric. While that is certainly true, it does not really address the issue of utility or effectiveness, for several reasons:

1. Having a true copy of your biometric does not equate to being able to falsely present it and have it accepted. This is the issue of spoofing (presentation attack) addressed below. High-end sensors, perhaps augmented with additional modalities or other protections, make presentation attack impractical in most cases.
2. Academic research has shown that it is theoretically possible to combine biometrics with other information, encrypt the result, and then store that as the reference "cancellable biometric." This cancellable reference can be matched in the encrypted domain, yet it remains impossible to recover the original biometric should the reference be compromised.
3. Ultimately, as previously stated, the real question is whether biometrics afford better protection than other alternatives. It is certain that usernames and passwords provide limited protection. Such compromises are frequently reported. Hundreds of millions of people employ fingerprints daily to access their personal mobile phones and have for years. No widespread problems have yet been reported.

What is spoofing?

Spoofing is the use of an artifact containing a copy of the biometric characteristics of a legitimate enrollee to fool a biometric system (Anil Jain, "Encyclopedia of Biometrics")

How can you prevent spoofing?

Different biometrics are associated with different spoofing techniques, sometimes called "presentation attacks". Biometric sensor/software makers are aware of the attack techniques and build features into their sensors and software to counter the most prevalent attack types.

1. Fake fingerprints (e.g. latex or "gummy bear" attacks) can be countered by sensor designs that include liveness detection (e.g. galvanic skin response, pulse, pulse-oximetry, and vein pattern detection among others).
2. Contact lens iris attacks can be countered by liveness detection such as measuring the pupillary response to light.
3. Facial recognition compromises based on photos, masks, or videos may be prevented by requiring the subject to blink or speak, or using 3D technology (like the new Apple iPhone X).

With all biometric modalities, there are multiple ways to detect liveness. The level of sophistication actually employed is based on the risk to any particular system, with the objective being to make the cost of spoofing (in time, money, and specialized skills needed) high enough so that it isn't worth pursuing. NIST has been investigating quantification of this concept through their work on SOFA or "Strength of Function for Authentication." This includes making multi-factor biometric authentication stronger than alternative means of authentication, while also making presentation attacks far less attractive to criminals.

Are all types of biometrics susceptible to spoofing?

Generally, there is at least a theoretical spoof and counter-spoof for all the biometric types, though some are stronger than others. For example, at present, it is very hard to spoof someone's DNA.

How does the effectiveness of biometrics vary with age? (How old must children be before biometrics are effective?)

For most modalities, biometric technology is effective throughout the lifespan. The age at which biometrics become effective depends on the modality in question.

1. Fingerprints form in the womb and remain similar throughout life unless injured or altered. According to research done by Anil Jain (see http://msutoday.msu.edu/news/2016/identifying-children-and-saving-lives-one-thumbprint-at-a-time/), fingerprints of children as young as 6 months can be correctly identified 99% of the time just based on their two thumbprints. This principle probably also applies to other so-called friction ridges like those on the palms or soles of the feet. (This is the reason for taking inked baby footprints or the Armed Forces taking footprints of all their pilots.)

2. Irises are very stable over time, like fingerprints (see Daugman, J, 2003. The Importance of Being Random: Statistical Principles of Iris Recognition. Pattern Recognition 36: 279– 291).
3. Faces are generally changing rapidly through youth, and the discriminating features we get as we age (e.g. blemishes, creases, lines, scars and cracks) are not prevalent on the relatively smooth faces of young people. For this reason, face recognition is generally not effective on very young children.
4. Speech can be very discriminating as a biometric, especially for adult subjects. Like faces, speech does change continuously through childhood, though it is discriminating for segments of time during maturation. However, automated speaker recognition, while studied for over 70 years, has yet to emerge as a reliable mechanism for individualization.
5. DNA is very discriminating from the formation of a fertilized egg throughout a lifetime. Note however that identical (technically monozygotic) siblings have the same DNA except for random mutations that occur over the course of a lifetime. There are about 3 sets of such children per 1,000 live births in the USA. Such children cannot be readily distinguished by their DNA although their fingerprints and irises are clearly distinguishable.

Are biometrics equally effective across different population groups?

Biometrics are effective across different groups, including ethnic groups. However, the statistics may differ in terms of true match rate (sometimes called "accuracy") and error rates. The most striking example of this is DNA. The frequency of the alleles used in human DNA identification, such as that used by the FBI's CODIS system, varies depending on the population group. (An allele is one of a pair of genetic markers – one from the father and one from the mother - that appears at a particular location on a particular chromosome.) These variances are listed in associated "pop stats" (population statistics) used in conjunction with DNA testing to determine the frequency of potential occurrence of a DNA sample within the sub-group in which the subject resides. This has implications in the legal system when DNA is used as evidence.

Similarly, there are variations in the depth of friction ridges and population activities that can render capturing and accurate matching of fingerprints problematic. For instance, in the India Aadhaar program where 1.21 billion people have been biometrically enrolled, the Unique Identification Authority of India (UIDAI) found that fingerprints alone were not adequate to uniquely identify the populace at this level. Friction ridges were sometimes thinner, either due to unique characteristics of the population sub-group or due to the agrarian nature of the society (wearing the fingerprints down) or both effects. Therefore, the UIDAI added iris as a secondary biometric factor to provide their required level of certainty at this population level. This is the case among other Asian populations, and especially women.

Facial recognition is often cited for variances across ethnic groups, and this has been referred to as algorithm 'bias'. This is semantically loaded terminology. Machines do not exhibit bias. Any bias would have to be included, intentionally or otherwise, by the human designers. There have in the past been variations in accuracy and error rate across sub-groups, which can be impacted among other factors by the training data used. For example, the relatively smooth faces of children are more difficult for machine algorithms to identify uniquely due to lack of discriminating features (as discussed before). In visible light, it is more difficult to discern skin blemishes and other features on dark-skinned people than it is on light skinned people (who exhibit greater contrast). It should be noted that the most recent independent testing showed that high performance face matching algorithms had no material differences in performance for different demographic groups of the same gender. Match performance was still slightly better for males than for females.

To achieve similar levels of accuracy across the many diverse populations, developers create and test algorithm accuracy in diverse populations. In addition, examiners and forensic specialists who work in this field have developed procedures designed to minimize uncertainly due to such variations while at the same time transparently acknowledging the normal variations between groups.

For example, if the gender and/or ethnicity of a probe subject is known, the examiner can select a subset search gallery that contains only people of the same gender or ethnicity. Smaller galleries yield faster and more assured results compared to unconstrained searches if the whole database is relatively large. If the search is in the context of an investigation, the search is set to return a number of candidate matches (2 or more) that are then used in conjunction with other information to develop investigatory leads.

Identity Matters

IBIA

International
Biometrics+Identity
Association

1090 Vermont Avenue, NW • 6th Floor
Washington, DC 20005

202.789.4452 x1309
IBIA.org