Guidance on Collection and Use of Biometric Data

Introduction

This guidance note is intended to assist data users[1] who wish to collect biometric data to comply with the Personal Data (Privacy) Ordinance (Ordinance). It should be read **BEFORE** data users decide on whether or not biometric data is to be collected, and if collected, be regularly referred to.

Biometric data includes the physiological data[2] with which individuals are born with and behavioural data[3] developed by an individual after birth. Biometric data is therefore data directly related to an individual. While it may not be reasonably practicable for a lay person to ascertain the identity of an individual by merely looking at the individual's fingerprint images or their numeric representations (i.e. the templates[4]), when the biometric data is linked with personal data in another database, a particular individual (also called "data subject" under the Ordinance) can be identified. For the purpose of this guidance note and for the reason above, biometric data is therefore considered to be personal data under the Ordinance[5]. As such, all those who collect and/or use biometric data are data users under the Ordinance. This guidance note seeks to recommend good practices in collecting and using biometric data.

The need for caution to handle sensitive biometric data

Biometric data could be sensitive data as it often contains an individual's intimate information relating to health, mental condition and/or racial origin[6], and it is often used for identification in criminal investigation[7] because of the uniqueness of the data. Any wrongful disclosure of biometric data could lead to unintended/unauthorised re-identification[8] of individuals, impersonation[9], or even discrimination due to unauthorised disclosure of intimate details of the individuals[10], which all entail grave consequences.

The appropriateness of the collection of biometric data and the precautions to be taken to protect such data collected vary with the level of sensitivity of the biometric data concerned.

1 As defined under the Personal Data (Privacy) Ordinance, a data user is a person who, either alone or jointly or in common with other persons, controls the collection, holding, processing or use of personal data.
2 For example, DNA samples, fingerprints, palm veins, hand geometry, iris, retina and facial images. Most physiological data cannot be changed.
3 For example, handwriting pattern, typing rhythm, gait and voice pattern. The behavioural data are prone to changes by the individual concerned either consciously or subconsciously.
4 Numeric representations or templates refer to information describing types and locations of major features of biometric samples/images (such as ridge ending, diversion, merger, etc. of a fingerprint) in relation to each other.
5 Under the Ordinance, "personal data" means any data (i) relating directly or indirectly to a living individual; (ii) from which it is reasonably practicable for the identity of the individual to be directly or indirectly ascertained; and (iii) in a form in which access to or processing of the data is reasonably practicable.
6 DNA is known to reveal the congenital conditions of an individual, together with gender and ethnic original, and increasingly believed to uncover mental health conditions and the inclination on personality. Retina images have been accepted as being able to tell the health condition of individuals. Some also believe that iris images can indicate individual's health and personality.
7 Fingerprints, DNA, facial images and handwriting have long been used by law enforcement agencies in criminal investigations for the purpose of identification.
8 Re-identification may happen when biometric data is leaked with or without other information. For example, if facial images of patients of a drug rehabilitation centre are leaked, it may either directly identify individuals who are famous or arouse interests in identifying the individuals.
9 For example, if the fingerprint templates are leaked, fake fingers with sufficient details can be produced which may be used to impersonate the owner of the template for gaining access to areas protected by fingerprint recognition system.
10 If DNA sequences or characteristics are leaked, it may lead to the assumption that the individual concerned has a higher probability of certain health or mental issues (but the individual has not been clinically diagnosed as such) and have services or job opportunities denied.

Keeping biometric data in their original formats may pose greater privacy risk than in their template form because the templates usually contain less details and offer little secondary use when compared with the original samples/images[11]. Data users should therefore, as soon as possible, derive biometric data templates from the original biometric samples/images for storage and subsequent use, and discard the original samples/images safely afterwards. The templates derived from biometric samples/images should be stored in such a form from which it is technically infeasible or difficult to convert back to the original graphical images.

Data users need to be aware of the sensitivity of the data concerned before deciding what data to collect and in what format they are to be kept. In this regard, the cost and the availability of biometric data readers and scanners should not be the prime consideration of the data users.

Good practices in collecting and using biometric data

Biometric data may be collected and used for various purposes. For example, in a bio-hazardous laboratory where access must be restricted to trained professionals, a retina or iris recognition system may be used for access control that does not involve any physical contact with the biometric scanner. Another example could be the use of palm-geometry recognition systems for access control and attendance recording by construction-site workers who have attained the necessary skills/safety certificates. In some cases, facial recognition or typing rhythm analyser may be deployed to continuously verify the identities of users of sensitive computer systems after the initial log on. Whether or not a particular type of biometric data could be collected depends on the purpose of their collection and the ways such data are collected.

(1) Necessity and proportionality

A data user should ensure that the collection of biometric data is for a lawful purpose related directly to its function and activity[12]. Examples of lawful purpose in this regard include the collection of DNA by law enforcement agencies for investigation of crime, facial images by immigration authorities for immigration control, or fingerprints by employers for control of access to high security and restricted areas.

The collection of biometric data must be "adequate but not excessive" for achieving such purpose[13] of investigation of crime, immigration control, or control of access to restricted areas, etc. Data users therefore have to consider whether it is feasible to collect less sensitive data to achieve the same purpose without compromising effectiveness.

In determining whether the use of biometric data is a proportionate measure to achieve the intended purpose, data users may make reference to the 4-stage proportionality test set out in *Hysan Development Co Ltd v Town Planning Board* [2016] HKCFA 66 (paragraphs 134 – 135), which determines the justification of a particular intrusion into fundamental rights and consider:

- whether the measure pursues a **legitimate aim**;
- whether the measure is **rationally connected** with advancing that aim;
- whether the measure is **no more than necessary** for advancing that aim; and
- whether a **reasonable balance** has been struck between the societal benefits of the encroachment of rights and the protected rights of the affected individuals, asking in particular whether pursuit of the societal interest results in an unacceptably harsh burden on the individuals.

(2) Data minimisation

Data minimisation is a demonstration of the "necessity and proportionality" principle. The level of privacy concerns varies with the amount of biometric data (including the amount of features of the biometric samples/images) to be collected. Minimum biometric data should be collected for achieving a purpose. For example, fingerprint data and facial images are the most common biometric data collected and used for the purposes of identification and verification, but the amount of features of the biometric samples/images required for identification and verification may vary.

[11] For example, the leakage of fingerprint images may allow a higher chance of re-identification, impersonation, and other usages than the leakage of their templates. The leakage of facial images instead of facial templates may more easily reveal the gender and ethnic origin of the data subjects.

[12] Data Protection Principle (DPP) 1(1)(a) in Schedule 1 of the Ordinance

[13] DPP 1(1)(b) and (c)

Identification

Identification involves the presenting of a live biometric sample and then asking the system to search and find a match from a database holding templates of many individuals. Because of the possible similarities in the various templates in the database, more reference points are usually needed in the template and from the sample in order to find a match in the database with certainty. For example, in the case of a facial recognition attendance system used in a company of 1,000 people, an employee arriving in the office needs to present his face to a camera, the system then needs to capture features of that employee and compare them with 1,000 templates in the database until only one person is confidently identified. In order to do this, both the features captured and the features stored in templates will have to be quite detailed so that no one else would be mistaken as the person arriving.

Verification

Verification, on the other hand, requires fewer reference points from the sample when compared with the identification process. Verification involves the presenting of a live biometric sample and then asking the system to verify whether or not it belongs to a specified person. In this process, the system merely needs to retrieve the template of the claimed individual from the database and confirms that it is the same or similar to the live sample. For a similar example, the employee arriving at work, apart from showing his face to the camera, also enters his staff number to tell the system who he is. In this case, the system only needs to capture fewer details of the face, retrieves the template of the declared employee, and confirm if they match. The system does not need to be overly concerned if there are other similar templates in the database and how similar those templates are when carrying out the comparison. As such, the level of details required to perform the verification will be less than the level of details required to identify the employee without first knowing who he is.

Very often, commercial organisations collect biometric data just for confirming the identities of individuals and as such they should choose verification biometric systems that operate in the way described above to minimise the number of biometric features to be collected.

(3) Privacy Impact Assessment ("PIA")

Given the wide range of sensitivity of biometric data, data users who intend to collect biometric data must first consider whether such collection is necessary at all[14]. To this end, they are encouraged to conduct a PIA, which is a systematic process that evaluates a proposal in terms of its impact on personal data privacy. Engaging a PIA could help avoid or minimise the adverse impact on the individuals concerned.

Below are some pointers to assist data users in conducting a PIA.

(i) The need for collecting biometric data

Data users should consider the following matters in order to determine whether collection of biometric data is necessary:

- What is the need for the collection of biometric data?
- If there is already a non-biometric system in place to serve the need and if it is not working adequately, can the inadequacy be remedied? If so, remedying the existing system is preferred to resorting to collection of biometric data.

(ii) Least intrusive option

Whenever there are different options for achieving the same purpose, the least privacy intrusive option should be adopted. In this regard, the collection of biometric data is usually more intrusive.

- If there is an alternative system that can be used to serve the same purpose as collecting biometric data, the alternative should be considered by evaluating its privacy intrusiveness.
- Less sensitive and/or less amount of biometric data should be collected to achieve the same purpose in order to minimise the privacy intrusiveness to the individuals concerned.

The above considerations would also help the data users justify the need for the collection of biometric data in the event of any legal challenge arising under the Ordinance.

[14] DPP 1(1)

Purpose and justification for collecting biometric data vary in different situations. While the Privacy Commissioner for Personal Data (Privacy Commissioner) will consider them on a case-by-case basis, some purposes are common and it would be useful to discuss them here for general guidance.

- Recording attendance: Attendance of staff or students is usually recorded by signing in personally or with the use of an access card held by the staff or students. Data users must have overriding reasons to justify the collection of biometric data instead of the less intrusive measures.
- Security control: While collection of biometric data may be justified by security reasons, e.g. to ensure that only authorised persons are permitted to enter restricted areas or to gain access to confidential information, the use of biometric data is not necessarily a better choice. Access to restricted areas or data may also be protected by passwords and access cards given to authorised persons. Installation of surveillance cameras monitoring restricted areas/ computer terminals with regular checks may further strengthen security.

Data users need to remember that the purposes of attendance recording and security control may often be achieved by other less privacy-intrusive methods, particularly when sufficient penalty for non-compliance of those methods is introduced.

Continuous and indiscriminative use of biometric scanners, such as installation of fingerprint scanners in all accessible areas including toilets, should be avoided as it would unlikely to be justified.

(iii) Whose biometric data should and could be collected

Strong justifications are required if biometric data of a large number of individuals are to be collected, as the potential damage caused by data breaches could be very serious.

Hence, where the collection of biometric data is to ensure only authorised entry, only the biometric data of those authorised persons should be collected.

Children of school age or individuals who are less capable of managing their own affairs are vulnerable and require stronger protection of their data privacy. Collection of biometric data from these groups, if challenged, will be critically examined by the Privacy Commissioner. In any event, it is objectionable for children of school age to be exposed to acts or practices that depreciate privacy, as they may as a result become less aware of the data privacy risks inherent in certain acts or practices that may have an adverse impact upon them later in life.

(iv) The extent of the data to be collected

It may well be unnecessary for data users to collect extensive or complete biometric data of an individual, so long as the data collected are sufficient for their purposes. For example, in the case of fingerprint data collection, it is probably unnecessary to involve more than two fingers for an individual.

Even if only a subset of the overall biometric characteristics is used to generate a template, the number of reference points should be kept to a minimum depending on circumstances. For example, the number of reference points a data user needs from a fingerprint to differentiate an individual from a population of 30 should be less than those needed for a population of 1,000 individuals.

(4) Transparency, explainability and informed choice

Data subjects should be provided with free and informed choice upon collection of their biometric data, together with a full explanation of the personal data privacy impact of the collection of such data. Transparency and explainability are important in this regard.

(i) Transparency

Data users should inform each data subject, on or before his biometric data is to be collected:

- whether provision of the biometric data is voluntary or obligatory[15];
- where provision of the biometric data is obligatory, what the consequences would be for failing to provide the data[16];
- the purpose(s) for which the biometric data is to be collected and used[17];
- who may access the biometric data, and under what circumstances may access be given;
- if the biometric data may be transferred to other persons, the classes of persons to whom the data may be transferred[18];
- whether the biometric data may be relied upon to take adverse actions against the individual; and
- the right to request access to or correction of the biometric data, and how the request should be made (name, post and contact particulars of the person who is authorised to handle the requests)[19].

(ii) Explainability

To enable data subjects to make informed choices and to build trust with them, data users should provide clear explanation on the use of biometric data. For example, clear explanation should be provided on:

- why it is necessary to use the biometric system for achieving the stated purpose;
- what impact there is on the rights and liberties of individuals; and
- what mitigating measures are in place to minimise any adverse impact.

(iii) Free choice and no undue pressure

Data users should exercise extra care in the collection of biometric data if there is disparity in the negotiation powers between the data users and the data subjects. For example, data users who wish to collect employees' biometric data should ensure that their employees are given a free and informed choice on the supply of the data. Assuming the collection of employees' biometric data is "adequate but not excessive," such collection must be by means that is fair in the circumstances[20], and collection of biometric data from employees who fear to be penalised if they are unwilling or unable to do so may not be fair collection.

Data users should make every effort to dispel any reasonable suspicion of undue pressure imposed on the data subjects. If the data subjects are given a choice to choose and do choose to allow their biometric data to be collected or processed, such choice will be respected. As such, the Privacy Commissioner will not interfere unless the choice is not voluntary or is made under undue pressure. An individual's consent, if any, should be recorded in writing to avoid future dispute.

To dispel any reasonable suspicion of undue pressure, a data user should, as far as practicable, provide each individual with the free choice of a less privacy-intrusive alternative to the collection of his biometric data, e.g., the option of using a smartcard with CCTV monitoring as an alternative to a fingerprint-based attendance system. The data user should adopt all practicable measures to protect the privacy of individuals' personal data and minimise any adverse privacy impact on the individuals. Evidence of such measures having been taken will be viewed favourably by the Privacy Commissioner, should a complaint against the data user be brought before him. Inconvenience to the data user is generally not an acceptable reason for denying such an alternative to the individuals.

15 DPP 1(3)(a)(i)
16 DPP 1(3)(a)(ii)
17 DPP 1(3)(b)(i)(A)
18 DPP 1(3)(b)(i)(B)
19 DPP 1(3)(b)(ii)
20 DPP 1(2)(b)

(5) Avoidance of covert data collection

The manifestation of the obligations of transparency and fair collection is that biometric data should not be collected covertly (unless there is a lawful basis that authorises covert data collection in specified circumstances).

While some biometric tools usually require the participation of the individuals in providing their data (such as the provision of their fingerprints or DNA samples), others may have the ability to collect data in a clandestine manner (such as facial recognition enabled cameras). Covert collection of biometric data is highly intrusive, and may have negative impact on individuals' dignity, privacy and other rights. Hence, the collection of facial biometric data by hidden cameras should not be conducted, unless there are strong justifications.

(6) Notice about automated decision-making and human intervention

The precision and accuracy in identifying a person vary amongst different biometric technologies. Some of them, such as facial recognition technology, are considered probabilistic – it seeks to provide only an alert that the target person is "likely" to match one of the individuals in the database. Some other established tools, such as fingerprints and DNA analyses, are more well-developed and considered more reliable. Furthermore, the accuracy of a biometric tool also depends on its own settings based on the required purpose (e.g. whether it is necessary to identify a large number of targets at high speed or to verify one person at a time). These settings create a tension between the likelihood of false positives and false negatives[21].

Since some biometric systems may not produce fully reliable identification of a person, it would not be advisable to adopt automated decision-making with the aid of such biometric systems without prior privacy impact assessment.

As a matter of good practice, if automated decision-making tools are indeed to be used in conjunction with biometric systems, then clear prior notice should be given to the affected individuals as to the existence and likely impact of such tools. Furthermore, individuals should be provided with an option to seek human intervention, where the automated decision-making is likely to produce significant or legal effects concerning them.

(7) Retention of biometric data

Data users should regularly and frequently purge biometric data which is no longer required for the purpose for which it is collected[22]. It therefore follows that if an employee's biometric data has been collected to control access to the employer's premises or computer systems, such data should be deleted as soon as the employment is terminated.

Retaining personal data for a period beyond what is necessary would not only contravene the requirements of the Ordinance, it also creates burden on the data user in safeguarding data security and assuming unnecessary risk of a data breach.

For the purpose of research or statistics, data users who want to keep personal data collected for longer than is necessary may apply anonymisation to the personal data collected so that it can no longer be used to identify individuals, and therefore is not regulated under the Ordinance. Data users should, however, consider seriously the implication of possible privacy impact of anonymised biometric data and whether it is genuinely possible to anonymise biometric data. For example, DNA samples or sequences, even when they are not associated with any names, may still reveal such information like race, physical or mental disability, family relationship with one another, etc, that may allow individuals to be re-identified under certain circumstances.

(8) Data accuracy

Data users are required to take all reasonably practicable steps to ensure that personal data held is accurate[23].

As biometric data collected can be used to take adverse action against an individual, accuracy of the data is of particular importance to the individual. Where an employee supplies biometric data on each working day to prove work attendance, any inaccuracy of the data collected may result in salary deduction or even termination of employment.

---

[21] False positive occurs when a system incorrectly reports a match. It is more likely to occur when the system is set to low precision (e.g. such monitoring crowds in public at high speed). False negative occurs when a system fails to report a match. It is more likely to occur when the system is set to require a high precision of match before it would spark an alert.

[22] DDP 2(2) and section 26 of the Ordinance

[23] DDP 2(1)

To ensure the accuracy of biometric recognition systems, data users must ascertain and be satisfied that the false positive rate and false negative rate of the biometric recognition systems are within reasonable limits, having regard to the size of the population monitored by the systems. Data users should also give the affected individuals a reasonable opportunity to explain the irregularity before deciding whether to take any adverse action against the individuals.

This consideration is closely related to the recommendation on giving notice for the use of automated decision-making.

(9) Use limitation and avoidance of function creep

Data users are not allowed to use personal data collected for a new purpose without the express and voluntary consent of the data subjects[24], unless any exemption under Part 8 of the Ordinance is applicable.

Some biometric data, such as DNA and retina images, may contain rich information about an individual in terms of physical health or even mental conditions. Data users collecting such data for one purpose must ensure that it is not used for another unrelated purpose without obtaining express and voluntary consent from the data subjects. For example, DNA samples collected or DNA tests carried out originally for an annual body check-up as part of the medical benefits offered by an employer should not be used by the employer to determine the long-term employability of the employees without the employees' consent. Doing so would also undermine trust.

(10) Data security

All reasonably practicable steps shall be taken to ensure that personal data held by a data user is protected against unauthorised or accidental access, processing, erasure, loss or use having particular regard to the kind of data and the harm that could result if any of those things should occur[25]. Given the sensitivity of biometric data, it is important that data users guard against any risk of compromising and thieving of the biometric data and that effective security measures are implemented as are reasonably practicable in the particular circumstances. Examples of worthy security measures are as follows:

- the information and communications systems which are used to store and process biometric data should be carefully and regularly evaluated to ensure that sufficiently effective security and privacy-protective measures are in place;
- biometric data should be encrypted at rest (in storage) and in transit; and
- data access should be restricted to authorised persons on a need-to-know basis and is protected by strong passwords (e.g. combination of letters, numbers and/or symbols) while all such accesses are recorded/logged.

(11) Written policy

Data users should devise privacy policies and procedures setting out clearly the rules and practices that are to be followed in collecting, holding, processing and using biometric data, and make them known to all parties concerned, such as employees, contractors and/or customers. Data users should draw the specific attention of the individuals affected to such policies and procedures, and make them publicly available for review[26].

(12) Staff training

Regular privacy compliance assessments and reviews should be conducted by data users to ensure that the acts done and practices engaged are in compliance with the Ordinance. Proper training, guidance and supervision have to be given to the staff responsible for the collection and management of biometric data. Employees who fail to properly carry out their duties in the handling of biometric data may be subject to appropriate disciplinary action.

In particular, data users should be mindful that some biometric technologies are still in their development stage. Staff training should raise awareness that such systems are prone to inaccuracy and mis-identification. System operators may be directed to use them with caution and only as an aide.

[24] DPP 3(1)
[25] DPP 4(1)
[26] DPP 5

(13) Use of contractors (data processors)

If contractors are engaged in the handling of personal data, data users must adopt contractual or other means to prevent personal data transferred to the contractors from being kept longer than necessary and from unauthorised or accidental access, processing, erasure, loss or use[27].

Data users should also note that they may be held liable for any personal data leakage or misuse resulting from a security failure on the part of the contractors[28].

It is, therefore, in the best interest of data users who engage contractors to observe recommendations given in the *Outsourcing the Processing of Personal Data to Data Processors Information Leaflet*[29] published by the Privacy Commissioner, in addition to any other relevant and applicable considerations in relation to data security.

(14) Audit and review

It is a good practice to conduct periodic, independent audits and evaluation of biometric systems, to assess whether they should be modified, improved or terminated, either because the use of the systems is ineffective in achieving their intended purposes, or because the initial purposes have since diminished in significance. The paramount consideration of "necessity and proportionality" should be revisited upon such audits.

---

27 DPP 2(3) and DPP 4(2)
28 Section 65(2) of the Ordinance
29 Available at www.pcpd.org.hk/english/resources_centre/publications/information_leaflet/files/dataprocessors_e.pdf

PCPD.org.hk
香港個人資料私隱專員公署
Office of the Privacy Commissioner for Personal Data, Hong Kong

PCPD website

Download this publication

Enquiry Hotline : **(852) 2827 2827**
Fax : **(852) 2877 7026**
Address : **Room 1303, 13/F, Sunlight Tower, 248 Queen's Road East, Wanchai, Hong Kong**
Email : **communications@pcpd.org.hk**

First published in July 2015
August 2020 (First Revision)