An Overview of Privacy-Enhancing Technologies in Biometric Recognition

PIETRO MELZI, Universidad Autonoma de Madrid, Madrid, Spain
CHRISTIAN RATHGEB, Hochschule Darmstadt, Darmstadt, Germany
RUBEN TOLOSANA, Biometrics and Data Pattern Analytics (BiDA) Lab, Universidad Autonoma de Madrid, Escuela Politecnica Superior, Madrid, Spain
RUBEN VERA-RODRIGUEZ, Universidad Autonoma de Madrid, Escuela Politecnica Superior, Madrid, Spain
CHRISTOPH BUSCH, Hochschule Darmstadt, Darmstadt, Germany

Privacy-enhancing technologies are technologies that implement fundamental data protection principles. With respect to biometric recognition, different types of privacy-enhancing technologies have been introduced for protecting stored biometric data, which are generally classified as sensitive. In this regard, various taxonomies and conceptual categorizations have been proposed and standardisation activities have been carried out. However, these efforts have mainly been devoted to certain sub-categories of privacy-enhancing technologies and therefore lack generalization. This work provides an overview of concepts of privacy-enhancing technologies for biometric recognition in a unified framework. Key properties and differences between existing concepts are highlighted in detail at each processing step. Fundamental characteristics and limitations of existing technologies are discussed and related to data protection techniques and principles. Moreover, scenarios and methods for the assessment of privacy-enhancing technologies for biometric recognition are presented. This article is meant as a point of entry to the field of data protection for biometric recognition applications and is directed toward experienced researchers as well as non-experts.

CCS Concepts: • **Security and privacy** → **Privacy protections**; *Database and storage security*; *Security requirements*; • **General and reference** → **Surveys and overviews**; *Evaluation;*

Additional Key Words and Phrases: Privacy-enhancing technologies, biometric recognition, data protection, generic framework

ACM Reference Format:
Pietro Melzi, Christian Rathgeb, Ruben Tolosana, Ruben Vera-Rodriguez, and Christoph Busch. 2024. An Overview of Privacy-Enhancing Technologies in Biometric Recognition. *ACM Comput. Surv.* 56, 12, Article 310 (October 2024), 28 pages. https://doi.org/10.1145/3664596

This work has in part received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 860813 - TReSPAsS-ETN and the German Federal Ministry of Education and Research and the Hessen State Ministry for Higher Education, Research and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE. R. Tolosana and R. Vera-Rodriguez are also supported by INTER-ACTION (PID2021-126521OB-I00 MICINN/FEDER) and Cátedra ENIA UAM-VERIDAS en IA Responsable (NextGenerationEU PRTR TSI-100927-2023-2).
Authors' Contact Information: Pietro Melzi, Universidad Autonoma de Madrid, Madrid, Spain; e-mail: pietro.melzi@uam.es; Christian Rathgeb, Hochschule Darmstadt, Darmstadt, Germany; e-mail: christian.rathgeb@h-da.de; Ruben Tolosana, Biometrics and Data Pattern Analytics (BiDA) Lab, Universidad Autonoma de Madrid, Escuela Politecnica Superior, Madrid, Spain; e-mail: ruben.tolosana@uam.es; Ruben Vera-Rodriguez, Universidad Autonoma de Madrid, Escuela Politecnica Superior, Madrid, Spain; e-mail: ruben.vera@uam.es; Christoph Busch, Hochschule Darmstadt, Darmstadt, Hessen, Germany; e-mail: christoph.busch@h-da.de.

ACM 0360-0300/2024/10-ART310
https://doi.org/10.1145/3664596

1 INTRODUCTION

Privacy is a broad concept that specifies the right of individuals to protect their freedom and private life from interference or intrusion. The scope of privacy encompasses several areas; for instance, it assumes sociological, economical, and political perspectives [19, 114] and it has been included in numerous documents that define human rights [20, 21, 71, 72]. With the rapid development of technologies related to big data, the internet of things, artificial intelligence, and cloud computing, among others, the trend has been to collect more and more personal data throughout the years and applications [57, 88]. As a consequence, the meaning and scope of privacy have evolved, with a progressively increasing focus on the fundamental right to exercise control over the collection and use of personal data [11]. Notably, this right was first articulated by Alan Westin in 1968, who defined privacy as "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" [113]. In 2016, the ***General Data Protection Regulation*** **(GDPR)** was introduced by the European Union to protect individuals with regard to the processing of their personal data [77]. The GDPR establishes the data protection principles to fulfil that serve to regulate the use of personal data, for instance, binding to the original purpose of processing the data and protecting data against unauthorised accesses. A successful implementation of data protection allows to enhance the privacy of individuals, safeguarding their rights and freedom with regard to the processing of personal data.

To comply with data protection principles, the GDPR demands that appropriate technical and organisational measures be implemented; for instance, pseudonymisation is one of the important concepts that is explicitly mentioned. Pseudonymisation means the processing of personal data in a way that they can no longer be attributed to specific individuals without the use of additional information, provided that such information is kept separately and is subject to technical and organisational measures to ensure that personal data are not attributed to an individual [77]. Another technical measure indicated by the GDPR is encryption, which may be implemented to ensure a level of security appropriate to the risk. In addition, the GDPR incorporates the concept of *data protection by design*, which consists in the implementation of technical and organisational measures to satisfy data protection principles both at the time of the determination of the means for processing and at the time of the processing itself [77]. A similar concept called ***privacy by design*** **(PbD)** was introduced by Ann Cavoukian, according to which the protection of data must be integrated in a system from its creation to achieve strong privacy protection without diminishing its functionality [13, 14]. Furthermore, with the notion of PbD, Cavoukian requires for the first time that privacy settings in an IT system are activated as default, similarly to the following definition of *data protection by default* outlined in Article 25 of the GDPR [77]. Cavoukian's definition of PbD can be regarded as a formidable contribution to the literature, bridging different disciplines. Notably, the same author proposed applying the PbD concept to systems that handle biometric data, citing concerns about potential data misuse and security vulnerabilities in such contexts [15, 16].

Biometric data are measurements of human characteristics with the purpose of recognising and describing individuals. They can be divided into two categories: (1) biological and (2) behavioural [44]. Biological data are related to individuals' bodies, such as fingerprints, faces, and irises, while behavioural data involve the individuals' actions (i.e., functions of the body), such as keystroke, gait, signature, and speech. Biometric data are widely used in recognition systems as they are unique to each individual and cannot be forgotten, lost, and transferred to other individuals, such that this authentication factor offers a clear advantage over traditional knowledge- and possession-based authentication systems [4]. Biometric recognition systems have been implemented in various application scenarios and with multiple purposes, such as automated border control, access control in high-security facilities, e-banking, healthcare, forensics for law enforcement, and

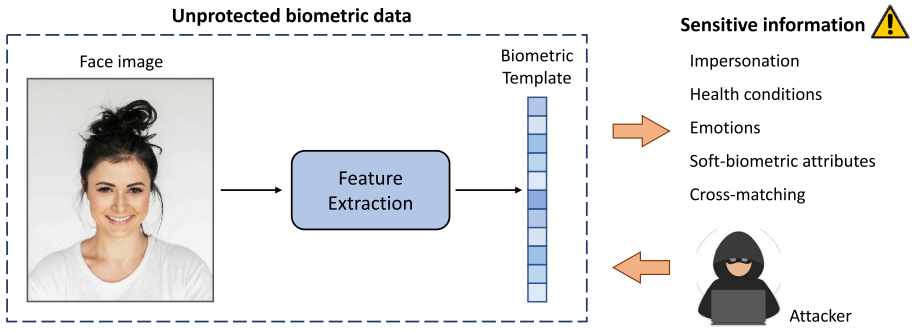

Fig. 1. Privacy concerns derived from the storage of unprotected representations of biometric data and unprotected biometric templates. In the figure we consider a face image. Similar concerns apply to other biometric traits.

smartphone unlocking [64, 81, 92, 110, 111]. At the time of enrolment, i.e., registration, biometric data are usually stored as biometric reference. Biometric data are commonly represented as so-called *templates*, i.e., sets of biometric features related to an individual and comparable directly to probe biometric features extracted during authentication, e.g., minutiae extracted from fingerprints [61], or binary iris-codes extracted from irises [7]. According to [41], with the term *biometric data* we refer to both the original representations of human characteristics and the features extracted from them. In fact, the concepts described in the article remain applicable to any type of biometric data.

The unprotected storage of biometric data raises privacy concerns about the ultimate use of them. If the original representation (i.e., captured samples like face images or fingerprint images) are stored, they could potentially leak out from the server.[1] An attacker who compromises a biometric database can obtain the biometric data of the enrolled individuals and eventually impersonate them to gain access to the corresponding authentication system. In addition, further information can be derived from the biometric data, including health conditions, emotions, soft-biometric attributes, and other personal aspects [24, 64]. Also, we note that storing processed biometric data (e.g., feature vectors) is not a protection level, as image representations can be easily reconstructed from biometric features in many cases [12, 32, 60].

Soft-biometric attributes consist in information contained in biometric data that increases the chances to recognise individuals [22]. Soft-biometric attributes, such as age, gender, ethnicity, and many more, can be automatically extracted from biometric data without the user's agreement and used for purposes that were not originally intended. Moreover, if multiple systems in which individuals have registered their biometric data are compromised, an attacker can cross-match biometric data across such systems to gain further profiling information about individuals [84]. In Figure 1 we summarise the privacy concerns in the case of face images. Similar concerns arise from other types of biometric characteristics, e.g., iris, fingerprint, gait, voice, and so forth. Following the definition of biometric data in [41], which reads biometric samples or aggregation of biometric samples at any stage of processing, e.g., biometric reference, probe, biometric feature, or biometric property, we consider the entire variety as sensitive data and formulate and review the application of **privacy-enhancing technologies (PETs)** to them to provide a solution for the privacy concerns discussed above.

We observe that some of the approaches to privacy enhancement investigated in the literature cannot be successfully implemented in the scenario of biometric recognition systems. For

[1] https://www.opm.gov/news/releases/2015/09/cyber-statement-923/

instance, personal data can be made anonymous in such a manner that the data subject is no longer identifiable [77]; i.e., anonymous information indicates information that does not relate to an identifiable person. Hence, anonymisation techniques provide strong privacy assurances but prevent the recognition of individuals, eliminating the utility of biometric data. Moreover, traditional encryption algorithms cannot be applied to biometric data because small changes in the original data, such as the unavoidable variances between multiple biometric data measurements from the same characteristic of the same individual, cause a drastic change in the encrypted data. Hence, encrypted biometric data need to be decrypted before they can be compared, introducing a risk of exploitation by potential adversaries [85]. Finally, it is important to remark that biometric data require a special effort compared to other data to protect the additional information, for instance, related to the health status of the captured subject, that can be easily derived from the data.

In this article, we refer to PETs suitable for biometric recognition systems with the term ***biometric recognition privacy-enhancing technologies*** **(BR-PETs)**, and to the data generated by BR-PETs with the term *protected biometric data*. We provide a summary of concepts with practical relevance, so that practitioners can attain a comprehensive overview of BR-PET concepts through this work and identify the best approach to achieve privacy enhancement in their application, and according to their needs. Existing surveys in this field do not fulfil this function of guideline and consider other categories of PETs applied to biometrics or limit their scope to specific (sub-)categories of BR-PETs. For instance, **cancelable biometrics (CBs)** and **biometric cryptosystems (BCs)** are investigated in [69, 78, 86], while the removal or concealing of specific information from biometric data is investigated in [62], together with PETs not suitable for biometric recognition systems. The present article aims to gather under a general framework different BR-PETs that are usually investigated in separate works in the literature.

In particular, the wide range of PETs applied to biometrics described in [62] focuses only on faces and aims to prevent the recognition of individuals through the application of deidentification, anonymisation, redaction, obfuscation, obscuration, soft-biometric privacy enhancement, and controllable privacy. The definition of biometric PETs provided by [62] refers to technologies that maintain the utility of biometric data for applications involving soft-biometric attributes, in contrast to our concept of BR-PETs. Moreover, their definition does not include traditional **biometric template protection (BTP)** schemes, i.e., CBs and BCs, as they do not allow to specify which sensitive information to remove and preserve. We acknowledge that BTP schemes protect the whole biometric data without paying attention to the specific information and attributes contained therein. However, their definition of biometric PETs is rather contradictory, as authors include **homomorphic encryption (HE)** in the survey, which does not reduce utility for recognition or allow for controllable privacy protection. We find it beneficial to define the term *BR-PETs* and clearly delineate the scope of our study to provide an overview of the PETs that must be considered to safely store biometric data in biometric recognition systems.

A general taxonomy for PETs has been proposed in [38]. Compared to this work, we exclusively consider PETs applied to biometric recognition systems (BR-PETs) and focus on the protection of the stored data over transmitted data, as the latter involves network protocols that are not exclusive to biometric data. It is important to emphasize that this study focuses on PETs within the context of traditional biometric recognition systems, which represent the most prevalent use case for biometric data [44]. Our scope does not extend to other scenarios, including unwanted or unlawful biometric recognition, or broader aspects like biometric fairness.

To sum up, these are the main contributions of our work:

— We consider PETs in the widespread application of biometric recognition, enabling a thorough analysis of the different requirements considered for privacy enhancement.

— We focus on a set of well-established categories of BR-PETs and investigate them through the definition of a general framework that highlights properties, purposes, and singularities of each category. In this way, the reader can identify the most suitable BR-PET for their application, according to the advantages and disadvantages depicted in this overview.
— We qualitatively evaluate existing categories of BR-PETs according to their ability to satisfy specific privacy requirements and prevent the extraction of sensitive information from the protected biometric data that they generate. During the evaluation, we consider attackers with different capabilities and knowledge of BR-PETs.

The remainder of the article is organised as follows. In Section 2, we introduce the fundamentals of biometric recognition systems and BR-PETs. Section 3 describes the proposed general framework for BR-PETs, providing details of the different categories of BR-PETs. Section 4 presents some metrics suitable for the evaluation of BR-PETs and performs an assessment of the considered categories of BR-PETs under different attack scenarios. Finally, Section 5 points out potential lines for future research in this field, and Section 6 draws the conclusions of this study.

2 FUNDAMENTALS

In this section, fundamental concepts of biometric recognition systems and BR-PETs are introduced to facilitate a full understanding of this work.

2.1 Nomenclature

To ensure a clear understanding of the concepts presented in this study and to avoid any confusion arising from the incorrect use of certain terms, we define below the meaning of some fundamentals terms utilised in the document:

— *Privacy-Enhancing Technologies (PETs):* The literature offers various definitions of PETs, but we adopt two complementary definitions. PETs have been defined as a coherent system of information and communications technology measures that protects privacy by eliminating or reducing personal data or by preventing unnecessary and/or undesired processing of personal data, all without losing the functionality of the data system [5]. Additionally, Wikipedia characterises PETs as technologies that embody fundamental data protection principles by minimising personal data use, maximizing data security, and empowering individuals [73, 115].
— *Data protection:* The implementation of appropriate techniques and organisational measures to ensure compliance with so-called "data protection principles." When effectively implemented, this enhances the privacy of individuals, safeguarding their rights and freedom with regard to the processing of personal data.
— *Data protection principles:* Seven principles relating to the processing of personal data, established in Art. 5 of the GDPR [77]. They include (1) lawfulness, fairness, and transparency; (2) purpose limitation; (3) data minimisation; (4) accuracy; (5) storage limitation; (6) integrity and confidentiality; and (7) accountability.
— *Data protection techniques:* Appropriate technical and organisational measures taken to provide data protection. Such measures could consist, among other things, of minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, monitoring of the data processing, and creation and improvement of security features [77].
— *Biometric data:* They are biological and behavioural measurements of human characteristics with the purpose of recognising and describing individuals [44]. Biometric data is defined

as a biometric sample or aggregation of biometric samples at any stage of processing (e.g., biometric reference, biometric probe, biometric feature, or biometric property) [41].
— *Biometric Recognition Privacy-Enhancing Technologies (BR-PETs):* PETs that implement suitable technical approaches (Section 3.2) to satisfy the privacy requirements (Section 2.4) associated with biometric data within the context of biometric recognition systems.

2.2 Biometric Recognition Systems

Biometric recognition systems perform an automated recognition of individuals based on their behavioural and biological characteristics. They are generally composed of four subsystems that allow to capture biometric samples of individuals and process and compare them to determine whether individuals are recognised or not [44]. We describe the subsystems in the following, keeping in mind that, as per the definition in [41], biometric data is defined as a biometric sample or aggregation of biometric samples at any stage of processing (e.g., biometric reference, biometric probe, biometric feature, or biometric property):

— *Data capture:* It captures biometric samples of individuals through capture devices.
— *Signal processing* and *feature extraction:* They process the captured biometric samples to extract a set of salient or discriminatory features (i.e., a feature vector) from them.
— *Comparison:* It performs comparisons between acquired biometric data x (i.e., biometric probe) and stored biometric data y (i.e., biometric reference) and generates similarity scores $s = S(x, y)$ according to some similarity functions S. Based on similarity scores, the system decides if two biometric feature vectors are from the same subject (match) or from different subjects (non-match).
— *Data storage:* It stores the biometric data (i.e., biometric reference) of enrolled individuals and provides them when necessary to perform comparisons.

In this work we focus on privacy issues related to the storage of biometric data; other threats such as attacks to capture devices and communication channels must be addressed in appropriate ways. Subsequent presentations leading to multiple representations (i.e., biometric data that are either samples or feature vectors) provided by the same individual usually exhibit some variance. Hence, it is essential to choose the proper threshold t to determine if two biometric data x and y belong to the same individual, i.e., if $S(x, y) \geq t$. High thresholds may result in false non-matches of mated comparison trials (i.e., genuine attempt), while low thresholds may result in false matches of non-mated comparison trials (i.e., impostor attempt). According to these error types, common metrics to measure the performance of biometric recognition systems are (1) **false match rate (FMR)**, i.e., the probability that an impostor is incorrectly accepted as genuine, and (2) **false non-match rate (FNMR)**, i.e., the probability that a genuine individual is incorrectly rejected as impostor.

2.3 Biometric Privacy-enhancing Technologies

We have already discussed the privacy concerns related to the collection, processing, and storage of biometric data and the consequent application of BR-PETs to address them. We observe that traditional BTP schemes are necessary but not sufficient to safely store biometric data. Hence, we recommend the use of the umbrella term *BR-PETs* to include and analyse under the same framework all the technologies that generate protected biometric data, satisfying the privacy requirements for biometric data specified in Section 2.4. In this way, both traditional privacy requirements, i.e., irreversibility and unlinkability, and the privacy of soft-biometric attributes can be considered at the same level. We denote the application of BR-PETs as a function f applied to biometric data x: $\tilde{x} = f(x, k)$, where $\tilde{x}$ are the resulting protected biometric data, and k represents optional parameters of f. We organise the different BR-PETs proposed in the literature in the following categories:

(1) *Cancelable Biometrics (CBs):* Consist of intentional, repeatable distortions of the original biometric data based on transformations that enable a comparison of biometric data in the transformed domain [84].
(2) *Biometric Cryptosystems (BCs):* Are designed to securely bind a digital key to biometric data or generate a digital key from biometric data [16].
(3) *Homomorphic Encryption (HE):* Allows to generate the encrypted result of operations performed on plaintexts directly computing operations on ciphertexts, i.e., without any intermediate decryption [30, 34].
(4) *Soft-Biometric Minimisation (SBM):* Identify soft-biometric attributes in the representations of biometric data, discard them, and generate new representations of biometric data excluding such soft-biometric attributes [6].
(5) *Soft-Biometric Protection (SBP):* Modify the representation of biometric data to prevent the extraction of soft-biometric attributes [25, 67, 109]. In this case, soft-biometric attributes are not discarded, but they are considered inaccessible in the new representations of biometric data.

2.4 Privacy Requirements for Biometric Data

The protected biometric data obtained from BR-PETs are required to satisfy specific privacy requirements to overcome the concerns related to the use of biometric data in recognition systems. Some requirements are long established [69], although not always properly evaluated. More recently, novel threats related to the possible extraction of soft-biometric attributes as well as other information from biometric data have emerged [107]. As a consequence, privacy requirements must continually evolve to ensure their ongoing comprehensiveness. In the following, we describe the main privacy requirements for biometric data that BR-PETs must ensure [40]:

— **Irreversibility:** It should be impossible or at least sufficiently difficult to reconstruct biometric samples similar to the original captured samples from the stored protected biometric data. Irreversibility can be achieved by applying irreversible transformations or transformations that make use of secret parameters to biometric data. We highlight the importance of BR-PETs, also considering the possibility of partial irreversibility of protected biometric data. In fact, partial reconstructions of the original data may reveal soft-biometric information of individuals and may allow attackers to access the system.

  Given a function g that attempts to reconstruct the original biometric data x (i.e., a sample) from the protected biometric data $\tilde{x}$, such that $x' = g(\tilde{x})$ is the reconstructed biometric data, irreversibility is achieved if $S(x, x') < t$, for any biometric similarity function S and any given threshold t.
— **Unlinkability:** It should be impossible or at least sufficiently difficult to determine if different representations of protected biometric data belong to the same individual or not. Unlinkability can be obtained by introducing some randomness with keys or random parameters in transformations that are protecting biometric data. Unlinkability shall prevent cross-matching attacks across multiple systems. When unlinkability is satisfied, compromised biometric data can be revoked and substituted with new protected representations.

  Given the protected biometric data $\tilde{x}_1 = f(x, k_1)$, $\tilde{x}_2 = f(x, k_2)$, $\tilde{y}_1 = f(y, k_1)$, and $\tilde{y}_2 = f(y, k_2)$ obtained from BR-PET f with different biometric data x and y and different parameters k_1 and k_2, unlinkability is achieved if $P(S(\tilde{x}_1, \tilde{x}_2) \geq t) = P(S(\tilde{x}_1, \tilde{y}_1) \geq t) = P(S(\tilde{x}_1, \tilde{y}_2) \geq t)$, for any given threshold t.
— **Privacy of soft-biometrics:** The extraction of soft-biometric attributes from biometric data for purposes different than the originally intended ones must be prevented. We observe that false-positive comparisons carried out in biometric recognition systems should not disclose

any information about the soft-biometric attributes of protected biometric data, which is usually not the case [75]. For instance, several large empirical studies have shown that face recognition false matches generally occur with higher probability within the same demographic groups [28, 35, 39].

Given a soft-biometric attribute A, the set of its possible values $\{a_1, a_2, \ldots, a_n\}$, and a soft-biometric classifier h trained to determine the soft-biometric attribute A from biometric data x, the following condition must be valid to ensure the privacy of soft-biometric attribute A: $P(h(\tilde{x}) = a_i) = P(h(\tilde{x}) = a_j), \forall a_i, a_j \in A$.

The compliance with these privacy requirements results in a sound protection of biometric data and other information that can be obtained from them. The degree to which privacy requirements are met depends on a variety of factors, including the type of biometric characteristic, the feature extraction algorithm, the BR-PET utilised, and the use of secret keys. In case of unlinkability, the relationship with the required technical approaches is straightforward: they must incorporate some sort of randomness to provide unlinkability. Differently, more technical approaches can provide irreversibility with encryption that relies on the strength of the key, and data disentanglement that ensures unrecoverable elimination of specific information.

While CBs and BCs are usually claimed to provide irreversibility and unlinkability, the fulfilment of these requirements may be overestimated during evaluation. For instance, the irreversibility of the transformation presented in [117] is overestimated, as the original paper's estimated time complexity for an attack can be significantly reduced by employing two alternative attack methods as outlined in [100]. Furthermore, the possibility of extracting soft-biometric attributes from biometric data is usually investigated by training classifiers with unprotected biometric data and evaluating them with protected biometric data. This is a common scenario for attackers attempting to infer soft-biometric attributes from protected data when they do not have access to the labels necessary for training classifiers with the protected data. Notably, an extensive set of soft-biometric classifiers has been trained with both protected and unprotected biometric data in [108]. Regardless of whether a classifier is trained on protected data, unprotected data, or a combination of both, its inability to learn the pattern of an attribute does not necessarily imply that the pattern does not exist [107]. Also, some patterns may differ between the unprotected and protected representations of biometric data. In Section 4 we discuss accurate ways to carry out the evaluation of privacy requirements for the different categories of BR-PETs.

We observe that BR-PETs, which are PETs designed specifically for biometric recognition systems, require privacy requirements that are tailored to the specific context of biometric data. These requirements adapt certain data protection principles to enhance the privacy of biometric information, as illustrated in Figure 2. In particular, the principle of *integrity and confidentiality* requires that data processing guarantees the adequate security of personal data [77]. In the case of biometric data, the concepts of *irreversibility* and *unlinkability* introduce additional, context-specific requirements, building upon the broader *integrity and confidentiality* principle. In ISO/IEC 24745, *confidentiality* is a biometric information privacy requirement on the same level of *irreversibility* and *unlinkability* [40]. However, its scope is limited to maintaining data confidentiality to protect against unauthorized access. This is a rather general requirement that applies to any kind of personal data. Furthermore, the techniques indicated to provide *confidentiality* involve specifications for physical data storage and encryption, which are common to *irreversibility* and *unlinkability*. In addition, the latter two require additional technical measures beyond mere encryption.

The principles of *purpose limitation* and *data minimisation* require that data must be collected for specified, explicit, and legitimate purposes in a format that is adequate, relevant, and limited to what is necessary. This is precisely what the *privacy of soft-biometrics* requirement aims to

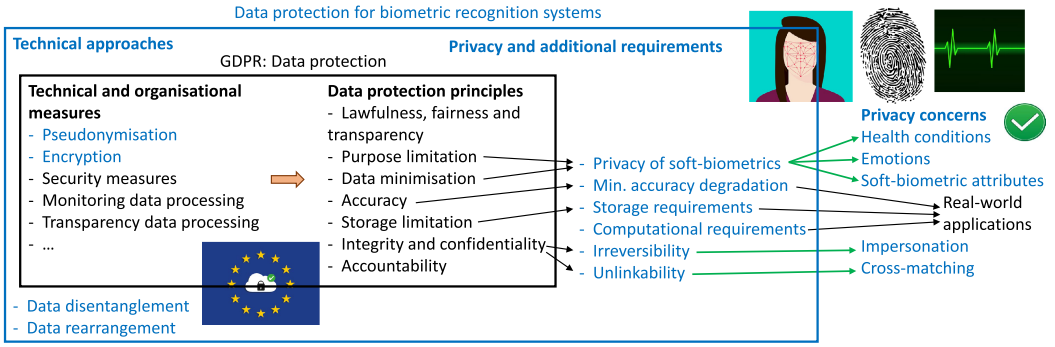

Fig. 2. Data protection for biometric recognition systems builds upon the foundational concept of data protection also outlined in the GDPR [77]. It involves the implementation of specific technical approaches and the specification of privacy and additional requirements, beyond the conventional technical and organisational measures and data protection principles. Data protection for biometric recognition systems is designed to address privacy concerns associated with the storage of biometric data and ensure practicality for real-world applications.

achieve, specifically preventing the extraction of soft-biometric attributes for purposes other than the original ones. Finally, data protection principles that do not require any adaptation to the context of biometric recognition must still be adhered to in accordance with their original intent.

3 GENERAL FRAMEWORK

The protection of biometric data is an essential aspect in biometric recognition systems, given the multiple privacy concerns that may arise from the unprotected storage and potential misuse of biometric data. Over the years, many BR-PETs have been proposed to enhance the privacy of biometric data, and they have been categorised according to certain properties. In this section, we introduce a general framework to outline the different components that constitute classes of BR-PETs, showing the common aspects and (dis)similarities between different categories of BR-PETs. Figure 3 provides an illustration of said general framework. With the introduction of this framework, we show that the categories of BR-PETs are not mutually exclusive, as it may appear from the literature, and that different BR-PETs can be combined together to improve the final protection of biometric data.

3.1 Architecture of the Framework

We describe the different components of the general framework and how they can be implemented in the different categories of BR-PETs, highlighting the distinctive traits of them. In this sense, Table 1 provides a comparison of the main characteristics related to the implementation of the considered categories of BR-PETs.

3.1.1 Input Data Format. The common procedure in biometric recognition systems consists in capturing biometric data through capture devices and subsequently extracting biometric features from the captured samples, i.e., numbers or data, into a format that can be easily processed to compare biometric probes and biometric references [41]. Feature extraction allows to reduce the size of biometric data and map them into a discriminative space, where different representations of individuals can be well separated [87]. Hence, BR-PETs can be applied to biometric data at different levels:

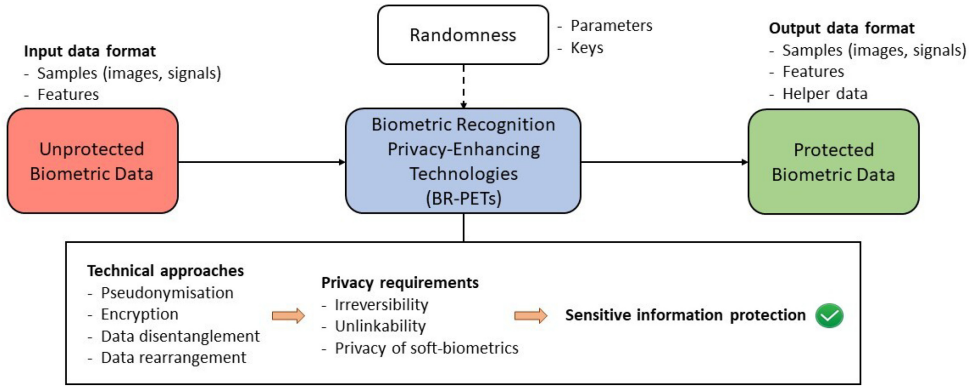

Fig. 3. Representation of the general framework and the technical approaches considered to generate protected biometric data. The dashed line indicates an optional component of the framework.

Table 1. Overview of the Different Categories of BR-PETs in Terms of Characteristics of Their Implementation

BR-PETs	Input data format	Randomness	Output data format	Comparison
CBs	Samples, features	Parameters	Samples, features	Standard
BCs	Features	Keys	Helper data	Key retrieval
HE	Features	Keys	Encrypted features	In encrypted domain
SBM	Samples, features	No	Samples, features	Standard
SBP	Samples, features	No	Samples, features	Standard

CBs = Cancelable Biometrics, BCs = Biometric Cryptosystems, HE = Homomorphic Encryption, SBM = Soft-Biometric Minimisation, SBP = Soft-Biometric Protection.

— *Sample level:* BR-PETs are applied directly to the collected images or signals, modifying the appearance of biometric data from a human point of view [56]. Biometric features can be subsequently extracted, and it is assumed that the protection introduced at the sample level is transferred to them.
— *Feature level:* BR-PETs are applied to the extracted features and relate to the form in which machines observe and process biometric data, trying to prevent the execution of automatic and unintended operations on biometric data [67].

All categories of BR-PETs provide solutions working at feature level. This is not the case of the sample level, where the bigger size of biometric data may pose a limit for some BR-PETs. Nevertheless, several CB technologies have been designed to transform data at the sample level [84, 122], and a list of BR-PETs directly applied to face images is presented in [37]. Many SBM and SBP technologies can be applied at the sample level [25, 62]. Finally, with the advent of generative models, numerous approaches have been proposed for face manipulation, allowing the modification of soft-biometric attributes such as age, gender, and ethnicity in face images [18, 98]. However, the primary goal of face manipulation methods is to generate visually realistic images. These methods have not taken into account the aspects of recognition utility and privacy protection [56].

3.1.2 Randomness. Randomness is an optional although widely employed component that can be provided as further input to BR-PETs in the categories of CBs, BCs, and HE. The incorporation

of randomness in BR-PETs allows the generation of multiple templates from the same biometric data, enhancing the desired properties of unlinkability and renewability required for protected biometric templates. In CBs, the random element consists of user-specific or application-specific parameters of the transform that must be secretly stored, as they are required during authentication and, if compromised, they may allow attackers to launch *linkage attacks*, as well as facilitating the reconstruction of the original biometric data [86]. For instance, BioHashing can rely on user-specific randomness generated from a seed stored in a USB token or smart card microprocessor [45]. In BCs, explicit random values and additional sources of randomness are combined in biometric data in multiple ways, for instance, in fuzzy embedders [9, 10]. To provide unlinkability in HE, random numbers can be combined together with biometric data, prior to encrypting them with the same public key [34]. The encryption key itself is also random.

Achieving unlinkability and renewability without relying on randomness is basically infeasible, unlike achieving the privacy of soft-biometrics. As observable in Table 1, both SBM and SBP do not require any randomness. BR-PETs utilising **deep neural networks (DNNs)** to transform biometric data in a privacy-protected version, according to some criteria, have been introduced. Typically, DNN architectures do not incorporate randomness. For example, DNNs have been successfully employed to secure soft-biometric attributes in [67]. An exception is represented by the randomised DNN proposed in [59] for face template protection, where two kinds of randomness have been considered: (1) random activation of DNN neurons and (2) random permutation and sign flip of extracted templates. That work provides an analysis showing that the proposed protected template satisfies the criteria of unlinkability.

3.1.3 Output Data Format. BR-PETs are applied to unprotected biometric data to generate protected biometric data that satisfy specific privacy properties and consequently can be stored in biometric recognition systems without raising certain privacy concerns. The format of protected biometric data varies according to the BR-PETs considered to generate them and the format of input data. Typically, protected biometric data come in the format of obscured biometric features, and they are obtained when BR-PETs are applied to unprotected biometric features [69] or directly to original biometric data [80]. However, it is also possible that the application of BR-PETs to images or signals generates protected biometric data that maintain the same format of the input, as in the case of image morphing [54].

Differently from the other categories of BR-PETs, the output of BCs is generally more complex to describe and assumes the name of *helper data*. It can be obtained according to numerous algorithms from biometric features, eventually combined with a secret key [10, 17, 49]. Helper data alone should not reveal information about the original biometric data and key. They may as well be unprotected; for instance, alignment information could be provided in plain format [103].

3.1.4 Data Comparison. Protected biometric data are employed for the recognition of individuals in biometric recognition systems. Individuals provide to the system their probe biometric data, which (in most cases) will be protected with the same BR-PET used during enrolment. Recognition can be carried out in two modes:

— *Verification:* Individuals also have to claim their identity. The recognition process consists in a single comparison between the probe biometric data and the previously enrolled biometric data that serves as a biometric reference for the claimed identity.
— *Identification:* The comparison between probe and enrolled biometric data is usually performed for each individual enrolled in the system (exhaustive search). It is important to note that many methods have been developed with the aim of reducing or optimising the

Table 2. Summary of the Technical Approaches Implemented by the Different Categories of BR-PETs

BR-PETs	PN	EN	DD	DR
Cancelable biometrics	✓	✓		✓
Biometric cryptosystems	✓	✓		✓
Homomorphic encryption	✓	✓		✓
Soft-biometric minimisation			✓	
Soft-biometric protection				✓

PN = pseudonymisation, EN = encryption, DD = data disentanglement, DR = data rearrangement.

computational workload of biometric identification systems and thereby speeding up their identification transactions [27]. Identification can be seen as a sequence of verifications.

For simplicity, in this study we consider the single comparison performed during verification.

The modality of comparison in biometric recognition systems depends on the format assumed by the protected biometric data. If BR-PETs do not modify the original format of biometric data, the comparison can be performed in the *transformed domain* with the same comparator of the original unprotected system, as in the case of most CBs [78]. Eventually, individuals have to present their secret parameters along with probe biometric data, if they are stored outside the biometric recognition system [45]. It is also possible that BR-PETs incorporate a randomisation step where the order of biometric features is altered to protect the contained information. This can necessitate additional operations to reorder the biometric features before comparison. For instance, in [109], a BR-PET designed for SBP generates templates in a mixed representation of minimal information units, where the pattern of soft-biometric attributes appears randomised. During the verification process, these units of probe templates are reordered according to a matching with the units of reference templates by solving an optimal best-matching problem to provide high recognition accuracy.

Finally, the comparison can be made in the encrypted domain when biometric data are protected with specific categories of BR-PETs that make use of secret keys. In case of BCs, at the time of authentication no secrets need to be presented, only biometric data. BCs are usually employed in verification scenarios, and the result of a comparison consists in the disclosure of the key to the individual or in a failure message. In case of HE, the comparison produces a comparison score that, once decrypted, is identical to what would be obtained if the computation was carried out in the unencrypted domain, at the cost of increased computation and communication overhead.

3.2 Technical Approaches for Privacy Enhancement

Privacy enhancement consists in the adoption of measures and precautions during the processing of personal data to increase their protection according to predefined principles without losing the system functionalities. In the case of privacy enhancement in biometric recognition, specific technical approaches are required to satisfy the privacy requirements for biometric data without affecting the performance of biometric recognition systems. This originates the well-known trade-off between privacy and utility of biometric data [69]. BR-PETs are designed to incorporate different technical approaches that provide privacy enhancement by executing suitable operations on biometric data. To be more effective, BR-PETs usually focus on specific technical approaches and can be considered more or less suitable according to the context of application. However, the distinction between existing BR-PETs is not sharp, with numerous BR-PETs that implement various technical approaches at the same time to protect biometric data, as reported in Table 2.

Technical approaches for privacy enhancement include data protection techniques defined in the GDPR for the protection of personal data, such as pseudonymisation and encryption [77]. These two techniques are adapted to the specific context of privacy enhancement for biometric data. The GDPR also outlines other technical and organisational measures, including monitoring and transparency of data processing, as well as the creation and improvement of security features. We refrain from providing the details of these measures, as they are required for the processing of personal data, regardless of whether it involves biometric data (Figure 2). Instead, we explore technical approaches tailored to the processing of biometric data, which demand more specific privacy requirements than general personal data. In the following, we describe the technical approaches implemented by BR-PETs:

— *Pseudonymisation:* Consists in the generation and use of **pseudonymous identifiers (PIs)** to identify individuals instead of their real identifiers, such as names and biometric data. A general architecture to obtain PIs from biometric data is described in ISO/IEC 24745 [40]. Differently from *anonymisation,* which is ideally irreversible and not suitable for the applications considered in this study, pseudonymisation allows to uniquely identify particular individuals while hiding their actual identity [38]. Additional information is required to attribute pseudonyms to specific individuals.
— *Encryption:* The encoding of human-readable information into a coded format that can only be interpreted by authorised parties, preventing unauthorized access to data. Encryption and decryption of data are performed with public algorithms that rely on secret keys. As we have previously noted, traditional encryption is not applicable to biometric data. However, specific encryption algorithms and suitable techniques to address the biometric variance can be successfully implemented, for instance, HE [26, 34].
— *Data disentanglement:* Directly implements the data protection principle of *data minimisation* described in the GDPR, according to which the processing of personal data should be limited to what is necessary to achieve the purposes of the system [77]. To address data minimisation, BR-PETs identify the sensitive information included in biometric data and not necessary for the system of interest, disentangle it from the biometric data, and discard it. The removal of information from biometric data can be considered as a preliminary approach to data minimisation: the reconstruction of original biometric data may be difficult in some cases, but sensitive information is still present in biometric features. Data disentanglement from biometric data is extensively discussed in the literature [6, 43, 70].
— *Data rearrangement:* This approach consists in data transformations intended to protect biometric data and prevent the extraction of information contained in biometric data. Data rearrangement entails altering the representation of biometric data to obtain features in a different representation, making certain information inaccessible. Differently from data disentanglement, where specific information is separated and discarded from biometric data, data rearrangement retains all the information within the biometric data, in a concealed form. Data rearrangement in biometrics is extensively discussed in the literature [2, 51, 74].

We note that BR-PETs may fail to provide inaccessibility to information. The limits of two technologies implementing respectively data disentanglement and data rearrangement have been shown in [75]. However, inaccessibility is not provided exclusively by these two technical approaches. Other approaches ensure inaccessibility to information while keeping certain elements secret, such as additional information in pseudonymisation or keys in encryption.

CBs and BCs generate protected biometric data that do not reveal significant information about the original data or the identity of their owner [83]. The recognition process carried out with these protected biometric data is considered fully pseudonymous, as the original biometric data are never

exposed during comparisons [86]. This is also the case of HE, with comparisons performed in the encrypted domain. Hence, we observe that there is a close relationship between pseudonymisation and encryption. When pseudonymisation is achieved through encryption, the privacy assurances completely rely on the security of the encryption process. Otherwise, in CBs, pseudonymisation can be obtained with parameterised irreversible transformations that provide different privacy assurances, but biometric data are still processed in a protected domain. Finally, BCs implement cryptographic algorithms with some error tolerance in order to generate protected biometric data.

We distinguish some categories of BR-PETs that are specifically designed to prevent the extraction of soft-biometric attributes from biometric data (i.e., SBM and SBP), differently from CBs, which protect the overall information contained in biometric data given the computational difficulty to recover the original biometric data from the transformed one [109]. These BR-PETs are not originally intended to pseudonymise or encrypt biometric data, with such characteristics that possibly result as a consequence of data transformations. On the contrary, the implementations of these BR-PETs follow the approaches of data disentanglement and data rearrangement. According to the former (i.e., data disentanglement implemented for SBM), sensitive information is identified in biometric data and disentangled from their representation, and a novel representation that does not embed such sensitive information is finally learned [6, 63, 106]. According to the latter (i.e., data rearrangement implemented for SBP), the representation of biometric data is modified so that the extraction of sensitive information is made impossible in the novel representation [109]. Several works have been proposed to conceal soft-biometric attributes through face mixing and other obfuscations [76, 119], adversarial perturbation [65, 66, 120], and similarity-sensitive noise transformations [104] while preserving the recognition utility. A novel BR-PET that could obfuscate facial attributes in the visual appearance of face images while preserving the identity discriminability has been proposed in [56].

Different categories of BR-PETs can be combined to provide hybrid BR-PETs that leverage the distinct properties provided by singular BR-PETs to provide an overall improved privacy enhancement. In the literature, hybrid BR-PETs that combine different BTP schemes have been proposed in response to the challenge of finding a single BTP method that can simultaneously provide both privacy requirements and performance [29, 50, 55, 116]. Novel instances of hybrid BR-PETs combine HE with CBs [3, 96]. HE ensures the preservation of recognition performance. However, its security completely relies on the secrecy of the decryption key, and the computations performed in the encrypted domain tend to be computationally intensive. To solve these challenges, CBs can be applied at first to biometric templates. They provide irreversibility even in the event of homomorphic key exposure and facilitate dimensionality reduction in biometric templates, enhancing computational efficiency within the encrypted domain [96]. While offering notable advantages, these hybrid BR-PETs do not incorporate any form of soft-biometric privacy enhancement when secret keys are compromised.

It is worth noting that legislation leaves completely open which exact protective measures are to be taken. The GDPR deliberately does not define which specific technical and organisational measures are considered suitable in each case in order to accommodate individual factors. Only a non-exhaustive list of measures that can be implemented is provided in Art. 32 of the GDPR [77]. One can use other standards, such as ISO standards. The text of law leads one to conclude that often several protective measures must be used with one another to satisfy statutory requirements.

3.3 Additional Requirements

In addition to privacy requirements, BR-PETs must fulfil additional requirements that are not directly related to enhancing biometric data privacy. These additional requirements are essential

to guarantee that privacy enhancement schemes do not compromise the system's functionality, aligning with the principles of PbD. Therefore, while these additional requirements may not directly address privacy concerns, they are equally vital in achieving practical privacy enhancement in biometric recognition systems. We describe such requirements in the following:

— *Minimisation of accuracy degradation:* Degradation in accuracy is defined as the difference between a metric used to measure the recognition accuracy (such as an FNMR at a fixed FMR) when template protection is not applied and the same metric when template protection is applied [102]. The performance of biometric recognition systems provided with unprotected biometric data should be maintained when processing protected biometric data. The degradation in accuracy may be expressed as a percentage. This degradation depends not only on the BR-PET applied but also on the way in which the biometric feature extraction algorithm is implemented [42].

Given the protected biometric data $\tilde{x}_1 = f(x_1, k), \tilde{x}_2 = f(x_2, k), \tilde{y}_1 = f(y_1, j)$, obtained from BR-PET f with biometric data x_1 and x_2 from the same individual, biometric data y_1 from a different individual, and different parameters k and j, minimisation of accuracy degradation is achieved if $P(S(x_1, x_2) \geq t) - P(S(\tilde{x}_1, \tilde{x}_2) \geq \tilde{t}) \leq \epsilon$, $P(S(x_1, y_1) < t) - P(S(\tilde{x}_1, \tilde{y}_1) < \tilde{t}) \leq \epsilon$, and $P(S(x_2, y_1) < t) - P(S(\tilde{x}_2, \tilde{y}_1) < \tilde{t}) \leq \epsilon$, for any biometric similarity function S, possibly different thresholds t and $\tilde{t}$, and a sufficiently small $\epsilon > 0$.

— *Computational requirements:* The processing of protected biometric data should not result in a significant increase in computational costs, specifically in terms of the number of operations required for biometric recognition, when compared to the original processing of unprotected biometric data.

— *Storage requirements:* Protected biometric data should not significantly increase storage requirements, defined as the number of bits required per enrolled individual. If a BR-PET necessitates storing protected templates along with additional parameters, the storage requirements encompass the sum of the bits from these various components [42].

Minimising accuracy degradation is essential because the information processing inequality [121] demonstrates that the processing applied to safeguard biometric data cannot increase their information content. Consequently, the application of BR-PETs cannot improve the recognition accuracy of biometric data and may possibly decrease it [42]. However, when BR-PETs employ user-specific secrets unknown to an adversary, they can potentially enhance recognition performance, as discussed in [93, 95, 96] in the context of CB schemes. Along with accuracy degradation, computational and storage requirements assume a significant importance for biometric recognition systems [23, 44].

In Section 2.4, we examined the adaptation of certain data protection principles to align with the specific privacy requirements of biometric data within biometric recognition systems. Similarly, the requirement of *minimisation of accuracy degradation* extends the data protection principle of *accuracy*, which requires that personal data remains accurate and updated. Additionally, the *storage requirements* expand upon the existing *storage limitation* principle, which stipulates that personal data should be retained only for as long as necessary for the processing purpose [77] (Figure 2).

4 EVALUATION OF THE FRAMEWORK

A standardised evaluation of BR-PETs is difficult to achieve, as numerous BR-PETs have been proposed in the literature and attacks are specifically developed to target the different BR-PETs. For this reason, ISO/IEC 30136 introduces general concepts and metrics for the evaluation of privacy requirements at an abstract level, while concrete evaluations are specific to the architectures considered [42, 83]. In this section, we discuss important aspects to consider for the evaluation of

privacy requirements. We observe that a comprehensive evaluation must encompass the consideration of specific attacks that could compromise individual BR-PETs. These particular vulnerabilities might not be adequately addressed by conventional metrics that evaluate the privacy requirements for biometric data [75].

4.1 Setup

The evaluation of BR-PETs can be carried out in two ways:

— *Theoretical*: It consists in a formal demonstration of the attack potential, or the advantage of an attacker over random guessing. It considers *information-theoretic* metrics related to entropy [69] and relies on statistical assumptions that may, however, result in an overestimation of privacy properties.
— *Empirical*: It consists in assessing the feasibility of implemented attacks in terms of computational complexity [68]. Indeed, even if protected biometric data expose no information at all, an attacker can use brute force to guess the original data.

The analysis of computational complexity may invoke a theoretical perspective. However, it is worth noting that this analysis is inherently empirical, as it pertains to concrete attacks, in contrast to theoretical evaluations, which rely on mathematical assumptions and proofs without practical experiments. In [90], both theoretical and empirical analyses of a BC are conducted. The former involves approximations and bounds for the complexity of attacks, while the latter determines how closely a practical attacker can approach these theoretical bounds. Moreover, [36] recommends focusing on empirical evaluation strategies to provide more tangible insights into the irreversibility and unlinkability of face BTP methods in practice. This work also includes a table that summarises whether the evaluation of irreversibility is provided theoretically or empirically for various BTP schemes. In conclusion, theoretical and empirical evaluations complement each other: the former usually shows whether an algorithm has potential vulnerabilities, while the latter shows if attackers can exploit them [102].

4.1.1 Threat Models. The evaluation of BR-PETs requires the specification of threat models to represent the expertise and a priori information at the disposal of attackers. The following threat models have been described in ISO/IEC 30136 [42]:

— *Naive model:* The attacker neither has information about the algorithms implemented by BR-PETs nor owns a large biometric database. They have only access to the attacked protected biometric data. We do not consider meaningful the evaluation of BR-PETs according to this threat model.
— *General model:* The attacker knows the algorithms implemented by BR-PETs and the statistical properties of biometric features and has access to the protected biometric data. Privacy protection relies on the presence of secret parameters, from which further threat models are built upon each other:
  – *Standard model:* The attacker cannot execute the submodules that make use of secrets.
  – *Advanced model:* The attacker can execute part of the submodules that make use of secrets.
  – *Full-disclosure model:* All the secrets are disclosed to the attacker, which can execute the entire system.

The different categories of BR-PETs are usually evaluated according to the *standard* and *full-disclosure* models, where the only difference between the two models consists in the knowledge of the secrets by the attacker. In Table 3 we provide an overview of the evaluation of BR-PETs that we describe in Section 4.2.

Table 3. Comparison of Privacy Requirements for the Different Categories of BR-PETs When Standard and Full-disclosure Models Are Considered for Attacks

BR-PETs	Standard model			Full-disclosure model		
	Irreversibility	Unlinkability	SB privacy	Irreversibility	Unlinkability	SB privacy
CBs	✓	✓	(✓)	(✗)	(✗)	(✗)
BCs	✓	✓	✓	✗	✗	✗
HE	✓	✓	✓	✗	✗	✗
SBM	(✗)	✗	✓	(✗)	✗	✓
SBP	(✗)	✗	✓	(✗)	✗	✓

We do not consider the advanced model, as the executable submodules are specific for each BR-PET. Parentheses indicate that the validity/non-validity of requirements depends on the goodness of the reconstructed biometric data. CBs = Cancelable Biometrics, BCs = Biometric Cryptosystems, HE = Homomorphic Encryption, SBM = Soft-Biometric Minimisation, SBP = Soft-Biometric Protection.

4.2 Evaluation of Privacy Requirements

In this section, we explore various critical aspects that must be considered for conducting a comprehensive evaluation of how well BR-PETs fulfil privacy requirements. While some implementations of BR-PETs may not completely align with the requirements of their respective categories [75, 82, 99], Table 3 outlines the intended requirements that different BR-PET categories should meet, according to the *standard* and *full-disclosure* threat models.

In general, BR-PETs capable of meeting the expected privacy requirements are provided for each BR-PET category. For instance, CB schemes that provide irreversibility and unlinkability, as demonstrated in the literature, include BioHashing [45], MLP-Hash [94], and IoM-Hashing [46]. Together with threat models, it is important to regard the target of attackers, as it can be many things (Figure 1). When reconstructions of biometric data are considered *sufficiently good*, it depends on the target of the attack. For instance, the target can be to achieve similarity scores above the recognition threshold [112] or to derive the soft-biometric attributes of the original biometric data [108].

4.2.1 Irreversibility. Numerous information-theoretic metrics have been proposed to measure the irreversibility of protected biometric data; for instance, *conditional entropy* $H(x|\tilde{x})$ and *mutual information* $I(x;\tilde{x})$ quantify the uncertainty in estimating the original data x from the protected biometric data $\tilde{x}$ [42, 69]. According to ISO/IEC 30136, quantified in this manner, irreversibility is measured as the number of bits of information about the original data x revealed to the adversary [42]. Metrics of entropy are difficult to compute theoretically, especially if biometric templates contain a high number of features [69]. The application of **Principal Component Analysis (PCA)** to reduce features and simplify the computation of entropy has been proposed in [95]. PCA retains the most significant information of biometric templates, with reduced matrices that are suitable to account for the partial reversibility of protected biometric data.

As metrics of entropy are difficult to compute theoretically, irreversibility is usually measured empirically according to the computational complexity of the best-known inversion attack. However, the attacker may come up with a better attack not known by the system designer [69]. *Privacy leakage* is proposed to evaluate irreversibility in [83] with the number of bits leaked about the original data when (part of) the protected biometric data is compromised. In this sense, privacy leakage can also assess the partial reconstruction of biometric data, which may lead to successful attacks even if the original data is not completely reversed.

To evaluate irreversibility, we consider an analogy with cryptography, where the security level of an algorithm is expressed in *bits*, with n-bit security meaning that the attacker must perform 2^n operations to break the system. However, the same evaluation does not directly apply to BR-PETs.

While cryptographic systems require exact inputs to provide the desired outputs, in biometric systems similar enough approximations of biometric data may suffice. Also, compared to cryptography, an educated guess based on the statistics of biometric features can facilitate the reconstruction of the original biometric data. This is due to *broad homogeneity*, according to which the biometric data of individuals of the same sex or ethnicity present similar characteristics [39]. In cryptography, a minimum security level of 100 bits is required to consider attacks impractical [1]. This means that brute force can always be applied to guess short biometric feature vectors. Finally, we observe that the application of BR-PETs at the feature level makes the recovering of original data more difficult, as both the reconstruction of biometric features and original data must be done.

Different aspects relate to the evaluation of irreversibility for the different categories of BR-PETs. CBs require the analysis of the computational effort to reverse transformations and approximate the original biometric data. BCs require keys with sufficient size and entropy so that the number of guesses necessary to retrieve the biometric data or the key itself is high [86]. In the case of BR-PETs that either discard or transform biometric features to prevent the extraction of soft-biometric attributes, it is possible to achieve sufficiently good approximation of the original biometric, even if soft-biometric attributes are protected. As a result, these technologies typically lack irreversibility. Certain algorithms, depending on their implementation, can achieve it [105], even if safeguarding soft-biometric privacy remains their primary objective. For more detailed information, the reader is referred to existing surveys on the different topics [52, 62, 78, 86, 118]. As evident from Table 3, in the *standard* threat model, the absence of attackers' knowledge about secrets ensures irreversibility for categories of BR-PETs such as CBs, BCs, and HE. Within this threat model, the inclusion of user-specific keys introduces randomness to the secure template, thereby providing template irreversibility [31, 59].

In conclusion, we highlight the difference between BR-PETs that apply encryption to protect biometric data, like HE, and BR-PETs that apply data disentanglement. For the former, irreversibility completely relies on the secrecy of the key: if known, the original data can be immediately obtained. For the latter, irreversibility does not depend on secrets, as information is discarded and it cannot be recovered. Between the two approaches, we have CBs, which apply irreversible transformations to protect biometric data. The assurance of irreversibility in CBs depends on the difficulty of obtaining accurate approximations of the original biometric data. In fact, as discussed in [96], in the *full-disclosure* threat model, where secrets are disclosed to an adversary, CBs can provide irreversibility, in contrast to HE.

4.2.2 Unlinkability. Compared to irreversibility, the evaluation of unlinkability has received less attention in the literature, and no metrics have yet been specified by ISO/IEC 30136. Common approaches consist in the definition of linkage functions to determine if multiple representations of protected biometric data belong to the same individual, and the consequent evaluation of these functions with traditional performance metrics, such as FMR and FNMR, to compare the performances obtained when biometric data are protected with the same or different keys [8, 102].

In [79], a comparison is conducted between the ROC curves obtained by plotting (1) the probability of matching different templates of the same biometric data (obtained from multiple transformations of the proposed CB scheme with different parameters) against the FNMR and (2) the FMR against the FNMR. The similar behaviour of the ROC curves indicates that matching templates transformed with different parameters may be as challenging as achieving false matches by randomly comparing biometric data from another individual. The analysis demonstrates that the CB scheme is capable of generating templates, from the same biometric data, that are as distinct as those generated from entirely separate original biometric data. Other empirical evaluations of unlinkability have been proposed in [101], with heuristics that exploit the information leaked by protected biometric data, and in [68], with attackers that match reversed biometric data.

A general framework for the evaluation of unlinkability has been proposed in [33], with the definition of two metrics for the quantitative measurement of unlinkability. The first metric is *score-wise* and represents the difference between the conditional probabilities of having cross-matching and non-cross-matching data given a specific similarity score s. It indicates if BR-PETs fail to provide unlinkability for specific values of s. The second metric is *global* and assesses in the entire score domain if the score distributions of cross-matching and non-cross-matching data overlap. The evaluation of unlinkability depends on the linkage function considered, as inaccurate functions fail to reveal threats of specific attacks. For instance, if biometric data are protected with permutations, linkage functions have to consider not only the inversion of permutations but also attacks computing simple statistics of protected data to link individuals.

Recently, a new metric for measuring unlinkability has been introduced in [97]. This method is founded on maximal leakage, a well-established measure that could be regarded as a generalisation of mutual information. The suggested *linkability* metric utilizes maximal leakage to gauge the extent of information disclosed by two templates generated by distinct BR-PETs concerning two potential hypotheses: (1) the templates are mated, and (2) the templates are not mated.

Finally, we observe that bad sources of randomness may prevent BR-PETs from providing unlinkability. In CBs the protected biometric data of the same individual require distant transformation parameters to be unlinkable, limiting the parameter space suitable for transformations [86]. BCs may require opportune randomness to hide the information about individuals that may be contained in helper data. BR-PETs fail to provide unlinkability if they do not use random keys or parameters. As in the case of irreversibility, we observe in Table 3 that unlinkability relies on the secrecy of keys and the difficulty to obtain good approximations of the original biometric data.

4.2.3 Privacy of Soft-biometrics. BR-PETs that prevent the extraction of soft-biometric attributes from biometric data are typically provided without any formal evaluation of the proposed techniques. To measure the validity of BR-PETs, classifiers of soft-biometric attributes are applied to both original and protected biometric data, and performance differences are reported [62]. However, these approaches assume that attackers have limited resources, but attackers that possess a database of protected biometric data labelled according to soft-biometric attributes can also train classifiers in the protected domain. Additionally, attackers may derive soft-biometric attributes when they attempt to revert or link protected biometric data and when they observe the similarity scores obtained for non-mated samples. In particular, facial recognition systems produce higher similarity scores and consequently more false matches for individuals with similar soft-biometric attributes. An attack that successfully exploits this effect to derive soft-biometric attributes of protected biometric data is presented in [75]. The study encourages considering the proposed attack in the evaluation of BR-PETs. When evaluating the suppression of soft-biometric attributes, it is also important to analyse if the recognition performance achieved with protected biometric data gets worse, as soft-biometric attributes generally facilitate the recognition of individuals.

Numerous BR-PETs have been proposed in the literature to enhance the privacy of soft-biometric attributes. Given the increasing adoption of these techniques in real-world applications, it is essential to understand the extent to which soft-biometric attributes can be recovered from privacy-enhanced biometric templates [89]. A standardised protocol to evaluate the privacy of soft-biometric attributes has been proposed in [108], considering the most critical scenario of attackers that know and adapt to BR-PETs. They are able to reproduce BR-PETs and train an extensive set of soft-biometric classifiers with both protected and unprotected biometric data. According to the study [108], this attack scenario requires more consideration than others, for instance, the manual investigation of the biometric data reconstructed from protected data, because the patterns of soft-biometric attributes should be easily detectable with multiple classifiers, also trained in the

Table 4. Comparison of Additional Requirements for the Different Categories of BR-PETs When Standard and Full-disclosure Models Are Considered for Attacks

BR-PETs	Standard model			Full-disclosure model		
	Minimisation of accuracy degradation	Computational requirements	Storage requirements	Minimisation of accuracy degradation	Computational requirements	Storage requirements
CBs	✓	✓	(✓)	(✗)	✓	(✓)
BCs	(✓)	(✓)	(✓)	(✗)	(✓)	(✓)
HE	✓	(✗)	(✗)	✓	(✗)	(✗)
SBM	(✓)	✓	✓	(✓)	✓	✓
SBP	(✓)	✓	✓	(✓)	✓	✓

Checkmarks and crossmarks indicate whether the requirement is satisfied or not. Parentheses indicate that the validity/non-validity of requirements is not a simple binary (yes or no) decision but can vary based on tradeoffs with privacy requirements. Computational and storage requirements do not depend on the threat model considered. CBs = Cancelable Biometrics, BCs = Biometric Cryptosystems, HE = Homomorphic Encryption, SBM = Soft-Biometric Minimisation, SBP = Soft-Biometric Protection.

protected domain. Recognition performance and estimation of soft-biometric attributes are evaluated with protected and unprotected biometric data and suitable metrics. Subsequently, they are combined in the ***privacy gain identity loss coefficient*** **(PIC)**, which weights the gain in privacy against the loss in recognition to determine the benefit of using the analysed BR-PETs. The soundness of PIC relies on the metrics used to quantify the recognition performance and the estimation of soft-biometric attributes.

While some BR-PETs have been specifically designed to protect soft-biometric attributes, and they succeed even when multiple classifiers are trained with protected biometric data [67], it can be assumed that CBs, BCs, and HE protect soft-biometric attributes only when secrets are unknown to attackers. Depending on the specific implementation, CBs may or may not effectively safeguard soft-biometric attributes. For example, distortion transforms applied at the sample level might prove inadequate for safeguarding soft-biometrics, even within the *standard* threat model [84].

4.3 Evaluation of Additional Requirements

In Table 4 we show the additional requirements that different BR-PET categories should meet, according to the *standard* and *full-disclosure* threat models. Recognition accuracy degradation is a common issue of BR-PETs, as the transformations applied to biometric data intend to protect privacy and not to increase their ability to distinguish individuals. Ideally BR-PETs should retain the recognition performance of the original recognition systems, but it is challenging to design data transformations that achieve it and at the same time satisfy privacy requirements [69]. Accuracy degradation is assessed by comparing the recognition performance achieved with original and protected biometric data, using a common metric for evaluation, which may be FNMR at fixed FMR. However, one cannot simply conduct two tests on a given system, one with protected biometric data and one with unprotected biometric data, and then compare the performance results. Instead, an experimenter may need to conduct a technology test to generate results for multiple conventional biometric recognition systems, thereby establishing a baseline for FMR/FNMR against which the performance of protected biometric data can be compared [42].

In CBs non-invertible transformations reduce the discriminability of biometric data, while in BCs the use of error correction schemes precludes the design of recognition systems with sophisticated comparators. A comprehensive analysis of the recognition performance for numerous BCs has been conducted in [52]. On the other side, most of the BR-PETs designed to prevent the extraction of soft-biometric attributes consist in deep neural networks trained to maintain the recognition performance [6, 67]. This creates a tradeoff between recognition performance and the protection of soft-biometric attributes. Notably, the utilization of user-specific secrets within CB schemes has demonstrated its ability to enhance recognition accuracy in the *standard model* [95, 96].

Computational complexity is evaluated with the number of operations or the runtime required to execute the algorithms of BR-PETs. We have already mentioned that HE allows to perform encrypted comparisons and obtain the same results of unencrypted comparisons. However, this can be achieved only with an increase of computational costs, in contrast to PbD, which requires not to affect the functionality of the system. Practical implementations of HE schemes present significant challenges. As an alternative, *somewhat homomorphic encryption* has been introduced, which permits only a limited set of operations in the encrypted domain. Consequently, integrating state-of-the-art systems into the proposed framework while maintaining a low verification time suitable for real-time applications might be difficult [34]. The computational costs introduced by the other BR-PETs are generally lower, due to the implementation of cryptographic algorithms (in BCs) and the training of neural networks, with the latter only affecting the implementation phase of BR-PETs. As shown in Table 1, the biometric comparison method employed by BR-PETs that do not employ cryptographic algorithms (i.e., CBs, SBM, and SBP) is equivalent to the one used in unprotected systems. The increase of computational cost requires particular attention when biometric recognition systems perform identification tasks, where the number of comparisons is equal to the number of individuals enrolled in the system, unless methods to mitigate the computational workload are implemented [27].

Finally, storage requirements must be considered, especially when biometric data need to be stored on portable devices with limited storage capacities or within barcodes. Storage requirements are evaluated according to the number of bits required to store the protected biometric data of an individual enrolled in the system. BR-PETs may require storing protected templates along with additional parameters, and these elements may not be located in the same place. Therefore, when reporting the storage requirements, the number of bits needed for each component of the protected biometric data should be reported separately [42]. For example, the storage requirement for biometric data protected with BCs is calculated as the number of bits needed to store a cryptographic hash of an individual-specific data and the number of bits required to store helper data, such as the set of genuine feature points and chaff points in case of a fuzzy vault [48]. In the case of CBs, the storage requirement is the number of bits required to store the transformed biometric data and the number of bits required to store the transformation parameters. The storage requirements of these BR-PETs are lower when compared to HE, due to the increased memory requirement for ciphertext [47].

5 FUTURE WORK

In the fast-evolving field of PETs for biometric recognition systems, numerous promising research directions for future work and development can be defined. Given the increasing demand for secure and privacy-conscious biometric solutions, these research areas hold significant relevance:

(1) *Synthetic data in biometrics:* Exploring the generation of synthetic biometric data that faithfully replicates real-world patterns while preserving privacy presents an exciting frontier. Research should orient toward the development and assessment of synthetic data generation techniques, primarily for training biometric recognition systems Reference [123–125]. This approach serves to mitigate evolving privacy issues, particularly during the training of biometric recognition systems. An essential research goal is to address the intricate balance between the privacy enhancements offered by synthetic data and the utility required for effective biometric recognition systems. Additionally, face manipulation techniques hold promise in safeguarding soft-biometric attributes. Existing methods can be adapted to the context of biometric recognition systems, with a particular emphasis on preserving the identities of individuals within the scope of face manipulation.

(2) *Hybrid privacy-enhancing technologies:* Hybrid BR-PETs are a promising avenue for enhancing privacy in biometric data. This integration of multiple BR-PETs, while not a new concept in the literature, has gained prominence in recent years. For instance, the inclusion of HE schemes in hybrid BR-PETs has significantly improved privacy assurance for protected biometric data [96]. As security threats and attacks evolve, staying ahead of these challenges is crucial to maintain the integrity of privacy-enhancing biometric systems. In this context, the combination of various BR-PET techniques within hybrid systems, especially when safeguarding soft-biometric attributes, can significantly enhance overall privacy protection.
(3) *Efficient cryptographic techniques:* Advancements in cryptographic techniques, particularly within the domain of fully HE, are of high importance. While fully HE offers robust security and privacy, its implementation is often computationally intensive. Typically, there is a need to strike a balance between security and computational efficiency to make these technologies practical for real-world applications. Ongoing research is expected to improve this tradeoff. In addition, advancements in post-quantum cryptography are expected, with certain HE schemes that achieve it [53].
(4) *Improved evaluation of privacy requirements:* Research should be directed towards enhancing the metrics employed to evaluate privacy requirements. Current metrics frequently encounter challenges, including the computational cost of calculating irreversibility or dependence on known attack models. The development and potential standardisation of metrics, such as the novel metric for unlinkability [97] and the framework for restoring soft-biometric information [89], are of great significance in evaluating the privacy requirements for biometric data in biometric recognition systems.
(5) *Mobile biometrics:* Given the central role of mobile devices in our lives, ensuring privacy on these platforms is crucial. Future research should prioritise secure on-device biometric recognition to enhance usability and user experience while minimising data exposure. Mobile devices collect extensive behavioral biometrics and offer the potential for continuous authentication based on user behavior in a non-invasive manner. However, it is important to note that mobile data collection often consists in over-collection, leading to privacy issues for mobile users' data [58, 91]. This underscores the necessity for BR-PETs and legal enforcement of privacy regulations concerning mobile data.
(6) *Interdisciplinary collaboration:* Encouraging interdisciplinary collaboration between computer science, cryptography, ethics, and law will be essential. In this article, we drew upon several aspects outlined by the GDPR, customizing them to the specific context of interest, namely biometric recognition systems. Research efforts should encompass not only technical aspects but also the ethical, legal, and social implications of privacy-enhancing biometric technologies.

6 CONCLUSION

The landscape of BR-PETs has seen various concepts emerge over the years, often accompanied by different terminologies. This diversity has contributed to a lack of clarity in the data protection domain related to biometric recognition systems, an important matter for stakeholders and practitioners in the field. This work summarised the properties of the most relevant BR-PETs within a comprehensive framework. Different categories of BR-PETs (i.e., CBs, BCs, HE, SBM, and SBP) are compared in detail at each processing step and their main properties are described, including technical approaches for privacy enhancement (i.e., pseudonymisation, encryption, data disentanglement, and data rearrangement) and privacy requirements (i.e., irreversibility, unlinkability, privacy of soft-biometrics, and additional requirements) in the realm of biometric recognition.

This contextualisation serves as a valuable resource, particularly for non-experts, aiding in the selection of suitable BR-PETs based on specific privacy requirements. Furthermore, this work outlines fundamental approaches for evaluating privacy requirements and addressing practical necessities concerning BR-PETs. In conclusion, the future of privacy-enhancing technologies in biometric recognition systems holds promise but also presents diverse challenges. Collaboration among researchers, practitioners, and policymakers is essential to address this dynamic landscape, ensuring privacy enhancement while delivering efficient and user-friendly solutions for biometric recognition.

REFERENCES

[1] Jean-Philippe Aumasson. 2019. *Too Much Crypto*. Technical Report 1492.

[2] Ahmed M. Ayoup, Ashraf A. M. Khalaf, Walid El-Shafai, Fathi E. Abd El-Samie, Fahad Alraddady, and Salwa M. Serag Eldin. 2022. Cancellable multi-biometric template generation based on arnold cat map and aliasing. *Computers, Materials & Continua* 72, 2 (2022), 3687–3703.

[3] Amina Bassit, Florian Hahn, Raymond Veldhuis, and Andreas Peter. 2022. Hybrid biometric template protection: Resolving the agony of choice between Bloom filters and homomorphic encryption. *IET Biometrics* 11, 5 (2022), 430–444.

[4] Debnath Bhattacharyya and Rahul Ranjan. 2009. Biometric authentication: A review. *Science and Technology* 2, 3 (2009), 22786.

[5] John J. Borking and Charles Raab. 2001. Laws, PETs and other technologies for privacy protection. *Journal of Information, Law and Technology* 1, 1 (2001), 2001.

[6] Blaž Bortolato, Marija Ivanovska, Peter Rot, Janez Križaj, Philipp Terhörst, Naser Damer, Peter Peer, and Vitomir Štruc. 2020. Learning privacy-enhancing face representations through feature disentanglement. In *Proc. 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG'20)*. 495–502.

[7] K. W. Bowyer and M. J. Burge. 2016. *Handbook of Iris Recognition*. Springer.

[8] Ileana Buhan, Jeroen Breebaart, Jorge Guajardo, Koen de Groot, Emile Kelkboom, and Ton Akkermans. 2009. A quantitative analysis of indistinguishability for a continuous domain biometric cryptosystem. In *Proc. 4th international Workshop, and 2nd International Conference on Data Privacy Management and Autonomous Spontaneous Security*. Springer-Verlag, Berlin, 78–92.

[9] Ileana Buhan, Jeroen Doumen, Pieter Hartel, Qiang Tang, and Raymond Veldhuis. 2008. Embedding renewable cryptographic keys into continuous noisy data. In *Proc. Information and Communications Security*, Liqun Chen, Mark D. Ryan, and Guilin Wang (Eds.). Springer, Berlin, 294–310.

[10] Ileana Buhan, Emile Kelkboom, and Koen Simoens. 2010. A survey of the security and privacy measures for anonymous biometric authentication systems. In *Proc. 2010 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. 346–351.

[11] France Bélanger and Robert E. Crossler. 2011. Privacy in the digital age: A review of information privacy research in information systems. *MIS Quarterly* 35, 4 (2011), 1017–1041. https://doi.org/10.2307/41409971

[12] R. Cappelli, D. Maio, A. Lumini, and D. Maltoni. 2007. Fingerprint image reconstruction from standard templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 9 (2007), 1489–1503. https://doi.org/10.1109/TPAMI.2007.1087

[13] Ann Cavoukian. 2010. Privacy by design: The definitive workshop. A foreword by Ann Cavoukian, Ph.D. *Identity in the Information Society* 3, 2 (2010), 247–251.

[14] Ann Cavoukian. 2009. Privacy by design: The 7 foundational principles. *Information and Privacy Commissioner of Ontario, Canada* 5 (2009), 12.

[15] Ann Cavoukian, Michelle Chibba, and Alex Stoianov. 2012. Advances in biometric encryption: Taking privacy by design from academic research to deployment. *Review of Policy Research* 29, 1 (2012), 37–61.

[16] Ann Cavoukian and Alex Stoianov. 2011. Biometric encryption. In *Encyclopedia of Cryptography and Security*, Henk C. A. van Tilborg and Sushil Jajodia (Eds.). Springer US, Boston, MA, 90–98.

[17] Yao-Jen Chang, Wende Zhang, and Tsuhan Chen. 2004. Biometrics-based cryptographic key generation. In *Proc. 2004 IEEE International Conference on Multimedia and Expo (ICME'04)*, Vol. 3. 2203–2206.

[18] Yunjey Choi, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo. 2018. StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*. 8789–8797.

[19] Roger Clarke. 2006. What's 'privacy'? In *Proc. Austral. Reform Comission Workshop (ALRCW'06)*, Vol. 28. http://www.rogerclarke.com/DV/Privacy.html

[20] Draft Committee. 1948. *Universal Declaration of Human Rights*. Retrieved April 21, 2022 from https://www.un.org/en/about-us/universal-declaration-of-human-rights
[21] European Convention. 2000. *Charter of Fundamental Rights of the European Union*. Retrieved April 19, 2022 from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT
[22] Antitza Dantcheva, Petros Elia, and Arun Ross. 2016. What else does your biometric data reveal? a survey on soft biometrics. *IEEE Transactions on Information Forensics and Security* 11, 3 (2016), 441–467.
[23] Shaveta Dargan and Munish Kumar. 2020. A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. *Expert Systems with Applications* 143 (2020), 113114.
[24] Paula Delgado-Santos, Giuseppe Stragapede, Ruben Tolosana, Richard Guest, Farzin Deravi, and Ruben Vera-Rodriguez. 2022. A survey of privacy vulnerabilities of mobile device sensors. *Computing Surveys* 54 (2022), 1–30.
[25] Paula Delgado-Santos, Ruben Tolosana, Richard Guest, Ruben Vera, Farzin Deravi, and Aythami Morales. 2022. GaitPrivacyON: Privacy-preserving mobile gait biometrics using unsupervised learning. *Pattern Recognition Letters* 161 (2022), 30–37.
[26] Pawel Drozdowski, Nicolas Buchmann, Christian Rathgeb, Marian Margraf, and Christoph Busch. 2019. On the application of homomorphic encryption to face identification. In *2019 International Conference of the Biometrics Special Interest Group (Biosig'19)*. IEEE, 1–5.
[27] Pawel Drozdowski, Christian Rathgeb, and Christoph Busch. 2019. Computational workload in biometric identification systems: An overview. *IET Biometrics* 8, 6 (2019), 351–368.
[28] Pawel Drozdowski, Christian Rathgeb, and Christoph Busch. 2021. The watchlist imbalance effect in biometric face identification: Comparing theoretical estimates and empiric measurements. In *Proc. IEEE/CVF International Conference on Computer Vision*. 3757–3765.
[29] Yi C. Feng, Pong C. Yuen, and Anil K. Jain. 2009. A hybrid approach for generating secure and discriminating face template. *IEEE Transactions on Information Forensics and Security* 5, 1 (2009), 103–117.
[30] Caroline Fontaine and Fabien Galand. 2007. A survey of homomorphic encryption for nonspecialists. *EURASIP Journal on Information Security* 2007 (2007), 1–10.
[31] Alwyn Goh and David C. L. Ngo. 2003. Computation of cryptographic keys from face biometrics. In *IFIP International Conference on Communications and Multimedia Security*. Springer, 1–13.
[32] Marta Gomez-Barrero and Javier Galbally. 2020. Reversing the irreversible: A survey on inverse biometrics. *Computers & Security* 90 (2020), 101700.
[33] Marta Gomez-Barrero, Javier Galbally, Christian Rathgeb, and Christoph Busch. 2018. General framework to evaluate unlinkability in biometric template protection systems. *IEEE Transactions on Information Forensics and Security* 13, 6 (2018), 1406–1420.
[34] Marta Gomez-Barrero, Emanuele Maiorana, Javier Galbally, Patrizio Campisi, and Julian Fierrez. 2017. Multi-biometric template protection based on homomorphic encryption. *Pattern Recognition* 67 (2017), 149–163.
[35] Patrick Grother. 2022. Face recognition vendor test (FRVT) Part 8: Summarizing demographic differentials. *National Institute of Standards and Technology (NIST)*, Gaithersburg, MD, USA, Interagency Rep. NISTIR, 8429.
[36] Vedrana Krivokuća Hahn and Sébastien Marcel. 2022. Biometric template protection for neural-network-based face recognition systems: A survey of methods and evaluation techniques. *IEEE Transactions on Information Forensics and Security* 18 (2022), 639–666.
[37] Vedrana Krivokuća Hahn and Sébastien Marcel. 2022. Towards protecting face embeddings in mobile face verification scenarios. *IEEE Transactions on Biometrics, Behavior, and Identity Science* 4, 1 (2022), 117–134.
[38] Johannes Heurix, Peter Zimmermann, Thomas Neubauer, and Stefan Fenz. 2015. A taxonomy for privacy enhancing technologies. *Computers & Security* 53 (2015), 1–17.
[39] John J. Howard, Yevgeniy B. Sirotin, and Arun R. Vemury. 2019. The effect of broad and specific demographic homogeneity on the imposter distributions and false match rates in face recognition algorithm performance. In *Proc. 2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS'19)*. 1–8.
[40] ISO/IEC JTC1 SC27 Security Techniques. 2022. *ISO/IEC 24745:2022. Information Technology—Security Techniques—Biometric Information Protection*. International Organization for Standardization.
[41] ISO/IEC JTC1 SC37 Biometrics. 2022. *ISO/IEC 2382-37:2022 Information Technology—Vocabulary—Part 37: Biometrics*.
[42] ISO/IEC JTC1 SC37 Security Techniques. 2018. *ISO/IEC 30136:2018. Information Technology—Performance Testing of Biometric Template Protection Schemes*. International Organization for Standardization.
[43] Anil K. Jain, Debayan Deb, and Joshua J. Engelsma. 2021. Biometrics: Trust, but verify. *IEEE Transactions on Biometrics, Behavior, and Identity Science* 4, 3 (2021), 303–323.
[44] Anil K. Jain, Arun Ross, and Salil Prabhakar. 2004. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology* 14, 1 (2004), 4–20.
[45] Andrew Teoh Beng Jin, David Ngo Chek Ling, and Alwyn Goh. 2004. Biohashing: Two factor authentication featuring fingerprint data and tokenised random number. *Pattern Recognition* 37, 11 (2004), 2245–2255.

[46] Zhe Jin, Jung Yeon Hwang, Yen-Lung Lai, Soohyung Kim, and Andrew Beng Jin Teoh. 2017. Ranking-based locality sensitive hashing-enabled cancelable biometrics: Index-of-max hashing. *IEEE Transactions on Information Forensics and Security* 13, 2 (2017), 393–407.

[47] Arun Kumar Jindal, Imtiyazuddin Shaik, Vasudha Vasudha, Srinivasa Rao Chalamala, Rajan Ma, and Sachin Lodha. 2020. Secure and privacy preserving method for biometric template protection using fully homomorphic encryption. In *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom'20)*. IEEE, 1127–1134.

[48] Ari Juels and Madhu Sudan. 2006. A fuzzy vault scheme. *Designs, Codes and Cryptography* 38 (2006), 237–257.

[49] Ari Juels and Martin Wattenberg. 1999. A fuzzy commitment scheme. In *Proc. 6th ACM Conference on Computer and Communications Security*. Association for Computing Machinery, New York, NY, USA, 28–36.

[50] Sanjay Kanade, Danielle Camara, Emine Krichen, Dijana Petrovska-Delacrétaz, and Bernadette Dorizzi. 2008. Three factor scheme for biometric-based cryptographic key regeneration using iris. In *2008 Biometrics Symposium*. IEEE, 59–64.

[51] Christof Kauba, Emanuela Piciucco, Emanuele Maiorana, Marta Gomez-Barrero, Bernhard Prommegger, Patrizio Campisi, and Andreas Uhl. 2022. Towards practical cancelable biometrics for finger vein recognition. *Information Sciences* 585 (2022), 395–417.

[52] Prabhjot Kaur, Nitin Kumar, and Maheep Singh. 2023. Biometric cryptosystems: A comprehensive survey. *Multimedia Tools and Applications* 82, 11 (2023), 16635–16690.

[53] Jascha Kolberg, Pawel Drozdowski, Marta Gomez-Barrero, Christian Rathgeb, and Christoph Busch. 2020. Efficiency analysis of post-quantum-secure face template protection schemes based on homomorphic encryption. In *2020 International Conference of the Biometrics Special Interest Group (BIOSIG'20)*. IEEE, 1–4.

[54] Pavel Korshunov and Touradj Ebrahimi. 2013. Using face morphing to protect privacy. In *Proc. 2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance*. 208–213.

[55] Lu Leng and Jiashu Zhang. 2011. Dual-key-binding cancelable palmprint cryptosystem for palmprint protection and information security. *Journal of Network and Computer Applications* 34, 6 (2011), 1979–1989.

[56] Jingzhi Li, Lutong Han, Ruoyu Chen, Hua Zhang, Bing Han, Lili Wang, and Xiaochun Cao. 2021. Identity-preserving face anonymization via adaptively facial attributes obfuscation. In *Proc. 29th ACM International Conference on Multimedia*. 3891–3899.

[57] Nguyen Cong Luong, Dinh Thai Hoang, Ping Wang, Dusit Niyato, Dong In Kim, and Zhu Han. 2016. Data collection and wireless communication in internet of things (IoT) using economic analysis and pricing models: A survey. *IEEE Communications Surveys Tutorials* 18, 4 (2016), 2546–2590.

[58] Michelle Madejski, Maritza Johnson, and Steven M. Bellovin. 2012. A study of privacy settings errors in an online social network. In *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*. IEEE, 340–345.

[59] Guangcan Mai, Kai Cao, Xiangyuan Lan, and Pong C. Yuen. 2021. SecureFace: Face template protection. *IEEE Transactions on Information Forensics and Security* 16 (2021), 262–277.

[60] Guangcan Mai, Kai Cao, Pong C. Yuen, and Anil K. Jain. 2019. On the reconstruction of face images from deep face templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41, 5 (2019), 1188–1202. https://doi.org/10.1109/TPAMI.2018.2827389

[61] Davide Maltoni, Dario Maio, Anil K. Jain, and Salil Prabhakar. 2009. *Handbook of Fingerprint Recognition*. Springer Science & Business Media.

[62] Blaž Meden, Peter Rot, Philipp Terhörst, Naser Damer, Arjan Kuijper, Walter J. Scheirer, Arun Ross, Peter Peer, and Vitomir Štruc. 2021. Privacy-enhancing face biometrics: A comprehensive survey. *IEEE Transactions on Information Forensics and Security* 16 (2021), 4147–4183.

[63] Pietro Melzi, Hatef Otroshi Shahreza, Christian Rathgeb, Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, Sébastien Marcel, and Christoph Busch. 2023. Multi-IVE: Privacy enhancement of multiple soft-biometrics in face embeddings. In *Proc. IEEE/CVF Winter Conference on Applications of Computer Vision Workshops*. 323–331.

[64] Pietro Melzi, Ruben Tolosana, Alberto Cecconi, Ancor Sanz-Garcia, Guillermo J. Ortega, Luis Jesus Jimenez-Borreguero, and Ruben Vera-Rodriguez. 2021. Analyzing artificial intelligence systems for the prediction of atrial fibrillation from sinus-rhythm ECGs including demographics and feature visualization. *Scientific Reports* 11, 1 (2021).

[65] Vahid Mirjalili, Sebastian Raschka, and Arun Ross. 2020. PrivacyNet: Semi-adversarial networks for multi-attribute face privacy. *IEEE Transactions on Image Processing* 29 (2020), 9400–9412.

[66] Vahid Mirjalili and Arun Ross. 2017. Soft biometric privacy: Retaining biometric utility of face images while perturbing gender. In *Proc. 2017 IEEE International Joint Conference on Biometrics (IJCB'17)*. 564–573.

[67] Aythami Morales, Julian Fierrez, Ruben Vera-Rodriguez, and Ruben Tolosana. 2021. SensitiveNets: Learning agnostic representations with application to face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 6 (2021), 2158–2164.

[68] Abhishek Nagar, Karthik Nandakumar, and Anil K. Jain. 2010. Biometric template transformation: A security analysis. In *Proc. Media Forensics and Security II*, Vol. 7541. SPIE, 237–251.
[69] Karthik Nandakumar and Anil K. Jain. 2015. Biometric template protection: Bridging the performance gap between theory and practice. *IEEE Signal Processing Magazine* 32, 5 (2015), 88–100.
[70] Hailong Ning, Xiangtao Zheng, Xiaoqiang Lu, and Yuan Yuan. 2021. Disentangled representation learning for cross-modal biometric matching. *IEEE Transactions on Multimedia* 24 (2021), 1763–1774.
[71] Council of the European Union. 1950. *Convention for the Protection of Human Rights and Fundamental Freedoms (ETS No. 005)*. Retrieved April 19, 2022 from https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatynum=005
[72] Council of the European Union. 1981. *Council of Europe Convention 108 for the Protection of Individuals with regard to Automatic Processing of Personal Data*. Retrieved June 7, 2022 from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02016R0679-20160504&qid=1532348683434
[73] Information Commissioner's Office. 2022. *Chapter 5: Privacy-Enhancing Technologies (PETs)*. Retrieved October 25, 2023 from https://ico.org.uk/media/about-the-ico/consultations/4021464/chapter-5-anonymisation-pets.pdf
[74] Jakub Oravec, L'uboš Ovseník, and Ján Turán. 2021. A plaintext-related image encryption algorithm usable in biometric systems. In *2021 31st International Conference Radioelektronika (RADIOELEKTRONIKA'21)*. IEEE, 1–6.
[75] Dailé Osorio-Roig, Christian Rathgeb, Pawel Drozdowski, Philipp Terhörst, Vitomir Štruc, and Christoph Busch. 2022. An attack on facial soft-biometric privacy enhancement. *IEEE Transactions on Biometrics, Behavior, and Identity Science* 4, 2 (2022), 263–275.
[76] Asem Othman and Arun Ross. 2015. Privacy of facial soft biometrics: Suppressing gender but retaining identity. In *Proc. Computer Vision (ECCV'14 Workshops): Zurich, Switzerland, September 6-7 and 12, 2014, Proceedings, Part II 13*. Springer, 682–696.
[77] European Parliament and Council of the European Union. 2016. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)*. Retrieved April 19, 2022 from https://bit.ly/3y0kDkD
[78] Vishal M. Patel, Nalini K. Ratha, and Rama Chellappa. 2015. Cancelable biometrics: A review. *IEEE Signal Processing Magazine* 32, 5 (2015), 54–65.
[79] Emanuela Piciucco, Emanuele Maiorana, Christof Kauba, Andreas Uhl, and Patrizio Campisi. 2016. Cancelable biometrics for finger vein recognition. In *Proc. 2016 1st International Workshop on Sensing, Processing and Learning for Intelligent Machines (SPLINE'16)*.
[80] Joao Ribeiro Pinto, Miguel V. Correia, and Jaime S. Cardoso. 2021. Secure triplet loss: Achieving cancelability and non-linkability in end-to-end deep biometrics. *IEEE Transactions on Biometrics, Behavior, and Identity Science* 3, 2 (2021), 180–189.
[81] Jannis Priesnitz, Rolf Huesmann, Christian Rathgeb, Nicolas Buchmann, and Christoph Busch. 2022. Mobile contactless fingerprint recognition: Implementation, performance and usability aspects. *Sensors* 22, 3 (2022), 792.
[82] Feng Quan, Su Fei, Cai Anni, and Zhao Feifei. 2008. Cracking cancelable fingerprint template of ratha. In *2008 International Symposium on Computer Science and Computational Technology*, Vol. 2. IEEE, 572–575.
[83] Shantanu Rane. 2014. Standardization of biometric template protection. *IEEE MultiMedia* 21, 4 (2014), 94–99.
[84] Nalini K. Ratha, Jonathan H. Connell, and Ruud M. Bolle. 2001. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal* 40, 3 (2001), 614–634.
[85] Christian Rathgeb and Christoph Busch. 2017. Biometric template protection: State-of-the-art, issues and challenges. *User-centric Privacy and Security in Biometrics* (2017), 173–191. https://digital-library.theiet.org/content/books/10.1049/pbse004e_ch8
[86] Christian Rathgeb and Andreas Uhl. 2011. A survey on biometric cryptosystems and cancelable biometrics. *EURASIP Journal on Information Security* 2011, 1 (2011), 3.
[87] Imad Rida, Noor Al-Maadeed, Somaya Al-Maadeed, and Sambit Bakshi. 2020. A comprehensive overview of feature representation for biometric recognition. *Multimedia Tools and Applications* 79, 7–8 (2020), 4867–4890.
[88] Yuji Roh, Geon Heo, and Steven Euijong Whang. 2021. A survey on data collection for machine learning: A big data-ai integration perspective. *IEEE Transactions on Knowledge and Data Engineering* 33, 4 (2021), 1328–1347.
[89] Peter Rot, Klemen Grm, Peter Peer, and Vitomir Štruc. 2023. PrivacyProber: Assessment and detection of soft-biometric privacy-enhancing techniques. *IEEE Transactions on Dependable and Secure Computing* (2023), 1–18. https://ieeexplore.ieee.org/abstract/document/10264192/authors#authors
[90] Enrique Argones Rúa, Davy Preuveneers, and Wouter Joosen. 2021. On the security of biometrics and fuzzy commitment cryptosystems: A study on gait authentication. *IEEE Transactions on Information Forensics and Security* 16 (2021), 5211–5224.

[91] Ira S. Rubinstein and Nathaniel Good. 2013. Privacy by design: A counterfactual analysis of Google and Facebook privacy incidents. *Berkeley Technology Law Journal* 28 (2013), 1333.
[92] Torsten Schlett, Christian Rathgeb, and Christoph Busch. 2021. Deep learning-based single image face depth data enhancement. *Computer Vision and Image Understanding* 210 (2021), 103247.
[93] Hatef Otroshi Shahreza, Vedrana Krivokuća Hahn, and Sébastien Marcel. 2021. On the recognition performance of biohashing on state-of-the-art face recognition models. In *2021 IEEE International Workshop on Information Forensics and Security (WIFS'21)*. IEEE, 1–6.
[94] Hatef Otroshi Shahreza, Vedrana Krivokuća Hahn, and Sébastien Marcel. 2023. MLP-Hash: Protecting face templates via hashing of randomized multi-layer perceptron. In *Proceedings of the 31st European Signal Processing Conference (EUSIPCO'23)*. IEEE.
[95] Hatef Otroshi Shahreza, Pietro Melzi, Dailé Osorio-Roig, Christian Rathgeb, Christoph Busch, Sébastien Marcel, Ruben Tolosana, and Ruben Vera-Rodriguez. 2023. Benchmarking of cancelable biometrics for deep templates. *arXiv preprint arXiv:2302.13286* (2023).
[96] Hatef Otroshi Shahreza, Christian Rathgeb, Dailé Osorio-Roig, Vedrana Krivokuca Hahn, Sébastien Marcel, and Christoph Busch. 2022. Hybrid protection of biometric templates by combining homomorphic encryption and cancelable biometrics. In *2022 IEEE International Joint Conference on Biometrics (IJCB'22)*. IEEE.
[97] Hatef Otroshi Shahreza, Yanina Y. Shkel, and Sébastien Marcel. 2023. Measuring linkability of protected biometric templates using maximal leakage. *IEEE Transactions on Information Forensics and Security* 18 (2023), 2262–2275.
[98] Yujun Shen, Ceyuan Yang, Xiaoou Tang, and Bolei Zhou. 2020. InterFaceGAN: Interpreting the disentangled face representation learned by GANs. *Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 4 (2020), 2004–2018.
[99] Sang Wook Shin, Mun-Kyu Lee, Daesung Moon, and Kiyoung Moon. 2009. Dictionary attack on functional transform-based cancelable fingerprint templates. *ETRI Journal* 31, 5 (2009), 628–630.
[100] Koen Simoens, Chi-Ming Chang, and Bart Preneel. 2010. Reversing protected minutiae vicinities. In *2010 4th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS'10)*. IEEE, 1–8.
[101] Koen Simoens, Pim Tuyls, and Bart Preneel. 2009. Privacy weaknesses in biometric sketches. In *Proc. 2009 30th IEEE Symposium on Security and Privacy*. 188–203.
[102] Koen Simoens, Bian Yang, Xuebing Zhou, Filipe Beato, Christoph Busch, Elaine M. Newton, and Bart Preneel. 2012. Criteria towards metrics for benchmarking template protection algorithms. In *5th IAPR International Conference on Biometrics (ICB'12)*. 498–505.
[103] Parul Sood and Manvjeet Kaur. 2014. Methods of automatic alignment of fingerprint in fuzzy vault: A review. In *Proc. 2014 Recent Advances in Engineering and Computational Sciences (RAECS'14)*. 1–4.
[104] Philipp Terhörst, Naser Damer, Florian Kirchbuchner, and Arjan Kuijper. 2019. Unsupervised privacy-enhancement of face representations using similarity-sensitive noise transformations. *Applied Intelligence* 49 (2019), 3043–3060.
[105] Philipp Terhörst, Marco Huber, Naser Damer, Florian Kirchbuchner, and Arjan Kuijper. 2020. Unsupervised enhancement of soft-biometric privacy with negative face recognition. *arXiv preprint arXiv:2002.09181* (2020).
[106] Philipp Terhörst, Naser Damer, Florian Kirchbuchner, and Arjan Kuijper. 2019. Suppressing gender and age in face templates using incremental variable elimination. In *Proc. 2019 International Conference on Biometrics (ICB'19)*. 1–8.
[107] Philipp Terhörst, Daniel Fährmann, Naser Damer, Florian Kirchbuchner, and Arjan Kuijper. 2020. Beyond identity: What information is stored in biometric face templates? In *Proc. 2020 IEEE International Joint Conference on Biometrics (IJCB'20)*. 1–10.
[108] Philipp Terhörst, Marco Huber, Naser Damer, Peter Rot, Florian Kirchbuchner, Vitomir Struc, and Arjan Kuijper. 2020. Privacy evaluation protocols for the evaluation of soft-biometric privacy-enhancing technologies. In *2020 International Conference of the Biometrics Special Interest Group (BIOSIG'20)*. 1–5.
[109] Philipp Terhörst, Kevin Riehl, Naser Damer, Peter Rot, Blaz Bortolato, Florian Kirchbuchner, Vitomir Struc, and Arjan Kuijper. 2020. PE-MIU: A training-free privacy-enhancing face recognition approach based on minimum information units. *IEEE Access* 8 (2020), 93635–93647.
[110] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, and Javier Ortega-Garcia. 2020. Biotouchpass2: Touchscreen password biometrics using time-aligned recurrent neural networks. *IEEE Transactions on Information Forensics and Security* 15 (2020), 2616–2628.
[111] Ruben Tolosana, Ruben Vera-Rodriguez, Carlos Gonzalez-Garcia, Julian Fierrez, Aythami Morales, Javier Ortega-Garcia, Juan Carlos Ruiz-Garcia, Sergio Romero-Tapiador, Santiago Rengifo, Miguel Caruana, Jiajia Jiang, Songxuan Lai, Lianwen Jin, Yecheng Zhu, Javier Galbally, Moises Diaz, Miguel Angel Ferrer, Marta Gomez-Barrero, Ilya Hodashinsky, Konstantin Sarin, Artem Slezkin, Marina Bardamova, Mikhail Svetlakov, Mohammad Saleem, Cintia Lia Szcs, Bence Kovari, Falk Pulsmeyer, Mohamad Wehbi, Dario Zanca, Sumaiya Ahmad, Sarthak Mishra, and Suraiya Jabin. 2022. SVC-onGoing: Signature verification competition. *Pattern Recognition* 127 (2022), 108609.

[112] Yazhou Wang, Bing Li, Jiaxin Wu, Qianya Ma, Guozhu Liu, and Yuqi Li. 2022. A secure biometric template protection mechanism against similarity-based attack. In *2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS'22)*. IEEE, 169–175.

[113] Alan F. Westin. 1968. Privacy and freedom. *Washington and Lee Law Review* 25, 1 (1968), 166.

[114] Alan F. Westin. 2003. Social and political dimensions of privacy: Social and political. *Journal of Social Issues* 59, 2 (2003), 431–453.

[115] Wikipedia. 2023. *Privacy-Enhancing Technologies*. Retrieved October 19, 2023 from https://en.wikipedia.org/wiki/Privacy-enhancing_technologies

[116] Wei Jing Wong, M. L. Dennis Wong, and Andrew Beng Jin Teoh. 2014. A security-and privacy-driven hybrid biometric template protection technique. In *2014 International Conference on Electronics, Information and Communications (ICEIC'14)*. IEEE, 1–5.

[117] Bian Yang and Christoph Busch. 2009. Parameterized geometric alignment for minutiae-based fingerprint template protection. In *2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*. IEEE, 1–6.

[118] Wencheng Yang, Song Wang, Hui Cui, Zhaohui Tang, and Yan Li. 2023. A review of homomorphic encryption for privacy-preserving biometrics. *Sensors* 23, 7 (2023), 3566.

[119] Lin Yuan, Linguo Liu, Xiao Pu, Zhao Li, Hongbo Li, and Xinbo Gao. 2022. PRO-Face: A generic framework for privacy-preserving recognizable obfuscation of face images. In *Proc. 30th ACM International Conference on Multimedia*. 1661–1669.

[120] Yushu Zhang, Tao Wang, Ruoyu Zhao, Wenying Wen, and Youwen Zhu. 2023. RAPP: Reversible privacy preservation for various face attributes. *IEEE Transactions on Information Forensics and Security* 18 (2023), 3047–3087.

[121] Jacob Ziv and Moshe Zakai. 1973. On functionals satisfying a data-processing theorem. *IEEE Transactions on Information Theory* 19, 3 (1973), 275–283.

[122] Jinyu Zuo, Nalini K. Ratha, and Jonathan H. Connell. 2008. Cancelable iris biometric. In *Proc. 2008 19th International Conference on Pattern Recognition*. 1–4.

[123] Pietro Melzi, Christian Rathgeb, Ruben Tolosana, Ruben Vera-Rodriguez, Dominik Lawatsch, Florian Domin, and Maxim Schaubert. 2023. GANDiffFace: Controllable generation of synthetic datasets for face recognition with realistic variations. In *Proc. IEEE/CVF International Conference on Computer Vision*. 3086–3095.

[124] Pietro Melzi, Ruben Tolosana, Ruben Vera-Rodriguez, Minchul Kim, Christian Rathgeb, Xiaoming Liu, Ivan DeAndres-Tame, et al. 2024. FRCSyn-onGoing: Benchmarking and comprehensive evaluation of real and synthetic data to improve face recognition systems. *Information Fusion* 107 (2024), 102322.

[125] Ivan DeAndres-Tame, Ruben Tolosana, Pietro Melzi, Ruben Vera-Rodriguez, Minchul Kim, Christian Rathgeb, Xiaoming Liu, et al. 2024. FRCSyn Challenge at CVPR 2024: Face recognition challenge in the era of synthetic data. *arXiv preprint arXiv:2404.10378* (2024).

Received 25 October 2022; revised 26 April 2024; accepted 4 May 2024