Computer Standards & Interfaces xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computer Standards & Interfaces

journal homepage: www.elsevier.com/locate/csi

How to implement EU data protection regulation for R&D in biometrics

Raul Sanchez-Reillo*, Ines Ortega-Fernandez, Wendy Ponce-Hernandez, Helga C. Quiros-Sandoval

Carlos III University of Madrid, University Group for Identification Technologies (GUTI), Av. Universidad, 30, Leganes, Madrid 28911, Spain

ARTICLE INFO

Keywords:
Anonymization
Biometrics
Coding
Data Protection
Databases
Privacy
User consent

ABSTRACT

Biometrics R&D has to deal with personal data. From the Universal Declaration of Human Rights, privacy of a human being shall be protected, and this is addressed in different ways in each region of the world. In the case of the European Union, Data Protection Directives, Laws and Regulations have been established, and interpreted in different ways by each European Member State. Such a diversity has pushed the European Union to generate an improved regulation that will be mandatory from May 2018. Biometric R&D shall not only comply with the current Directive, but also has to adapt its work to the new Regulation. This work is intended to describe the situation and provide a recommended procedure when having to acquire personal data. The recommended procedure is illustrated by the implementation of a Biometric Data Acquisition Platform, used to acquire fingerprints from nearly 600 citizens using different sensors.

1. Introduction

Data Protection Regulations are essential to guarantee the privacy of citizens, in particular in current society, where our personal data is extremely exposed to a large number and variety of attacks and abuse. This was initially included in the Universal Declaration of Human Rights, back in 1948, in its Article 12. But this article has evolved differently in different parts of the world. One of the regions where this article has been strongly articulated is the European Union (EU), where in 1995 a Directive was issued as to regulate data protection in all EU Member States (95/46/EC) [1]. Although the ideas behind such Directive where very clear, the Directives where understood and implemented differently in each Member State. The evolution of Information Technology, and the growing number of cases of identity theft, as well as personal data abuse, have led to the EU to develop a new version of this Directive, but in this case not as a Directive, but as a proper Regulation. This regulation, referred as 2016/679 [2], will be of mandatory application from May 25, 2018.

This kind of regulations, and in particular the new one, create major challenges to activities where personal data has to be considered [3]. In particular, Biometrics require to acquire personal data from citizens, as to both, develop the technology and evaluate the results achieved [4]. Such evaluation does not only include performance, but also security and usability, and not only in prototypes, but also in final products. The relationship between Biometrics and Data Protection has been studied in depth in the last decade [5,6]. But special emphasis has to be applied to R&D activities, where biometric and other personal data has to be acquired, stored and kept. In addition, the transmission of such personal data to third parties is under major controversy, which also make selling or acquiring databases a serious legal problem.

The present work presents the problem of privacy and data protection when related to activities requiring the recognition of human beings. In order to better understand the problem, the current data protection directive from the EU will be studied in Section 3, introducing the most important concepts and requirements. From such a knowledge, the application of such directive to Biometrics R&D will be shown in Section 4. This section will also include the proposal of a recommended procedure, which will be illustrated by explaining the particular case of a major evaluation of fingerprint technology. The new coming scenario, created by the new EU Regulation, will be analysed in Section 5, leading to proposed future work and conclusions.

2. Privacy vs. recognition

Human Recognition can be performed in three different ways, or any combination of those three:

- By **what the user knows** (e.g. passwords). The advantages of this method are based on the lack of addition infrastructure, plus the possibility of changing such a knowledge, and therefore updating

* Corresponding authors.
E-mail addresses: rsreillo@ing.uc3m.es (R. Sanchez-Reillo), inortega@ing.uc3m.es (I. Ortega-Fernandez), wponce@ing.uc3m.es (W. Ponce-Hernandez), hquiros@ing.uc3m.es (H.C. Quiros-Sandoval).

https://doi.org/10.1016/j.csi.2018.01.007
Received 20 August 2017; Received in revised form 31 December 2017; Accepted 31 January 2018

the credential. On the other side, knowledge can be forgotten, or even easily copied by eavesdropping.
- By **what the user has** (e.g. cards). Again the token used for recognition can be changed, although that will require some additional cost. But also the token can be stolen or lost.
- By **what the user is** (e.g. Biometrics). The credentials are expected to be unique (if the discriminative power of the algorithm is high), but they cannot be changed. If Biometrics is well developed and deployed, it can be a comfortable way to identify the individual. But credentials are usually publicly available (e.g. by taking a photo), and they may be spoofed. Therefore Presentation Attack Detection (PAD) mechanisms should be in place.

In any of these cases, the credentials are linked to the real identity of the subject within the system. As identity is a piece of personal data, both the identity and the link shall be protected, as to preserve citizens' privacy. Therefore many applications involve the acquisition and use of Personal Data (i.e. administrative data), for several reasons:

- For registering the participation of the citizen in the system
- To avoid duplicated entries in the system
- For future communications between the system and the citizen
- For allowing the citizen to claim his rights

Obviously, any system using a Human Recognition technique shall integrate a set of security mechanisms so as to provide the best conditions to preserve the privacy of the system. Therefore the evaluation of the security of the recognition procedure [7] should be mandatory.

In addition, it has to be considered that Biometric Data is a piece of personal data, which is physically linked to the user. Currently it is considered as of the same level as administrative data (e.g. name or address). In addition, for some modalities, the link between the data and the person can be direct (e.g. face recognition or even signatures). Such direct link may be considered as an evidence by a trial court.

As there is a direct relationship between Biometrics and Personal Data, there is a huge concern about privacy:

- How can the citizen be sure that his personal data is not used outside the claimed purpose?
- How feasible is for the system provider to use such data for other means?
- Is it possible for the user to belong to a system without providing his administrative data?

So in few words, if we need to recognize human beings, but handling such data may attack his/her privacy, how can this be handled? This question led to the definition of data protection laws and directives, such as the ones above mentioned.

3. Personal data protection: the EU directive

The present section will describe the current EU Directive, as a background to understand the changes that the new regulation will bring.

3.1. Background

Article 12 of the Universal Declaration of Human Rights (10 December 1948) reads as: *"No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation. Everyone has the right to the protection of the law against such interference or attacks"*. In addition, some countries also declare laws that limit the use of computer science to guarantee the honor and personal privacy of citizens.

With that in mind, some countries developed laws and acts regarding the automatic process of personal data (e.g. LORTAD in Spain in 1992 [8]). These laws should always be sustained by a regulation stating the rules for the treatment of the automated files using personal data.

In 1995, the EU approved the 95/46/EC Directive on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1]. It shall be clear that such directive only applies when data processing is automated or when data is stored in a structured file so that access to such data is simplified.

The Directive is not a law, but a strong recommendation to Member States. Therefore such Directive was implemented in laws and regulations in each of the countries of the EU. Depending on the Member State, the level of coverage of the Directive changed, achieving laws of different strength levels. Some researchers from some Member States provided strategies to support the local Data Protection law with security mechanisms [9].

In any case, there are two major principles behind the Directive:

- The citizen is entitled to preserve the full control on his personal data: who is collecting them, for what, and where are they going
- The data collector shall implement a relevant policy to preserve the citizen rights, declaring the filing of that data and the person responsible for keeping the policy.

3.2. Concepts and requirements

Going more in depth on the Directive, the following concepts and requirements are defined:

- The citizen is the one **deciding which part of his personal data is to be provided**.
- The citizen has the **right to declare his consent** towards the data collection act.
- The citizen has the **right to be informed** about **who** is collecting the data, **what** is the reason for collecting such data and **which** processes are going to be applied to his data.
- The citizen has the **right to deny** the collection of his personal data.

The Directive also defines the concept of Data Quality, including rules such as:

- Data should be relevant, adequate and non-excessive.
- Data cannot be used for a purpose different to the one declared when being collected.
- Data shall be maintained exact and accurate.
- Data shall be cancelled when no longer needed.

The Information Right that the citizen has during collection of his data, allows to him to ask about the existence, finality and recipients of the data, the optionality and mandatory character of each data collection, if there are some consequences in not providing some of the data, and how he can execute his rights. Additionally he has to know the identity and address of the file responsible.

In order to allow the traceability of such Information Right, a user consent shall be signed, being extremely important that such user consent shall be unambiguous and revocable.

Certain data may be subject to a higher level of protection. Typically 3 levels are defined, from a basic protection to the strongest one. For example:

- Level 1: Administrative data (including Biometrics)
- Level 2: Health data
- Level 3: Political, religious and ideological data, race, sexual orientation, etc.

The Security of the data collected is implicit within the Directive, but there is not an explicit regulation. The only statement is that the file