The Relevance of Untraceable Biometrics and Biometric Encryption:

A Discussion of Biometrics for Authentication Purposes

**Information and Privacy Commissioner
Ontario, Canada**

European Biometrics Group

August 2009

Ann Cavoukian, Ph.D.
on the term
Untraceable Biometrics

Max Snijder
on the term
Anonymous Biometrics

"*Untraceable Biometrics* (UB) is a term I developed to define a new class of emerging privacy enhancing technologies, such as BE, that seek to irreversibly transform the biometric data provided by a user. UB has five distinguishing features:

- There is no storage of a biometric image or conventional biometric template;

- The original biometric image/template cannot be recreated from the stored information, thereby rendering it untraceable;

- A large number of untraceable templates for the same biometric may be created for different applications;

- Untraceable templates from multiple applications cannot be linked together; and

- An untraceable template may, however, be revoked or cancelled.

These features embody standard fair information practices, providing for the time-honoured privacy principles of user control, data minimization, and a high degree of data security."

"*Anonymous biometrics*, in my view, refers to: a system where the biometric data are not being connected to any personal data; and, furthermore, where the biometric data is prevented from being taken to another system that is capable to connect the biometric data to personal information. This enables biometrics for authentication only, e.g., where a trusted third party (TTP) authenticates the biometrics on request of a user or a service provider. This is a crucial element for using biometrics in electronic environments (e.g., for e-services) and requires a special management of the biometric data, from capturing to storage, retrieval and matching. Anonymity might be associated with a lack of accountability, but it is a basic human right on the other hand. UB is a means to protect that right."

The Relevance of Untraceable Biometrics and Biometric Encryption:

A Discussion of Biometrics for Authentication Purposes

Context

The initial usage of biometrics dates from a time in which the storage medium of any record was ink and paper. Under this system, samples taken from a person — so-called 'live' data — were compared, by hand, to a reference sample in order to determine a match. The manual nature of this system largely precluded 'one-to-many' matching, in which a sample was compared against many or all stored records; instead, biometrics were generally used in a 'one-to-one' matching, to confirm that the sample came from a known individual. Additionally, as the biometric data were all stored in a physical means, the records (and the privacy thereof) could be protected by relatively simple physical access controls.

Currently, however, biometric data are commonly being integrated into a range of large and complex ICT systems and processes, and access to this data is becoming virtual rather than physical. This digital storage of data makes records, at least conceptually, accessible from anywhere in the world, making it difficult to control and verify authorized access. The digitization of data also broadens the usage possibilities, allowing for easy database-wide matching against live samples. Such developments make digitally-stored biometric data significantly more vulnerable to manipulation, theft and loss of control.

The need for a securely stored, non-vulnerable means of verification/authentication is becoming a key topic for discussion in the deployment of biometric systems for both government and commercial services. This is the context for this vision paper, which focuses on the need for secure storage and privacy protection for biometric reference data.

State of the Art: Biometric Encryption (BE)

There is a class of privacy-enhancing technologies that seeks to provide better security and privacy protection than conventional biometrics, by irreversibly transforming the biometric data provided by the user. These are referred to as *Untraceable Biometrics*, of which Biometric Encryption (BE) is a leading approach. It is this emerging area of privacy-enhancing biometric technologies that challenges the traditional zero-sum paradigm (privacy must be traded for security, and vice versa), by making possible the enhancement of *both* privacy and security, in a positive-sum manner.

BE is a process that securely binds a PIN or a cryptographic key to a biometric, or generates a key from the biometric; in either case, neither the key nor the biometric can be retrieved from the stored reference data. Instead, the key is recreated only if the correct live biometric sample is presented on verification. BE is sometimes mistakenly viewed as a technique to encrypt biometric images or template data. This is not the case. With BE, there is no storage (and therefore no potential for de-encryption or loss) of a biometric image or conventional biometric template.

Research on BE started in the mid-1990s and is now sufficiently mature to warrant consideration as a new facet of biometric verification and authentication schemes. The co-authors of this paper agree that the benefit of BE lies not only in the manner in which the biometric is rendered anonymous within an authentication scheme, but also in the fact that it can be rendered untraceable. An illustration of this extension from anonymous to untraceable is as follows:

Consider a transaction that may require biometric verification or authentication; for instance, the presentation of a biometrically protected ticket for an event. It is possible, with conventional biometrics, to anonymously show that the ticket holder is indeed the rightful owner of the ticket, and that the ticket has not been forged. If this ticket were to provide access to several events, however, an individual could be traced through the comparison of biometric templates, even while remaining anonymous at each individual event. With BE, however, no biometric template is stored. Instead, separate individual and unconnected PINs or keys are generated for each event; though they have been derived from, or encrypted by, the same biometric, they are randomly generated and therefore completely unlinkable and untraceable.

Max Snijder's early work on anonymous biometrics and e-Identity and Dr. Cavoukian's seminal work in the area of BE have made the co-authors excited about the prospect of future considerations as BE comes onto the biometric scene. Their discussion is meant to touch on the distinguishing features of the algorithmic process defined by BE, the terminology that best describes this process, and to provide a springboard for further thought leadership on the policies and guidelines needed to distinguish between conventional and untraceable biometrics schemes.

Ann Cavoukian on the term *Untraceable Biometrics*

Biometric data, such as fingerprints and facial images, are unique and permanent identifiers that are widely understood to be private, sensitive forms of personal information, worthy of legal protection. But some biometric data are also arguably semi-public information — our faces may be viewed everywhere we go, and our fingerprints are left behind on everything we touch.

This dual nature of biometric data makes it ill-suited for use as secret passwords (e.g., for access control) and for storage in centralized, networked databases, where it can be misused in many ways that can profoundly impact the individual. Yet identification and verification are the primary purposes for which biometric data are being used today. Both depend on the storage and routine matching of biometric data.

As the use and deployment of biometrics becomes more widespread across society, so will the privacy and security risks associated with the growing collection, use, disclosure and retention of personal biometric data. These risks include: loss of individual control over one's personal information; unauthorized cross-matching; unauthorized secondary uses; the prospect of surveillance, profiling and discrimination based upon biometric data; and the loss, theft, misuse and abuse of this data resulting in identity fraud, identity theft and other negative impacts on the individual. At stake is the trustworthiness of biometric systems and the behaviour of organizations that collect biometric data.

Techniques such as BE can directly address most of these risks at the source, simply because an individual's actual biometric data need never be disclosed, collected or retained to begin with. BE embodies three core privacy practices:

1. **Data minimization:** no retention of biometric images or templates, minimizing the potential for unauthorized secondary uses, loss, or misuse;

2. **Maximum individual control:** individuals may restrict the use of their biometric data to the purpose intended, thereby avoiding the possibility of unauthorized secondary uses (function creep); and

3. **Improved security:** authentication, data transmission and data security are all enhanced.

Instead of presenting actual biometric data that can be lost or misused, BE transforms the individual's primary biometric image into a virtually unlimited — and non-reversible — range of unique single-purpose data strings that can be used without fear of loss, correlation or misuse. These strings can then serve as passwords, encryption keys, or as identifiers for use anywhere that traditional biometrically-enabled systems are presently used (e.g., authentication, access control).

Better still, if one's unique BE identifier ever becomes compromised in any way, BE techniques allow it to be easily revoked and a completely new identifier generated from the same biometric, much like a new credit card number is issued when an old card is lost or stolen. With traditional biometrics, it is impossible to change one's fingerprints or face if one's biometric data is compromised. Using BE, this problem is mitigated because the original biometric data is never retained but only used to generate other identifiers. Thus, security is vastly improved.

I have been a long-standing proponent of BE technologies since the mid-1990s and devote considerable time and effort on promoting BE. I continue to press for strong privacy protections in the development and deployment of interoperable biometric technologies.

Untraceable Biometrics (UB)

Untraceable Biometrics (UB) is a term that I developed to define this class of emerging privacy-enhancing technologies such as BE, that seek to irreversibly transform the biometric data provided by a user. UB has five distinguishing features:

- There is no storage of a biometric image or conventional biometric template;

- The original biometric image/template cannot be recreated from the stored information, thereby rendering it untraceable;

- A large number of untraceable templates for the same biometric may be created for different applications;

- Untraceable templates from multiple applications cannot be linked together; and

- An untraceable template may, however, be revoked or cancelled.

These features embody standard fair information practices, providing for the time-honoured privacy principles of user control, data minimization, and a high degree of data security.

UB include two major groups of emerging technologies: BE and Cancelable Biometrics (CB). BE technologies securely bind a digital key to a biometric, or generate a key from the biometric, so that neither the key nor the biometric can be retrieved from the stored information (the latter is often called "helper data"). The key is recreated only if the correct biometric sample is presented up on verification; therefore, the output of BE verification is either a key or a failure message. On the other hand, CB technologies apply a transform (usually kept secret) to the original biometric and store the transformed template. The transform can be either invertible or, preferably, not. On verification, the same transform is applied to a fresh biometric sample, and the matching is done between two transformed templates. The output of CB verification is a Yes/No response, as in conventional biometrics.

Although I had initially considered using the term "Anonymous Biometrics" for BE, I favoured instead using the term "untraceable." Post 9/11, the term "anonymous" is often associated with a lack of accountability (e.g., an anonymous user makes defamatory comments on a website, or, worse, a terrorist organization abuses anonymity to raise funds or other illicit purposes). In addition, BE can be used in both anonymous and identifying one-to-one modes. Even in the latter

case, however, BE still provides significant privacy and security benefits, such as data minimization (no way to reconstruct the original biometrics), and the ability to change/revoke one's template. Conventional biometrics can also operate anonymously, if no personal data are attached to the biometric template. However, those templates appearing in different databases can be linked together, making this kind of anonymity largely illusive.

I agree that biometric "traceability" may refer both to leaving physical traces (e.g., latent fingerprints) and to tracing back to personal data. It seems that the second meaning of traceability which I have used, i.e., referring to the data connectivity involved, is the more commonly used definition. Clearly, the terminology in this area has not been well established. For example, the "Harmonized Biometric Vocabulary" ISO standard does not address these issues.

I also note with interest that there has been some work done on template-free biometric key generation. There is a company whose owner claims to have invented "traceless" biometrics — a 4-digit PIN is extracted from the biometric without storing any helper data (i.e., there are no traces in the storage).

For the physical traceability problem, I prefer to use the terms "covert/overt" biometrics, as defined by NIST (http://www.biometrics.gov/Documents/glossary.pdf) or IBG (http://www.bioprivacy.org/bioprivacy_main.htm). "Overt" means that end users are aware that biometric data is being collected and used, and acquisition devices are in plain view.

Indeed, leaving a physical trace (a latent fingerprint or a DNA sample) does not necessarily constitute the highest risk to privacy/security. It is much easier to take a person's photo than to lift a latent fingerprint. Stored photos are also widely available in digital format, e.g., on Facebook. In other words, a facial image is much more

accessible than a fingerprint or DNA. Moreover, the technology is constantly improving, such that it has become possible to take pictures of one's iris covertly, at a distance. It will likely be possible, in the near future, to also covertly obtain a finger/palm vein pattern. Thus, there will ultimately be no "traceless" biometrics that could not be captured covertly. I think that this problem should be considered in the application context, i.e., whether the system is designed to operate covertly or overtly.

I agree that untraceable biometrics no longer relates exclusively to "who you are," (as with conventional biometrics) but rather "what you have" (e.g., smart card, biometric or encryption key storage media) is now tied intrinsically to "who you are." UB elevates token and password-based systems to a new, higher level of security without compromising privacy — positive-sum, not zero-sum.

Max Snijder on the term *Anonymous Biometrics*

Historically, we associate biometrics with on-site identification: taking fingerprints with ink, storing them on paper and then relying on human inspection of these latent biometric identifiers at a physical, usually well protected, secure location. However, in the wake of the Information Technology Revolution of the 1990s, we are now moving from a "physical" environment to a new "virtual" domain where biometrics exist in a digital universe. It is now possible to store reference data at any location (worldwide) and search for millions of matches in a matter of seconds.

Traditionally, biometrics have entailed the use of fingerprints to establish the identity of a person about whom there is little or no information, and this is still the most common use of biometrics around the world. Nonetheless, biometric technology is evolving into its next phase –

"verification" – and in the near future we will be seeing this at border control checkpoints with the increased use of electronic passports. While the biometric data will still be stored at a protected physical location and controlled by the rightful owner of the passport, the passport holder's identity will be verified by a one-to-one check of the biometrics stored in the protected chip. This is a situation characterised by the physical aspects of the process: you have to cooperate by offering your passport to a border control officer.

Yet the link between the biometrics and the identity of the person is less evident. The passport's biometric is indirectly used to verify identity and not to create a direct link between biometrics and identity. In other words, the biometric reference data are made available to facilitate a single, one-off check but are still retained by the passport holder.

Even more encouraging, the next evolutionary step after "verification" involves the use of biometrics that have absolutely no link to identity at all – the sole purpose being for authentication only. Even in the most extreme case, a service provider wouldn't need to know which biometrics or whose biometric data are being used. Why? Because biometrics can be converted into secret keys.

But how secret can biometrics be? A face is public information. Using a face finder on a Web-search engine can reveal a person's identity within minutes, if not seconds. The same process applies to typing in a name, which could generate multiple pictures of a face, including plenty of personal information. As for fingerprints, these can be traced and "stolen" because they can be left anywhere (e.g., doorknobs, drinking glass, keypads, etc.).

There are now enough freely available databases where unintended links to a given identity can be established. New solutions are required for this problem. The choice of biometrics used

(traceable? detectable?) is one of them, as well as the way that reference data is stored and the performance level of live detection by sensors (automated? human supervision?).

Some biometrics, such as the human iris, obviously leaves no fingerprint-like trace or is very difficult to "spoof" or fake. The newly available optical vein pattern recognition technologies are another good example of a "difficult-to-spoof" technology. But the matter of storing reference data and performing the matching function still remains. Currently, we need to revert the biometric reference back from whichever encrypted domain into the original image or biometric template in order to use it for matching. That means that the original biometric is not only being exposed to the sensor, but also to the wider system behind that sensor – and that is where user control starts to become questionable.

New privacy-enhancing technologies will not only strengthen the protection of biometric data but also make it possible for their owner to control the storage and matching of his or her biometric data. From a single biometric (e.g., one fingerprint) these technologies can generate a non-biometric derivative and multiple unique (disposable) biometric "keys" for carrying out the matching function.

This in turn allows for biometrics to be revocable (in case a key is lost or stolen) and to be protected against being used as "leads" to other databases. This "biometric encryption" introduces new horizons for the correct and beneficial use of biometrics in more complex operating environments such as the Internet and large federated systems.

Biometric encryption has the potential of making both biometric data and the matching function anonymous by not exposing an individual's original biometric information to fraud or manipulation. Technology that

should accommodate this function is already commercially available.

Indeed, we can now clearly see the two emerging faces of biometrics: "identifying" biometrics for establishing or verifying identity where the link to identity is key; and "un-identifying" biometrics — or anonymous biometrics — where a link to identity is strictly avoided.

Given that these two "faces" will exist in parallel to one other and could get easily mixed up with each other, a new level of thinking about biometrics is needed - one that defines what policies and guidelines are required so that both kinds of technology can be used, without one compromising the other.

Anonymous Biometrics

Anonymous biometrics, in my view, refers to: a system where the biometric data are not being connected to any personal data; and, furthermore, where the biometric data is prevented from being taken to another system that is capable of connecting the biometric data to personal information. This enables biometrics for authentication only, such as a trusted third party (TTP) that authenticates the biometrics at the request of a user or a service provider. This is a crucial element for using biometrics in electronic environments (e.g., e-services) and requires special management of the biometric data, from capturing to storage, retrieval and matching. Anonymity may very well be associated with a lack of accountability, but on the other hand, privacy is also considered a basic human right. Untraceable biometrics is a means to protect that right.

Some observations/remarks on the term *Untraceable Biometrics*:

- Leaving physical traces is not the key element. Rather, the prime factor is the extent to which people are aware that they are submitted to a biometric check;

- With an eye on ongoing technical developments, any biometric characteristic can be taken covertly, either physically (e.g., latent fingerprint) or with contactless sensors (cameras, X-ray etc.). In that view, taking a fingerprint from a person's drinking glass without his or her knowledge is essentially the same as covertly taking an image from a person's iris. The difference is that the fingerprint stays (for a while) even though the individual may have left the scene;

- Biometrics that are used for establishing/ verifying identity have the purpose of connecting the biometrics to an identity, either to the root identity or an alias. If stored biometrics (or encrypted) are connected to personal data, these personal data can lead to the original biometrics (e.g., through Facebook or through conventional systems that use biometrics for identification, like the central repository of a public institution). So, in order to be totally untraceable, there should be no link to personal data;

- If a biometric reference is untraceable and not connected to any personal data, the biometric is no longer "who you are", but rather "what you have" and "what you know", (i.e., a secret key).

I do see a value in the term "untraceable" and I totally agree with Commissioner Cavoukian's definition. The use of the term "untraceable" refers to the reference data from a storage point of view, whether the data are remotely stored in a database or on a portable drive (e.g., smart card, USB stick, cell phone etc.).

About the Authors

Ann Cavoukian, Ph.D., Information and Privacy Commissioner of Ontario, Canada

Dr. Ann Cavoukian is recognized as one of the leading privacy experts in the world. Noted for her seminal work on Privacy Enhancing Technologies (PETs) in 1995, her concept of *Privacy by Design* seeks to embed privacy into the design specifications of technology, thereby achieving the strongest protection. An avowed believer in the role that technology can play in protecting privacy, Dr. Cavoukian's leadership has seen her office develop a number of tools and procedures to ensure that privacy is strongly protected, not only in Ontario, but around the world. She has been involved in a number of international committees focused on privacy, security, technology and business, and endeavours to focus on strengthening trust and confidence in emerging biometric applications. As a member of the International Biometric Advisory Council (IBAC), Dr. Cavoukian provides advice and expert opinion to the European Biometrics Forum (EBF), its members and partners, on the most pertinent privacy issues facing biometrics around the world today. In 2007, Dr. Cavoukian co-authored a white paper entitled *Biometric Encryption: A Positive-Sum Technology that Achieves Strong Authentication, Security AND Privacy*. Recently reappointed as Commissioner for an unprecedented third term, Dr. Cavoukian intends to "grow" *Privacy by Design* and make it go viral.

Max Snijder, CEO, European Biometrics Forum and Chairman of the International Biometrics Advisory Council (IBAC)

Max Snijder is one of the leading independent biometrics experts in Europe. His practical experience and integrated approach has lead to an overall view on the biometrics business. He is CEO of the European Biometrics Forum and chairman of the International Biometrics Advisory Council (IBAC). He is coordinator of BioTesting Europe. For the EC JRC/IPTS he produced the report, *Security and Privacy in Large Scale Biometric Systems*. Today, Max Snijder is involved in the key areas of the biometrics business. On a European level he is involved in numerous workshops, committees and expert groups, such as the Consortium on Security and Technology of the EastWest Institute, The Porvoo Group, the CEN Working Group on Integrated Border Management, and the CEN Biometric Focus Group. He is Founding Member of the IFIP Working Group on Identity Management (WG11.6) and member of the ePractice Working Group on eID. He recently became member of ThinkTrust, a European think tank on 'Investigating Security, Dependability, Trust, Privacy and Identity from ICT and social perspectives,' funded by the European Commission.

European Biometrics Group
Pr. Willem van Oranjelaan 4
1412 GK The Netherlands
Telephone: +31 (0)35 6321162
Website: www.eubiometricsgroup.eu
E-mail: info@eubiometricsgroup.eu

European Biometrics Group

Information and Privacy Commissioner of Ontario, Canada
2 Bloor Street East, Suite 1400
Toronto, Ontario   M4W 1A8
CANADA

Telephone: (416) 326-3333
Toll-free: 1-800-387-0073
Fax: (416) 325-9195
TTY (Teletypewriter): 416-325-7539
Website: www.ipc.on.ca
E-mail: info@ipc.on.ca