PANELFIT

PARTICIPATORY APPROACHES TO A NEW ETHICAL AND LEGAL FRAMEWORK FOR ICT

**Guidelines on Data Protection Ethical and Legal Issues in ICT Research and Innovation.**

**GENERAL INTRODUCTION TO DATA PROTECTION – UNDERSTANDING DATA PROTECTION**

# 1 Understanding data protection: the EU regulation in a nutshell

*Bud P. Bruegger (ULD)*

*Acknowledgements: The author thankfully acknowledges advice, input and feedback on drafts from Jörg Pohle, Harald Zwingelberg, Martin Rost, Iñigo de Miguel Beriain, and Josh Bruegger.*

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document reflects only the author's view and the Agency is not responsible for any use that may be made of the information it contains.*

## 1.1 Data protection in the law

The following attempts to provide a concise introduction to the principles of data protection from a European perspective. The protection of personal data in Europe is a **fundamental right** as stated by Article 8 of the *Charter of Fundamental Rights of the European Union*[1]. It has been operationalized by *the General Data Protection Regulation* (**GDPR**)[2].

## 1.2 Purpose of this introduction

The GDPR spans some 99 Articles that in turn are usually divided into several paragraphs which again can contain multiple points. In addition, there are 173 Recitals that help in the interpretation of the articles. Of the eleven chapters of the GDPR, the first four are directly relevant to any party who wants to process personal data. Without recitals, they span a total of 43 articles that fill 28 pages of legal text in the official PDF version[3]. It is therefore not surprising that many people who need to comply with the GDPR, but are not versed in the reading and interpretation of legal text, find the learning curve to be rather steep.

The present introduction attempts to ease this difficulty. It does so not only by giving an overview of the most relevant content, but attempts to present the GDPR as a single consistent system. It does not limit itself to stating *what* the requirements are, but proposes a way to also understand *why* each requirement is there and how it is a necessary part in the whole system. It is hoped that this approach does not only help to get a good overview, but beyond that provides a deeper level of understanding. This is hoped to support practitioners when they have to translate abstract requirements into concrete measures or have to decide at what level measures provide sufficient protection and safeguards.

## 1.3 The problem that data protection addresses

In order to present the GDPR as a system, it is assumed that data protection is concerned with a single problem. Evidently, this assumption is not part of the law, nor was the GDPR systematically created to solve a single stated problem. The basic problem that is postulated here may not even find general consensus. Nevertheless, the postulated base problem is suited to explain the GDPR in a systematic way as a single system. This is the only purpose this base problem has in this introduction. Alternative base problems and ways to systematically explain the GDPR may well exist.

No general consensus exists on what problem data protection actually addresses[4]. The thesis of an "influential minority"[5] is that data protection is concerned with power[6]. This

[1] The Charter of Fundamental Rights was ratified on 7 December 2000.

[2] The GDPR went into effect on 25 May 2018.

[3] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN (last visited 7/5/2020)

[4] See pages 104-105 in Pohle, Jörg. (2018). Datenschutz und Technikgestaltung : Geschichte und Theorie des Datenschutzes aus informatischer Sicht und Folgerungen für die Technikgestaltung, Berlin, Germany:

introduction adopts this thesis to address the basic problem. In particular, the basic problem addressed by data protection is thus to limit the power that organizations gain over individuals by processing their personal data[7]. The well-known phrase *knowledge is power* expresses this very idea. Namely, the possession of information about an individual provides power over that person.

Practically, processing such personal information can influence a person's behavior in itself (for example through *chilling effects*[8]), can help to predict a person's behavior, can make it easier to manipulate a person to act a certain way (for example through targeted advertisement), or can in extreme cases even allow to force a person to a certain behavior (for example through blackmail). The Facebook-Cambridge Analytica data scandal illustrates how far power based on personal information can reach in as far as it may threaten the basic values of democracy. The use of personal information by totalitarian surveillance states to exert power over its citizens may be the ultimate illustration of the problem.

Gaining power over individuals through personal information was always possible. In the past, however, the limited technical capabilities have typically restricted who had access to such power[9] and how much information could actually be collected and processed. With the advent of electronic data processing, the situation has drastically changed. Storing, finding, combining, and analyzing data has become ever more inexpensive and accessible to everyone. The advent of personal devices and ubiquitous sensors have drastically increased the ease of collecting personal data. Data protection is the response to the increasing risk to individuals that comes with this situation.

In the same way as data protection legislation can be seen as a remedy for power imbalance between individuals and organizations in the context of data processing, anti-trust legislation can be seen to address power imbalances in the market place[10].

## 1.4 What is the basic structure of the GDPR?

In principle, gaining power over individuals through the processing of their data is undesirable since it can impede the rights and freedoms of individuals. Forbidding the processing of personal data altogether would be excessive, however. In particular, it could infringe on other fundamental rights and freedoms, such as the freedom to

Humboldt-Universität zu Berlin, DOI http://dx.doi.org/10.18452/19136, https://edoc.hu-berlin.de/handle/18452/19886 (last visited 11/03/2020), (in German).

[5] Personal communications with Jörg Pohle.

[6] See for example Austin, Lisa M., Enough About Me: Why Privacy is About Power, Not Consent (or Harm) (January 1, 2014). Forthcoming in Austin Sarat, ed., A World Without Privacy?: What Can/Should Law Do.. Available at SSRN: https://ssrn.com/abstract=2524512.

[7] Note that processing of personal data for private purposes in a household is excluded from data protection (see Article 2(2)(b) GDPR).

[8] A chilling effect is the inhibition or discouragement of the legitimate exercise of a right or freedom due to data processing, such as video surveillance.

[9] Access was for example limited though a high cost.

[10] Reiner Rehak, Was schützt eigentlich der Datenschutz?, presentation at the 35th Chaos Communication Congress (35C3), Leipzig, Germany, 28/12/18, Slide 18, https://mirror.netcologne.de/CCC/congress/2018/slides-pdf/35c3-9733-was_schutzt_eigentlich_der_datenschutz.pdf (last visited 24/04/2020).

conduct business[11]. For this reason, the data protection legislation has to find a balance between all these rights.

Therefore, the GDPR uses the following basic structure:

- Only processing for certain kinds of **purposes** are allowed;
- and then only under certain conditions on how the processing is **implemented**.

This is explained in more detail in the sequel.

In this context, the **implementation of processing** determines among others what data is collected, what human and technical resources (computing hardware and infrastructure) process the data and how (software and procedures) and for how long, and to whom the data are disclosed.

## 1.5 For which purposes is processing allowed?

In principle, the GDPR forbids the processing of personal data, unless it is conducted for **legitimate and lawful purposes**[12].

A **purpose** describes a concrete objective that shall be reached by the processing.

**Legitimate** means compliance with the letter of the law (not limited to the GDPR), the spirit of the law (e.g., without exploiting legal loopholes), the values of society (as for example expressed in the European Charter of Fundamental Rights), and the principles of ethics. In certain areas of research, compliance with ethics may be verified in formal procedures such as approval by a research ethics committee.

**Lawfulness** is defined in Article 6 GDPR. In particular, for processing to be lawful, its purposes must fall into one of six foreseen categories that are called *legal basis*[13]. Controllers are only allowed to process personal data if they can present a valid legal basis.

In terms of the problem addressed by data protection, this means that gaining power over individuals is only then permitted when it serves legitimate purposes of the kinds foreseen in the GDPR.

## 1.6 What are the conditions for the implementation of processing?

Processing of personal data for legitimate and lawful purposes is thus allowed, but only under certain conditions on its implementation. The following describes these conditions in more detail.

The basic rationale of these conditions is to **limit and balance the power** gained by the organization who processes personal data (so called ***controllers***) over the affected individuals (so called ***data subjects***).

As an overview, this is achieved in the following ways:

[11] See Article 16, European Charter of Fundamental Rights.
[12] See Article 5(1)(a) and (b) GDPR.
[13] See Article 6(1) GDPR.

- Accountability of the controller,
- empowerment of data subjects,
- power balance through a supervisory authority,
- restricting the controllers to use the gained power solely for reaching the declared legitimate purposes,
- limitation of the gained power to what is minimally necessary to fulfill the legitimate purposes,
- protection of the data subjects' investments and assets,
- prohibition of processing that fails to be fit for purpose.
- The individual bullet points are discussed in more detail in the sequel.

### 1.6.1 Controllers are fully accountable

A first measure to limit the power of controllers is to hold them fully accountable for the whole processing activity. This is one of the key principles of the GDPR (see Art. 5(2)). It goes beyond just mandating controllers to make their processing **transparent**[14] (to data subjects and supervisory authorities) by obliging controllers to be able to actually **demonstrate compliance** with the GDPR. Evidently, this opens the processing to oversight. Also, it clearly assigns the "burden of proof": It is not the data subjects or supervisory authorities who need to demonstrate a violation of the GDPR; non-transparency that hides non-compliance is in itself a violation.

To practically achieve this, in a first step, the GDPR makes sure that the full **responsibility** is clearly in the hands of the (joint) controller(s) who determine(s) the purposes and means of processing[15]. This is done, for example, by mandating controllers to exercise control over their **employees**[16] and stipulating contracts[17] with possible external computing services (so called ***processors***) that guarantee control up to the right of on-premise audits by the controller[18].

Once the responsibility is clarified, controllers are obliged to be fully **transparent** about the processing. This includes to proactively **inform data subjects** about the existence and major characteristics of the processing[19] and provide other kinds of information upon request[20]. For the latter purpose, controllers usually also have to designate a *Data Protection Officer*[21] whose contact details are part of the mandatory information[22] and who serves as contact point for data subjects[23].

[14] Note that transparency is also a principle of the GDPR as stated in Art. 5(1)(a).
[15] See Art. 4(7) GDPR.
[16] See Art. 29 and 32(4) GDPR.
[17] See Art. 28(3) GDPR.
[18] See Art. 28(3)(h) GDPR.
[19] See Art. 13 and 14 GDPR.
[20] See for example Art. 15 12(3) and 19 GDPR.
[21] See Art. 37 GDPR.
[22] See Art. 13(1)(b) and 14(1)(b) GDPR.
[23] See Art. 38(4) GDPR.

Controllers further have to notify data breaches to both, the competent **supervisory authority**[24] and (if likely exposed to high risk) the data subjects[25]. In addition, for supervisory authorities, controllers have to maintain records of all processing activities that concern personal data[26] and be able to present a *Data Protection Impact Assessment* for processing activities that are likely to result in a high risk to the rights and freedoms of data subjects[27]. The latter is a prime instrument to *demonstrate* compliance with the GDPR.

### 1.6.2 Empowerment of data subjects

Since there is a power imbalance in data processing, the GDPR empowers the weaker party, i.e., the data subjects. This transforms data subjects from powerless observers of processing to stakeholders who can defend their rights and freedoms through intervention.

The GDPR empowers data subjects mostly through so-called ***data subject rights***[28]. They include the following[29]:

- The *right of access*[30] to the data about the data subjects that is processed,
- the *right to rectification*[31] that permits to correct inaccurate personal data and supplement incomplete data,
- the *right to erasure*[32] that is also called the *right to be forgotten*,
- the *right to restriction of processing*[33] that permits data subjects to demand the suspension of processing of their data in certain circumstances[34].
- the *right to object*[35] that permits data subjects to demand the termination of processing of their data in certain circumstances.
- the *right not to be subject to a decision based solely on automated processing* which produces legal effects concerning them or similarly significantly affects them[36] which includes the *right to obtain human intervention on the part of the controller*[37].

Beyond these rights, data subjects also have:

[24] See Art. 33 GDPR.

[25] See Art. 34 GDPR.

[26] See Art. 30 GDPR.

[27] See Art. 35 GDPR.

[28] See Chapter 3 GDPR that comprises Articles 12 through 23.

[29] Note that the right to data portability is discussed in the section on the protection of the data subject's assets.

[30] See Art. 15 GDPR.

[31] See Art. 16 GDPR.

[32] See Art. 17 GDPR.

[33] See Art. 18 GDPR.

[34] These circumstances are listed in Art. 18(1) GDPR.

[35] See Art. 21 GDPR.

[36] See Art. 22 GDPR.

[37] See Art. 22(3) GDPR.

- the *right to withdraw consent at any time*[38] in the case where the legal basis of processing is consent[39],
- the right to be informed by the controller about the propagation of data subject right invocations to all recipients[40].

### 1.6.3 Balancing power through the institution of supervisory authorities

While data subjects are empowered by the above rights, their resources may be insufficient to enforce them. In particular, they may seem unable to make use of their *right to an effective judicial remedy against a controller or processor*[41] on their own. For this reason, the GDPR grants data subjects the ***right to lodge a complaint with a supervisory authority***[42].

In other words, the GDPR provides data subjects with an ally whose power is comparable to or above that of the controller and thus sufficient for enforcing the data subjects' rights.

The GDPR therefore grants according powers to supervisory authorities[43]. These range from investigative powers[44], such as on-premise audits[45], to corrective powers[46], such as imposing administrative fines[47], ordering the suspension of data flows to recipients[48], and banning the processing altogether[49].

### 1.6.4 Restricting the controllers to use the power solely for reaching the declared legitimate purposes

By demonstrating that the purposes are legitimate and lawful, a controller has justified the gain of power that comes with the processing activity. It is evident that using this power for any other purposes would lack justification. In other words, the permission to process is limited to the declared purposes for which the data is collected.

The GDPR calls this principle **"purpose limitation"** (see Art. 5(1)(b)).

The way to technically and organizationally implement this principle is through **separation** of distinct processing activities.

As a second line of defense, even if data from different processing activities came together anyhow, measures such as pseudonymization can render it more difficult to actually combine them by linking data records pertaining to the same person.

[38] See Art. 7(3) GDPR.
[39] See Art. 6(1)(a) and 9(2)(a) GDPR.
[40] See Art. 19 GDPR, second sentence.
[41] See Art. 79 GDPR.
[42] See Art. 77 GDPR.
[43] See Art. 58 GDPR.
[44] See Art. 58(1) GDPR.
[45] See Art. 58(1)(b) and (f) GDPR.
[46] See Art. 58(2) GDPR.
[47] See Art. 58(2)(i) GDPR.
[48] See Art. 58(2)(j) GDPR.
[49] See Art. 58(2)(f) GDPR.

Note that this rule also prevents the **accumulation of power** by combining the data from different processing activities. Such a combination would typically lead to a deeper insight in the life of data subjects, covering more aspects, or in a wider coverage of knowledge comprising a larger number of data subjects. In both cases, it can be argued that the combined power is greater than the sum of its parts.

### 1.6.5 Minimization of power to what is necessary to fulfill the declared purposes

While the demonstration of legitimacy and lawfulness of purposes has justified the processing as such, it has to be implemented in a way to minimize the power gain to what is minimally necessary to fulfill these purposes. This minimization of power concerns the following three aspects:

- Information content of the personal data,
- degree of association of the data with the data subject, and
- limitation of recipients who have access to power.

These are described in further detail in the following.

#### 1.6.5.1 Minimization of information content (i.e., power)

Since knowledge is power, the minimization of power means that the personal data that are collected have to be minimized. Only the data that can be shown to be necessary for fulfilling the declared purposes can be legitimately collected.

The GDPR calls this principle **"data minimization"** (see Art. 5(1)(c)). Specifically, it requires the collected data to be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed". When looking at it over time, it also requires to store the data no longer than necessary for the purposes. In the case of more complex processing with multiple phases, every phase should have only the data that is really necessary and information content shall be reduced between phases.

#### 1.6.5.2 Minimizing the association to the data subject

The ease with which power over the data subject can be exercised depends on the degree to which the data subject can be associated with the data. The strength of the association between data and its data subject should therefore be minimized.

The GDPR distinguishes three kinds of data with different degrees of association:

- Fully identifying data,
- pseudonymized data, and
- anonymized data.

The first permits “**direct identification**”[50] of the data subject by use of “an “**identifier”** such as a name, an identification number, location data, [or] an online identifier”[51]; **pseudonymized data** permits **identification only with the use of “additional**

[50] This term is introduced in Art. 4(1) GDPR.

[51] This wording is extracted from Art. 4(1) GDPR.

**information"**[52]; and **anonymous data** where "**the data subject is not or no longer identifiable**"[53].

In analogy to data minimization, the data shall be collected with the minimal degree of association with the data subject. Considering the temporal aspect, "personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes"[54]. In the case of more complex processing with multiple phases, every phase should have only the minimal degree of association that is really necessary and pseudonymization or anonymization should be used between phases.

The GDPR calls this principle **"storage limitation"** (see Art. 5(1)(e)).

#### 1.6.5.3 Limitation of the access to power

Power is in the hands of persons and organizations. If knowledge is power, this power is available only to parties to whom the personal data is disclosed. The GDPR calls such parties *recipients*[55]. They can be either employees of the controller or processor, intended third-party recipients, or unintended parties such as attackers.

The access to power has to be limited to what is necessary to fulfill the declared purposes. The GDPR calls this principle **"confidentiality"**[56].

Confidentiality has two aspects:

- Preventing access by unauthorized parties, and
- restricting access by authorized parties.

The former protects to a large degree against external attackers with measures such as encryption of data at rest or communications and firewalls. The latter is usually called **access control**. It makes sure that the party accessing the data is indeed authorized (authentication), restricts the access to data that is needed (access rights) and may restrict access to the times when it is necessary.

### 1.6.6 Protection of the data subject's assets

In many kinds of processing activities, the personal data stored by the controller is also of significant value to the data subject. Prime examples are cloud-based photo collections and office suites and document management systems but also medical data residing with a patient's physician. We call such data *assets*.

These assets may be of much lower value to the controller who may be reluctant to investing significantly in their protection. Also, one way a controller can exert power over a data subject is to make access to a data subject's assets dependent on certain conditions.

To prevent such exertion of power, the GDPR mandates controllers to protect data subjects' assets. In particular, it requires to protect these assets against:

[52] Note that this term is used in Art. 4(5) GDPR that provides the definition for *pseudonymization*.
[53] This wording is extracted from the 5th sentence of Recital 26 GDPR.
[54] This wording is extracted from Art. 5(1)(e) GDPR.
[55] See Art. 4(9).
[56] See Art. 5(1)(f).

- accidental loss, destruction or damage[57], and
- refusal to let the data subject use the assets independently of the controller.

The former kind or protection is also known as **availability** and **resilience**[58]. The latter is called **data portability** and is one of the data subject's rights[59].

### 1.6.7 Prohibition of processing that fails to be fit for purpose

Gaining power through any processing that is unfit to fulfill the declared purposes is evidently illegitimate.

The GDPR uses two principles to enforce fitness for purpose:

- **Integrity** (see Art. 5(1)(f)) and
- **accuracy** (see Art. 5(1)(d)).

The former mandates to protect data against accidental damage and unauthorized modification; the latter mandates that data are kept up to date and accurate and that where this is not the case the data are erased or rectified without delay.

## 1.7 The notion of risk

*Risk* is an important concept in the GDPR[60]. The presented view that data protection is about mending the power imbalance between controller and data subject clarifies also the notion of risk:

The main risk is that the processing of personal data indeed results in a power imbalance that restricts the rights and freedoms of the affected individuals. From this point of view, it becomes clear that the risk is not that some undesirable event occurs (such as an attack or a natural disaster), but rather that the controller exerts excessive power over data subjects.

Note that this understanding of risk is very different from risk in cybersecurity. There, the controller is typically seen as the "good guy" defending against predominantly external "attacks". In data protection in contrast, the controller’s behavior, i.e., the processing activity, is the source of risk. The likelihood that this occurs is 100%. Unlike in cybersecurity, controllers now have to protect the weaker data subject from risk resulting from their own processing. Controller are thus no longer automatically the good guys, but have to make explicit efforts to not become bad guys themselves.

For people mostly familiar with cybersecurity, understanding data protection may require a significant mental shift. Understanding this difference is a pre-requisite to being able to comply with the GDPR. For further reading we recommend an article[61] about eight different types of risk.

[57] See Art. 5(1)(f) GDPR.

[58] See Art. 32(1)(b) and (c) GDPR.

[59] See Art. 20 GDPR.

[60] See for example Art. 24(1), 35(1) and Recitals 75 and 84.

[61] Martin Rost, Risks in the context of data protection, http://www.maroki.de/pub/privacy/Rost_Martin_2019-02_Risk:_8types_v1.pdf (last visited 8/5/2020).