D3.1 General Legal Aspects

Gerald Spindler (GU), Anna Zsófia Horváth (GU), Lukas Dalby (GU)

The project SODA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731583.

Project Information

Scalable Oblivious Data Analytics

Project number:	731583
Strategic objective:	H2020-ICT-2016-1
Starting date:	2017-01-01
Ending data:	2019-12-31
Website:	https://soda-project.eu/

Document Information

Title:	General Legal Aspects				
ID:	D3.1.	Type:	R	Dissemination level:	PU
		Month:	M12	Release date:	December 30, 2017

Contributors, Editor & Reviewer Information

Contributors (person/partner)	Gerald Spindler (GU) Anna Zsófia Horváth (GU) Lukas Dalby (GU)
Editor (person/partner)	Anna Zsófia Horváth (GU)
Reviewer (person/partner)	Claudio Orlandi (AU)

Release History

Release number	Date issued	Release description / changes made
1.0	December 30, 2017	First release to EU

SODA Consortium

Full Name	Abbreviated Name	Country
Philips Electronics Nederland B.V.	PHI	Netherlands
Alexandra Institute	ALX	Denmark
Aarhus University	AU	Denmark
Göttingen University	GU	Germany
Eindhoven University of Technology	TUE	Netherlands

Table 1: Consortium Members

Executive Summary

Utilizing Big Data could substantially reduce costs and improve efficiency in the healthcare sector. The main purpose of the SODA project is to develop privacy-preserving technologies for large scale data analysis, that can contribute to a more effective research of health data and at the same time, helps to encourage individuals to provide their data for such processing. To achieve this, new cryptographic techniques offering secure processing with de-identified data are being developed within the frames of SODA.

To ensure the compliance of these new technologies with European data protection law, as the first deliverable of the WP3 legal section, the present SODA Deliverable 3.1 intends to provide a thorough legal analysis of the current privacy law in the EU, with emphasis placed on the GDPR, which shall apply from the 25th May 2018. In doing so, we put special emphasis on regulations related to the utilisation of Big Data, as well as on the genuine conflict between Big Data and privacy.

We target the relevant legal issues in the four main sections of the Deliverable. In the first section (section 2.2), we assess what anonymous data mean under the GDPR, and whether encrypted or otherwise de-identified data can be treated as anonymous data. This overview includes a detailed discussion on the latest developments of the legal evaluation of privacy preserving technologies. In the second section (section 3.1), we continue by delineating the conditions for the lawfulness of processing of personal data as well as the special conditions of the processing of special categories of personal data e.g. health, genetic or biometric data. Here we discuss here the specific rules applying for research as well. In the third section (section 3.2), we outline the concepts of data controller and processor, and address the responsibilities and obligations placed on them by the GDPR. The rights of the data subject are discussed together with the provisions as well as in the final section (section 3.3).

About this Document

Role of the deliverable

This deliverable aims to give a thorough legal analysis of the current privacy law in the EU related to Big Data analytics, with emphasis placed on the GDPR, which shall apply from the 25th May 2018.

Relationship to other SODA deliverables

Deliverable D3.5 will contain a legal analysis of the use cases based on the general framework described in this deliverable.

Relationship to other versions of this deliverable

This deliverable only has a single version.

Structure of this document

Section 2 describes the relation between Big Data and the European Data Protection Law. Section 3 discusses the requirements for the lawful processing of personal data.

1 Table of Contents

2 Big Data and the European Data Protection Law

2.1 Introduction

With multiple terabyte information generated every single day[1], we have undoubtedly reached the era of Big Data. This vast amount of available information represents a huge economical value on a global scale, triggering continuous innovation and productivity. However, inherent risks also come along with the big benefits. Recent technological developments have reduced significantly the time and cost required to collect and analyse data, thereby causing a strong incentive to gather data whenever there's a chance to, even without a clear vision of utilization.

Hence it can be particularly difficult to achieve the required compliance with the data protection law whilst conducting a Big Data analysis. Chapter 2 of this deliverable aims to outline the core relations of the European data protection law to Big Data *de lege lata* as well as *de lege ferenda.*

2.2 Legal Framework

In order to fully comprehend the complexity of the problems arising from the application of Big Data, light must be shed on the underlying legal framework. To solve the puzzle, it is inevitable to discuss the core functioning and key definitions and principles of the existing as well as the upcoming data protection system of the EU.

The current data protection legal framework roots in early documents.[2] The European Convention of Human Rights introduces the protection of private and family life, home and correspondence. However, with the evolution of technology being a powerful driving force behind the interpretation of privacy law and data protection, these minimum requirements have become outdated over time. As the protection of personal data began to be understood as a separate fundamental right, different from the right to privacy, the right of information self-determination was born, i.e., the right to have a say in how and to which extent data relating to oneself are processed.[3]

This chapter mainly focuses on data protection in this strict sense, and evaluates the latest developments related to data protection law in the European Union.

2.2.1 The Data Protection Directive 95/46/EC (DPD)

The Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data (hereafter: the Directive)[4] was adopted in 1995 and has served as statutory framework for the processing of personal data within the European Union in the last two decades. As a directive, it is part of the secondary sources of the EU law. The definition, and scope of a directive according to Article 288 of the Treaty on the Functioning of the European Union are as follows:[5]

[1] The digital universe of opportunities: Rich data and increasing value of the internet of things" Veron Turner, John F. Gantz, David Reinseland, Stephan Minton, Report from IDC for EMC April 2014.

[2] e. g. The European Convention of Human Rights (ECHR), adopted 04.11.1950, available at: http://www.echr.coe.int/Documents/Convention_ENG.pdf ; OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, adopted 23.09.1980, updated in 2013, available at: http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf ; Convention on the Protection of Individuals with regard to Automatic Processing of Personal data adopted 28.01.1981, available at: https://rm.coe.int/1680078b37 ; Charter on Fundamental Rights of the European Union,

[3] Kuner, European Data Protection Law: Corporate Regulation and Compliance, Second edition, 2007, p. 3

[4] Directive 95/46/EC of the European Parliament and of Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movements of such data, Official Journal of the European Communities, L 281, pp. 31-50

[5] Treaty on the Functioning of the European Union, Official Journal of the European Union, C 202, pp. 47-200

> "Article 288 (ex Article 249 TEC): […] shall be binding, as to the result to be achieved, upon each Member State to which it is addressed, but shall leave to the national authorities the choice of form and methods […]"

This essentially means that each directive defines its goals, and the Member States are obliged to accomplish these within a given period of time. This implementation period provided by the Directive was three years from the date of adoption. Since directives are generally, but not directly applicable, they create a legislative obligation for Member States to implementation by incorporating them into national law. If necessary, Member States have to adapt their national laws in order to provide the most efficient implementation (*effet utile*), however, they are given some leeway as how to do so.

Directives weren't originally designed for giving right to the individuals directly, but it was later established by the Court of Justice (hereafter: ECJ), that vertical direct effect (i.e. against the state) on directives does exist. However, only if they state rights for citizens and failed to be transposed adequately.[6]

Laying down only an intended outcome without imposing specific form and methods, directives are appropriate legal acts to approximate different national legal systems, with varying degrees of harmonization.

Concerning the Directive, it goes substantially beyond a mere *de minimis* harmonization level, as it was also underlined by the ECJ.[7] It aims at nothing less than a fully harmonized data protection regime in the EU, where the same level of protection has to be ensured in all Member States. Thus, the Directive intends to ensure the free movement of personal data while guaranteeing a high level of protection for the rights and interests of the individuals.[8]

2.2.1.1 Territorial Scope of the Directive

When it comes to Big Data analyses, multiple undertakings are generally involved in the processing. In order for all affected parties to be able to fully comply, the scope of application must be clarified. The Directive contains several provisions addressing applicable law issues, first and foremost Article 4. This defines the territorial scope of the Directive, which determines which national data protection system may be applicable to the processing of personal data.

According to Art. 4 (1) a) the Directive applies to processing of personal data if it is carried out in the context of the activities of an establishment of the controller on the territory of the Member State. This means that organizations are subject to EU data protection law if they have an establishment in the EU and process data in the context of the activities of that establishment – regardless of whether or not the actual processing takes places in the EU. In two of its recent decisions[9], the ECJ provided an in-depth explanation on both of these criteria.

The underlying problem related to the condition of having an establishment of the controller on the territory of the Member State is, that the word "establishment" is not precisely defined. Recital 19 of the Directive states that an establishment implies "the effective and real exercise of activity through stable arrangements". The decisive factor is thereby, whether or not there is effective and real exercise of activity, and not the legal form of the establishment as such. Further indication can be found under the Art. 50 TFEU (former Art. 43 TEC) on the freedom of establishment and particularly in its interpretation by the ECJ. According to this, the establishment "is of a certain minimum size and both human and technical resources necessary for the provision of the services are permanently present".[10]

---

[6] ECJ, decision of 4/12/1974 – C 41/74 Yvonne Van Duyn v. Home Office

[7] ECJ, decision of 24/11/2011 – C-468/10 ASNEF/ FECEMD v. Administration del Estado, paragraph 29

[8] ECJ, decision of 6/11/2003 – C-101/01, Göta Hovrätt v. Lindqvist, paragraph 96, upheld by ECJ decision of 24/11/2011 – C-468/10

[9] ECJ, decision of 13/05/2014 – C-131/12 Google Spain SL/ Google Inc. v. AEPD/Mario Costeja Gonzales; ECJ, decision of 1/10/2015 – C-230/14 Weltimmo s. r. o. v. Nemzeti Adatvédelmi és Információszabadság Hatóság

[10] ECJ, decision of 4/06/1985 – Case 168/84 Berkholz v. Finanzamt Hamburg-Mitte-Altstadt, paragraph 18

The question of what exactly qualifies as "effective and real exercise of activity" was one of the key issues addressed in the Weltimmo case. The ECJ followed the observations of the Advocate General on a flexible definition by stating:

> "[…]in order to establish whether a company, the data controller, has an establishment, within the meaning of Directive 95/46, in a Member State other than the Member State or third country where it is registered, both the degree of stability of the arrangements and the effective exercise of activities in that other Member State must be interpreted in the light of the specific nature of the economic activities and the provision of services concerned. This is particularly true for undertakings offering services exclusively over the Internet.
>
> In addition, in order to attain that objective, it should be considered that the concept of 'establishment', within the meaning of Directive 95/46, extends to any real and effective activity — even a minimal one — exercised through stable arrangements.
>
> [...] Article 4(1)(a) of Directive 95/46 must be interpreted as permitting the application of the law on the protection of personal data of a Member State other than the Member State in which the controller with respect to the processing of those data is registered, in so far as that controller exercises, through stable arrangements in the territory of that Member State, a real and effective activity — even a minimal one — in the context of which that processing is carried out;
>
> in order to ascertain, in circumstances such as those at issue in the main proceedings, whether that is the case, the referring court may, in particular, take account of the fact (i) that the activity of the controller in respect of that processing, in the context of which that processing takes place, consists of the running of property dealing websites concerning properties situated in the territory of that Member State and written in that Member State's language and that it is, as a consequence, mainly or entirely directed at that Member State [...]"[11]

With the rejection of a restrictive interpretation of the term "establishment", the Court departed from a mere formalistic approach.[12] This resulted in a potentially broader territorial application of the Directive, since it not only considers where the establishment of the controller in question is registered, but also evaluates in which Member States it performs its actual economic activity. In other words, if a controller engages in "real and effective activity" in a Member State other than where it is registered, it will presumably be subject to the data protection law of that Member State.

In its Google Spain case the Court shed light on the phrase "context of the activities", the second condition triggering the applicability of the Directive. In relation to a search engine, operated by Google, conducting data processes related to data subjects within the EU, prior to the Court's decision, the Advocate General presented in his opinion firstly, that a subsidiary established by a company in any Member State undoubtedly constitutes an establishment according to Art. 4 (1) a) of the Directive.[13] Secondly, that if this subsidiary, under certain coordination of the parent company, carries out data processing, and in doing so, serves as a "bridge for the referencing service" to the market of the Member State, the processing of personal data takes place in the context of the controller's establishment, regardless of where the technical data processing operation is situated. He proposed, that the Court should conclude that:

[11] ECJ, decision of 1/10/2015 – C-230/14 Weltimmo s. r. o. v. Nemzeti Adatvédelmi és Információszabadság Hatóság, paragraphs 29, 31, 41

[12] ECJ, decision of 1/10/2015 – C-230/14 Weltimmo s. r. o. v. Nemzeti Adatvédelmi és Információszabadság Hatóság, paragraph 25; Revolidis, ZD-Aktuell 2016, 05016; Karg, ZD 2015, 580

[13] Opinion of Advocate General Jääskinen, delivered on 25/06/2013, Case C-131/12 – Google Spain SL/ Google Inc. v. AEPD/Mario Costeja Gonzales, paragraph 64.

> "[…] processing of personal data is carried out in the context of the activities of an 'establishment' of the controller within the meaning of Article 4 (1) a) of the Directive when the undertaking providing the search engine sets up in a Member State for the purpose of promoting and selling advertising space on the search engine, an office or subsidiary which orientates its activity towards the inhabitants of that Member State."[14]

The Court has opted for an approach similar to the Advocate General's arguments. It stated that the processing does not necessary have to be performed by the establishment itself, but only in the context of the activities of the establishment. The reason behind the notion of an extensive interpretation is, that "it cannot be accepted, that the processing of personal data carried out […] should escape the obligations and guarantees laid down by the Directive 95/46, which would compromise the Directive's effectiveness and the effective and complete protection of the fundamental rights and freedoms of natural persons which the Directive seeks to ensure".[15]

It is to be noted, that in this decision, the Court proposed a new indicator to decide on whether or not the data processing is carried out in the context of the activities of the controller. The EU data protection law can be triggered even in cases, where the establishment does not have any role in the data processing whatsoever:

> "[…]in the light of the inextricable link between the activity of the search engine operated by Google Inc. and the activity of Google Spain, the latter must be regarded as an establishment of the former and the processing of personal data is carried out in context of the activities of that establishment. [...].
>
> In such circumstances, the activities of the operator of the search engine and those of its establishment situated in the Member State concerned are inextricably linked since the activities relating to the advertising space constitute the means of rendering the search engine at issue economically profitable and that engine is, at the same time, the means enabling those activities to be performed."[16]

As a result, the possibility rises that if a case-by-case assessment reveals an inextricable link between the activities of an establishment in the EU and the processing of personal data carried out by a controller in a third country, the latter will be subject to the Directive, regardless whether the establishment took part in the processing directly or not. It results in a significantly broad territorial reach of the Directive.[17]

Article 4 (1) b) addresses a less common case, in which the national law of the Member State is triggered by virtue of international public law. In these cases international public law may determine the criteria for the extension of the application of national data protection law beyond the national boundaries.

Article 4 (1) c) serves as a residual and limited criterion, as it becomes relevant only when a controller has no presence, or at least no establishment with relevant activities in the territory of the EU. In accordance with Recital 20 it states that when a non-EU controller "for purposes of processing personal data makes use of an equipment situated on the territory of a Member State" the national provisions of that Member State shall apply, "unless such equipment is only used for purposes of transit". This essentially means that EU data protection law can be triggered even when merely substantial processing

[14] Opinion of Advocate General Jääskinen, delivered on 25/06/2013, Case C-131/12 – Google Spain SL/ Google Inc. v. AEPD/Mario Costeja Gonzales, paragraph 68

[15] ECJ, decision of 13/05/2014 – C-131/12 – Google Spain SL/ Google Inc. v. AEPD/Mario Costeja Gonzales, paragraph 53

[16] ECJ, decision of 13/05/2014 – C-131/12 – Google Spain SL/ Google Inc. v. AEPD/Mario Costeja Gonzales, paragraphs 47, 56

[17] Art.-29 Working Party Update of Opinion 8/2010, WP 179 update, p. 5; Kartheuser/Schmidt, ZD 2016, 155 (156)

infrastructure is located in the EU. The Art.-29 Working Party suggests a broad interpretation of the word "equipment", which should therefore include human as well as technical intermediaries. However, this approach might have unwelcome consequences, e. g. if controllers established in multiple countries conduct analysis on the very same dataset based on personal data collected in one of the Member States, those controllers have to comply with the data protection law of that Member State.[18]

2.2.1.2 Objectives and Material Scope of the Directive

It is crucial to understand the subject matter of the Directive in order to determine whether European data protection law applies to a specific data processing activity or not.

2.2.1.2.1 Objectives of the Directive

Art. 1 of the Directive sums up the aim and objectives of the Directive:

> "1. […] Member States shall protect the fundamental rights and freedoms of natural persons, and in particular their right to privacy with respect to the processing of personal data.
>
> 2. Member States shall neither restrict nor prohibit the free flow of personal data between Member States for reasons connected with the protection afforded under paragraph 1."

This provision mirrors a basic trade-off which is so inherent in data protection law. This is the need for weighing up of interest of the various parties involved, namely the stakeholders and market participants on the one side, and individuals and their fundamental rights on the other. EU data protection law aims to govern the processing of personal data by preventing the balance winging away from one in favour of the other.[19] With that the Directive intends to contribute to the economic and social progress and trade in the Union[20].

2.2.1.2.2 Material Scope of the Directive

The material scope of application is laid down in Art. 3 (1) of the Directive:

> "[…] shall apply to the processing of personal data wholly or partly by automatic means, and to the processing otherwise than by automatic means of personal data which form part of a filing system or are intended to form part of a filing system."

Having said that, the Directive protects the processing of personal data of natural persons, whether the processing is carried out by automatic means or not.[21] Personal data mean "any information relating to an identified or identifiable natural person", as outlined in Art. 2 lit a). Since the right to protection of personal data is a human right, not only EU citizens are entitled to it, but virtually anybody whose data is being processed.[22] Corporate enterprises, and thus the protection of confidential business information and trade secret as such are specifically excluded from the scope of the Directive. The definition of "processing" according to Art. 2 lit b) is just as wide, and it covers all kinds of operation or set of operation which is performed upon personal data.

On the other hand, Art. 3 (2) offers some exceptions. According to this, activities which "fall outside the scope of the Community law" as well as activities of a Member State in areas of criminal law fall

[18] Art.-29 Working Party Opinion 8/2010, WP 179, p. 20
[19] Ernst, in Paal/Pauly, Datenschutz-Grundverordnung, Art. 1. Rn. 1; Recitals 3, 8 of the Directive
[20] Recital 1 of the Directive
[21] Note, however, that solely the requirements for automatic/automated means of processing is discussed within the frames of this paper, due to the attributes of Big Data
[22] Ernst, in Paal/Pauly (supra note 19), Art. 1. Rn. 7; Recital 2 of the Directive

outside the scope of the data protection law[23]. Probably more relevant for online services is the third group of exceptions, the so called "household-exceptions", which refer to activities by natural person exclusively in the course of "purely personal or household" purposes.

2.2.1.3 Core Principles of the Directive

Art. 6 of the Directive lists the foundational principles relating to data quality. These principles serve as benchmark to controllers and processors when working with personal data. These are transparency, legitimate purpose, and principles related to proportionality. One of the most important foundations of data protection law is however not defined in this paragraph, namely the principle of prohibition with the reservation of authorisation.

2.2.1.3.1 Principle of Prohibition with the Reservation of Authorisation

According to this principle, the processing of personal data is not permitted by principle, except when there is a legitimate ground for processing.[24] Art. 7 contains the catalogue of criteria that make data processing legitimate. Processing of personal data is lawful only if the controller ensures the compliance with at least one of the legal bases given in Art. 7. Since in any other case, the processing activity is prima facie unlawful, this authorisation serves as a precondition for compliance.

Art. 8 imposes stricter conditions for the processing of special categories of data, that is to say, data "revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life". Art. 8 (1) explicitly prohibits the processing such data, unless one of the conditions offered in Art. 8 (2) serve as ground for a lawful processing. Most important conditions are the explicit consent as well as the vital interests of the data subject.

2.2.1.3.2 Transparency

Art. 6 (1) a) states that personal data must be processed fairly and lawfully, however, it does not mention transparency itself. This requirement is extensive. It covers the responsibility to process data only on a lawful basis, as well as the obligation of the controller to provide the necessary information to the data subjects[25]. Art. 10 and 11 regulate the information to be given to the data subject in cases of collection of data from the data subject and from another source, respectively. In both cases the controller must provide accurate and full information on its identity and the purposes of the processing, before starting the actual processing.

2.2.1.3.3 Legitimate Purpose

The concept of purpose limitation takes the right of informational self-determination into account and intends to make it possible for the affected individual to understand how widely its data is being collected.[26] It is a key principle to preserve trust and legal certainty. The specification of the purpose is a prerequisite for applying other data quality requirements.[27]

---

[23] Recital 13, 16 of the Directive
[24] Recital 30 of the Directive
[25] Recital 38 of the Directive
[26] Eskens, Profiling the European Citizen in the Internet of Things: How will the General Data Protection Regulation Apply for this Form of Personal Data Processing, and How Should it? University of Amsterdam, Institute for Information Law (IViR) 22.03.2016, pp. 50, 51, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2752010
[27] Art.-29 Working Party, Opinion 03/2013, WP 203, p. 4

According to Art. 6 (1) b) in line with Recital 28, data must be collected "for specified, explicit and legitimate purposes" and the data collected for one purpose shall not be used for new, incompatible purposes.

2.2.1.3.4 Proportionality

The principles outlined by Art. 6 (1) c)-e) can be derived from the two principles described above. Data minimisation, accuracy and rules on data retention periods are closely linked, as they all ensure the proportionality of the data processing. A processing activity is proportional, if it is suitable and necessary to achieve the specific purpose, and is reasonable, considering the different interests of the affected parties.

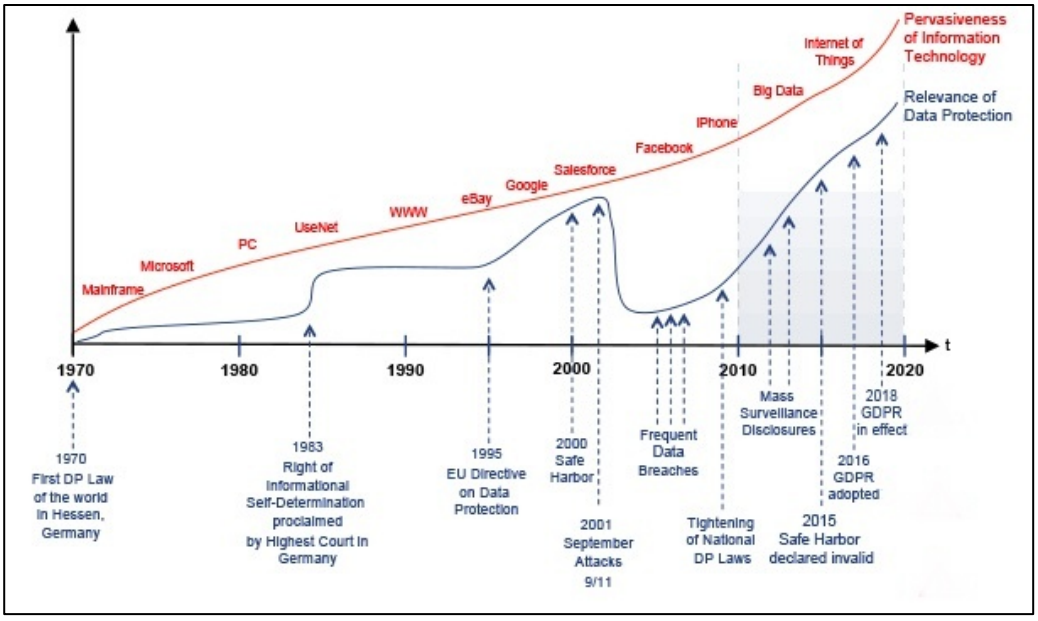
Data minimisation, as stated in Art. 6 (1) c) means that data must be "relevant and not excessive in relation to the purposes". This principle implies that only those data shall be processed, which are actually needed in order to achieve the processing purposes.

The principle of accuracy shall prevent the risks arising from processing inaccurate data. Therefore Art. 6 (1) d) obligates controllers to keep the data up to date, and to take "every reasonable step" in order to further process, erase or rectify the inaccurate or incomplete data.

Art. 6 (1) e) outlines the principle of data retention periods. The underlying idea is that personal data shall not be retained for longer than necessary in connection with the purposes for which they were originally collected, or for which they are further processed. Appropriate safeguards have to be laid down for personal data stored for longer periods for historical, statistical or scientific use.

2.2.2 The General Data Protection Regulation (EU) 2016/679 (GDPR)

Significant changes have occurred in the ways in which information has been collected and processed since the Directive was drafted. Many tools that are widespread today, were unknown at the time. The interconnected world, within which Big Data as such arose, hardly existed, let alone served as basis for many businesses and researches. In addition, the similar but not identical implementation of the Directive by the Member States resulted in a rather divergent data protection law with a whole range of compliance requirements across the EU.[28] For these reasons, the Directive clearly needed to be revised.



[28] Kühling/Martini, EuZW 2016, 448 (448)

Figure 1. Evolution of European Data Protection Regime[29]

After years of negotiations[30] the European Parliament adopted the General Data Protection Regulation[31] (hereafter: GDPR) on 14th April 2016. It was published in the official journal of the European Union on 04th May 2016, and entered in force 20 days later.[32] The enforcement of the GDPR will only begin on 25th of 2018, this transitional period should provide enough time for the affected companies to ensure compliance with the requirements of the GDPR. Among the final provisions in Art. 94 (1) the GDPR declares that the Directive is repealed with effect from 25th May 2018. On the other hand, Art. 94 (2) states that references to the repealed Directive "shall be construed as references to this Regulations", furthermore "references to the Working Party on the Protection of Individuals with regard to the Processing of Personal Data established by Art. 29 of Directive 95/46/EC shall be construed as references to the European Data Protection Board established by this Regulation". According to these transitional arrangements the guidelines set out by the Art.-29 Working party will continue to apply under the GDPR.[33]

The most important difference between a directive and a regulation is, that regulations have a direct binding legal force. They must be applied in their entirety across the EU without transposition and implementation by the Member States. As Art. 288 (2) TFEU states:

> "A regulation shall have general application. It shall be binding in its entirety and directly applicable in all Member States."

This makes a regulation an ideal tool for a full harmonisation of the data protection law in the EU, which is exactly what the GDPR aims to achieve.[34] However, the GDPR includes many opening clauses, which give Member States discretion to modify, concretize or even restrict the Article in which the clause resides via national legislation.[35]

2.2.2.1 Territorial Scope of the GDPR

After witnessing the tendency to expand the territorial reach of the Directive, it can hardly be a surprise, that the transition to the GDPR introduces significantly broader territorial application of the EU data protection law.[36]

Art. 3 and the corresponding Recitals 22-24 define the rules on applicability of the GDPR. Art. 3 offers three alternatives of application, namely for processing of personal data by establishments within

[29] source: Wilhelm, A brief history of the General Data Protection Regulation, available at: https://iapp.org/resources/article/a-brief-history-of-the-general-data-protection-regulation/
[30] Albrecht, CR 2016, 88 (89)
[31] 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), Official Journal of the European Union, L 119, pp 1-88.
[32] Art. 99 (1) GDPR
[33] Jenny, in Plath, BDSG-DSGVO Kommentar, Art. 94, Rn. 2; Pauly, in Paal/Pauly (supra note 19) Art. 94 Rn. 3; Kühling/Raab, in in Kühling/Buchner Datenschutz-Grundverordnung Kommentar, Art. 94 Rn. 2 ff.
[34] Eckhard/Kramer/Mester, DuD 2013, 623 (630)
[35] Laue, ZD 2016, 463 (463); e. g. new German Federal Data Protection Act, Gesetz zur Anpassung des Datenschutzrechts an die Verordnung (EU) 2016/679 und zur Umsetzung der Richtlinie (EU) 2016/680 – Datenschutz-Anpassungs- und –Umsetzungsgesetz EU (DSAnpUG-EU), 30 June 2017.
[36] Plath, in Plath (supra note 33) Art. 3 Rn. 2, 3

the EU, as well as by non-EU establishments if data subjects within the EU are affected, lastly by virtue of public international law.[37]

Art. 3 (1) echoes the provisions of Art. 4 (1) a) of the Directive:

> "This Regulation applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not."

According to this first option the GDPR applies when there is a processing "in the context of the activities" of an establishment within the EU, disregarding the actual location of this processing. This means that the criteria laid down in the Directive stay intact, and the decisive questions are whether or not there is an establishment, and if the reply to this question is affirmative, whether or not the specific processing takes place "in the context of the activities" of that establishment. Somewhat remarkable, the GDPR leaves the word "establishment" undefined, and merely reiterates the phrase known from the Directive, that an establishment implies "the effective and real exercise of activity through stable arrangements".[38] As discussed above (see 2.2.1.1) the ECJ has dealt with the territorial scope of the European data protection law in two of its latest decisions, and explicitly stated that the provisions of Art. 4 (1) of the Directive shall "not be interpreted restrictively".[39] Although these decisions were made prior the GDPR entered in force, it shall be assumed, that the new approach outlined by the ECJ will by analogy hold for the GDPR.[40]

The GDPR addresses the processors as well, therefore when a processor within the EU processes personal data for a controller outside the EU, Art. 3 (1) only applies to the activity of the processor (and not for the controller outside the EU), since a processor may not be considered as an establishment of the controller.[41] Consequently, the processor has the obligation to comply with the GDPR in each phase of the processing activity. In a reverse situation, when an EU-based controller assigns processing of personal data to processors outside of the EU, only the controller will be subject to the GDPR, unless of course the processor falls under the territorial scope of the GDPR on the basis of Art. 3 (2).

Moreover, one of the most significant reform of the GDPR is the introduction of the market principle in Art. 3 (2)[42]:

> "This Regulation applies to the processing of personal data of data subjects who are in the Union by a controller or processor not established in the Union, where the processing activities are related to:
>
> (a) the offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union; or
>
> (b) the monitoring of their behaviour as far as their behaviour takes place within the Union."

With this provision, the GDPR abandons the Directive's approach of requiring some sort of direct connection with the EU through an establishment and/or use of equipment, and focuses on whether or not

[37] Art. 3 (3) is not amended by the GDPR. Circumstances in which the applicability of the Regulation is defined by virtue of international public law are extremely rare in relation to a Big Data analysis, and so this issue is unlikely to affect it.

[38] Recital 22, GDPR

[39] ECJ, decision of 13/05/2014 – C-131/12 – Google Spain SL/ Google Inc. v. AEPD/Mario Costeja Gonzales, paragraph 53, upheld by ECJ, decision of 1/10/2015 – C-230/14 Weltimmo s. r. o. v. Nemzeti Adatvédelmi és Információszabadság Hatóság, paragraph 25

[40] Since the GDPR does not amend the essential provisions of Art. 4 (1) a) of the Directive, Klar, in Kühling/Buchner (supra note 33), Art. 3 Rn. 2; Albrecht, CR 2016, 88 (90)

[41] Klar, in Kühling/Buchner (supra note 33), Art. 3. Rn. 38

[42] Kühling/Martini, EuZW 2016, 448 (450)

a company offers its products in the EU single market. Thus, the applicability of the GDPR depends on the relevant targeting of individuals.[43] This "service oriented approach" was previously offered by the Art.-29 Working Party as an additional criterion for processing when the controller is located outside the EU.[44] It is important to note however, that the way to Art. 3 (2) is only opened, if there is no relevant establishment of the controller or processor in the Member States whatsoever.

The first alternative set out in Art. 3 (2) extends the territorial scope of the GDPR to processing of personal data to non-EU based controllers and processors when they carry out processing activities related to offering goods or services affecting data subjects within the EU[45]. This provision is to be interpreted in the light of Recital 23:

> "In order to ensure that natural persons are not deprived of the protection to which they are entitled under this Regulation, the processing of personal data of data subjects who are in the Union by a controller or a processor not established in the Union should be subject to this Regulation where the processing activities are related to offering goods or services to such data subjects irrespective of whether connected to a payment. In order to determine whether such a controller or processor is offering goods or services to data subjects who are in the Union, it should be ascertained whether it is apparent that the controller or processor envisages offering services to data subjects in one or more Member States in the Union. Whereas the mere accessibility of the controller's, processor's or an intermediary's website in the Union, of an email address or of other contact details, or the use of a language generally used in the third country where the controller is established, is insufficient to ascertain such intention, factors such as the use of a language or a currency generally used in one or more Member States with the possibility of ordering goods and services in that other language, or the mentioning of customers or users who are in the Union, may make it apparent that the controller envisages offering goods or services to data subjects in the Union."

It should be highlighted that in order to be subject to this provision, there has to be an intentional offer to data subjects in the Union, that is, the willingness to engage in business must be "apparent". Recital 23 contains an indicative list of examples to help companies to comply. However, other apparent means of envisaging this will have to be considered, such as using a Top-Level-Domain.[46] This virtually means that if a non-EU based organisation uses a website where the user interface is offered in local language, possibly but not necessarily with local Top-Level-Domain to collect data from data subjects within he Union, this organisation will be subject to the GDPR in the course of this processing, even if it has no operations on the territory of the EU.

Art. 3 (2) b) introduces the possibility for extraterritorial application of the GDPR for cases, where non-EU based controllers and processors process data in order to monitor the behaviour of data subjects, if this behaviour takes place in the EU. This means that the applicability of this provision is tied to the observed behaviour and not to the residence or citizenship of the affected data subjects. It does not play any role either, whether the activity of a website is oriented towards the EU or not.[47] The question, what

---

[43] Härting, Datenschutz-Grundverordnung, 2016, Otto Schmidt Verlag Köln, p. 58 R. 220,

[44] Art.-29 Working Party Opinion 8/2010, WP 179, p. 31, see also Dammann, ZD 2016, 307 (309)

[45] The GDPR does not specify the attributes of „goods", and „services", which means that it operates with the guiding definitions laid down in Art. 28 and Art. 57 TFEU respectively, available at: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012E/TXT&from=hu ; additionally to the concept of "services" see Art. No. 1 of the Directive 2006/123/EC of the European Parliament and of the Council of 12/12/2006 on services in the internal market, available at: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32006L0123&from=EN ; these definitions are plausible within the meaning of the GDPR, with the modification that Art. 3 (2) a) of the GDPR includes goods, and services in return of payment as well as free of charge.

[46] Klar, in Kühling/Buchner (supra note 33) Art. 3 Rn. 84; ECJ decision of 7/12/2010 – C-585/08 and C-144/09 Pammer und Alpenhof, paras. 90-93

[47] Spindler, GRUR 2013, 996 (1003); Spindler GRUR-Beilage 2014, 101 (107)

exactly constitutes "monitoring" has to be determined on a case-by-case basis. It may include tracking online behaviour on the internet via cookies[48], or use of data processing techniques to profile individuals in order to analyse their personal preferences, as laid down by Recital 24. At any rate, the provision requires a certain level of intensity. A continuous long-term surveillance is not needed.[49] This could potentially place compliance obligations on a researcher outside the EU when monitoring data subjects e.g. by web analytics within the EU without actually targeting them.

Together with the extensive interpretation of Art. 3 (1) suggested by the ECJ, this shift in paradigm presented in Art. 3 (2) results in a rather low threshold for triggering the application of the GDPR. Although one can argue that the initiative of a board and potentially extraterritorial application was imminent[50], these changes expose the controllers and processors to considerable new compliance burdens. Those who are not currently affected by the Directive will be subject to the entire palette of obligations presented by the GDPR in relation to the relevant data processing activities.

This wide territorial scope might facilitate data subjects in exercising their rights, but at the same time it also raises the question of an effective enforcement. Art. 51 (1) states that the "Member States shall provide for one or more independent public authorities to be responsible for monitoring the application of this Regulation". According to Art. 51 (2), these supervisory authorities "shall cooperate with each other and the Commission", however, only within the EU. Since they are not empowered with extraterritorial jurisdiction, it is yet to clarify, how they will be able to enforce sanctions imposed on controllers based outside the EU in practice. Article 27 might shed some light on this issue, by stating, that every controller and processor outside the Union, who are subject to Art. 3 (2) GDPR shall designate a representative in the Union in writing.[51]

Given the above outlined significantly broader territorial scope under the GDPR and the numerous unclear and unambiguous issues which come with it, it is strongly suggested for companies to utilize techniques like the one developed within the SODA project, especially encryption, secret sharing and secure multi-party computation in order to keep the processing of personal data to a necessary minimum and thus possibly escape the long-arm territorial reach of the GDPR.

2.2.2.2 Subject Matter and Scope of the GDPR

Essentially, many of the underlying principles of the Directive remain intact. The goals set by the Directive and the GDPR are closely aligned, however, the GDPR is intended to abandon the similar but not identical data protection law system, and to lead to a more harmonised data protection law[52] and thus higher level of legal certainty across the EU. This approach is intended to facilitate the free flow of personal data in the digital single market.

2.2.2.2.1 Subject Matter of the GDPR

When defining its objectives, the GDPR even has the same structure as seen in the Directive. According to Art. 1, the main aims of the GDPR are the protection of the fundamental rights and freedoms of natural persons, in particular their right to the protection of their personal data, as well as enabling the free movement of personal data within the EU and so to contribute to the economic and social progress.[53] This reflects the same trade-off as the one previously outlined in the Directive (see 2.2.1.2.1). According to Art. 1 (3) GDPR it is even forbidden to restrict or prohibit the free movement of personal data for reasons connected to the protection of natural persons.

[48] Art.-29 Working Party Opinion 04/2012 WP 194, 1 ff.
[49] Zerdick, in Ehmann/Selmayr, Datenschutz-Grundverordnung Kommentar, Art. 3 Rn. 19
[50] Roßnagel/Richter/Nebel, ZD 2013, 103 (104); Klar, ZD 2013, 109 (114)
[51] Kühling/Martini, EuZW 2016, 448 (450)
[52] Buchner, in Kühling/Buchner (supra note 33), Art. 1 Rn. 18
[53] Recital 2-7 GDPR

2.2.2.2.2 Material Scope of the GDPR

The GDPR's material scope is regulated in Art. 2 (1):

> "[…] applies to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system."

The wording of this provision is almost identical to Art. 3 (1) of the Directive. The two most important definitions of this provision, namely that of the personal data and processing, are defined in Art. 4 (1), (2) respectively. According to Art. 4 (1) personal data means:

> "any information related to an identifiable natural person (data subject) […]".

The GDPR follows the black-and-white approach of its predecessor, according to which information is either personal or not.[54] It operates with a broad notion of personal data, which essentially means that even the smallest reference to a natural person can trigger the application of the GDPR, regardless the nature, content, and format of the data.[55] "Any" in this context means practically anything what is available on a person, publicly available information included. One of the innovations of the GDPR is that it clarifies that EU data protection law does not apply to the data of deceased persons.[56] "Processing" in terms of Art. 4 (2) includes:

> "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction".

This provision implies a rather extensive interpretation of data processing.[57] It was an on-purpose decision that the GDPR does not contain any further clarification of what "automated means" of processing are. In order to stay technologically neutral, it is left undefined, given the rapid technological developments.[58] This notion means that "the protection of natural persons should be technologically neutral and should not depend on the techniques used" and is outlined in Recital 15.

In Art. 2 (2) the GDPR gives a list of activities excluded of its material scope. Thus, amongst other, it does not apply to processing of personal data carried out by a natural person "in the course of a purely personal or household activity".[59]

A more important exception is mentioned in Recital 26:

> "the principles of data protection should […] not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes".

---

[54] Forgó, International Data Privacy Law, 2015, Issue I. 54 (59)
[55] Art. 29-Working Party Opinion 04/2007, WP 136, 6 ff.
[56] Recital 27, 158, 160 GDPR
[57] Kühling/Raab, in Kühling/Buchner (supra note 33) Art. 2. Rn. 13
[58] Kühling/Raab in Kühling/Buchner (supra note 33) Art. 2 Rn. 15; Ernst, in Paal/Pauly (supra note 19) Art. 2. Rn.4
[59] Art. 2 (2) lit c) GDPR, Gola/Lepperhof, ZD 2016, 9 (10)

Anonymity is presented here as an opposite of references of persons.[60] As long as anonym or anonymized data are processed, and thus no reference to any natural person can be established, the processing activity does not fall under the scope of the GDPR. To what degree data have to be de-identified in order to count as anonym or anonymized, will be discussed in the next chapter.

2.2.2.3 Principles of the GDPR

The principles of the GDPR provide guidance on the conditions on which the processing of personal data is permitted. They are more than mere abstract proclamations,[61] if the parties taking part in the data processing cannot satisfy the principles, such processing will be unlawful.[62] The changes in the GDPR are not revolutionary, but they do consolidate and concretize the meaning of certain principles. Article 5 of the GDPR contains the principles governing the processing of personal data.

2.2.2.3.1 Lawfulness, Fairness, and Transparency

According to Art. 5 (1) a) of the GDPR, personal data shall be "processed lawfully, fairly and in a transparent manner in relation to the data subject". This provision mentions several different principles, which together form the requirement of lawfulness, an indispensable basis of data processing.

The principle of lawfulness can be interpreted in two different ways.[63] Pursuant to a narrow understanding, it strengthens the principle of prohibition with the reservation of authorisation, as it implies that without consent of the affected person or another justifiable legal basis, the processing of personal data is principally forbidden and unlawful.[64] This explicit reference to Art. 6 (1) of the GDPR appears in Recital 40 as well, which states that in order to ensure the lawfulness of the processing of personal data, "personal data should be processed on the basis of the consent of the data subject concerned or some other legitimate basis, laid down by law, either in this Regulation or in other Union or Member State law as referred to in this Regulation". Following a broad understanding, the reservation of lawfulness has a subsidiary catch-all function, as it covers those legal requirements too, that cannot be subsumed under another principle. However, it should be noted that the lax and shapeless nature of the principle of lawfulness speaks against a broad interpretation.[65]

Fairness, as regulated in Art. 5 (1) a) of the GDPR, requires the guarantee of a fair processing. In doing so, the "reasonable expectations"[66] of the data subject has to be taken in consideration during the processing activity.[67] Nonetheless, it does not justify any decision based on equity.[68]

Unlike the Directive, the GDPR explicitly takes transparency up in its principle catalogue, thereby imposing an additional compliance burden on the controllers and processors. Recital 39 states that transparency "requires that any information and communication relating to the processing of those personal data be easily accessible and easy to understand, and that clear and plain language be used. That principle concerns, in particular, information to the data subjects on the identity of the controller and the purposes of the processing and further information to ensure fair and transparent processing in respect of the natural persons concerned and their right to obtain confirmation and communication of personal data concerning them which are being processed". Recital 58 describes it even more precise:

[60] Härting (supra note 43) Rn. 291
[61] Frenzel, in Paal/Pauly (supra note) Art. 5 Rn. 2
[62] Unless exemptions or derogations apply for the given processing
[63] Herbst, in Kühling/Buchner (supra note 33) Art. 5 Rn 8, 9; Frenzel, in Paal/Pauly (supra note 19) Art. 5 Rn. 14 ff.
[64] Heberlein, in Ehmann/Selmayr (supra note 49) Art. 5 Rn. 8
[65] Frenzel, in Paal/Pauly (supra note 19) Art. 5 Rn. 16; Herbst, in Kühling/Buchner (supra note 33) Art. 5 Rn. 10
[66] Recital 47 of the GDPR
[67] Härting (supra note 43) Rn. 89
[68] Frenzel, in Paal/Pauly (supra note 19) Art. 5 Rn. 20

> „The principle of transparency requires that any information addressed to the public or to the data subject be concise, easily accessible and easy to understand, and that clear and plain language and, additionally, where appropriate, visualisation be used. Such information could be provided in electronic form, for example, when addressed to the public, through a website. This is of particular relevance in situations where the proliferation of actors and the technological complexity of practice make it difficult for the data subject to know and understand whether, by whom and for what purpose personal data relating to him or her are being collected, such as in the case of online advertising. Given that children merit specific protection, any information and communication, where processing is addressed to a child, should be in such a clear and plain language that the child can easily understand."

These recitals outline two aspects of transparency. The retrospective component is that it obliges controllers and processor to carefully record and keep track of the processing activity step-by-step. On the other hand, prospectively, it gives clear instructions on how controllers have to fulfil their obligations laid down among the rights of the data subject in Chapter III of the GDPR. This duality ensures the clarity of the data processing itself, and helps data subjects to receive all the information they are entitled to according to Art. 12 ff. GDPR.[69]

In practice, controllers participating in a Big Data analysis should pay close attention to the obligations related to the transparency principle. It can be difficult to provide the affected persons with precise information about a processing in such an environment, where data is obtained from several different sources, and these small inputs are later aggregated to produce a whole dataset.[70] Especially in cases where the data subjects does not have the full control over the destiny of their data, that is to say processing where personal data have not been obtained from the data subject. A core provision in this context is Art. 14 (2) g) of the GDPR, which states that in such case information is to be provided on "the existence of automated decision-making, including profiling, […] meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject". This is reasonable and doable when the mapping between the input and the output is clear. It is less clear however, what this means for automated decision-making when the underlying algorithms emerge out of advanced machine learning techniques.[71]

2.2.2.3.2 Purpose Limitation

The GDPR brings limited changes to the principle of purpose limitation. This concept of purpose limitation is one of the founding principles of current data protection law,[72] and it essentially states that personal data collected for one purpose should not be used for a new, incompatible one. According to Art. 5 (1) lit b), personal data

> "shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89 (1), not be considered to be incompatible with the initial purposes."

The first component of the provision is the purpose specification. Recital 39 gives further explanation on what this requires by stating that "the specific purposes for which personal data are processed should

---

[69] Frenzel, in Paal/Pauly (supra note 19) Art. 5 Rn. 21
[70] Richards/King, Three Paradoxes of Big Data, 66 Stan. L. Rev. Online 41, 42-43 (2013-2014)
[71] Leenes, Accounability and transparency in Big Data Land, DSC/t Blog, 2016, available at: https://www.tilburguniversity.edu/research/institutes-and-research-groups/data-science-center/blogs/data-sience-blog-ronald-leenes/
[72] Dammann, ZD 2016, 307 (311)

be explicit and legitimate and determined at the time of the collection of the personal data". The prerequisites 'specified' and 'explicit' not only serve the purpose limitation principle, but are also closely connected to transparency, data minimisation and protection of the data subject's rights. For the sake of user control, the purpose should be precise and clear enough to predict how and to what extent the controller handles the data in question. For a purpose to be specified, it must be sufficiently defined to delimit the scope of the processing operation. Secondly, the purpose must be unambiguous and clearly expressed, without any hidden purpose, such as secret algorithms or hidden profiling.[73] Purpose specification requires an internal assessment carried out by the data controller, prior to, and in an event, not later than, the time when the collection of personal data occurs.[74] The specification of purpose can be made in alternative or complementary ways, e. g. by public declarations, information to the data subject, legislation, administrative decrees and licenses provided by the supervisory authority.[75] Expressing the purpose in writing and adequate documentation will also help the controller to demonstrate the compliance with the requirements of Art. 5 (1) b).

In addition, the purpose of the processing must be legitimate. This refers to Art. 6 of the GDPR, since processing is only legitimate, if there is either consent from the data subject, or another prerequisite provided by Art. 6 (1). This results from the principle of prohibition with the reservation of authorisation. However, the requirement of legitimacy means more than a simple cross-reference with Art. 6, as it also requires that the purpose must be in accordance with all provisions of applicable data protection law.[76] Personal data can of course be collected for more than one purpose. In this case, in order to ensure compliance with Art. 5 (1) b), each purpose must be specified enough. If personal data are processed for several purposes, all requirements of Art. 5 (1) b) apply to each purpose separately.

The second building block of this principle is the compatible use of personal data in case of further processing.[77] The legislator opted for a double negation on this matter by stating that personal data "shall not be further processed in a manner that is incompatible with those purposes".[78] It is important to notice that a different purpose does not necessarily and automatically results in incompatibility, this needs to be assessed in a case-by-case basis.[79] In order to decide whether or not the further processing after change of purpose is lawful, a multi-factor compatibility assessment is to be carried out.[80] The criteria, which need to be evaluated within the frames of this assessment were developed by the Art.-29 Working Party[81] and with little modification became part of the GDPR.[82] Art. 6 (4) of the GDPR lists up these factors to be taken into account by the controller "in order to ascertain whether processing from another purpose is compatible with the purpose for which the personal data are initially collected". Art. 6 (4) a) focuses on the relationship between the purposes of the collection and the purposes of the further processing, and states that the focus should be on the substance of this relation rather only be seen as a textual issue. Generally, the greater the distance between the purposes, the more problematic this would be. Art. 6 (4) b) concentrates on "the context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller". When evaluating the

[73] Art.-29 Working Party, Opinion 03/2013, WP 203, p. 12, 69
[74] Herbst, in Kühling/Buchner (supra note 33) Art. 5. Rn. 31
[75] Härting (supra note43) Rn. 95; R. 54 of the Explanatory Memorandum to the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, available at: http://www.oecd.org/sti/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm#memorandum , quoted by: Art.-29 Working Party, Opinion 03/2013, WP 203, p. 18
[76] Art.-29 Working Party, Opinion 03/2013, WP 203, p. 19
[77] Recital 50 (1) s. 1 GDPR
[78] It is unclear, what exactly qualifies as "further processing". The Art.-29 Working Party represents the opinion, that any processing following collection must be considered "further processing" and must therefore meet the requirement of compatibility.Art.-29 Working Party, Opinion 03/2013, WP 203, p. 21; whereas Herbst does not share this opinion, Herbst, in Kühling/Buchner (supra note 33) Art. 5 Rn. 39 ff.
[79] Frenzel, in Paal/Pauly (supra note 19) Art. 5 Rn. 30
[80] Albrecht, CR 2016, 88 (92); Plath, in Plath (supra note 33) Art. 6 Rn. 38
[81] Art.-29 Working Party, Opinion 03/2013, WP 203, p. 23 ff.
[82] Herbst, in Kühling/Buchner (supra note 33) Art. 5 Rn. 45

context of the data collection, account should be taken on the reasonable expectations of the data subject as to the further use of the collected data based on that context. The more unexpected the further use, the less likely it would be considered as compatible. Art. 6 (4) c)-d) focus on protecting the affected person against the consequences of improper or excessive use of their personal data. It is therefore essential to assess the nature of the data, especially if the further processing involves special categories of personal data according to Art. 9. The more sensitive the data is, the narrower the scope of a compatible use would be.[83] Relevant can be at this point the way in which data are further processed, such as whether or not by a different controller in another context, whether or not large amounts of personal data are processed or combined with other data, particularly if such operations were not foreseeable at the time of collection.[84] The last factor, given in Art. 6 (4) e) focuses on safeguards applied by the controller to ensure the safe and fair processing, and to guarantee confidentiality and security of the data. This might require technical and organisational measures, such as full anonymization, pseudonymisation, aggregation of data or encryption, to ensure functional separation, but also additional steps e. g. increased transparency.[85]

The above outlined criteria together serve as standard to decide on the issue of incompatibility with the initial purpose. However, the nature of assessment carried out by the controller depends on the specific processing activity. A formal method, basically a mere comparison between the initial and new purposes seems to be objective and neutral for the first sight, but might be too rigid. A rather substantive approach, that goes beyond formal statements and takes into account the context and other factors is more flexible. It is safe to say, the greater the distance between the initial purpose specified at the original collection and the purpose of the further use, the more thorough and comprehensive the analysis will have to be. It is strongly advised to include additional safeguards to compensate for the change of purpose in such situations.[86]

It must be noted at this point that there are two conditions, in which the compatibility of the initial purpose of the processing is irrelevant and not required for the further processing. These are cases, where the processing for a purpose other than that for which the personal data have been collected is based on either the data subject's consent, or Union or Member State law, so Art. 6 (4) of the GDPR.[87] If this is not the case, the controller always has to carry out the compatibility assessment.

Art. 5 (1) b), in line with Recital 156, contains another important provision which it states that "further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes". With that the GDPR creates a group of privileged secondary purposes and establishes a legal fiction of compatibility of these purposes with the initial ones.[88] However, providing an overall exception from the requirement of compatibility does not mean a general authorisation to further processing of personal data for the above-mentioned purposes.[89] Art. 89 (1) itself highlights that such processing should be "subject to appropriate safeguards". Thus, all relevant circumstances must be taken into account when deciding what safeguards can be considered appropriate and sufficient.

[83] Art.-29 Working Party, Opinion 03/2013, WP 203, p. 25
[84] Art.-29 Working Party, Opinion 03/2013, WP 203, p. 26
[85] Art.-29 Working Party, Opinion 03/2013, WP 203, p. 46
[86] Art.-29 Working Party, Opinion 03/2013, WP 203, p. 22
[87] Plath, in Plath, BDSG-DSGVO Kommentar, Art. 6, Rn. 31
[88] Herbst, in Kühling/Buchner (supra note 33) Art. 5 Rn. 50
[89] Art.-29 Working Party, Opinion 03/2013, WP 203, p. 28

2.2.2.3.3 Data Minimisation

This principle complements and adds more value to the principle of purpose limitation. The core idea behind data minimisation is that subject to limited exceptions, controllers should only process the personal data – and not more as – they need to process in order to achieve their processing purposes. This ensures that the collection of data does not exceed the required level in light of the specific processing.

Art. 5 (1) c) of the GDPR contains the provisions regarding the principle of data minimisation:

> "Personal data shall be […] adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed."

The three conditions stipulated by this provision can hardly be defined distinctly, but together they form one single requirement, that of an appropriate and reasonable use of personal data.[90] Personal data are adequate, if they are inevitable for the fulfilling of the purposes of the processing. The criterion 'relevant' stands for appropriate usage, and serves as an objective requirement, which means that the use of the personal data in question must be necessary for the processing based on an impartial and unbiased test, not only labelled as such by the controller.[91] The phrase "limited to what is necessary" is new in the GDPR, and means a more restrictive obligation than the "not excessive" condition in the Directive did. This last criterion is particularly important in those cases, where the collected personal data is indeed adequate and relevant, however, the purpose of processing can be achieved without, in fact, using them.

Controllers processing personal data in order to carry out Big Data-related analysis typically collect personal data and later decide on the purpose for which they wish to use these data. With the amendment of the principle of data minimisation, the GDPR tightens the restrictions on such processing even further. However, Art. 25 (1) might serve as a silver lining for companies planning to engage in such analysis. Art. 25 (1) contains the rules on data protection by design, and also concretises the data minimisation principle. It obliges controllers to implement appropriate technical and organisational measures, which are designed to implement data protection principles, such as the data minimisation. One of these is the pseudonymisation, which "can reduce the risks to the data subjects concerned and help controllers and processors to meet their data-protection obligations".[92] Furthermore, with anonymization, the data processing would fall outside the scope of the GDPR, consequentially would not be subject to the data minimisation principle anymore.[93] Thus, if the purpose of the processing can be achieved in such manner, it is strongly suggested to adopt measures like pseudonymisation or anonymization.

2.2.2.3.4 Storage Limitation

The storage limitation in Article 5 (1) e) GDPR is an important principle of the GDPR and is closely linked to the data minimisation of Article 5 (1) c) GDPR and the purpose limitation of Article 5 (1) b) GDPR. It constitutes a barrier to the excessive storage of personal data. On the one hand, the storage limitation contains a requirement of reservation. This means that personal data may only be stored as long as the purpose requires it. On the other hand, it obligates the responsible party to justify the longer retention periods.[94] This corresponds to the principle of accuracy in Article 5 (1) lit d) GDPR. However, according to Article 5 (1) lit e), data may be stored for a longer period of time, if they are used for

[90] Herbst, in Kühling/Buchner (supra note 33) Art. 5 Rn. 57
[91] Frenzel, in Paal/Pauly (supra note 19) Art. 5, Rn. 37
[92] Recital 28. GDPR
[93] Herbst, in Kühling/Buchner (supra note 33) Art. 5 Rn. 58
[94] Frenzel, in Paal/Pauly (supra note 19) Art. 5 Rn. 43 ff.

archiving purposes in the public interest, they are processed for scientific or historical research purposes or for statistical purposes in accordance with Article 89 (1) GDPR.[95]

2.3 Personal Data and Privacy Preserving Technologies

A growing danger of intruding in privacy comes along with the huge potential Big Data holds, whether in a structured or unstructured form. The issue to protect individuals against the increasing usage of personal data is becoming more important as the techniques used for processing develop. However, privacy preserving technologies may show a pathway to a lawful and secure utilisation of personal data. They are therefore of paramount importance for SODA-technologies, whose main application area is to share highly sensitive data between different organisations, e. g. hospitals and insurance companies.

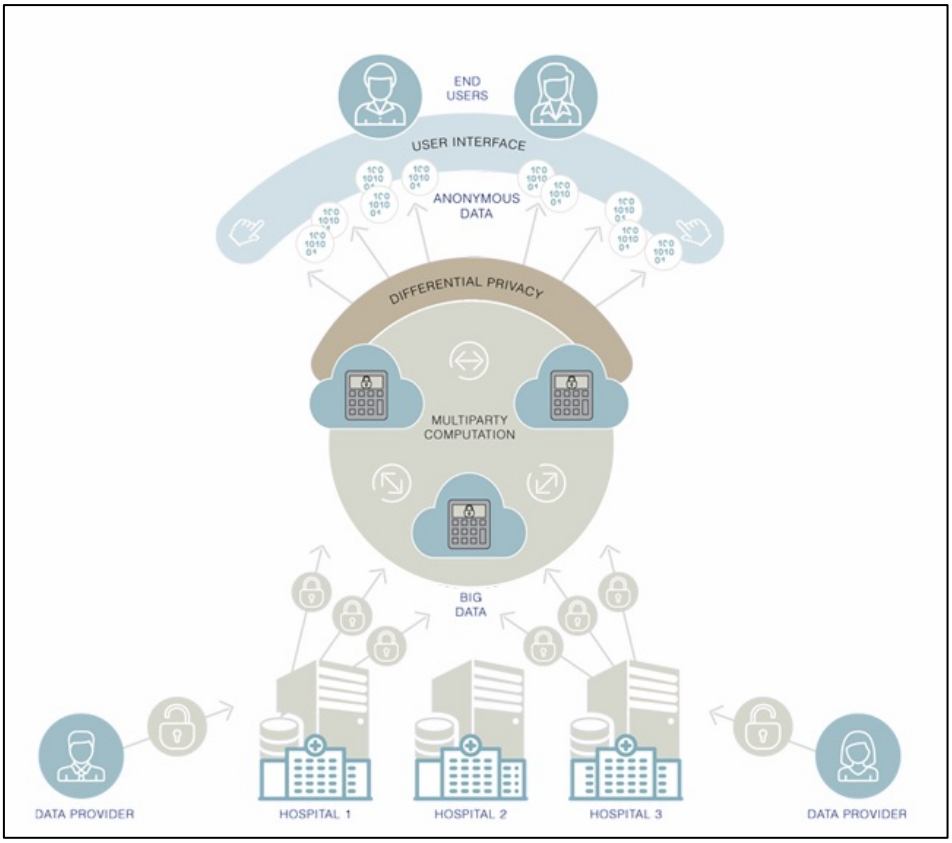


Figure 2. The SODA Model[96]

When exchanging their data, these actors must guarantee that the individuals will not be identified from the combined datasets. Generally speaking they have two ways to achieve analysis without compromising security. Data must either rendered anonymous, i. e. completely stripped from the natural person. In this case data cease to be personal data and may be processed as other “feature data”. The other method is to encrypt the data with secure cryptographic techniques which are compliant with the GDPR. One of the most suitable way to minimise the chance of identification is the so called secret sharing, i. e. dividing data into pieces (shares) which are unreadable without the other shares. Implementing more privacy preserving strategies at the same time can result in a high-level security. This is precisely what

[95] Chassang: The impact of the EU general data protection regulation on scientific research, DOI: 10.3332/ecancer.2017.709, avaiable at: https://www.researchgate.net/publication/312251732_The_impact_of_the_EU_general_data_protection_regulation_on_scientific_research

[96] source: https://www.soda-project.eu/

SODA aims by combining secret sharing with other strategies, such as differential privacy. In order to decide whether secret-shared or otherwise de-identified data are secure enough for the – partial – non-applicability of the GDPR or the prevention of unwanted identification, the legal background of "identifiability" and privacy-preserving methods" as well as their implications must be assessed.

As discussed in the previous section, the two pillars of the material scope of European data protection law are the (a) processing of (b) personal data.[97] If one of these conditions is not fulfilled, the European data protection law does not apply.[98] It is important to see that depending on the interpretation of personal data and the context of the particular processing activity, the effects such methods have on data as well as their legal assessment may be different.

2.3.1 The Concept of "Identified" and "Identifiable" Natural Person

Defining the term "personal data" is the focal point of the European data protection law. According to Art. 2 (a) of the Directive personal data mean

> "any information relating to an identified or identifiable person (data subject); an identifiable person is one who can be identified, directly or indirectly, in particular by referencing an identification number, or one or more factors specific to his physical, physiological, mental, economic, cultural or social identity."

The GDPR imposes minor changes, but for the most part, the definition in Art. 4 No. 12 GDPR echoes the definition of the Directive:

> "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."

As outlined above, the wording "any" implies a broad scale of information, ranging from undisclosed data to publicly available information about a natural person (see 2.2.2.2.2.). Both the Directive and the GDPR require additionally that the information must relate to a natural person, that is identified or identifiable. This is a decisive, yet controversial element of the definition; therefore, it is crucial to understand the conception of these terms.

When a natural person is identified, is not further elaborated on. Generally, someone is identified, if he or she is precisely distinguished from the others and detected in a group of people.[99] Unlike the Directive, the GDPR's definition distinguishes between identifiability on the basis of identifiers[100], such as name or identification number, or by reference to factors specific for the individual. The identification happens with the help of identifiers or specific factors, without the need of involving other information from different source.[101] It is to be noted, that the same information may identify someone in one case, but not in another, since the contextual element always has to be taken into account.[102] A common family name is unlikely to identify a particular individual in a country, but may be enough within a smaller group.[103]

---

[97] Esayas, European Journal of Law and Technology, Vol 6, No 2 (2015), p. 2
[98] cf. Klabunde, in: Ehmann/Selmayr (supra note 49) Art. 4 Rn. 17
[99] Art.-29 Working Party Opinion 4/2007, WP136, p. 12
[100] Urgessa, The Protective Capacity of the Criterion of 'Identifiability' under EU Data Protection Law, European Data Protection Law Review (EDPL), 2016 Vol. 4, 521 (522); Klabunde, in: Ehmann/Selmayr (supre note 49) Art. 4 Rn. 12
[101] Klar/Kühling, in: Kühling/Buchner (supra note 33) Art. 4 Nr. 1 Rn. 19; Härting (supra note 43) Rn. 275
[102] Art.-29 Working Party Opinion 4/2007, WP136, p. 13
[103] ECJ case C-28/08, European Commission v The Bavarian Lager Co. Ltd., ECLI:EU:C:2010:378

The second alternative concerns an "identifiable" natural person. A natural person is identifiable, when there is a possibility of identification, but it has not happened yet.[104] The Directive offers a rather laconic guidance on the circumstances of identifiability in its Recital 26:

> "Whereas the principles of protection must apply to any information concerning an identified or identifiable person; whereas, to determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person; whereas the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable […]"

A more detailed explanation is provided by Recital 26 GDPR:

> "The principles of data protection should apply to any information concerning an identified or identifiable natural person. [...] To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments. The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable."

The GDPR amends Recital 26 of the Directive by outlining specific factors that help decide on the issue of identifiability based on a case-by-case evaluation, as well as by refining the conditions for anonymous information.

However, both the Directive and the GDPR are rather ambiguous, and therefore subject to different interpretations. There are two concurring concepts for determining the threshold of identifiability, an absolute, and a relative, more context-sensitive approach.[105]

The absolute approach takes into account all possibilities and means available for the data controller and third parties to be able to identify one particular data subject. This objective position assumes an "overall knowledge"[106]. As a result, it includes even theoretical, unlikely possibilities of combining data which may lead to identification.[107] If the identifiability is indeed assessed according to this approach, it is sufficient enough that there is somewhere someone, who holds the additional information to identify a data subject. This would mean that there would virtually be no more anonymous data at all, since it is in actual fact impossible to prove that nobody has any chance to relate to a natural person.[108] In terms of privacy preserving methods, as long as there is *any* chance, that *anyone* in the world is able to decrypt and obtain the dataset, the processing of the data would always be subject to data protection legislation, regardless of the used technique.

In contrast, a relative approach is based on the assessment of the realistic chances of the data controller to identify the data subject. Relevant factors in determining whether a data controller has the

---

R. 68; in line with appeal case T-194/04 of the Court of First Instance (Third Chamber). ECLI:EU:T:2007:334 R. 67, 104

[104] Esayas (supra note 96) p. 2; Urgessa (supra note 99) p. 521, who defines identifiability as a qualifying factor

[105] for a thorough summary on the different opinions see: Bergt, ZD 2015, 365 (365 ff.); Specht/Müller-Riemenschneide, ZD 2014, 71

[106] Schantz, NJW 2016, 1841 (1842 f.)

[107] Brink/Eckhardt, ZD 2015, 205 (206); Spindler/Nink, in: Spindler/Schuster, Recht der elektronischen Medien, § 11 TMG, Rn. 7; Esayas (supra note 96) p. 6

[108] Härting, (supra note 43) Rn. 267; Nink/Pohle, MMR 2015, 563 (565); Keppeler, CR 2016, 360 (361)

means, possibility and knowledge must be considered, but not a merely hypothetical opportunity.[109] The basis of defining when information is personal data is therefore a result of a careful evaluation of the necessary endeavour required by the controller. Accordingly, in situations where the controller has, or with reasonable chances is able to obtain the decryption key of an encrypted dataset, these data are personal data, and thus data protection law applies to the processing.[110]

Although there is a tendency leaning towards the less restrictive relative approach, the absolute approach has not been explicitly rejected.

2.3.1.1 The Article 29 Data Protection Working Party's Approach

In its opinion about "the concept of personal data" the Article 29 Data Protection Working Party outlines its stance regarding the elements of the Directive's "personal data" definitions, inter alia the identifiability.[111]

Beside the quite clear situation of direct identifiability the opinion faces the case of indirect identifiability, as a "phenomenon of unique combinations" as well. It points out that a mere hypothetical possibility of recombination is not sufficient, yet there are circumstances where the probability of identification is higher than in other cases, in particular, when "the combination of details on categorical level (age category, regional origin, etc…) may also be pretty conclusive in some circumstances, particularly if one has access to additional information of some sort."[112]

In line with Recital 26 of the Directive, the opinion examines the issue whether the means used by the controller or by any other person to identify a given person are likely reasonably to be used. Considering not only the means of the controller but also that of any other persons may be an indication of an absolute approach. Then again, the Working Party recognises such situations, where particular information may be considered personal data with respect to one party but not another, which suggests a relative notion.[113] This apparent discrepancy derives from the fact, that the Working Party highlights the importance of the contextual elements, that is the circumstances of the specific processing activity rather than the personal perspective. Thus, even if the means by the controller and any other person are relevant and need to be assessed, these "means" are limited to those which are likely reasonably to be used in each specific case.[114] The proposed factors to be taken into account are the cost of conducting identification, the purpose of processing, the way the processing is structured, the anticipated advantage on the data controller's site, the interest at stake for the individuals and the risk of organisational dysfunctions as well as technical failures.[115]

Another important point to see in the opinion is that in cases where the identification is not the purpose of the processing, the technical measures preventing the identification may make the difference to consider that a given person is not identifiable. In such cases, the implementation of those measures is rather a condition for the information not to be considered to be personal data. This would lead to a processing which were not subject to the data protection law.[116]

All things considered, the opinion tends to facilitate a relative approach[117], especially in case of processing via state-of-the-art cryptographic techniques.

---

[109] Härting, (supra note 43) Rn. 264; Klar/Kühling, in: Kühling/Buchner (supra note 33) Art. 4 Rn. 25; Gola, in: Gola, DS-GVO – Datenschutz-Grundverordung – Kommentar, Art. 4 Rn. 18; Brink/Eckhardt, ZD 2015 205 (211); Roßnagel/Scholz, MMR 2000, 721 (723); Meyerdierks, MMR 2009, 8 (8 ff.), Knopp, DuD 2015, 527 (529)

[110] Spindler/Schmechel, JIPITEC 2016, 163 (165 f.)

[111] Art. 29-Working Party, Opinion 04/2007 on the concept of personal data, WP 136

[112] Art. 29-Working Party, Opinion 04/2007, WP 136, p. 13

[113] Art. 29-Working Party, Opinion 04/2007, WP 136, p. 15

[114] Art. 29-Working Party, Opinion 04/2007, WP 136, pp. 15, 21

[115] Art. 29-Working Party, Opinion 04/2007, WP 136, p. 15

[116] Art. 29-Working Party, Opinion 04/2007, WP 136, p. 17

[117] Cf. Eckhardt, CR 2011, 339 (341, 343); Stiemerling/Hartung, CR 2012, 60 (63), who argue that the opinion follows a rather absolute concept

2.3.1.2 The ECJ's Approach

In its landmark Breyer case[118], the ECJ provided further clarification as to the qualification of personal data. The significance of the case lies in the interpretation of "identifiability" expressed by the Data Protection Directive.

The case was referred to the Court as a request for preliminary ruling procedure[119] by the German Federal Court of Justice (BGH) for guidance on the interpretation of the dispute whether dynamic IP-addresses can be considered as personal data.[120] After referring to the disagreement relating to the absolute or relative criterion, the German Federal Court of Justice opted for a relative approach in its request, and argued, that the consideration of the means a third party has to identify an individual does not necessary exclude a relative approach, inasmuch as only the means realistically be used are taken into account.[121]

There was, on the other hand, a strongly voiced opinion of the European Commission along with several Member States regarding this case, leaning towards an absolute approach.[122]

Campos Sánchez-Bordona, the Advocate General assigned to the case (henceforth: AG),[123] released his opinion on May 12, 2016.[124] In his opinion, the AG acknowledges the relevance of the fact that additional data, which may enable the identification, is in possession of a certain third party.[125] However, when assessing whether such additional data, which combined with the data of the data controller can lead to identification, may be personal data he contradicts an interpretation of "means likely reasonably to be used […] by any other person" in a way that it would be sufficient that any third party might obtain additional data[126], since such an "overly strict interpretation would lead, in practice, to the classification as personal data of all kinds of information, no matter how insufficient it is in itself to facilitate the identification of a user".[127] The AG pointed out, that the inherent problem of such restrictive approach is, that "it would never be possible to rule out, with absolute certainty, the possibility that there is no third party in possession of additional data which may be combined with that information".[128] Besides, the AG argues, that a systematic interpretation of R. 26 DPD would be "the means likely reasonably to be used", which will not occur, when contacting those certain third parties "is, in fact, very costly in human and economic terms, or practically impossible, or prohibited by the law".[129] Therefore, the proposition of the AG can be considered as a shift towards a relative approach.

---

[118] ECJ judgement of 19 October 2016, Case C-582/14, Patrick Breyer v Bundesrepublik Deutschland, E-CLI:EU:C:2016:779

[119] For the provisions of the preliminary ruling procedure see: Article 267 of the Treaty on the Functioning of the European Union (TFEU)

[120] German Federal Court of Justice (BGH), decision of 28/10/2014 - VI ZR 135/13 = MMR 2015, 13; regarding the classification of dynamic IP addresses as personal data for access providers judged by the EJC, judgement of 24 November 2011, Case C‑70/10 – Scarlet Extended SA v Sabam, Recital 51, which states that "[IP] addresses are protected personal data because they allow those users to be precisely identified"

[121] ECJ, Case C-582/14 (supra note 117), Recital 21, 25; German Federal Court of Justice (BGH), decision of 28/10/2014 - VI ZR 135/13 = MMR 2015, 131 (133 f.) R. 28

[122] Bergt, IP-Adressen: EU-Kommission gibt BGH Nachhilfe in Sachen Grundrechte, available at: http://www.cr-online.de/blog/2015/09/13/ip-adressen-eu-kommission-gibt-bgh-nachhilfe-in-sachen-grundrechte/

[123] For the provision on the advocate general see: Art. 252 TFEU

[124] Opinion of Advocate General Campos Sánchez-Bordona, delivered on 12 May 2016, Case C-582/14 – Patrick Breyer v Bundesrepublik Deutschland.

[125] Opinion of the Advocate General (supra note 122) Recitals 63, 67

[126] Opinion of the Advocate General (supra note 122) Recital 61

[127] Opinion of the Advocate General (supra note 122) Recital 65, S. 1

[128] Opinion of the Advocate General (supra note 122) Recital 65, S. 2

[129] Opinion of the Advocate General (supra note 122) Recitals 67, 68

In its judgement of October 16, 2016, the ECJ based its argumentation on the expansive scope of Art. 2 (1) of the Directive, which clearly states that a person can be identified directly or indirectly.[130] By taking the indirect identifiability into account, the reasoning invokes Recital 26 of the Directive, which refers to "all the means likely reasonably to be used either by the controller or by any other person to identify the said person".[131] This implies that in order to treat information as personal data within the provision of Art. 2 (1) of the Directive, it is not required that all the information enabling the identification of the data subject must be in the hands of the same person.[132] In this case, this other person is the internet service provider. Since the online media service provider would necessarily need to collaborate with the internet service provider in order to access the additional information to identify the data subject, this may be a step towards an absolute understanding of identifiability.[133] However, the ECJ continued by discussing whether the possibility of combining a dynamic IP address with additional data held by an internet service provider constitutes a mean likely reasonably to be used to identify the data subject. In line with the AG's opinion, the ECJ concluded that this would not be the case "if the identification of the data subject was prohibited by law or practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power", and thus "the risk of identification appears in reality to be insignificant".[134] By including a concept of proportionality the ECJ recognizes the importance of the controller's willingness and ability to achieve a certain result. Furthermore, it emphasizes the legal possibility of the provider to enforce the third party releasing the information.[135] In so far as the ECJ does not take the illegal means or any kind of disproportional effort into account, this argumentation is close to a relative approach.[136] After exploring the above mentioned aspects, the ECJ ruled, that "(…) a dynamic IP address (…) is personal data within the meaning of that provision, in relation to that provider, where the [provider] has the legal means which enable it to identify the data subject with additional data which the internet service provider has about that person".[137]

Thus, the ECJ acknowledged the fact, that a specific information, in this case a dynamic IP-address, is not personal data *per se*, but it can become personal data for the certain controller under specific circumstances. The Court defined these circumstances and clarified that despite of the existence of additional data dynamic IP addresses will not be personal data, when the linking of the IP addresses with those data is prohibited by the law or the acquisition is disproportionate as such.[138]

This progressive approach can be interpreted as a rather relative approach, however, not without any doubts. The ECJ included some absolute criteria in its judgement (such as the potential knowledge

---

[130] ECJ, Case C-582/14 (supra note 117), Recital 41; Reid, Alan S., The European Court of Justice Case of Breyer, p. 5, available at: https://journals.winchesteruniversitypress.org/index.php/jirpp/article/download/32/14

[131] El Khoury, Alessandro, Dynamic IP Addresses Can be Personal Data, Sometimes. A Story of Binary Relations and Schrödinger's Cat, EJRR 8 (2017), p. 191 (191)

[132] ECJ, Case C-582/14 (supra note 117), Recital 43

[133] Bergt, Das Ende der Rechtssicherheit im Datenschutzrecht, available at: http://www.cr-online.de/blog/2016/10/19/das-ende-der-rechtssicherheit-im-datenschutzrecht/; different opinion Stadler, EuGH entscheidet zum Personenbezug von IP-Adressen, available at: http://www.internet-law.de/2016/10/eugh-entscheidet-zum-personenbezug-von-ip-adressen.html

[134] ECJ, Case C-582/14 (supra note 117) Recital 46; in favour of auch an "unreasonableness" of using illegal means Spindler/Nink in: Spindler/Schuster (eds.), Recht der elektronischen Medien, 3rd Ed. 2015, § 11 TMG Recital 8; Brisch/Pieper, CR 2015, 724 (728), who argue that the word „reason" is indeed not consistent with the use of illegal means, however, is against for a complete exclusion of "the illegal means" and calls for a case-by-case consideration; El Khoury, EJRR 8 (2017), p. 191 (195 f) and Kelleher, In Breyer decision today, Europe's highest court rules on definition of personal data, available at: https://iapp.org/news/a/in-breyer-decision-today-europes-highest-court-rules-on-definition-of-personal-data/ are rather critical about what may be "prohibited by law"

[135] ECJ, Case C-582/14 (supra note 117), Recital 47

[136] Stadler (supra note 132); El Khoury EJRR 8 (2017), p. 191 (196) supports the view of "double relativity", based on the phrase "means likely reasonably to be used" of the Directive as well as the phrase "disproportional" in the ECJ's judgement.

[137] ECJ, Case C-582/14 (supra note 117), Recital 65

[138] Reid, Alan S., (supra note 129) p. 6

of 'any' other party), thus, it was left undecided in the end whether the identifiability of a natural person should be decided based on a relative or an absolute approach.

2.3.1.3 The Implications of the Absolute Concept of Identifiability on Big Data Analytics

The absolute perspective of identifiability is based on an extensive interpretation of the definition of personal data, whereas it handles the exceptions rather restrictively. This results in a radically far-reaching material scope of the data protection law. Such an objective regime would leave virtually no space for non-personal data, since even a mere theoretical possibility of data recombination which may lead to the identification of the individual would trigger the application of data protection law.[139] Notably, this approach is not limited to the consideration of legal means, but includes also the possibility of obtaining data in an unlawful way.[140]

For Big Data analytics using encryption or other cryptographic techniques this would mean, that no matter which technical measures are implemented, it would not change the basic character of the personal data itself, only the access by unauthorised parties would be more difficult. This sole obstruction of access does not exclude the theoretical chance of obtaining the decryption key, and acquiring the data.[141] For example, in case of transferring raw data to a storage provider using state-of-the-art encryption, the storage provider would be a processor, even though he never actually possesses the decryption key. The data controller (the party transferring the data) holds the additional knowledge, the decryption key. In this case, the existence of key that in theory would allow the identification of certain individuals excludes non-personal data *per se*, and the processor would be subject to legal and contractual obligations of the data protection law.[142]

According to this approach, encryption serves as a technical measure to secure the processing, but not a condition for information not to be considered personal data. It is merely supposed to ensure that data are protected against unauthorised access, use, or disclosure.

2.3.1.4 The Implications of the Relative Concept of Identifiability on Big Data Analytics

The most important difference between the absolute and the relative approach of identifiability is the context-sensitivity of the latter, making it much more favourable for future Big Data related research. Unlike the objective concept, it focuses on those means of identification, which are under given circumstances likely reasonably to be used. With that it accepts the existence of certain low risks, and at the same time, it leaves those low-probability events of possible identification out of the equation when assessing the question of identifiability.[143]

Essentially this approach emphasises the plausible and rational means a controller or third party may use to identify the data subject, in particular when it is economically and legally[144] feasible.[145]

---

[139] Keppeler, CR 2016, 360 (364)

[140] Klabunde, in: Ehmann/Selmayr (supra note 49) Art. 4 Rn. 13 who sees this as a necessary limitation in the age of Big Data

[141] Note, that the absolute approach considers also the non-legal means of data-access, which would be in practice impossible to rule out.

[142] Marnau, DuD 2016, 428 (430); Nink/Pohle, MMR 2015, 563 (566)

[143] Esayas, (supra note 96), p. 6; Roßnagel/ Scholz, MMR 2000, 721 (726)

[144] Klar/Kühling, in: Kühling/Buchner (supra note 33) Art. 4 Nr. 1 Rn. 28 f; Non-disclosure agreements may be one way to guarantee the impediment of identification, since such contractual clauses are legal prohibitions of recombining data from different providers. Note, that a non-disclosure agreement does not mean data processing agreement between the controller and the processor according to Art.28 (3) GDPR, it is merely an agreement by which the parties are bound not to disclose certain information, Waen/ Van Essen/ Wellens: Confidentiality agreements are not data processing agreements, available at: https://www.lexology.com/library/detail.aspx?g=e1d5ccfb-f0a0-4c32-aa59-a65864af1acd

[145] Spindler/Nink, in: Spindler/Schuster, Recht der elektronischen Medien, § 11 TMG, R. 8

Following this logic, data do not relate to a natural person anymore, when the de-identification[146] is almost impossible and so the reference can no longer be established.[147] However, due to the wide spectrum of data processing activities, it is hardly attainable – or beneficial – to precisely define what constitutes a reasonable effort of re-identification, and thus a case-by-case evaluation must be carried out for each and every processing activity.

Generally speaking, when the cloud user holds the decryption key, but not the provider, the data cannot be rendered legible by the provider, provided that the provider is not able to obtain the decryption key and to identify the data subject. Whether or not the de-identification is effective enough, depends on the chosen encryption or other privacy-preserving technique by the user. However, data do not cease to be personal data for the party who holds the decryption key.[148] Taking the same example of transferring state-of-the-art encrypted data to a storage provider, according to a relative approach of identifiability the storage provider processes anonymous data, considering that he neither has access to the decryption key himself, nor can technically bypass the encryption system.[149] Therefore, the storage provider would not be a processor, and were not governed by the data protection rules.

However, in cases the encryption still implies personal information data protection law continues to apply. Having said that, it would be a mistake to disregard the fact that any personal identifier (e. g. IP-addresses) may under specific circumstances be qualified as personal data.[150]

Apart from the proportionality constituted by the phrase "likely reasonably" to be used[151], another paramount aspect of the relative concept is the focus on state-of-the-art measures.[152] With the rapidly developing technology the possibilities of identification may change over time. Consequently, every expectable development of the foreseeable future must be carefully assessed when deciding on the nature of data or dataset.[153] Furthermore, the implemented technique must be adaptable to the new circumstances, and the assessment whether the identification is still in fact, practically impossible should be dynamic and continuous.[154]

For this reason, data controllers who engage in processing activities using privacy preserving technologies must continuously monitor and repeatedly verify their implemented technology. Equally, encryption and other cryptographic operators must regularly check the state-of-the-art of their method. Such evaluation of their own method supposed to ensure the compliance of the data protection law, since from that moment on when a new technology appears that increases the probability of a potential re-identification the data will be considered personal, and will therefore be subject of data protection.

2.3.1.5 The GDPR's Approach

As an updated version of the Directive, the new set of laws of the GDPR shall, among others, address the gap between data protection law and modern technology. However, the GDPR does not give a full

[146] Tene/Polonetsky, Stanford Law Review Online 2011-2012, 63 (65) uses de-identification as a collective term for „various methods of de-identification (anonymization, pseudonymization, encryption, key-coding, data sharding)"
[147] Klar/Kühling, in: Kühling/Buchner (supra note 33) Art. 4 Nr. 1 Rn. 32
[148] Cf. Hon/Millard/Walden, The Problem of "Personal Data" in Cloud Computing – What Information is Regulated?, Queen Mary University of London – Legal Studies Research Paper No. 75/2011, pp. 25 f, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1783577
[149] Marnau, DuD 2016, 428 (430)
[150] Spindler/Nink, in: Spindler/Schuster, Recht der elektronischen Medien, § 11 TMG Recital 11 ff; ECJ, Case C-582/14 (supra note 117)
[151] El Khoury, EJRR 8 (2017), p. 191 (196)
[152] Roßnagel/Nebel/Richter, ZD 2015, 455 (456)
[153] Spies, MMR-Aktuell 2011, 313727
[154] Art.-29 Working Party Opinion 4/2007, WP136, p. 15, cf. Roßnagel/Scholz, MMR 2000, 721 (723).

clearance on the issue of identifiability, and leaves the question open for several possible interpretations.[155]

As mentioned above, the two decisive factors for the applicability of the GDPR are the existence of "personal data" and the "processing" of these data. Art. 4 No. 1 of the GDPR defines an identifiable natural person as someone, who

> "can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;"

Art. 4 No. 1 states, that identifiers themselves shall contain some sort of personal reference, which may imply, that they would per se be personal information.[156] In connection with online identifiers Recital 30 states:

> "natural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers such as radio frequency identification tags […]."

Accepting the fact that identifiers per se qualify as "personal data" would veer to a rather absolute concept. However, it is yet to see, whether or not the identifiers qualify as personal data irrespective of the particular circumstances quasi automatically.[157]

Regarding the question of indirect identifiability, Recital 26 GDPR outlines the circumstances that need to be taken into account when assessing the nature of the data:

> "To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly."

Similar to Recital 26 of the Directive, in line with the Art.-29 Working Party's opinion, Recital 26 of the GDPR underlines the importance of the "means reasonably likely to be used" when deciding on the identifiability.

Recital 26 expressly declares "singling out" as one mean of indirect identification.[158] Although the GDPR does not provide a definition, this implies an identification of the data subject with different methods without disclosing their names or identities.[159]

It also states that not only the means of the controller must be taken into account, but also such means by "another person". Consequently, if a third person, i. e. any person in the world would have the means reasonably likely to be used, it would be sufficient for data to be "personal data" within the provision of Art. 4 No. 1 GDPR.[160] This lowers the threshold of identification and facilitates an absolute

---

[155] Spindler, DB 2016, 937 (937 ff.)

[156] Buchner, DuD 2016, 155 (155 ff); opposite opinion: Schantz, NJW 2016, 1841 (1843)

[157] cf: Härting (supra note 43) Rn. 279; ECJ Case C-582/14 (supra note 117) advocates against such a broad understanding, since in this case the ECJ laid down the criteria that need to be considered when deciding the question of identifiability, and concluded, that IP addresses only constitute personal data when those criteria are fulfilled.

[158] in line with the LIBE Proposal

[159] Hon/Kosta/Millard/Stefanatou, Tilburg Law School Legal Studies Research Paper Series No. 07/2014, p. 9, who suggest a definition for singling out as "action or inaction detrimentally affecting that individual materially"; Marnau, DuD 2016, 428 (430); Regarding singling out people without knowing their names (for behavioural targeting) see Zuiderveen Borgesius, Computer Law & Security Review 2016, 251 (256 ff.)

[160] Cf. Zuiderveen Borgesius, Computer Law & Security Review 2016, p. 256 (267) who for this reason suggests an objective interpretation for R. 26 GDPR

interpretation.[161] The reference to another person is of special relevance in cases, when data that has no personal reference for the controller whatsoever are transferred to a third party, who has the additional knowledge and the means reasonably likely to be used in order to identify the data subject.[162] This would extend the scope of the GDPR to a degree, where basically every simple feature data may potentially qualify as personal data.[163] Anyone could hypothetically cross-reference available datasets and deduct on the identity of a specific person, the reference would be given, and the data of both datasets would be considered personal data, even if independently they do not relate to the given person.

Against this background it is indifferent whether data are encrypted or otherwise de-identified, since from an absolute point of view every piece of information that can be associated with an individual is personal data. On the other hand, Recital 26 takes the required expense in consideration, which serves as a silver lining for secure privacy-preserving technologies.[164]

With emphasising the evaluation of means likely reasonably to be used, the GDPR abandons a sheer absolute approach, and includes a rather relative element.[165] The Art.-29 Working Party's opinions also support such a relative interpretation, since they already stated that a mere hypothetical possibility to single out the individual is not enough to consider the person as 'identifiable".[166] This view is supported by the fact, that in case of a zero-risk tolerance no existing technique could achieve a required level of anonymity in the era of Big Data.[167]

Recital 26 offers an illustrative list of factors, that are decisive for the interpretation of the means used to identify the data subject:

> "To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments."

These factors set out the frames of the proportionality established by the phrase "reasonably likely"[168], and thus they point in the direction of a relative concept.[169] Significant factors are the general expenses of the party motivated to obtain the data and to identify the data subject, as well as the state of science and technology, including foreseeable future technical development.[170] The word "objective" as provided in Recital 26 means that the relevance of the factors in each case is decided by general economic factors as unbiased standard, not by discretion of the involved controller or the another person.[171] The wording of the new definition of personal data in Art. 4 No. 1 GDPR as well as the matching Recital 26 are very similar to the old ones in the Directive. Hence it can be assumed that the interpretation outlined by the ECJ judgement in the Breyer Case will continue apply.[172] The argumentation the ECJ opted is another rather strong incentive for a relative approach.

---

[161] Klabunde, in: Ehmann/Selmayr (supra note 49) Art. 4 Rn. 13; Schantz, NJW 2016, 1841 (1843)
[162] Klar/Kühling, in Kühling/Buchner (supra note 33) Art. 4 Nr. 1 Rn. 27; cf. Bergt ZD 2015, 365 (369)
[163] Härting, CR 2013, 715 (719); Hullen, PinG 2015, 210 (211);
[164] Klar/Kühling, in Kühling/Buchner (supra note 33) Art. 4 Nr. 1 Rn. 26; see also the opinion of the Advocate General in the ECJ Case C-582/14 (supra note 117)
[165] Härting (supra note 43) Rn. 282; Gola, in Gola (supra note 109), Art. 4 Rn. 17 f.
[166] Art. 29-Working Party, Opinion 04/2007, WP 136, 15; Art. 29-Working Party, Opinion 05/2014, WP 216, 8 ff.
[167] Esayas, (supra note 96) p. 6.
[168] Urgessa, (supra note 99) 522
[169] Spindler, DB 2016, 937 (937 f.)
[170] Härting (supra note 43) Rn. 284; Ernst, in: Paal/Pauly (supra note 19) Art. 4 Rn. 10
[171] Marnau, DuD 2016, 428 (430);
[172] Kelleher, In Breyer decision today, Europe's highest court rules on definition of personal data, available at: https://iapp.org/news/a/in-breyer-decision-today-europes-highest-court-rules-on-definition-of-personal-data/

From the perspective of Big Data analytics and the 'Internet of Things', it is crucial to understand that even when following a relative approach, data without any personal relevance may become personal data, when they are combined with other data:

> "[…] This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them."[173]

This clearly underlines that non-personal data can become personal instantly from the moment they enable a realistic chance of identification, directly or indirectly.[174] Therefore, data controllers and other parties engaging in Big Data analytics, are strongly advised to constantly monitor the status of the data they work with.[175]

Consequently, following a more favourable relative approach, the GDPR may still apply for data processed with privacy preserving methods, if there is a viable risk of identification by means used by the respective controller and a third person - but only if those means are reasonably likely to be used.[176] If the data is not reasonably likely to be disclosed by decryption or other technical measures, the data could be considered non-personal, as long as the affected person would not be identifiable.

2.3.1.6 Conclusion

The European mindset of data protection can best be thought of as a right of "informational self-determination", one fundamental right guaranteed every EU-citizen.[177] Safeguarding those fundamental rights in today's information society is a key issue in the current transition of the European data protection law.

An absolute approach would extend the scope of the GDPR essentially without any real or considerable boundaries.[178] The evaluation of the nature of the data would be definite and abstract, and as a result data would more frequently be qualified as personal data.[179]

Therefore, from the perspective of the authors of this report, in line with the prevailing majority of authors, strong arguments support the acceptance of a relative approach.[180] Following a relative concept does not result in a protection loophole, since it does not change the protected purpose, the right of informational self-determination and integrity of the data subject.[181] This pragmatic approach considers a reconciliation of interest of different stakeholders, accordingly, it simply imposes limitations on the scope of the data protection law, leaving the essence of data protection intact.[182] This would not contradict the purposes of the GDPR, since in scenarios where no realistic or reasonable chances to identify the data subject exist – in other words, the residual risk of identification is minimal – with respect to the processing in question, the protection offered by the GDPR is not affected at all. Hence it is unnecessary to apply the restrictions of data protection law for such processing activities.[183]

The GDPR itself does not settle this issue, however, there is a noticeable tendency of leaning on elements relativizing a strict absolute approach. With the reference for "another person" in Recital 26

---

[173] Recital 30 GDPR
[174] see more detailed in Spindler/Schmechel, JIPITEC 2016, 163 (168)
[175] Spindler, Medizinrecht 2016, 691 (695)
[176] Lang, K&R 2012, 145 (146).
[177] see 3.2 Legal Framework
[178] Meyerdierks, MMR 2009, 8 (10).
[179] Brink/Eckhardt, ZD 2015, 205 (206)
[180] Härting, (supra note 43) Rn. 264; Klar/Kühling, in: Kühling/Buchner (supra note 33) Art. 4 Rn. 25; Gola, in: Gola (supra note 109) Art. 4 Rn. 18; Brink/Eckhardt, ZD 2015 205 (205); Roßnagel/Scholz, MMR 2000, 721 (723); Meyerdierks, MMR 2009, 8 (8 ff.), Knopp, DuD 2015, 527 (529); Marnau, DuD 2016, 428 (428 ff.)
[181] Eckhard, CR 2011, 339 (343); Kroschwald, ZD 2014, 75 (76); Brink/Eckhard, ZD 2015, 205 (208)
[182] Brink/Eckhard, ZD 2015, 205 (208)
[183] Cf. Eckhard, CR 2011, 339 (342); Härting, ITRB 2009, 35 (37); Maisch, ITRB 2011, 13 (14).

the GDPR categorically refuses a one-sided relative concept, which would entirely exclude the relevance of the potential knowledge of a third party.[184] On the other hand, by including means reasonably likely to be used by the motivated parties, it acknowledges the importance of the particular context and the unique circumstances of a specific case.[185]

Nevertheless, controllers must bear in mind that in the era of Big Data and "Internet of things" the level of interconnectivity constantly carries the chance of identification through re-combining "harmless data".[186] Even if at the beginning of the processing activity the data had no personal relevance, controllers and processors should regularly check whether the data they use is still non-personal.[187] Also, with reasonable effort data originally relating to things can be brought into a direct ratio with a natural person, consequently, it may end up as personal data as well.[188]

The GDPR demands high requirements for the assessment.[189] Following the ECJ's as well as the Art.-29 Working Party's standpoint it is likely that strong de-identified data will not be considered personal data for those who do not possess the decryption key or other auxiliary information and do not have means reasonably likely to use those for identification. Thus, following a relative approach, state-of-the-art de-identification technique offering an adequate level of security may lead to a partial non-applicability of the GDPR.

2.3.2 Anonymity and Anonymous Data

Despite its tremendous practical relevance and strong incentive for its utilisation, neither the Directive nor the GDPR have a specific provision for "anonymous information".[190] It is merely mentioned in the recitals, respectively:

> "[…] whereas the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable;"[191]

> "principles of data protection should […] not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes."[192]

Anonymous information is presented as an opposite pole to personal data, as data with no personal reference whatsoever both in the Directive and in the GDPR.[193] Consequently, data protection law does not apply to information or combination of information that does not relate to or identify a natural person, and to the processing of such data.[194] Then again, the effective anonymization depends on the sole understanding of what constitutes personal data. As mentioned above, Recital 26 GDPR does not

[184] Brink/Eckhard, ZD 2015, 205 (209)
[185] in line with Art.-29 Working Party Opinion 04/2007, WP 136, p. 25
[186] Cf. Boehme-Neßler, DuD 2016, 419 (422); Raabe/Wagner, DuD 2016, 434 (435)
[187] Spindler, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 116.
[188] Spindler, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 116; Urgessa, (supra note 99) 530
[189] Marnau, DuD 2016, 428 (430);
[190] Härting (supra note 43) Rn. 290; Spindler/Schmechel, JIPITEC 2016, 163 (170)
[191] Recital 26 of the Directive
[192] Recital 26 s. 4, 5 GDPR
[193] Karg, DuD 2015, 520 (523); Knopp, DuD 2015, 527 (529); ebenso Klar/Kühling, in Kühling Buchner (supra note 33) Art. 4 Nr. 1 Rn. 31; cf. Schreiber, in Plath (supra note 33) Art. 4 Rn. 22
[194] Art. 29-Working Party, Opinion 05/2014, WP 216, p. 5 Gola, in Gola (supra note 109) Art. 11 Rn. 7; cf. Klabunde, in: Ehmann/Selmayr (supra note 49) Art. 11 Rn. 13; Oostveen: Identifiability and the Applicability of Data Protection to Big Data, International Data Privacy Law, 2016, 299 (307)

put an end to the dispute about the binary interpretation. However, by including a statement that it does not apply to anonymous data, it seemingly accepts the possibility that anonymous data, in fact, exists, which can be interpreted as another step towards a relative understanding of identifiability.[195]

When discussing the relation between anonymity and the GDPR, first and foremost, distinction must be made between 'anonymous' and 'anonymised" data, since the latter means such anonymous data that previously referred to an identifiable person, but where that identification is no longer possible.[196] When entering the Big Data value chain, data usually have some sort of personal reference, and therefore it is less common, that data processing falls outside the scope of the GDPR from the very beginning.[197] Essentially, the only way the acquisition of data falls outside of the scope of the GDPR is when a third party processes a dataset that was previously anonymised and released by the original data controller – or in that less likely situation, when data was collected anonymously.[198]

This gives rise to the question of whether anonymization, i. e. the de-identification process itself is a form of "further processing".[199] Art. 4 No. 2 GDPR offers a definition for "processing", however, it does not include the term "anonymization", or any technical requirements, how the anonymization have to be carried out. It focuses rather on the result of the de-identification process.[200] This could mean that if an art of de-identification cannot be subsumed under the term "processing", that specific technique will not be considered as "processing" within the provision of Art. 4 No. 2GDPR. This might be the case when data is de-identified with secret sharing, since 'alteration' within the provision of Art. 4 No. 2 GDPR refers to a change in content of the information, not in its appearance.[201]

The Art.-29 Working Party categorically handles anonymization as further processing: "anonymization constitutes a further processing of personal data; as such, it must satisfy the requirement of compatibility by having regard to the legal grounds and circumstances of the further processing".[202] This means that the process of anonymising must comply with the test of compatibility with the original purposes. The Art.-29 Working party considers anonymization "compatible with the original purpose of the processing, but only on condition the anonymization process is such as to reliably produce anonymised information".[203] In other words, anonymising personal data to reuse for purposes not compatible with the original purpose would be a violation of data privacy law, unless there are other legitimate grounds for processing.[204] It should be noted that with respect to statistical and research purposes, Art. 5 (1) b) GDPR offers an exception from the purpose limitation principle by stating that "further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89 (1), not be considered to be incompatible with the initial purposes".[205] De-identification methods are precisely those safeguards, what Art. 89 (1) requires. Therefore, anonymizing the dataset might be a "safe harbour" from the burdens of the GDPR.

Following the interpretation opted by the Art.-29 Working Party, anonymization must be in the current state of technology as permanent as erasure, i. e. making it impossible to process personal data.[206] Irreversibility is the key to anonymity.[207] Therefore, if the controller does not delete the original

---

[195] Härting (supra note 43) Rn. 291
[196] Art. 29-Working Party, Opinion 04/2007, WP 136, p. 21; cf. Esayas (supra note 96) p. 4
[197] Art.-29 Working Party Opinion 05/2014, WP 216, p. 8; cf. Gola, in Gola (supra note 109) Art. 4 Rn. 40
[198] Art.-29 Working Party Opinion 05/2014, WP 216, p. 10
[199] See El Emam/Álvarez, International Data Privacy Law 2015, p. 73 (79); Esayas, (supra note 96), p. 4 ff.
[200] cf. Herbst, in Kühling/Buchner (supra note 33) Art. 4 Nr. 2 Rn. 25; Klabunde, in Ehmann/Selmayr (supra note 49) , kArt. 4 Rn. 16; Hammer/Knopp, DuD, 2015, 503 (505); for a summary on basic anonymisation methods see: Angiuli/Blitzstein/Waldo: How to De-identify Your Data?, Privacy and Rights, 2015, Vol. 13 Issue 8, available at: https://dl.acm.org/citation.cfm?id=2838930
[201] Spindler/Schmechel, JIPITEC 2016, 163 (176)
[202] Art. 29-Working Party, Opinion 05/2014, WP 216, p. 3, 7
[203] Art. 29-Working Party, Opinion 05/2014, WP 216, p. 7
[204] Esayas (supra note 96), p. 4
[205] for more details (see section 2.2.2.3.2.)
[206] Art. 29-Working Party, Opinion 05/2014, WP 216, p. 6
[207] Kroschwald, ZD 2014, 75 (75); cf. El Emam/Álvarez, International Data Privacy Law 2015, p. 74

version of raw data, even the de-identified dataset stays personal data.[208] However, due to various reasons, e. g. data retention periods as legal duty, the controller cannot destroy the original data at "event-level".[209] This calls for a more flexible, relative method for managing the risks of identification, like the one emphasised by the Art.-29 Working Party, the "likely reasonably" used means.

On the other hand, over the past years it became clear that anonymized datasets can be re-identified, and the identity of specific individuals can be revealed.[210] In order to determine which anonymization methods are most likely able to "reliably produce anonymised information", the Art.-29 Working Party provided a guide on efficiency of the most common anonymization techniques.[211] According to this, the anonymization must be robust against three versions of identification threats:

1. Singling out: possibility to isolate records of an individual in the dataset
2. Linkability: ability to link, at least, two records concerning the same data subject or group of data subjects in the same database or in two different data bases
3. Inference: the possibility to deduce, with significant probability, the value of an attribute from the values of a set of other attributes.[212]

According to the thorough evaluation provided in the opinion, differential privacy is robust against all of the threats above, and therefore it might be the best method to prevent identification.[213] Apart from inference, re-identification and combining datasets which allow the emergence of patterns related to a single individual or a specific group another concern emerges from the sole understanding of Big Data analytics itself. Big Data Analytics mean the whole data management lifecycle, the acquisition, analysis and application of the data.[214] The same data or set of data may be anonymous during one phase of the processing, yet not in another. As mentioned above, there are limited cases, where the collection of data already falls outside of the scope of the GDPR. It is likely that data will be anonymized for the analytics phase, i. e. storage, computation, but the reverse is also possible. If a controller obtains datasets containing non-personal data and combines them, originally non-identifiable data might become identifiable.[215]

It is clear that anonymization is a beneficial way to exploit the value of Big Data. Nonetheless, the GDPR itself does not provide a comprehensive regulation on what exactly anonymization means. Recital 26 focuses on the result, not the technology itself, therefore a wide spectrum of anonymization techniques can be used by the controllers. As long as anonymization is state-of-the-art and secure enough to prevent re-identification, it can serve as a safe harbour from the regulations of the GDPR in its entirety. It is crucial however, that companies developing and utilizing anonymization techniques

[208] Art. 29-Working Party, Opinion 05/2014, WP 216, p. 9; Klabunde, in Ehmann/Selmayr (supra note 49) Art. 4 Rn. 16; critical El Emam/Álvarez, International Data Privacy Law 2015, p. 81
[209] El Emam/Álvarez, International Data Privacy Law 2015, p. 81 ff.
[210] Ohm, Broken promises of privacy: responding to the surprising failure of anonymization", 57 UCLA Law Review 2010, 1701 ff.
[211] Art. 29-Working Party, Opinion 05/2014, WP 216, p. 11 ff.; for a technical perspective on risks related to anonymising see: El Emam/Rodgers/Malin: Anonymising and sharing individual patient data, BMJ, 2015, 350 (350 ff.) available at: http://www.bmj.com/content/350/bmj.h1139
[212] Art. 29-Working Party, Opinion 05/2014, WP 216, p. 11, 12; see also: Ernst, in Paal/Pauly (supra note 19) Art. 4 Rn. 51; Hintze/El Emam: Comparing the Benefits of Pseudonymisation and Anonymisation Under the GDPR, p. 11, available at: https://iapp.org/resources/article/comparing-the-benefits-of-pseudonymization-and-anonymization-under-the-gdpr/
[213] Art. 29-Working Party, Opinion 05/2014, WP 216, p. 24
[214] ENISA, Privacy and Data Protection by Design, 2015, p. 11; cf. Oostveen: Identifiability and the Applicability of Data Protection to Big Data, International Data Privacy Law, 2016, 299 (231 ff.), who calls this a „three-phase-model" of Big Data
[215] Oostveen: Identifiability and the Applicability of Data Protection to Big Data, International Data Privacy Law, 2016, 299 (307)

constantly evaluate their method and the status of the data, since the contextual elements are different for each processing activity, and they might, in essence, influence whether data is personal or not – and thus the applicability of the GDPR.

2.3.3 Pseudonymisation and Encryption

Neither pseudonymisation nor encryption were regulated in the Directive as privacy preserving methods. However, the Art.-29 Working Party included pseudonymous and key-coded data in its opinion on personal data. One of the path-breaking innovations of the GDPR is that it included provisions on pseudonymisation and encryption. With that, it facilitates new methods at the time of growing need for secure data processing and ensuring a privacy-friendly utilization of Big Data.

2.3.3.1 Pseudonymisation

Recital 28 GDPR introduces the pseudonymisation as a method applied on personal data to "reduce the risk to the data subject and help controllers and processors to meet their data protection obligations". This shows the dual role pseudonymisation plays during the data processing activity. It is an appropriate tool for risk management[216], which aims to help controllers fulfil the obligations imposed by the GDPR.[217] However, the explicit introduction of pseudonymisation does not mean the preclusion of other privacy-enhancing measures.[218]

Unlike the Directive, the GDPR provides the definition of pseudonymisation in Art. 4 No. 5:

> "Pseudonymisation means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person."

When considering the use of pseudonymisation it must be considered that data treated with pseudonymisation does not cease to be personal, instead, it stays indirectly identifiable.[219] Therefore, pseudonymised information "cannot be equated to anonymised information".[220] The main difference between the two techniques is the existence of an assignment rule for pseudonymous data. This is underlined in Recital 26 as well, which states, that "personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person." Consequently, the GDPR's protective provisions continue to apply for pseudonymised data.

Three conditions have to be met cumulatively for a successful pseudonymisation. Firstly, pseudonymisation should exclude the assignment of data to an individual without involving the additional information. Secondly, the additional information, i. e. the assignment rule must be kept separately – technically and spatially[221] – from the data. The secure management of the additional information is the key of the provision. Thirdly, technical and organizational measures must ensure the safekeeping of the additional information and the non-assignability.[222]

---

[216] Härting (supra note 43) Rn. 299; Marnau, DuD 2016, 428 (430)
[217] Klar/Kühling, in Kühling/Buchner (supra note 33) Art. 4 No. 5 Rn. 2; Schreiber, in Plath (supra note 33) Art. 4 Rn. 21; Klabunde, in Ehmann/Selmayr (supra note 49) Art. 4 R.
[218] Recital 28 GDPR; see also: Klar/Kühling, in Kühling/Buchner (supra note 33) Art. 4 No. 5 Rn 4; Ernst, in Paal/Pauly (supra note 19) Art. 4 Rn. 40
[219] Art. 29-Working Party, Opinion 04/2007, WP 136, p. 18
[220] Art. 29- Working Party, Opinion 05/2014, WP 216, p. 10; Spindler/Schmechel, JIPITEC 2016, 163 (170)
[221] Klar/Kühling, in Kühling/Buchner (supra note 33) Art. 4 No. 5 Rn. 6
[222] Ernst, in Paal/Pauly (supra note 19) Art. 4 Rn. 42 ff.

The application of pseudonymisation also affects the compliance of the controller with the GDPR, since in some provisions it appears as a requirement for the lawfulness of the processing.[223] It is listed as a compliance criterion in Art 6 (4) e) GDPR what controllers must take into account while carrying out the compatibility assessment[224]; Art. 25 (1) imposes the obligation of implementing privacy by design on the controllers and names pseudonymisation as one method to do so; Art. 32 (1) lit a) along with Recital 78 consider pseudonymisation as a measure to ensure the required level of security; and within the provision of Art. 89 (1) pseudonymisation is also considered to be an appropriate safeguard. According to Art. 40 (2) d) associations and other bodies representing categories of controllers and processors may prepare codes of conduct in order to contribute to the proper application of the GDPR.[225] On the other hand, pseudonymisation may also free controllers from the obligation to communicate personal data breach to the data subject if they can eliminate its risk by using technical organizational measures, according to Art. 34 (2) a) GDPR, which is another incentive for controllers to apply pseudonymisation.[226]

2.3.3.2 Encryption

The GDPR does not offer a legal definition of encryption.[227] Encryption can be best explained as a technical security measure, the process of changing plaintext in to unintelligible code.[228] Also the GDPR mentions encryption in those provisions, which in one way or another regulate the security of processing. It is supposed to mitigate the inherent risks as well as to enhance the integrity and confidentiality of the processing.[229]

According to Art. 32 (1) a) GDPR in consistence with Recital 83, encryption serves as a technical organizational measure to ensure the security of processing. Art. 34 (3) a) GDPR lists encryption as one of the exemptions from the data breach notification obligation towards the data subject, provided that it renders "the personal data unintelligible to any person who is not authorized to access it". This could encourage controllers to encrypt their data, since they can directly benefit from it. Furthermore, Art. 6 (4) e) GDPR includes encryption – along with pseudonymisation – in the factors controllers have to take into account while carrying out a compatibility assessment when they intend to use the data for a purpose other than that the data have been originally collected for.

All of the abovementioned provisions treat encryption as a compliance requirement, within the scope of the GDPR, precisely as a method to fulfil the obligations of data protection law. A systematic interpretation points to this direction as well, hence regulating encryption in the GDPR suggests that the GDPR continues to apply for encrypted data.

An important question in this respect is whether encrypted data can be classified as pseudonymous data or anonymous data.[230] The basic idea behind encrypting personal data is to render plaintext into ciphertext in a way that the result, the encrypted text is unreadable without the proper key to decrypt

[223] Härting (supra note 43) Rn. 305
[224] see section 2.2.2.3.2.
[225] Marnau, DuD 2016, 428 (431)
[226] Klar/Kühling, in Kühling/Buchner (supra note 33) Art. 4 No. 5 Rn. 13; Reif, in Gola (supra note 109) Art. 34 Rn. 8
[227] Article 4 No. 2b of the proposal of the European Parliament for a GDPR (LIBE proposal) defined encrypted data as "personal data, which through technological protection measures is rendered unintelligible to any person who is not authorised to access it", Within the meaning of this article, encrypted data is a subcategory of personal data, consequently, according to the Parliament's standpoint, encrypted data shall not lose its personal reference
[228] Esayas (supra note 96) p 10; Knopp, DuD 2015, 542 (542); Martini, in Paal/Pauly, Datenschutz-Grundverordnung, Art. 32 Rn. 34
[229] Tinnefeld/ Buchner/ Petri, Einführung in das Datenschutzrecht, 2012, 5. edition, München, p. 435; Jandt, in Kühling/Buchner (supra note 33) Art. 32 Rn. 19
[230] Pordesch/Steidle, DuD 2015, 536 (538)

it.[231] This scheme fits in with Art. 4 No. 5 GDPR, insofar as the "additional information" it requires is the decryption key, which has to be kept separately and protected by appropriate technical and organisational measures.[232] Secure key management and adequate encryption technique are crucial in order to "ensure that the personal data are not attributed to an identified or identifiable natural person".[233] For this reason, i. e. the existence of an assignment rule in form of a decryption key, encryption is generally considered to be one type of pseudonymisation.[234] This would contradict the classification of encryption as anonymization, since the Art.-29 Working Party explicitly expressed in its opinion, that a "specific pitfall is to consider pseudonymised data to be equivalent to anonymised data."[235]

However, this approach does not take into account that the personal reference of the data is always context-related, and depends on the means and knowledge of the specific controller. In other words, with accepting a rather relative approach – or, as argued above, at least an absolute approach with relative elements[236] – determining the nature of encrypted data is much more complex than that. Whether or not encrypted data stays personal for the party unable to decrypt it, depends on how "identifiability" is defined.[237] Therefore, despite assuming that encryption is a subcategory of pseudonymisation, encrypted data may still be regarded non-personal.

As long as the decryption key to the original data – plaintext – is available, the possibility of identifying, even if merely theoretically, but exists.[238] No encryption system can eliminate the identifiability aspect of the information in its entirety.[239] Consequently, following an absolute approach, encrypted data will never cease to be personal, since even in case of a state-of-the-art, secure encryption, there will be at least one person – the key holder – always being able to decrypt the data. Following an absolute interpretation, the role of encryption, by definition, is to secure the processing within the frames of the GDPR. Encryption in this sense is a requirement of compliance with data protection law, but not a safe harbour from the obligations imposed by it.

On the other hand, the relative approach is more permissive towards privacy preserving techniques, and implies, that when a given encryption guarantees a sufficient level of security, encrypted personal data can be considered anonymous. As discussed above, Recital 26 GDPR states, that what needs to be taken into account are "all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly".[240]

Nevertheless, this concept does not deny the fact, that the party possessing the auxiliary information has sufficient economic resources and computing power to decrypt the data. Therefore, even following this logic, encrypted data should remain personal data at least to the party who holds the decryption key.[241] The relevant question here is, whether encrypted data is still personal to other parties, who have no access to the decryption key whatsoever.

There are several objective factors to consider when assessing whether the applied privacy preserving technique can be regarded as computationally secure or not. The three most relevant are the strength of the encryption method used, the length of the encryption key (the longer the key, the safer the encryption is against attacks), and the management of the decryption key, i. e how securely the key is

[231] Spindler/Schmechel, JIPITEC 2016, 163 (174 ff.) and Tinnefeld/ Buchner/ Petri (supra note 227) p. 435 ff. provide a summary on encryption techniques
[232] Spindler/Schmechel, JIPITEC 2016, 163 (171)
[233] Art. 4 No. 5 GDPR
[234] Esayas (supra note 96) p 10; Marnau, DuD 2016, 428 (431); Spindler/Schmechel, JIPITEC 2016, 163 (171); Art.-29 Working Party, WP216, p. 20; cf. Kühling/Raab, in Kühling/Buchner (supra note 33) Art. 4 No. 5 Rn. 10; different opinion: Knopp, DuD 2015, 542 (543) who unequivocally rejects the equating of encryption to pseudonymisation
[235] Art.-29 Working Party, WP216, p. 10
[236] see section 2.3.1
[237] see section 2.3.1
[238] Art.-29 Working Party, WP216, p. 29;
[239] Esayas,(supra note 96), p 10
[240] see section 2.3.1
[241] Hoeren (ed.): Big Data und Recht, 2014, München, p. 80; Spindler/Schmechel, JIPITEC 2016, 163 (171)

stored, how many people have access to it, etc.[242] These together must prevent any unauthorised access to data.

The most obvious ways of decryption are exhaustive key search or brute-force attack, that is, to try all possible keys by trial and error until eventually guessing correctly.[243] With an effective encryption however, the possibility that a dataset can be decrypted this way, does not seem to be particularly likely.[244] There are several legal ways for accessing the original data in plaintext, e. g. via court order, extracting the key from a software or hardware, or by using accidental errors or systematic backdoors implemented in the algorithm.[245] These methods are considered likely reasonably to be used, if the motivated party has computational power and other resources proportional to the necessary expenses.

Aside from these, if a controller can effectively encrypt the data by taking appropriate technical-organisational measures to prevent re-identification, the data may be treated as anonymous data in the hands of a third party – despite the theoretical possibility of re-identification.

Using privacy-preserving technologies developed by SODA are promising to result in anonymous data for the processor or third parties. In a Multi-Party Computation[246] scenario, data are encrypted via secret sharing by the input holders, before the transmission to other parties occur. By computing on the input shares, information cannot be extracted during the analysis phase about the data or the computation output.[247] Since in this case data are transmitted in an encrypted form, the party performing the computation – research institutions or other stakeholders – does not have access to the raw data nor to the decryption key. This may remove them from the data protection legislation provided that the encryption was performed to industry standards and to the best of knowledge by the controller prior to transmission, and the controller implements secure key management methods.[248] However, it must be kept in mind that when previously encrypted data are reconstructed or transmitted back to the controller holding the decryption key after computation, data will be "personal data" again.[249]

Anonymity is relative in terms of time as well; de-identified data will only stay anonymous for a period of time. Thus, according to Recital 26 GDPR, controllers must constantly monitor the available encryption technologies and technical developments in order to keep their technique up to date.[250] This obligation includes also evidentially foreseeable future technical developments.[251] In cases where controllers acquire fully encrypted dataset, they have to obtain further information regarding whether the original dataset included personal data, and if so, they are obliged to regularly check the state-of-the-art of the encryption technique.[252]

---

[242] Hon/Millard/Walden, Queen Mary University of London – Legal Studies Research Paper No. 75/2011, pp. 9, 23

[243] See Gürses/Preenel, in: van der Sloot/Broeders/Schrijvers (eds.), Exploring the Boundaries of Big Data, 2016, Part I, 3, Cryptology and Privacy in the Context of Big Data, 49 (62) for more examples

[244] Spindler/Schmechel, JIPITEC 2016, 163 (172)

[245] Spindler/Schmechel, JIPITEC 2016, 163 (172); Gürses/Preenel, in: van der Sloot/Broeders/Schrijvers (eds.), Exploring the Boundaries of Big Data, 2016, Part I, 3, Cryptology and Privacy in the Context of Big Data, 49 (62)

[246] for the technical functioning of Multi-Party Computation see: Lindell/Pinkas: Secure Multiparty Computation for Privacy-Preserving Data Mining, Journal of Privacy and Confidentiality: 2009, Vol. 1 : Iss. 1 , Article 5. Available at: http://repository.cmu.edu/jpc/vol1/iss1/5

[247] Yakoubov/Gadepally/Schear/Shen/Yerukhimovich: A Survey of Cryptographic Approaches to Securing Big-Data Analytics in the cloud, available at: http://www.ieee-hpec.org/2014/cd/index_htm_files/finalpapers/28.pdf ; Gürses/Preenel, in: van der Sloot/Broeders/Schrijvers (eds.), Exploring the Boundaries of Big Data, 2016, Part I, 3, Cryptology and Privacy in the Context of Big Data, 49 (60)

[248] Hon/Millard/Walden, Queen Mary University of London – Legal Studies Research Paper No. 75/2011, p. 25

[249] see section 3.3.2

[250] Article 29 Data Protection Working Party, WP216, p. 29 outlines that „chosen encryption scheme with a given key size is designed to ensure confidentiality for a given period (most of the current keys will have to be resized around 2020)"

[251] Spindler/Schmechel, JIPITEC 2016, 163 (173)

[252] Article 29 Data Protection Working Party, WP 216, p. 10

Hence, encrypted data may be considered as anonymised data under the GDPR for anybody other than the key holder, provided that the controller implies appropriate technical measures to prevent the disclosure of the decryption key and the original data. Nevertheless, the possibility of re-identification always has to be considered, on a case-by-case basis, if the means used for identification are reasonably likely.

3 Requirements for the Lawful Processing of Personal Data

3.1 Requirements for the Lawful Processing

3.1.1 The Definition of Processing

Art. 4 No. 2 GDPR defines processing as:

> "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction."

The GDPR introduces minor changes to the wording of the Directive's definition of "processing". It adds "structuring" to the list of examples and changes "blocking" for "restriction". This term for "processing" is noticeably broad, and encompasses essentially any activity that is done to or with personal data. The provided list of examples is not exhaustive, which means that other, in Art. 4 No. 2 not specified forms may also appear as "processing". The definition focuses on the data processing activity as a whole instead of the separate steps of the processing.[253] From the moment the data had been collected, the GDPR applies to every single step of the processing activity.

The only option for controllers to escape the applicability of the GDPR is to render the data "non-personal". Stripping the data from the individual may result in "feature data", which are not subject to data protection law. Anonymisation technically means the alteration of the data. Alteration falls under the GDPR only, if it means changing the content of the information, but not its appearance.[254] If the data controller uses methods where the de-identification leaves the content of the data intact, by dividing the data into shares, or adding random noise to it, the anonymization process itself may fall outside the scope of the GDPR.[255] Another way to avoid data protection law from the very beginning is to obtain already anonymised data.

Therefore, applying privacy preserving methods that render the data anonymous, such as the one being developed in SODA can lead to a complete or partial non-applicability of the GDPR.

3.1.2 Legitimate Grounds of Processing

The principle of prohibition with the reservation of authorisation is of outstanding importance in the GDPR.[256] It demands that each and every processing activity carried out by the data controller concerning personal data must have a legitimate basis.[257] Without a legitimate basis, the processing of personal data is prima facie unlawful.

[253] Härting (supra note 43) Rn. 331
[254] Herbst, in Kühling/Buchner (supra note 33) Art. 4 Nr. 2 Rn. 25; Spindler/Schmechel, JIPITEC 2016, 163 (176)
[255] see section 3.3.2
[256] Buchner/Petri, in Kühling/Buchner (supra note 33) Art. 6 Rn. 11 ff; Schulz, in Gola (supra note 109) Art. 6 Rn. 2; Heberlein, in: Ehmann/Selmayr (supra note 49) Art. 6 Rn. 1; Frenzel, in Paal/Pauly (supra note 19) Art. 6 Rn. 1
[257] cf. Recital 40 GDPR

3.1.2.1 Consent or Explicit Legal Permission

The GDPR does not introduce significant changes, Art. 6 (1) GDPR essentially corresponds with Art. 7 of the Directive.[258] It enumerates six legal bases:

a. "the data subject has given **consent** to the processing of his or her personal data for one or more specific purposes;

b. processing is necessary **for the performance of a contract** to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;

c. processing is necessary **for compliance with a legal obligation** to which the controller is subject;

d. processing is necessary in order to **protect the vital interests** of the data subject or of another natural person;

e. processing is necessary **for the performance of a task carried out in the public interest** or **in the exercise of official authority vested in the controller**;

f. processing is necessary for **the purposes of the legitimate interests** pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child."

The list is definitive and exhaustive[259], the processing may only be justified on the basis of one of the enlisted permissions. Data controllers cannot invoke any other circumstance. In accordance with the principle of purpose limitation laid down in Art. 5 (2) irrespective of the lawful basis, processing is only permitted if, and to the extent that it is necessary for certain purposes. The order, especially regulating consent on the first place, does not mean any hierarchic or privileged relation, all the lawful bases are of equal rank.[260]

Concerning the relation between the lawful grounds, it must be noted that one processing activity cannot be based on multiple lawful bases. Data controllers must identify the appropriate lawful basis in advance, and are not allowed to "swap" between lawful bases.[261]

Pursuant to Art. 6 (1) lit a) consent continues serving as a lawful basis for processing personal data, however, under the GDPR, a valid consent becomes more difficult to obtain. Art. 6 (1) lit b) permits the processing if it is a contractual necessity, or if it is necessary "for the intention to enter into a contract".[262] It covers all kinds of claims between private entities, yet not the statutory claims. Art. 6 (1) lit c) can be invoked in cases where personal data is processed on the basis that the controller has a legal obligation to perform the processing. Art. 6 (1) lit d)-e) allow the processing on the basis of vital interest of the data subject or of another natural person, and for performances of tasks carried out by an official authority or private organisations for the public interest, respectively. The GDPR extends the condition "vital interest" to other individuals too (e. g. family members of the data subject), although it does not

[258] for an overview and comparison see: Ursic/Custers, Legal Barriers and Enablers to Big Data Reuse, 2 European Data Protection Law Review, 2016, 209 (211 ff)
[259] Buchner/Petri, in Kühling/Buchner (supra note 33) Art. 6 Rn. 1
[260] Schulz, in Gola (supra note 109) Art. 6 Rn. 10
[261] Art.-29 Working Party, Guidelines on Consent under Regulation 2016/679, adopted on 28 November 2017, WP 259, p. 22
[262] Recital 44 GDPR

define "vital interest".[263] According to Art. 6 (3) in line with Recital 45 GDPR, the processing should have a basis in European or Member State law, when the processing of personal data takes place on the bases of Art. 6 (1) lit c) and lit e).

Art. 6 (1) lit f) stipulates that personal data may be processed in cases where the controller has a legitimate interest in processing those data. This is the broadest permission clause in the GDPR, the only limitations are the overriding interest as well as the fundamental rights and freedoms of the affected data subjects.[264] It is based on an extensive balancing between the interests of the stakeholders on the opposite sides, the controller or a third party and the data subject.[265] Recital 47 gives an exemplary list of what constitutes legitimate interest, which means that other legal, economic or non-material interests may also justify the processing within the provisions of Art. 6 (1) lit f).[266] The ECJ addressed a similar conflict of interest in its recent decision, where a Spanish citizen brought a suit against Google Spain and Google Inc. to remove his personal data from its search hit list and to prevent access to those data in the future.[267] The court concluded that a pure economic interest of a search engine cannot justify such processing, hence the rights of the data subject override the economic interest of the search engine as well as the interests of the general public.[268] Although this case exclusively dealt with personal data processed by a search engine, the reasoning the ECJ that financial interest itself would not outweigh the rights and freedoms of the data subject my apply on a wider spectrum. Moreover, it should be noted, that Recital 47 outlines:

> "At any rate the existence of a legitimate interest would need careful assessment including whether a data subject can reasonably expect at the time and in the context of the collection of the personal data that processing for that purpose may take place. The interests and fundamental rights of the data subject could in particular override the interest of the data controller where personal data are processed in circumstances where data subjects do not reasonably expect further processing."

The reasonable expectations of the data subject can also influence the lawfulness of the processing on the basis of the legitimate interest of the controller. The wording "could in particular override" calls for a case-by-case assessment, but does not automatically imply unlawfulness.[269]

In all likelihood, the GDPR will essentially permit the processing of personal data for Big Data analytics if the data subject has consented to such processing, or on the basis that the controller has a legitimate basis provided that this is not overridden by the rights and freedoms of the data subject. These are most suitable for justifying data collection – primary – for research purposes. Furthermore, although research is not specifically mentioned, Recital 157 sketches the potential benefits of personal data research as obtaining "essential knowledge about the long-term correlation of a number of social condi-

[263] Roßnagel/Nebel/Richter; ZD 2015, 455 (457); cf. Schulz, in Gola (supra note 109) Art. 6 Rn. 45 who argues that vital interest encompasses existential interest related to healthcare, and is not limited to life-threatening danger.
[264] Tikkinen-Piri/Rohunen/Markkula, EU General Data Protection Regulation: Changes and implications for personal data collecting companies, Computer Law & Security Review 2017, doi: 10.1016/j.clsr.2017.05.015
[265] Frenzel, in Paal/Pauly (supra note 19) Art. 6 Rn. 31; cf. Buchner, DuD 2016, 155 (159), who highlights the legal uncertainty caused by the fact that the GDPR does not provide criteria to take into account while carrying out the test
[266] Schulz, in Gola (supra note 109) Art. 6 Rn. 51
[267] ECJ, Judgment from the 13th May 2014 in Casec-131/12 – Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez.
[268] ECJ, Judgment from the 13th May 2014 in Casec-131/12 – Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez, par. 97.; cf. Buchner/Petri, in Kühling/Buchner (supra note 33) Art. 6 Rn. 171
[269] cf. Schulz, in Gola (supra note 109) Art. 6 Rn. 57

tions". This implies that in absence of the data subject's consent, research itself may under certain circumstances furnish a legitimate basis for the processing as "legitimate interest".[270] Moreover, Recital 47 considers that also the processing of personal data for direct marketing purposes may be regarded as carried out for a legitimate interest of the controller or third party.

3.1.2.2 Requirements of a Valid Consent

The GDPR does not change the principle that consent may provide a lawful basis for data processing, however, it makes substantially more difficult for data controllers to obtain valid consent by intensifying the conditions. The requirements should be strict, since the data subject's consent is a direct manifestation of the right of informational self-determination, which is anchored in the Art. 7, 8 of the Charter of Fundamental Rights of the European Union.[271]

Unlike the GDPR, the Directive merely demanded that data subject must "signify" the consent.[272] Beside amending this definition, the GDPR outlines the conditions for a valid consent in an additional provision. Art. 7, 8 in accordance with Recitals 32, 33, 42 and 43 provide guidance as how controllers must act in order to comply with the consent requirements.

3.1.2.2.1 Conditions for Consent

Art. 4 No. 11 GDPR defines consent as:

> "any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her."

The core of the element "freely given" is the fact that the data subjects must have a genuine choice to accept or decline the terms of processing offered to them.[273] If there is any sign of compulsion, undue pressure or if negative consequences arise from the declination, the consent will not be valid.[274] This is highlighted also in Recital 42 which clarifies that "consent should not be regarded as freely given if the data subject has no genuine or free choice or is unable to refuse or withdraw consent without detriment". Recital 43 addresses the situation where there is a clear imbalance between the data subject and the controller. This is often a case whenever a controller is a public authority, since in such cases the data subject does not have any realistic alternative to accepting the terms.[275] Additionally, Recital 43 states that consent "is presumed not to be freely given if it does not allow separate consent to be given to different personal data processing operations despite it being appropriate in the individual case". Such a granularity of consent is extremely important in case if a processing activity involves multiple processing operations for more than one purpose. All processing activities carried out for the same purpose or purposes should be covered by consent, as outlined by Recital 32. When the processing has multiple purposes, consent should be given to all of them, respectively. Data subject must be granted the possibility to choose to accept or refuse the processing rather than having to consent to a "package deal". This is also closely related to the requirement of specificity.

Another important provision in connection to the freedom of choice is the prohibition of conditionality established by Art. 7 (4):

---

[270] cf. Maldoff, How GDPR changes the rules for research, available at: https://iapp.org/news/a/how-gdpr-changes-the-rules-for-research/

[271] Art.-29 Working Party, WP 259, p. 4; Buchner, DuD 2016, 155 (158); Buchner/Kühling, in Kühling/Buchner (supra note 33) Art. 4 Nr. 11 Rn. 5;

[272] Art. 2 lit. h) in accordance with Art. 7 lit. a) of the Directive

[273] Art.-29 Working Party, WP 259, p. 4, 6; Schaar, ZD 2017, 213 (214)

[274] Art.-29 Working Party, 15/2011 Opinion on the definition of consent, WP187, p. 12

[275] Art.-29 Working Party, WP 259, p. 7; Ziegenhorn/von Heckel, NVwZ 2016, 1585 (1587); Buchner, DuD 2016, 155 (158)

> "When assessing whether consent is freely given, utmost account shall be taken of whether, *inter alia*, the performance of a contract, including the provision of a service, is conditional on consent to the processing of personal data that is not necessary for the performance of that contract."

Two scenarios are regulated here, the "bundling" consent with acceptance of terms and conditions and "tying" the provision of a contract or service to a request for consent to process personal data that are not necessary for the performance of that contract or service.[276] Recital 43 backs this provision by adding that "if it does not allow separate consent to be given to different personal data processing operations despite it being appropriate in the individual case, or if the performance of a contract, including the provision of a service, is dependent on the consent despite such consent not being necessary for such performance." Consequently, wherever it is possible, data controllers should avoid making the performance of the contract conditional upon the data subject's consent, hence the consent will be invalid if it is tied to the performance of a contract which is not necessary for the original purpose of the consent. The prohibition of such linking applies not only vertical, but also in horizontal relations too, e. g. where the consent cannot be obtained to different processing operations although it should have been obtained.[277]

Essentially, if the processing was based on the consent of the data subject, withdrawal or achieving the purpose of processing are the only way to terminate the processing of personal data.[278] The GDPR does not specify any time limitation for how long the given consent will last.[279] Therefore, a freely given consent, by nature, must be capable of being withdrawn.[280] It will likely result in an invalid consent, if the data controller does not permit the refusal or withdrawal of consent without detriment.[281] This right of the data subject is regulated in Art. 7 (3) GDPR. Data subject must be able to exercise this right at any time, in a way as easy as to give consent. However, this right is not retrospective, as it shall not affect the lawfulness of processing based on the same consent before the withdrawal.[282] Data subject shall be informed about their right of withdrawal ex ante, prior to giving consent. A safe way to ensure this could be for controllers to include a phrase "*valid until revoked*" in the consent form.[283]

A further element for a consent to be valid is that it must be "specific", that is, it can be given to "one or more specific purposes". The GDPR itself does not explain this term further. It is closely linked to the principle of purpose limitation in Art. 5 (1) lit b) on the one hand, and to the consent requirements "informed", "freely given", and "granular" on the other.[284] Due to this condition, blanket consents and the use of catch-all phrases will result in invalidity.[285] In order to be specific, consent must be intelligible, that is, clear and precise about the scope and consequences of the processing.[286] The more serious the interference with someone's private sphere is, the more precise the purpose of processing shall be explained.[287] Controllers must fulfil three conditions for providing a specific consent, namely the purpose specification, granularity, and clear separation of information related to obtaining consent from

[276] Art.-29 Working Party, WP 259, p. 9
[277] Gola, in Gola (supra note 109) Art. 4 Rn. 68 in conjunction with Schulz, in Gola (supra note 109) Art. 7 Rn. 22
[278] Schulz, in Gola (supra note 109) Art. 7 Rn. 58
[279] Art.-29 Working Party, WP 259, p. 20
[280] Schaar, ZD 2017, 213 (214)
[281] Recital 42 GDPR; Art.-29 Working Party, WP 259, p. 11
[282] Schulz, in Gola (supra note 109) Art. 7 Rn. 54
[283] Schulz, in Gola (supra note 109) Art. 7 Rn. 58
[284] Art.-29 Working Party, WP 259, p. 12
[285] Ernst, ZD 2017, 110 (113)
[286] Art.-29 Working Party, 15/2011 Opinion on the definition of consent, WP187, p. 17; Tikkinen-Piri/Rohunen/Markkula, (supra note 262) p. 6
[287] Ernst, ZD 2017, 110 (113)

information about other matters.[288] To that end, Art. 7 (2) requires that in cases if the data subject's consent is given in the context of such a written declaration that also concerns different matters, the request for consent shall be presented in a manner which is "clearly distinguishable from the other matters".

With respect to certain areas of scientific research, including certain analytics on Big Data, it should be noted that it is often not possible to fully set out the purpose of processing personal data at the time of the data collection. Therefore, the GDPR acknowledges that "data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research".[289] In order to comply with the GDPR controllers must do their best to ensure that the "essence" of the consent requirements are served best. In research, consent for subsequent steps can be obtained before the next stage begins, provided that this is in line with the relevant ethical standards.[290] However, data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose.[291]

The third element for a valid consent is that the consent must be "informed". The GDPR requires controllers to take significant steps in order to ensure that data subjects are provided with sufficient information. Data subjects must be provided with information prior to obtaining their consent. The two requirements data controllers must accomplish to ensure appropriate information are the quality of the information as well as the accessibility and visibility of the information.[292] The first addresses the issue of what kind of information must be provided, while the second deals with the conditions on how to provide information.

As far as the minimum content requirements are concerned, Recital 42 clarifies that "for consent to be informed, the data subject should be aware at least of the identity of the controller and the purposes of the processing for which the personal data are intended". In practice, the elements that are crucial for data subjects in order to understand what they give consent for are:

1. the controller's identity – in case of multiple/joint controllers, all controllers should be named
2. the purpose of each of the processing operations for which consent is sought
3. the type of data that will be collected and used
4. the existence of the right to withdraw consent
5. information about the use of the data for decisions based solely on automated processing, including profiling, in accordance with Art. 22 (2)
6. the possible risks of data transfer to third countries in the absence of an adequacy decision and appropriate safeguards.[293]

Transparency itself is not defined in the GDPR. Article 12 is the key article in the GDPR that regulates the framework of transparency obligations. It outlines the material requirements on transparency first, followed by rules on modalities.[294] In addition, Recital 39 states that the processing "should be transparent to natural persons that personal data concerning them are collected, used, consulted or otherwise processed and to what extent the personal data are or will be processed. The principle of transparency

[288] Art.-29 Working Party, WP 259, p. 12, which basically repeats the Working Party's opinion outlined in Art.-29 Working Party, 15/2011 Opinion on the definition of consent, WP187, p. 17
[289] Recital 33 GDPR; Klabunde, in Ehmann/Selmayr (supra note 49) Art. 4 Rn. 36
[290] Art.-29 Working Party, WP 259, p. 28
[291] Recital 33 GDPR
[292] Art.-29 Working Party, 15/2011 Opinion on the definition of consent, WP187, p. 20
[293] Art.-29 Working Party, WP 259, pp. 13, 14
[294] Bäcker, in Kühling/Buchner (supra note 33) Art. 12 Rn. 5 ff

requires that any information and communication relating to the processing of those personal data be easily accessible and easy to understand, and that clear and plain language be used. That principle concerns, in particular, information to the data subjects on the identity of the controller and the purposes of the processing and further information to ensure fair and transparent processing in respect of the natural persons concerned and their right to obtain confirmation and communication of personal data concerning them which are being processed."

Required information or method of communication	Relevant provision	Appropriate action
Requirements for transparency		
concise, transparent, intelligible	Art. 12 (1) s. 1 Recital 39	- efficient and succinct presentation to avoid information fatigue - understandable language for an average member of the intended audience - offering a description of consequences, especially in case of complex processing
clear and plain language	Art. 12 (1) s. 1 Recital 39	- in as simple manner as possible - avoid wording such as "may", "might", "some", "possibly", "often" - precisely, but no legal jargon
clear and plain language for children	Art. 12 (1) s. 1 Recital 38	- consider the vocabulary, tone and style of a child-friendly language
visualisation tools	Art. 12 (7); Art. 12 (8) Recital 60	- use of standardised icons in combination with the information, not instead of it - electronically or in a machine-readable form (e. g. QR code, physical paperwork)
Modalities		
easily accessible form	Art. 12 (1) s. 1	- making the information immediately apparent (e. g. providing directly, linking, FAQ, chatbot) - in a form compatible with all browsers
in writing or by other means, including where appropriate, by electronic means	Art. 12 (1) s. 2; Art. 12 (3) s. 4	- default position is in writing - by electronic form means where appropriate or where the data subject requests in a clear and concise way - in a method appropriate to the circumstances, e. g. how the controller and data subject interact

		- unspecified "means" may include: "just-in-time" contextual pop-up notices, videos, smartphone apps or voice alerts, infographics
may be provided orally	Art. 12 (1) s. 3	- on request, provided that the identity information is proven by other - person-to-person basis - automated oral information, e.g. pre-recorded message - provide opportunity to re-listen – imperative for visual impaired data subjects
obligation to provide information to the data subject on request	Art. 12 (3)	- provide information without undue delay, but in any event in 1 month - in case of complexity or large number of requests exceptionally in 2 months
obligation to provide information to the data subject on request if the controller does not take action	Art. 12 (4)	- provide information without delay but not later than in 1 month of receipt - explain the reasons for not taking action - refer to the possibility of lodging complaint and seeking judicial remedy
free of charge	Art. 12 (5)	- transparency requirements cannot be made conditional upon payment, extra charge or purchase of goods or services

Table 2. Art. 12 GDPR on transparency requirements[295]

The information to be provided to the data subject and the corresponding obligations of the data controller are regulated in Article 13 and Article 14 GDPR. Art. 13 GDPR lists the categories of data to be provided in cases where the data are collected directly from the data subject, while Art. 14 refers to processing of personal data where the data are obtained from another source. The GDPR brings substantial changes and additions regarding the information provision to the data subject.[296] Article 13 (1) extends the corresponding Article 10 of the Directive, and clarifies that "data controllers shall at the time when personal data are obtained[297], provide the data subject with:

a) "the identity and the contact details of the controller and, where applicable, of the controller's representative;

[295] based on Art.-29 Working Party, WP 259, pp. 13, 14; Art.-29 Working Party Guidelines on transparency under Regulation 2016/679 WP 260 pp. 7 ff.; Bäcker, in: Kühling/Buchner (supra note 33) Art. 12 Rn. 9 ff.; Franck, in Gola (supra note 109) Art. 12 Rn. 13 ff.

[296] Tikkinen-Piri/Rohunen/Markkula, (supra note 262) p. 7

[297] Art.-29 Working Party, WP260, p. 14 points out that this generally refers to a "reasonable period" after obtaining the personal data,

b) the contact details of the data protection officer, where applicable;

c) the purposes of the processing for which the personal data are intended as well as the legal basis for the processing;

d) where the processing is based on point (f) of Article 6 (1), the legitimate interests pursued by the controller or by a third party;

e) the recipients or categories of recipients of the personal data, if any;

f) where applicable, the fact that the controller intends to transfer personal data to a third country or international organisation and the existence or absence of an adequacy decision by the Commission, or in the case of transfers referred to in Article 46 or 47, or the second subparagraph of Article 49 (1), reference to the appropriate or suitable safeguards and the means by which to obtain a copy of them or where they have been made available."

Besides, in order to provide a fair and transparent processing, according to Art. 13 (2) controllers must provide information on:

a) "the period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period;

b) the existence of the right to request from the controller access to and rectification or erasure of personal data or restriction of processing concerning the data subject or to object to processing as well as the right to data portability;

c) where the processing is based on point (a) of Article 6 (1) or point (a) of Article 9 (2), the existence of the right to withdraw consent at any time, without affecting the lawfulness of processing based on consent before its withdrawal;

d) the right to lodge a complaint with a supervisory authority;

e) whether the provision of personal data is a statutory or contractual requirement, or a requirement necessary to enter into a contract, as well as whether the data subject is obliged to provide the personal data and of the possible consequences of failure to provide such data;

f) the existence of automated decision-making, including profiling, referred to in Article 22 (1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject."

Together, Recital 62 and Art. 13 (3) state that the only exceptions to the obligation to provide information are processing activities where the data subject is already aware of the information, or where the recording or disclosure of the personal data is expressly laid down by law or where the provision of information to the data subject proves to be impossible or would involve a disproportionate effort.

As mentioned before, it is common that data controller use data acquired from sources other than the data subject, e. g. from data brokers, third party controllers, publicly available sources or other data subjects.[298] However, this does not affect the obligation to provide the data subject with information listed in Art. 14 (1) -(2) GDPR:

a) "the identity and the contact details of the controller and, where applicable, of the controller's representative;

[298] Art.-29 Working Party, WP260, p. 14

b) the contact details of the data protection officer, where applicable;

c) the purposes of the processing for which the personal data are intended as well as the legal basis for the processing;

d) the categories of personal data concerned;

e) the recipients or categories of recipients of the personal data, if any;

f) where applicable, that the controller intends to transfer personal data to a recipient in a third country or international organisation and the existence or absence of an adequacy decision by the Commission, or in the case of transfers referred to in Article 46 or 47, or the second sub-paragraph of Article 49 (1), reference to the appropriate or suitable safeguards and the means to obtain a copy of them or where they have been made available.

In addition to the information referred to in paragraph 1, the controller shall provide the data subject with the following information necessary to ensure fair and transparent processing in respect of the data subject:

a) the period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period;

b) where the processing is based on point (f) of Article 6 (1), the legitimate interests pursued by the controller or by a third party;

c) the existence of the right to request from the controller access to and rectification or erasure of personal data or restriction of processing concerning the data subject and to object to processing as well as the right to data portability;

d) where processing is based on point (a) of Article 6(1) or point (a) of Article 9 (2), the existence of the right to withdraw consent at any time, without affecting the lawfulness of processing based on consent before its withdrawal;

e) the right to lodge a complaint with a supervisory authority;

f) from which source the personal data originate, and if applicable, whether it came from publicly accessible sources;

g) the existence of automated decision-making, including profiling, referred to in Article 22 (1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject."

On the other hand, Art. 14 (5) includes notable derogations for research. If the provision of the required information proves to be impossible or would involve a disproportional effort, the transparency obligations no longer apply.[299] Recital 62 also mentions that such obligation is no longer necessary "where processing is carried out for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes." A proportionality assessment has to be carried out in order to decide whether the conditions for the exception are fulfilled.[300] Recital 62 names some indication on what needs to be considered during the evaluation. These are the number of data subjects, the age of the data and any appropriate safeguards adopted.

[299] Schaar, ZD 2017, 213 (216)
[300] Franck, in Gola (supra note 109) Art.14, Rn. 23; Schaar, ZD 2017, 213 (216);

Additionally, it should be noted that the change in purpose results in renewed information provision obligation.[301] Both Art. 13 (3) GDPR and Art. 14 (4) oblige controllers who intend to further process personal data for a purpose other than that for which the personal data were collected to provide the data subject with information on that other purpose and any other relevant information prior to that further processing. This obligation refers to the compatibility assessment outlined in Art. 6 (4) GDPR, since "data controllers should provide data subjects with further information on the compatibility analysis carried out under Article 6 (4) where a legal basis other than consent or national/ EU law is relied on for the new processing purpose (in other words an explanation as to how the processing for the other purposes is compatible with the original purpose). This is to allow data subjects the opportunity to consider the compatibility of the further processing and the safeguards provided and to decide whether to exercise their rights e.g. the right to restriction of processing or the right to object to processing amongst others."[302]

Some of these obligations raise serious concerns in relation to Big Data analytics. It is hardly possible for controllers to provide all the required information. When using Big Data, the phases of the processing activities do not necessarily follow each other in a linear order, in many cases the processing operations occur simultaneously. Also, the recipients of a dataset may not be known at the time the information provision obligation is due, since it may depend on the result of the data analysis to whom – if – the data will be transferred. Then again, if controller de-identified in a secure way, they may be able to escape those transparency obligations.[303]

3.1.2.2.2 Method of Obtaining Consent

Art. 4 No. 11 in adherence with Recital 32 GDPR clarifies that consent must be the unambiguous indication of the data subject's wishes, which must have the form of a statement or a clear affirmative act.[304]

The GDPR acknowledges the validity of a wide range of commonly used methods of obtaining consent and affirms the principle that any appropriate method should be allowed to use. Consent can be provided by a written statement, including by electronic means, or an oral statement. Controllers should in any event careful thought to ensure that their consent mechanism is appropriate to the nature of the consent. Referring to this recital 32 highlights that consent can be collected "by any appropriate method". At the same time, it makes it crystal clear, that acquiescence does not equal consent. According to Recital 32 GDPR, silence, pre-ticked box or inactivity cannot amount to consent; when the data subject says nothing when given the opportunity to object, will not result in a valid consent. Also, the possibility of hiding statements in terms and conditions or using opt-out-techniques will no longer be allowed.[305]

At the same time, Recital 32 provides examples on what constitutes a valid consent, such as "ticking a box when visiting an internet website" as well as "choosing technical settings for information society services.[306]" – based on this example some argue that using the technical settings of an internet browser can be interpreted as a valid consent[307] – or "another statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data".

Finally, concerning the lawfulness of those consents under the GDPR which were obtained under the conditions of the legal framework of the Directive, Recital 171 GDPR stipulates that "where processing is based on consent pursuant to Directive 95/46/EC, it is not necessary for the data subject to give his or her consent again if the manner in which the consent has been given is in line with the

[301] Art.-29 Working Party, WP260, p. 20
[302] Art.-29 Working Party, WP260, p. 21
[303] Schaar, ZD 2017, 213 (217)
[304] Ernst, ZD 2017, 110 (113 f.)
[305] Spindler, DB 2016, 937 (940); Härting (supra note 43) Rn. 374 ff.
[306] for the scope of "information society service" see Art. 4 No. 25 GDPR which refers to Directive (EU) 2015/1535
[307] Härting (supra note 43) Rn. 364; different opinion Spindler, DB 2016, 937 (940) who doubts whether default settings of a browser are sufficient for a valid consent; cf. Art.-29 Working Party, WP 259, p. 17

conditions of the GDPR." Consequently, controllers operating on the basis of consent obtained before the GDPR effective date are not automatically required to entirely renew their consent relations.[308] On the other hand, it is important that controllers review their current consent policies, and make sure, that their methods are compliant with the new regulatory regime on consent under the GDPR.

3.1.3 Additional Protection for Special Categories of Personal Data

The differentiation between personal data as such and special categories of personal data results in a layered regulation, which is the direct manifestation of the so-called risk-based approach followed by the GDPR.[309] According to this approach, processing of sensitive data, processing activities that affect vulnerable individuals as well as large-scale processing are *per se* a processing associated with certain level of risk. Cumulating these factors will result in "high-risk" processing, such as large-scale processing of sensitive data, or processing sensitive data with newly introduced technical methods.[310]

Activities	Example of processing
High Risk	
large scale processing of sensitive data referred to in Art. 9 (1)	Big Data analytics on genetic or health data, e.g. genome mapping, gene testing; clinical trials
systematic and extensive evaluation of personal aspects relating to a natural person based on automated processing	the use of a camera system to monitor publicly accessible area
new data processing technologies	Any processing involving innovative use or applying newly introduced technological or organisational solutions
other activities ""likely to result in a high risk for the rights and freedoms of individuals"	calls for case-by-case evaluation
Risk	
processing of sensitive data referred to in Art. 9 (1)	hospital information system hospital processing patients genetic and health data
processing concerning vulnerable individuals	organisation monitoring its employees' activities, e. g. work station, pathway

[308] Art.-29 Working Party, WP 259, p 29

[309] for more details on the risk-based approach see: Maldoff, The Risk-Based Approach in the GDPR: Interpretation and Implementation, available at: https://iapp.org/resources/article/the-risk-based-approach-in-the-gdpr-interpretation-and-implications/

[310] Art. 35 (3) b) GDPR; Art.-29 Working Party, WP 259, p. 7;

processing where the data subjects are prevented from exercising control	placebo-controlled trials
large scale processing	data processing by social network sites, insurance companies

Table 3: Risks related to processing of sensitive personal data[311]

Therefore, a thorough evaluation of the lawfulness of processing special categories of personal data must be carried out.

3.1.3.1 Definition of Special Categories of Personal Data

Art. 9 (1) defines special categories of personal data as

> "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation, and explicitly prohibits the processing of those data."

It was a conscious decision to keep the scope of Art. 9 (1) wide, therefore this provision includes not only virtual data in a narrow sense, but also indirect indications or references to one of the characteristics.[312] The context of processing may also influence the nature of a certain information, e. g. a passport photo with glasses, address of a clinic for delivery service.[313] Compared to the Directive, the GDPR made substantial progress to enhance the protection of several categories by introducing new legal definitions, in particular related to the health domain.[314] It notably increases the types of data that are included in the definition of "data concerning health", which according to Art. 4 No. 15 means

> "personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status."

This is a much broader understanding than the term 'medical data', and it does not always necessary mean "ill health".[315] Essentially, health data refers to any kind of information that may reveal details about one's general health condition. To that end, Recital 35 GDPR clarifies that health data should include "all data pertaining to the health status of a data subject which reveal information relating to the past, current or future physical or mental health status of the data subject." A wide range of personal data may fall into the category of health data, making this category the most complex area of sensitive

[311] Recital 75 GDPR, Art. 35 GDPR; Art. 29 Working Party Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, WP 248

[312] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 20

[313] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 22; see also Frenzel, in Paal/Pauly (supra note 19) Art. 9 Rn. 8

[314] Chassang (supra note 95), p. 5; Zarsky, Incompatible: The GDPR in the age of Big Data, 47 Seton Hall Law Review, 2017, 995 (1012)

[315] Härting (supra note 43) Rn. 538; Art.-29 Working Party, ANNEX – health data in apps and devices, pp. 1 ff., available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf ; cf. Ohm, Sensitive Information, Southern California Law Review, Vol. 88, 2015, 1125 (1130)

data.[316] Aside from medical records, data on medication and diseases, information such as the fact related to a broken leg, a person's intellectual and emotional capacity, smoking or drinking habits, allergies disclosed to public bodies or private entities or membership in a patient support group are all data concerning the health of individual data subjects.[317] Recital 35 GDPR itself lists " information derived from the testing or examination of a body part or bodily substance, including from genetic data and biological samples; and any information on, for example, a disease, disability, disease risk, medical history, clinical treatment or the physiological or biomedical state of the data subject independent of its source, for example from a physician or other health professional, a hospital, a medical device or an in vitro diagnostic test" as examples for health data. Besides, it also includes information on time spent in hospital, health or wellness clinic, data exchanged between patient and caretaker or doctor via telecommunication services, browser history, downloaded smartphone or wearables applications as well as other relevant data emerging from the "Internet of Things".[318] Additionally, in some cases of medical research using Big Data, cases may be included where the controller uses any personal data – health data or not – for the purpose of identifying disease risk.[319]

Unlocking Big Data, new forms of analytics and systematic data mining technologies challenge the ability of differentiating between sensitive data and other categories. Health data can be deducted from a number of datasets, such as shopping databases, banking data, or bank account statement. For this reason, some argue that Big Data may eventually undermine the entire distinction between different categories of personal data.[320]

Apart from health data, the GDPR established the definition for genetic data as well as biometric data in Art. 4 No. 13 and No. 14 respectively:

> "genetic data' means personal data relating to the inherited or acquired genetic characteristics of a natural person which give unique information about the physiology or the health of that natural person and which result, in particular, from an analysis of a biological sample from the natural person in question"
>
> "biometric data' means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data."

3.1.3.2 Legitimate Basis for the Processing of Special Categories of Personal Data

Special categories of personal data are generally considered to be the most private details about an individual, therefore they must be subject to additional protection.[321] Art. 9 (1) GDPR expressively states that the processing of those data is by principle prohibited. The processing of sensitive data is allowed exceptionally and exclusively in case one of the legitimate bases outlined in Art. 9 (2) applies. Art. 9 (2) itemizes the exceptional circumstances that allow the processing of special categories of personal data. This reflects in Recital 51, which states that "such personal data should not be processed, unless processing is allowed in specific cases set out in this Regulation, taking into account that Member States law may lay down specific provisions on data protection in order to adapt the application of the

[316] Art.-29 Working Party, Advice paper on special categories of data ("sensitive data") p. 10, available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/others/2011_04_20_letter_artwp_mme_le_bail_directive_9546ec_annex1_en.pdf
[317] Art.-29 Working Party, ANNEX – health data in apps and devices (supra note 313) p. 2,
[318] cf. Weichert, in Kühling/Buchner, (supra note 33) Art. 9 Rn. 39
[319] Art.-29 Working Party, ANNEX – health data in apps and devices (supra note 313) p. 3
[320] Zarsky (supra note 312), p. 1013, who, on the other hand, accepts that "the enhanced protection of specific categories could nonetheless be justified in the Big Data age, even if drawing the actual distinctions prove impossible"; see also: Ohm (supra note 313) p. 1145
[321] cf. Ohm (supra note 313), p. 1126, 1169

rules of this Regulation for compliance with a legal obligation or for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller. In addition to the specific requirements for such processing, the general principles and other rules of this Regulation should apply, in particular as regards the conditions for lawful processing". It is "*lex specialis*" in relation to Art. 6 (1), therefore the processing of sensitive data shall only be permitted if one of the authorization of Art. 9 (2) applies.[322] Controllers cannot in any case justify the processing of sensitive data by referring to a lawful basis in Art. 6 (1), especially not by evoking their legitimate interest, since the Art. 9 (2) excludes the possibility of such a wide-ranging interest assessment. Art. 9 (2) provides a definitive list, processing of sensitive data is only permitted under these special conditions:

a) the data subject has given **explicit consent** to the processing of those personal data for one or more specified purposes, except where Union or Member State law provide that the prohibition referred to in paragraph 1 may not be lifted by the data subject;

b) processing is necessary for the purposes of carrying out the obligations and exercising specific rights of the controller or of the data subject in the **field of employment and social security and social protection law** in so far as it is authorised by Union or Member State law or a collective agreement pursuant to Member State law providing for appropriate safeguards for the fundamental rights and the interests of the data subject;

c) processing is necessary to protect **the vital interests** of the data subject or of another natural person where the data subject is physically or legally incapable of giving consent;

d) processing is carried out **in the course of its legitimate activities with appropriate safeguards by a foundation, association or any other not-for-profit body with a political, philosophical, religious or trade union aim** and on condition that the processing relates solely to the members or to former members of the body or to persons who have regular contact with it in connection with its purposes and that the personal data are not disclosed outside that body without the consent of the data subjects;

e) processing relates to personal data which are **manifestly made public by the data subject**;

f) processing is necessary for the **establishment, exercise or defence of legal claims** or whenever courts are acting in their judicial capacity;

g) processing is necessary **for reasons of substantial public interest**, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject;

h) processing is necessary **for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services** on the basis of Union or Member State law or pursuant to contract with a health professional and subject to the conditions and safeguards referred to in paragraph 3;

i) processing is necessary **for reasons of public interest in the area of public health**, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices, on the basis

[322] Schulz, in Gola (supra note 109), Art. 9 Rn. 1; cf. Härting (supra note 43) Rn. 528, 531

of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy;

j) processing is necessary **for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89 (1)** based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

The consent of the data subject is the safest way for controllers to justify processing special categories of personal data. The general requirements for a valid consent[323] continue to apply, however, consent within the meaning of this provision must additionally be "explicit".[324] An implicit statement, implied declaration or an opt-out consent policy will not result in a valid consent.[325] Data collection for the purpose of further anonymization does not eliminate the obligation for obtaining consent, unless the data acquisition can be justified by another provision of Art. 9 (2).[326]

Aside from the explicit consent given by the data subject, processing may be carried out on the basis of employment law, within the context of the activities of foundations or NGOs, to establish exercise or defend legal claims as well as for the vital interest of the data subject (or another individual) where the data subject is incapable of giving consent. In this latter case, the hypothetical will and intentions of the data subject must be considered as well.[327] If the data subject has a representative, legal guardian as well as in case of a power of attorney, Art. 9 (2) lit c) will not apply.[328] In addition, according to Recital 46, vital interest is rather an '*ultima ratio'*, because the processing based on vital interest "should in principle take place only where the processing cannot be manifestly based on another legal basis."

Art. 9 (2) lit g) GDPR allows the processing based on reasons of substantial public interest. Art. 8 (4) of the Directive allowed a similar lawful basis, however, it used merely used the wording "important". This was amended by the GDPR and changed for "substantial", which implies a rather higher threshold for assessing the required level of public interest.[329] However, a general public interest is not sufficient to justify the processing of sensitive data. The interest must be substantiated with respect of sensitive data. [330]

Art. 9 (2) lit h) in accordance with Recital 53 enables the processing of sensitive data for medical diagnoses and treatment[331] on the one hand and the management of health or social care systems and services on the other. Prerequisite for this exception is the particular importance for the data subject, for third parties or for society in general.[332] This provision limits the otherwise wide scope of this basis.[333] Processing within the provision of "medical diagnoses and treatment" must relate to the treatment itself, or to the management and administrative issues of that.[334] An important requirement for this is, according to Art. 9 (3), that the data "are processed by or under the responsibility of a professional subject to the obligation of professional secrecy under Union or Member State law or rules established

[323] see section 3.1.2.2. on conditions for consent
[324] Spindler, MedR 2016, 691 (697); Härting (supra note 43) Rn. 552
[325] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 47 ff; Schiff, in Ehmann/Selmayr (supra note 49) Art. 9 Rn. 28
[326] Schulz, in Gola (supra note 109) Art. 9 Rn. 15, 34; This argument implies that processing personal data with implementation of anonymisation methods only lead to a partial inapplicability of data protection law.
[327] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 66
[328] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 68
[329] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 88; cf. Schulz, in Gola (supra note 109) Art. 9 Rn. 28
[330] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 91
[331] Härting (supra note 43) Rn. 543
[332] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn 92
[333] Schulz, in Gola (supra note 109) Art. 9 Rn. 30
[334] Schulz, in Gola (supra note 109) Art. 9 Rn. 29

by national competent bodies or by another person also subject to an obligation of secrecy under Union or Member State law or rules established by national competent bodies". The other scenario regulated in Art. 9 (2) lit h) is processing for management of health and social care systems, which includes the entire organisational institution for performing of health services. In contrast to the Directive, the new concept of the GDPR is broader and encompasses procedures and contracts under private law as well, such as coverage of costs by insurance companies.[335] It must be noted that Art. 9 (2) lit h) GDPR neither enables the processing of sensitive data for medical research purposes, nor legitimizes processing of such data for reasons of public interest in the area of public health.[336]

The main difference between Art. 9 (2) lit h) and lit i) is that the former refers to the individual interests of the data subject whereas the latter focuses on the public interest and health. "Public health" should be interpreted as defined in Regulation (EC) 1338/2008, and it includes "all elements related to health, namely health status, including morbidity and disability, the determinants having an effect on that health status, health care needs, resources allocated to health care, the provision of, and universal access to, health care as well as health care expenditure and financing, and the causes of mortality." [337] Given this broad interpretation, even activities of social media or online platforms may qualify as public health research.[338] Referring to this, Recital 54 clarifies that within the provision of Art. 9 (2) lit i) the processing carried out by private organisations generally cannot be derived from such processing for reasons of public interest.[339] On the other hand, Recital 54 lists third parties such as employers or insurance and banking companies as example. This implies that the processing by public research bodies, e.g. universities, or by research institutes may still be justified if they contribute to the establishment and further development of a higher level of human health protection.[340]

The change introduced by the GDPR with establishing the exception of Art. 9 (2) lit j) has notable positive implications especially in the medical research sector, since it provides additional legitimate ground on which sensitive personal data may lawfully be processed.[341] However, it should be pointed out that Art. 9 (2) lit j) only refers to independent research. An external influence on the scientific findings or deferring those findings to sheer economic purposes must be precluded. On the other hand, the fact that a project is financed by third-party-funds does not necessarily exclude the applicability of Art. 9 (2) lit j).[342] On the other hand, it covers both research proposals initiated and carried out by the same controller and external projects, i. e. where the access to sensitive data is granted by the controller, but the implementation of the project, the analysis of those data is outsourced to one or more processors.[343]

Finally, Art. 9 (4) – in line with Recital 51 –offers an opening clause allowing Member States to maintain or introduce further conditions, including limitations regarding the processing of special categories of personal data. For that reasons, the possible exemptions under national law must be assessed.

---

[335] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 105
[336] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 93
[337] Recital 54 GDPR
[338] Maldoff, How GDPR changes the rules for research, available at: https://iapp.org/news/a/how-gdpr-changes-the-rules-for-research/
[339] Härting (supra note 43) Rn. 546
[340] Since this is a strategic objective of the EU, established in Art. 168 TFEU; see also: Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 117, 130; cf. Härting (supra note 43) Rn. 549, who is rather sceptical and argues, that based on Recital 159 in line with Art. 89 and Art. 9 (2) lit j) GDPR research by pharmaceutical companies or biochemical research does not benefit from the privilege established in Art. 89
[341] unlike the GDPR, the Directive did not contain corresponding provision to Art. 9 (2) lit j); Härting (supra note 43) Rn. 557
[342] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 129
[343] Schulz, in Gola (supra note 109) Art. 9 Rn. 35

3.1.4 Safeguards relating to the Processing for Scientific Research

The Digital Single Market tackles the issue of improving data sharing and dataflow across the EU, which among others intends to facilitate health care and research. This intention appears in Art. 89 GDPR, which establishes safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, thus making them secondary privileged processing purposes.[344] This corresponds with the above-mentioned primary goal of SODA to develop secure computation methods in order to enhance health care systems.

The GDPR recognises the freedom of research, which is enshrined in Art. 179 (1) TFEU as one aim of the Union of "strengthening its scientific and technological bases by achieving a European research area in which researchers, scientific knowledge and technology circulate freely, and encouraging it to become more competitive, including in its industry, while promoting all the research activities". Art. 13 of the Charter of Fundamental Rights of the European Union also guarantees the freedom of science.[345] However, these rights are not granted without limitation, the necessity of a comprehensive data protection system including research-specific safeguards is inevitable to compensate for the loss of control in data-sensitive health research.[346]

The GDPR adopts a broad concept of research. Each type of research is discussed separately in the recitals. Recital 159 defines scientific research "in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research". According to Recital 162, statistical research means "any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results" within the frames of the GDPR. Additionally, it highlights that "the statistical purpose implies that the result of processing for statistical purposes is not personal data, but aggregate data, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person." This clarifies that profiling is not covered by the privilege for research. Otherwise, statistical purposes are defined rather widely, which can permit the use of data for statistical purposes to be repurposed for Big Data utilisation.[347] Public health research is treated in the GDPR as an art of scientific research and is therefore subject to the same exemptions.[348] Concerning this matter, especially in cases where the controller processes genetic, biometric or health data Recital 53 states that "Union or member state law should provide for specific and suitable measures so as to protect the fundamental rights of the personal data of natural persons".

When discussing research, distinction must be made between primary research, i. e. where the acquisition of data occurs directly for research purposes, and secondary research, where personal data were collected originally for another purpose, and later re-used for research purposes.[349] In the pilot application cases within the frames of SODA research is typically secondary research. In one case data acquired from patients for individual reasons are later aggregated, secret-shared and analysed within or between hospitals. In the other case, similarly acquired data of hospitals are combined with client data of insurance companies. None of these organisations collect these data primary for research purposes. They reuse already existing datasets instead.

---

[344] Rumbold/Pierscionek, The Effect of the General Data Protection Regulation on Medical Research, Journal of Medical Internet Research, 2017 Feb; 19 (2): e47, doi: 10.2196/jmir.7108

[345] Pauly, in Paal/Pauly (supra note 19) Art. 89 Rn. 3

[346] Mostert/Bredenoort/van der Sloot/van Delden, European Journal of Health Law Vol. 24, 2017, 1 (1, 10); Pauly, in Paal/Pauly (supra note 19) Art. 89 Rn. 3

[347] Mayer-Schöneberger, Regime Change? Enabling Big Data Through Europe's New Data Protection Regulation, The Columbia Science & Technology Law Review, Vol. XVII 2016, 315 (326 f.)

[348] Maldoff, How GDPR changes the rules for research, available at: https://iapp.org/news/a/how-gdpr-changes-the-rules-for-research/

[349] Raum, in Ehmann/Selmayr (supra note 49) Art. 89 Rn. 18

Art. 89 (1) GDPR does not provide a separate legitimate basis for processing, that is, without an authorisation resulting either from Art. 6 (1) or from Art. 9 (2) primary research will be unlawful.[350] This raises the question of what could serve as a legitimate basis for processing personal data primarily for research, i. e. where the purpose of the original collection of personal data is research. Obtaining the consent of the data subject seems to be the evident and most secure way. However, resulting from the difficulty in identifying the research purposes prior to the actual processing, providing up front information can be challenging. This is particularly the case where Big Data and data mining algorithms are used, since in such cases the researcher may not be aware of the scope of the research until after computing and analysing the dataset. Recital 33 recognises this by allowing data subjects to "give their consent to certain areas of scientific research", provided that the „recognised ethical standards for scientific research" are respected.[351]

Apart from the data subject's consent, the processing might be justified on the basis of the legitimate interest of the controller according to Art. 6 (1) lit f), provided that the dataset involved does not contain special categories of personal data. If the controller intends to use sensitive data for research purpose, it must base the processing on one of the bases established by Art. 9 (1). Within this provision, controllers may refer to Art. 9 (1) lit h), lit i) or lit j) respectively, given the specific circumstances of the processing they intend to engage in.[352]

In order to decide whether further processing of personal data for research purposes in cases of secondary research is lawful, account should be taken to the compatibility assessment regulated in Art. 6 (4) GDPR. This provision allows subsequent processing operations for purposes that are compatible with the original purpose without a separate legitimate ground. Referring to this, Recital 50 is of paramount importance, as it specifies that "further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes should be considered to be compatible lawful processing operations", thereby establishing a legal presumption. Consequently, should data be further processed for scientific research purposes, the compatibility is presumed according to Art. 5 (1) lit b) in conjunction with Art. 89 (1). In such cases neither a compatibility assessment nor a new, separate legitimate basis is required.[353] This privilege is thereby an exception from the strict interpretation of the principle of purpose limitation regulated in Art. 5 (1) lit b) GDPR.

Another prerequisite for the research exemption within the provision of Art. 89 (1) is the implementation of "appropriate safeguards". Explicitly mentioned is here only pseudonymisation. However, a wide range of technical and organisational measures may be implemented, as long as they "ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation", such as encryption or other methods of access-control.[354]

One of the main aims of the SODA project is to develop such privacy preserving methods, that provide a secure computing environment for data controllers to carry out scientific research or to process data for scientific purposes. With anonymising or encrypting the data they guarantee the principle of data minimisation as well as an adequate level of security.

Controllers must take into account that the privilege outlined in Art. 89 does not exempt them from certain notification obligation towards the data subject. First of all, they are obligated to provide the data subject with appropriate information if they intend to further process data for different purpose, including for research.[355] On the other hand, they may be exempt from the notification obligation, if

[350] Pauly, in Paal/Pauly (supra note 19) Art. 89 Rn. 1; Raum, in Ehmann/Selmayr (supra note 49) Art. 89 Rn. 22; Buchner/Tinnefeld, in Kühling/Buchner (supra note 33) Art. 89 Rn. 1
[351] Mostert/Bredenoort/van der Sloot/van Delden, European Journal of Health Law Vol. 24, 2017, 1 (10)
[352] Weichert, in Kühling/Buchner (supra note 33) Art. 9 Rn. 130
[353] Schulz, in Gola (supra note 109) Art. 6 Rn. 193; Plath, in Plath (supra note 33) DSGVO Art. 6 Rn. 30 ff.
[354] Buchner/Tinnefeld, in Kühling/Buchner (supra note 33) Art. 89 Rn. 18 ff.
[355] see detailed in section 3.1.2.2.

they acquired the dataset from another source, and the provision of information would involve a "disproportional effort" according to Art. 14. This in particular could be the case in the research context.[356]

Resulting from the research privilege controllers are also exempted from certain obligations related to the rights of the data subject. Art. 89 (2) stipulates that in case of processing personal data for scientific research or statistic purposes, "Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes."[357] Furthermore, the Art. 17 (3) lit d) GDPR directly provides exemption from the right of erasure insofar as it is "likely to render impossible or seriously impair the achievement of the – research – objectives. Additionally, controllers engaging in research might override the data subjects' right to object, granted to them by Art. 21, if the processing "is necessary for the performance of a task carried out for reasons of public interest" pursuant to Art. 21 (6).[358]

Even though the GDPR established heightened obligations of compliance relating to the processing of personal data, by creating specific exemptions and derogations for research purposes it also enhances the dataflow and the further development of research across the Digital Single Market.

3.2 Responsible Party (the *Controller*) and the Processing on behalf of the Controller

Although there are several stakeholders in a processing activity through the whole Big Data value chain, not every motivated party is a controller or a processor. On the other hand, in every occasion where an organisation processes personal data, it acts either as controller or as processor. These roles impose different responsibilities and obligations on the parties respectively. For this reason, it is crucial for organisations to be able to identify the scenarios and operations in which they act as a controller – or as a processor –, to understand the obligations that apply to the controllers, and to comply with those obligations. While the Directive in general only imposed direct compliance obligations on the controller, the GDPR imposes certain obligations on both the controller and the processor.[359] In case of non-compliance, they both will face direct enforcement and penalties under the GDPR.

3.2.1 The Responsible Party (the Controller)

3.2.1.1 The Concept of Controller

In case of "in-house" processing, that is when data controller processes the data itself without outsourcing it to processors and provided the processing of the data as well as the decision-making about the means and purposes are carried out internally, there are no legal challenges in this regard. Otherwise, the primary function of defining the identity of the controller is to allocate responsibility. The concept of controller stays unchanged. Art. 4 No. 7 GDPR defines data controller as:

> "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data."

The rule of a controller is not assigned exclusively to one party, rather it shifts from one party to another in any case respectively, therefore, a pragmatic approach has to be adopted.[360] The two chief elements

[356] Recital 62 GDPR
[357] see also Chassang (supra note 95) p. 10
[358] Herbst, in Kühling/Buchner (supra note 33) Art. 17 Rn. 81 and Art. 21 Rn. 53
[359] Härting (supra note 43) Rn. 570 ff.
[360] Art. 29-Working Party, Opinion 01/2010 on the concepts of "controller" and "processor", WP 169 p. 9; Klug, in Gola (supra note 109) Art. 28 Rn. 5

of the definition are the "purposes and means" on the one hand and "alone or jointly with others" on the other. As to the first one, the controller is the party who takes the essential decision and determines the purposes for which and the means by which personal data is processed. This must be assessed based on the factual control, with respect to the circumstances of the given processing activity. If the effective influence does not reflect the conditions laid down in the contractual terms, the issue must be decided with taking into consideration the real-life situation as well as the reasonable expectations of the data subject.[361] Regarding the determination of the means it should be noted that the term "mean" covers a wide variety of elements. Despite the fact that the controller must without doubt act as "lord of the data"[362] processors can be allowed to have some degree of discretion. It is likely, that specific technical and organisational measures will be determined exclusively by the processor, in particular when the controller has no expertise in those technologies whatsoever. If the processor takes over the decision regarding the details on how the data analysis should be carried out, the chosen means should represent a "reasonable way" of achieving the purposes set out by the controller.[363] The essential elements, which cannot in any event be delegated to the processor are e. g.:

- determining and choosing the legal basis for collecting data
- the obligation to obtain consent, consent policies
- which items of personal data to collect
- which individuals to collect the data about
- from which sources to collect the data from
- purpose or purposes the data is used for
- ensuring that the data subjects are able to exercise their rights
- decision on how long to retain data

[361] Art. 29-Working Party, Opinion 01/2010, WP 169, p. 13
[362] Härting (supra note 43) Rn. 571
[363] Art. 29-Working Party, Opinion 01/2010, WP 169, p. 14

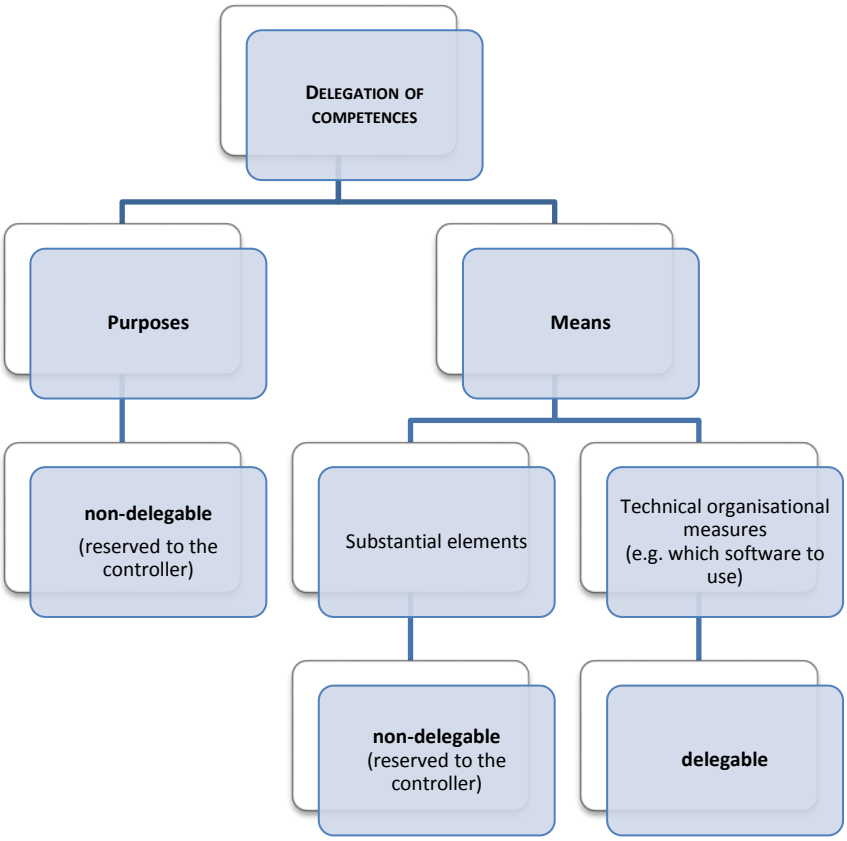


Figure 3. Delegation of competences

The second element of the definition is that controllers may act "alone or jointly with others". A complete assessment of all specific circumstances is advised for each actor in order to decide whether decisions should be made jointly as a separate controller.[364] Thus, the collaboration between actors involved in processing of personal data is not necessarily a controller – processor relation, it is also possible that multiple actors interact in the processing of personal data as controllers.[365] However, the fact that more controllers dispose of the same dataset does not equal joint controllership *per se*. Such data controllers in common would simply share a pool of personal data that they process independently of each other. Entering Big Data, there are potentially many parties involved in separate processing activities concerning the same particular set of data, e. g. multiple health care facilities sharing their data and using this combined dataset. Therefore, it is always necessary to evaluate the different degrees in which multiple parties are linked.[366] It is possible that the involved parties divide different tasks in a way that each and every processing operation appears to be independent and executed by only one controller. However, when assessing the whole set of operations on a "macro-level", the actors can also be regarded as joint controllers. This result can be derived from mutually determined purposes and a cooperatively set framework that determines the essential means or whether the decisions relating to both questions are taken together.[367] The GDPR regulates the joint controllership in Art. 26 as a scenario, where "two or more controllers jointly determine the purposes and means of processing". They

[364] Art. 29-Working Party, Opinion 01/2010, WP 169, 18
[365] Van Alsenoy, Liability under EU Data Protection Law: From Directive 95/46 to the General Data Protection Regulation, JIPITEC, Vol. 7, 2016, 271 (279)
[366] Art. 29-Working Party, Opinion 01/2010, WP 169, p. 19
[367] Art. 29-Working Party, Opinion 01/2010, WP 169, 20.

must determine the distribution of control in a "transparent manner" by "means of an arrangement between them", and provide a summary of the arrangement available to the data subject.[368] This clarification is of utmost importance firstly because of the clear allocation of responsibilities in order to comply, secondly concerning liability issues.[369]

2.2.1.2 Responsibilities and obligations of the controller

In general, controllers bear the primary responsibility for ensuring that the data processing activities are compliant with the GDPR. Recital 74 stipulates that the "responsibility [...] of the controller for any processing of personal data carried out by the controller or on the controller's behalf should be established".

The first category of responsibilities of the controller can directly be derived from the principle of accountability in conjunction with the principles of the data protection, both outlined in Art. 5 GDPR. They manifest in obligations during the processing activity. Art. 5 states that the controller shall be responsible for, and be able to demonstrate compliance with the principles relating to the processing of personal data during the processing activity. Therefore, these 6 principles are the first set of queries the data controller must pay attention to:

1. processing lawfully, fairly and in a transparent manner
2. collecting only for specified, explicit and legitimate purpose
3. ensuring that the processing is adequate, relevant and limited to what is necessary
4. ensuring that the data is accurate, and where's necessary, up to date
5. retaining data only as long as necessary
6. processing in an appropriate manner to maintain security

The second package of obligations serve the security of processing either by preventive measures or while carrying out the processing activity. These include firstly the obligation of the controller to implement appropriate data protection policies where it is proportionate in relation to the processing activity, according to Art. 24 (2).

The essence of this category of obligations is the provision in Art. 28 (1) that obliges controllers to implement appropriate technical and organisational measures in order to ensure the security – and confidentiality – of the processing. There are two core elements of this obligations, firstly to ensure a level of security during the processing which is appropriate to the risk, and secondly, the newly introduced principles of Privacy by Design and by Default.

[368] Art. 26 (2) GDPR
[369] Art. 29-Working Party, Opinion 01/2010, WP 169, 22

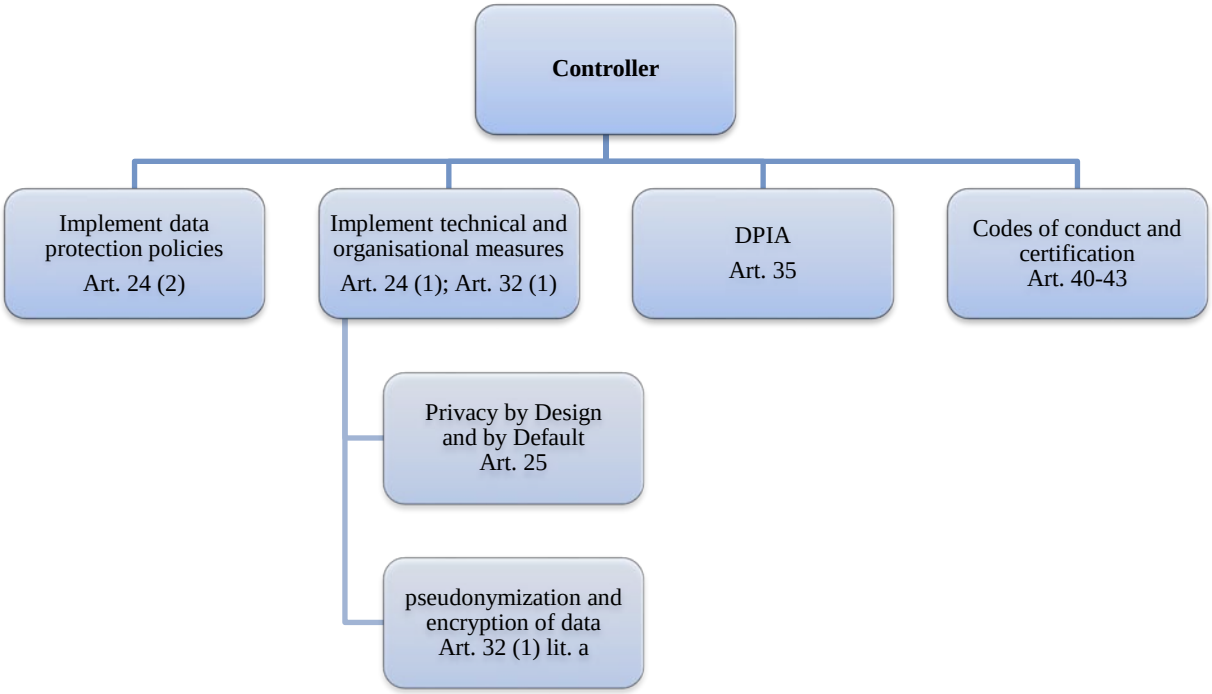


Figure 4. Obligations of the controller relating the security of processing

After evaluating the risk of their activity, in order to maintain the security of processing, controllers must implement adequate technical and organisational measures. Art. 32 (1) lists the factors which shall be considered when evaluating which of those measures is appropriate for a certain processing. These are:

1. the state of the art
2. the costs of implementation
3. the nature, scope, context and purposes of the processing
4. the risk of varying likelihood and severity for the rights and freedoms of natural persons

Thus, the GDPR predicates on the conditions of a proportionality test, and leaves it to the controllers to decide which measures to apply, considering the result of such assessment.[370] However, it does not provide guidance on what the term "state of the art" means. "State of the art" within this provision implies such technologies that are already made available to the public and are applied by the practice, and consequently, proved to be appropriate and suitable. It does not include brand-new introduced techniques.[371]

Art. 32 (1) then continues by giving an exemplary list of the wide range of relevant measures:

a) the pseudonymisation and encryption of personal data;

b) the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;

[370] Piltz, in Gola (supra note 109) Art. 32 Rn. 9, 13
[371] Hladjk, in Ehmann/Selmayr (supra note 49) Art. 32 Rn. 5; Piltz, in Gola (supra note 109) Art. 32 Rn. 15, 16

c) the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;

d) a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing.

Outlined in Art. 25, the concept of Privacy by Design and Privacy by Default are one of the main innovations introduced by the GDPR. Instead of adapting the product to data protection principles after the development and design process, Privacy by Design in the GDPR is a strategy that requires developers, producers or data controllers to respect data protection issues from the very beginning of the development on the one hand, and consequently the implementation and integration of privacy enhancing features from the first step.[372] Such a design process promotes a concept to take the entire life circle of the data into consideration, from the collection through the processing to deletion, systematically focusing on comprehensive safeguards relating accuracy, confidentiality, integrity, physical security and deletion of personal data.[373] Privacy by Design also intends that systems must be designed and constructed to reduce or avoid the amount of personal data being processed.[374] Accordingly, Art. 25 GDPR states that the controller "shall (…) implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing". For the evaluation of what those means should exactly be, Art. 25 (1) provides factors such as "the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing". Recital 78 GDPR adds that "such measures could consist, *inter alia*, 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, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features". Notable is the contradiction at first sight that despite the aforementioned target group of Privacy by Design Art. 25 only mentions the controller. Recital 78 comes as aid by outlining that it encourages producers of products, services and applications that process personal data to take into account the right to data protection when developing and designing such products, services and applications and to make sure that controllers and processors are able to fulfil their data protection obligations. An encouragement does not equal a binding obligation for developers to comply with the principles of Privacy by Design. However, achieving the requirements of Art. 25 (1) GDPR is also their best interest. Not being able to provide a service compatible with the GDPR's new provisions could eventually lead to the fact, that controllers will not use their services anymore to avoid violations of the requirements of the GDPR.[375] This could be a strong incentive for following the Privacy by Design strategy.

The contribution of the SODA project to privacy in Big Data analytics is to demonstrate feasibility of Big Data analytics and make it easier to develop privacy-friendly Big Data application. It provides a privacy preserving alternative to the current practice of distributed research. In doing so, it implements technical measures (described above in sections 2.3.) already in the planning, structuring and software development phases. Therefore, when controllers decide to use techniques developed by SODA, they will comply with the GDPR, since SODA-technologies are designed for privacy-preserving computation.

[372] Schaar, Privacy by Design, IDIS 2010, Vol. 3, Issue 2, p. 266
[373] European Union Agency for Network and Information Security, Privacy and Data Protection by Design – from policy to engineering, p. 11.
[374] Schaar, Privacy by Design, IDIS 2010, Vol. 3, Issue 2, p. 267
[375] Plath, in: Plath (supra note 33) DSGVO Art. 25 Rn. 7

Along with the principle of Privacy by Design comes Privacy by Default. Its aim is to establish a default setting that allows the customer to use the given product without having to fear unintentional disclosure of personal data. The pre-set should be designed in such a way that processing of personal data is minimized. Leaving it to the decision of the user to consent to more processing[376]. Privacy by default is regulated in Art. 25 (2) GDPR:

> "The controller shall implement appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility. In particular, such measures shall ensure that by default personal data are not made accessible without the individual's intervention to an indefinite number of natural persons."

To involve the process of Privacy by Design and by Default in the design and construction process in compliance with data protection, various concept and design strategies have been formed. Distinguished are data oriented strategies and strategies focusing on the processing itself,[377] such as:

1. Minimizing the amount of personal data being processed is the first strategy. Making sure to only process data proportionally meaning no collection of unnecessary personal data.
2. Hiding personal data and their interrelations from plain view. This should mitigate data abuse.
3. Processing personal data "in a distributed fashion"[378]. Known as "separate" strategy. Whenever possible personal data should be processed in separate compartments.
4. Following the "aggregate" strategy, the amount of detail that remains in personal data is to be restricted by processing at the highest level of aggregation.
5. Guaranteeing for data subjects to sufficiently be informed whenever their data is being processed. This ensures transparency of data process for data subjects.
6. Enabling appropriate level of control so "data subjects should be provided agency over the processing of their personal data"[379].
7. Enforcing a privacy policy, that is compliant with legal requirements.
8. Being able to demonstrate the compliance with data protection requirements, to be in control of the data process.

There is substantial overlap between these strategies and the objectives of the SODA project. The latter aims precisely such distributed computation, where data are processed in separate shares. Moreover, in SODA's Multi Party Computation based technology, data processors do not actually need direct access to consumer and personal data anymore, while providing the same outcome as without this technology. Furthermore, a special consent control component helps to ensure that data subjects are more in charge and therefore more confident to let their data processed.

[376] European Union Agency for Network and Information Security, Privacy and Data Protection by Design – from policy to engineering, p. 11.
[377] European Union Agency for Network and Information Security, Privacy and Data Protection by Design – from policy to engineering, pp. 18 ff
[378] European Union Agency for Network and Information Security, Privacy by Design in Big Data, p. 22.
[379] European Union Agency for Network and Information Security, Privacy by Design in Big Data, p. 22.

Consequently, the design of SODA-technologies not only helps controllers to comply with the requirement of Privacy by Design of GDPR, it also offers insurance for data subjects that their data will be used in a secure way.

Another obligation for the controller is to carry out data protection impact assessment, which is a direct manifestation of the risk-based approach of compliance opted by the GDPR.[380] According to Art. 35 (1) GDPR the impact assessment refers to the fact that some types of processing personal data, in particular when using new technologies, are likely to result in a high risk to the rights and freedoms of natural persons. Therefore, the controller shall, prior to the processing – that is, prior to the first processing activity, which is in most cases the collection – carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. Art. 35 (3) mentions 3 exampled cases, when the impact assessment shall in particular be required, while Art. 35 (7) lays down the minimum assessment requirements for the controller.

Additionally, Recital 84 highlights that an impact assessment is implemented to promote the compliance with the GDPR[381] as the evaluation resulting from the origin, nature, particularity and severity of risks should be taken into account when determining the appropriate measures. Since the GDPR does not provide any explanation on what constitutes "high-risk" or "risk", it is recommended for organisations to undertake the assessment whenever there is a reasonable chance for a risk to the rights and freedoms of natural persons.

Furthermore, data controllers oblige to a breach notification obligation. Pursuant to Art. 4 No. 12 a data breach within the provisions of the GDPR means

> "a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed".

It is therefore a wider interpretation than a simple unauthorised disclosure or data leakage. The rules of notification are different towards the supervisory authority and towards the data subjects. According to Art. 33 (1) GDPR the controller has to inform the supervisory authority without undue delay, or exceptionally, where feasible, not later than 72 hours after having become aware of it, unless the personal data breach is unlikely to result in a risk to the rights and freedoms of natural persons. The minimum content requirements of the notification as laid down in Art. 33 (3) include a description of the nature of the breach, the approximate number and categories of the affected data as well as data subjects, the likely consequences, contact details and finally the countermeasures taken. However, it might be hard to tell how many data subjects or data records have been lost if a server or more servers that were used for the analysis were compromised, due to the random distribution of data shares during the analytics.

Art. 34 (1) regulates the communication of data breach to the data subject when the personal data breach is likely to result in a high risk to the rights and freedoms of natural persons. This must happen without undue delay, in a plain, comprehensible language. Art. 34 (3) on the other hand states that controllers are exempted from the notification obligation if they implemented appropriate technical and organizational measures, "in particular those that render the personal data unintelligible to any person who is not authorised to access it, such as encryption".

Consequently, controllers may escape the obligation to communicate a personal data breach to the data subject, when they implement state-of-the-art encryption technologies.

[380] see section 3.1.3 for more details on risk-based approach concerning sensitive data
[381] Hansen, DuD 2016, 587 (588).

3.2.2 Processing on behalf of the Controller (Processor)

Controllers have the opportunity not to carry out the processing themselves, but to appoint another actor, a processor to do so. This is general practice for organisations whose main business profile is outside the IT-sector. According to the definition established in Art. 4 No. 8 GDPR, data processor is:

> "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller."[382]

The two key conditions to qualify as a processor are firstly to be a separate legal entity with respect to the controller,[383] secondly, processing personal data on the controller's behalf. The lawfulness of his processing activity depends on the mandate given by the controller.[384] This has a dual meaning. On the one hand, processors must act in the controller's best interest, and not for their own purposes, and on the other hand, they are bound to the controller's instructions, that is to say, "shall not process data except on instructions from the controller".[385]

The data controller obliges to several obligations when appointing a data processor. Art. 28 (1) requires the controller to make sure to use only such processors who can provide "sufficient guarantees to implement appropriate technical and organisational measures in such a manner that processing will meet the requirements of this Regulation and ensure the protection of the rights of the data subject." Furthermore, a contract or legal act defined by Art. 28 (3) must govern the processing by the processor. According to Recital 81 GDPR, individual contract or standard contractual clauses which are adopted either directly by the Commission or by a supervisory authority in accordance with the consistency mechanism and then adopted by the Commission may be chosen. Art. 28 (3) also regulates the minimal content to be set out in the contract. These are in particular:

a) "processes the personal data only on documented instructions from the controller, including with regard to transfers of personal data to a third country or an international organisation, unless required to do so by Union or Member State law to which the processor is subject; in such a case, the processor shall inform the controller of that legal requirement before processing, unless that law prohibits such information on important grounds of public interest;

b) ensures that persons authorised to process the personal data have committed themselves to confidentiality or are under an appropriate statutory obligation of confidentiality;

c) takes all measures required pursuant to Article 32;

d) respects the conditions referred to in paragraphs 2 and 4 for engaging another processor;

e) taking into account the nature of the processing, assists the controller by appropriate technical and organisational measures, insofar as this is possible, for the fulfilment of the controller's obligation to respond to requests for exercising the data subject's rights laid down in Chapter III;

f) assists the controller in ensuring compliance with the obligations pursuant to Articles 32 to 36 taking into account the nature of processing and the information available to the processor;

---

[382] Since the processor processes data exclusively on behalf of the controller, this constellation is also referred to as „order-processing"
[383] otherwise, it would be „in-house" processing
[384] Art. 29-Working Party, Opinion 01/2010, WP 169, 22
[385] Art. 29 GDPR

g) at the choice of the controller, deletes or returns all the personal data to the controller after the end of the provision of services relating to processing, and deletes existing copies unless Union or Member State law requires storage of the personal data;

h) makes available to the controller all information necessary to demonstrate compliance with the obligations laid down in this Article and allow for and contribute to audits, including inspections, conducted by the controller or another auditor mandated by the controller.

With regard to point (h) of the first subparagraph, the processor shall immediately inform the controller if, in its opinion, an instruction infringes this Regulation or other Union or Member State data protection provisions."

Should the processor either reach further in processing the data as to what the controller intended to or use the data for its own purposes, it will be considered as controller relating to that processing operations.[386] This shift to determining and controlling any data processing leads to a re-qualification of the processor now as a data controller – with all its obligations. For example, if the original processor starts using the stored customer data in order to provide commercial advertising in a manner not intended by the original controller, with respect to this new processing action, the former processor becomes controller, since he is the one setting the new purpose.[387]

Regarding the processing established in the mandate, no permission is needed for data transfer between the controller and the processor.

It is also possible, that different phases of the processing are outsourced to multiple different processors. It is perfectly likely for two separate organizations to be data processors of the same data, e. g. one of them runs the analytics whereas the other stores the data – both are data processors of the data. They can either have a direct contractual relation with the controller, but processors may also engage a sub-processor or contract a data processing service provider, provided that this occurs with the knowledge and authorization of the controller. Pursuant to Art. 28 (2) GDPR the processor „shall not engage another processor without prior specific or general written authorisation of the controller. In the case of general written authorisation, the processor shall inform the controller of any intended changes concerning the addition or replacement of other processors, thereby giving the controller the opportunity to object to such changes". This should ensure, that the controller stays the determining party. Article 28 (4) elaborates on the case where a processor engages a sub-processor stating that "the same data protection obligations as set out in the contract or other legal act between the controller and the processor (…) shall be imposed on that other processor by way of a contract or other legal act under Union or Member State law, in particular providing sufficient guarantees to implement appropriate technical and organisational measures (…).Where that other processor fails to fulfil its data protection obligations, the initial processor shall remain fully liable to the controller for the performance of that other processor's obligations."

Additional obligations imposed on the processor by the GDPR are:

- designating a representative in the Union according to Art. 27 (1) GDPR
- maintaining a record of all categories of processing activities carried out on behalf of a controller according to Art. 30 (2) GDPR
- cooperating with the supervisory authority according to Art. 31 GDPR
- implementing technical and organisational measures according to Art. 32 (1) GDPR

[386] Art. 28 (10) GDPR
[387] Art. 29-Working Party, Opinion 05/2012 on cloud computing, WP 196, p. 14.

- data breach reporting duty without undue delay towards the controller, without exemption according to Art. 33 (2) GDPR

- designating a data protection officer according to Art. 37 GDPR

- complying with the principles for personal data transfer to third countries according to Articles 44 ff. GDPR

3.2.3 Liability of the Controller and Processor

The GDPR introduced substantial changes relating to the liability regime of the data controller and processor. The liability remains strict, "no fault" liability, i. e. the controller cannot escape by demonstrating the absence of "personal fault". [388] The burden of proof lies on the processor or the controller, therefore, according to Article 30 GDPR they are both obliged to fulfil their duty to document the processing. This might be advantageous to the affected data subject claiming compensation for damages as it is the responsibility of the processing parties to prove that they are not liable. On the other hand, the affected person must still succeed in proving the given performance as unlawful and the causation of the unlawful processing for the damages. It has been criticized that this might not be possible for the affected persons because they will not have insight into, or be able to document, the controller's or the processor's internal procedures.[389] The other side of the coin is, that data subjects do not need to demonstrate that an unlawful act was committed personally by the controller, they must only.[390]

According to Article 82 GDPR, if data has been processed unlawfully, the data subject has the right to claim compensation, including for non-pecuniary damages. Notably the GDPR establishes a dual system of liability, since compensation claim can be directed towards the controller and the processor as well:

> "Any person who has suffered material or non-material damage as a result of an infringement of this Regulation shall have the right to receive compensation from the controller or processor for the damage suffered."

This results in a cumulative liability. In order to ensure effective compensation for data subjects, controllers and processors that **are involved in the same processing and are responsible for any damage** caused, **each shall be held liable for the entire damage**. However, a processor or controller that is held liable to pay compensation on this basis is entitled to recover from other relevant parties, with keeping in mind that the part of the compensation should correspond to their part of the responsibility for the damage. [391]

The GDPR incorporates the possibility to avoid liability for damages. However, unlike the Directive, the GDPR does not mention force majeure as an exemption, meaning that controllers may bear the risk in force majeure cases:

> "A controller or processor shall be exempt from liability under paragraph 2 if it proves that it is not in any way responsible for the event giving rise to the damage."

If several processors or joint controllers caused the damages, this would serve as a further advantage for the affected person as illustrated below:

---

[388] for a thorough assessment of the liability see: van Alsenoy, JIPITEC Vol. 7, 2016, 271 ff.
[389] Roßnagel/Richter/Nebel, ZD 2013, 103 (108).
[390] the principle of accountability this does not alter the burden of proof placed upon the data subject, being able to demonstrate does not equal an actual demonstration, therefore, controllers have to be ready to demonstrate compliance only when called upon
[391] van Alsenoy, JIPITEC Vol. 7, 2016, 271 (282)

> "Where more than one controller or processor, or both a controller and a processor, are involved in the same processing and where they are, under paragraphs 2 and 3, responsible for any damage caused by processing, each controller or processor shall be held liable for the entire damage in order to ensure effective compensation of the data subject."

Joint controllers are liable jointly and individually. Concerning the external relation towards the data subject the GDPR makes joint controllers fully liable and allows that the data subject may claim the compensation from any of the controllers. This means e. g. that if one of the controllers tries to avoid its obligation action can be brought against the other even if it is the first who lead to the data breach. In their internal relation to one another, once "full compensation" has been paid to the affected data subject, joint controllers may recover damages from one another.

However, the GDPR offers exemption to processors who are not responsible for the damage caused by the processing of a controller:

> "Any controller involved in processing shall be liable for the damage caused by processing which infringes this Regulation. A processor shall be liable for the damage caused by processing only where it has not complied with obligations of this Regulation specifically directed to processors or where it has acted outside or contrary to lawful instructions of the controller."

This exception above is a positive provision for cloud computing providers who act as processors. The processor shall be exempted from liability if they are able to prove that they are not in any way responsible for the damage However, the processor will only be exempted from liability if the instructions of the controller have been lawful; if, in the processor's opinion, an instruction infringes the GDPR or other Union or Member State data protection provisions and he or she would according to Article 28 Par. 3 GDPR be obliged to immediately inform the controller about this issue, the processor will be liable and not be able to refer to the instruction given by the controller.[392]

If a controller or processor is liable for the damage, they can claim back parts of the compensation from the other responsible party in accordance to Art. 82 (5):

> "Where a controller or processor has (...) paid full compensation for the damage suffered, that controller or processor shall be entitled to claim back from the other controllers or processors involved in the same processing that part of the compensation corresponding to their part of responsibility for the damage, in accordance with the conditions set out in paragraph 2."

3.3 Rights of the Data Subject

3.3.1 Right of Access

Art. 15 GDPR regulates the right of access by the data subject. The data subject has the right to obtain confirmation from the controller as to whether or not personal data concerning him or her are being processed, and, where that is the case, access to the personal data and to the information as displayed in Art. 15 (1) lit. a) to h) GDPR. Art. 15 (2) GDPR provides the right to be informed of the appropriate safeguards where personal data are transferred to a third country or to an international organization.[393] The provision of a copy of the personal data undergoing processing is regulated in Art. 15 (3) GDPR; whereby Art. 15 (4) GDPR offers restrictions regarding that copy.

[392] Becker, in: Plath (supra note 33) Art. 82 Recital 6.

[393] Art. 15 (1) and (2) not only include the certain data being processed, but they also refer to metadata, cf. Bäcker, in Kühling/Buchner (supra note 33) Art. 15 Rn. 10

Art. 15 (1) GDPR splits the right of access into two steps. The first step requires the controller to inform the data subject, due to request, if personal data are being processed. This also means, in case personal data is not being processed, the controller should give a negative response. Once the answer is given the data subject can, as the second step, require information of which data is being processed. This information includes all data the controller presently has; not data the controller has processed in the past. Consequently, the time of the request serves as orientation point which data the data subject has to be informed about. In addition to the processed data also meta-information is included in the right of access, which the data subject should already been informed about due to Art. 13 or 14 GDPR. This does not change the fact that the right of access also applies for such information. Especially since meta-information can change or increase. Therefor the controller has to update this information before informing the data subject in accordance with Art. 15 GDPR. More precisely the controller is required to provide the purposes of the processing; the categories of personal data concerned and the recipients or categories of recipient to whom the personal data have been or will be disclosed, in particular recipients in third countries or international organizations – that is, information has to be given regardless of where the recipient is. Furthermore, the envisaged period for which the personal data will be stored, or, if not possible, the criteria used to determine that period should be provided to the data subject as well as in accordance with Art. 15 (1) lit. e) the right of rectification (Art. 16) or erasure of personal data or restriction of processing of personal data concerning the data subject (Art. 18) or to object to such processing (Art. 21). Art. 15 (1) lit. f) makes sure the data subject is informed about the right to lodge a complaint with a supervisory authority. In case information has not been obtained from the data subject the controller has to provide any available information as to their source. Lastly, Art. 15 (1) lit h) provides a regulation to inform of the existence of automated decision-making, including profiling (Art. 22) and meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

Art. 15 (2) regulates that the data subject shall have the right to be informed of the appropriate safeguards pursuant to Art. 46 relating to the transfer, where personal data are transferred to a third country or to an international organization.

Further the controller shall provide a copy of the personal data undergoing processing in accordance to Art. 15 (3). The copy includes the information provided by the controller due to Art. 15 (1) and must be provided without any cost. For any further copies requested by the data subject though, the controller may charge a reasonable fee based on administrative costs. “Further copies” refer to the copy of Art. 15 (3) s. 1. The copy can also be provided in electronic form unless otherwise requested by the data subject.

Art. 15 (4) regulates that the right to obtain a copy referred to in paragraph 3 shall not adversely affect the rights and freedoms of others. This does not mean the controller can refuse any information based on the reason other rights and freedoms are going to be affected. As rights and freedoms of others Recital 63 displays: trade secrets or intellectual property and in particular the copyright protecting the software. Not included is the right of others of an obligation of professional secrecy.

3.3.2 Right of Rectification

According to Art. 16 GDPR, the data subject shall have the right to obtain from the controller without undue delay the rectification of inaccurate personal data concerning him or her. Taking into account the purposes of the processing, the data subject shall have the right to have incomplete personal data completed, including by means of providing a supplementary statement.

3.3.3 Right of Erasure – “Right to be Forgotten”

One of the major innovations of the GDPR is the right to erasure, regulated in Article 17 GDPR, which takes up the so-called “right to be forgotten” established by the ECJ in its Google Spain decision in

2014 for search engines to remove links to webpages that appear when searching a person's name.[394] The GDPR now expands this right to all data controllers.

According to Art. 17 (1) GDPR personal data concerning the data subject shall be erased "without undue delay" by the controller if Art. 17 (1) lit. a) to f) applies, e. g. if the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed or if the data subject withdraws consent and where there is no other legal ground for the processing or if the personal data have been unlawfully processed.

Art. 17 (2) refers to processing data online, and states that controllers who made the data public are obliged to take all reasonable steps to inform other controllers which are processing the data that the data subject has requested erasure of. The erasure shall include any links to, or copy or replication of those personal data. The test of reasonability refers to available technology as well as the cost of implementation.

On the other hand, Art. 17 (3) offers a few exceptions from the obligation to erase the data, such as when processing of the personal data is necessary

a) for exercising the right of freedom of expression and information;

b) for compliance with a legal obligation which requires processing by Union or Member State law to which the controller is subject or for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;

c) for reasons of public interest in the area of public health in accordance with points (h) and (i) of Article 9 (2) as well as Article 9 (3);

d) for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89 (1) in so far as the right referred to in paragraph 1 is likely to render impossible or seriously impair the achievement of the objectives of that processing; or

e) for the establishment, exercise or defence of legal claims.

3.3.4 Right to Restriction of Processing

Art. 18 grants the data subject the right of the person responsible to require the restriction of processing. The data subject may invoke the facts listed in Art. 18 (1) such as the accuracy of the data is disputed, the processing is unlawful, the data is no longer needed by the processor and the data subject is required for the prosecution of legal claims and if the person concerned has objected to the processing referred to in Art. 21 (1).[395]

Art. 18 (2) sets out the exceptions in which the right to limit processing cannot be exercised. This includes processing by the consent of the data subject, processing for the purpose of pursuing legal claims, processing being used to protect the rights of another natural or legal person or for reasons of significant public interest.[396] Art. 18 (3) then concludes with a procedural rule to remove the restriction of processing, which justifies a duty of disclosure of the person responsible. The legal consequences of a breach of duty on the part of the controller stem from fines under Article 83 (3) lit b) and possible liability claims under Art. 82.[397] Should the personal data have been anonymised or pseudonymised in the course of processing, then the question arises as to whether the data must first be converted back

[394] ECJ, Judgment of 13 May 2014, Case C-131/12 – Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez; see regarding the "right to be forgotten" Mantelero, Computer Law & Security Review 2013, 229 ff.; Tamò/George, JIPITEC 2014, 71 ff.

[395] Herbst, in: Kühling/Buchner (supra note 33) Art. 18, Rn. 10-24

[396] Herbst, in: Kühling/Buchner (supra note 33) Art. 18, Rn. 35-44

[397] Gola, (supra note 109) Art. 18, Recital 19

into a personal data record and whether there is also a balance of interests here. If the conditions of Art. 18 (1) are fulfilled, the question is superfluous because only the processing is restricted, i. e. the options for dealing with the data are restricted.

4 Bibliography

Albrecht, Jan Philipp: Das neue EU-Datenschutzrecht – von der Richtlinie zur Verordnung - Überblick und Hintergründe zum finalen Text für die Daten- schutz-Grundverordnung der EU nach der Einigung im Trilog. In CR 2016, 88

Angiuli, Olivia; **Blitzstein**, Joe; **Waldo**, Jim: How to De-identify Your Data? In Privacy and Right2015, Vol. 13 Issue 8, available at: https://dl.acm.org/citation.cfm?id=2838930

Art. 29-Working Party: Opinion 04/2007 on the concept of personal data, WP 136, 20/06/2007. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf.

Art. 29-Working Party: Opinion 03/2010 on the principle of accountability, WP 173, 13/07/2010. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2010/wp173_en.pdf

Art. 29-Working Party: Opinion 01/2010 on the concepts of "controller" and "processor", WP 169, 16/02/2010. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2010/wp169_en.pdf

Art. 29-Working Part: Opinion 08/2010 on applicable law, WP 179, 16/12/2010. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2010/wp179_en.pdf.

Art. 29-Working Part: Advice paper on special categories of data. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/others/2011_04_20_letter_artwp_mme_le_bail_directive_9546ec_annex1_en.pdf

Art. 29-Working Party: Opinion 15/2011 on the definition of consent, WP 187, 13/07/2011. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2011/wp187_en.pdf.

Art. 29-Working Party: Opinion 05/2012 on Cloud Computing, WP 196, 01/07/2012. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp196_en.pdf.

Art. 29-Working Party: Opinion 03/2013 on purpose limitation, WP 203, 02/04/2013. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf

Art. 29-Working Party: Opinion 05/2014 on Anonymisation Techniques, WP 216, 10/04/2014. Availabe at: http://www.cnpd.public.lu/de/publications/groupe-art29/wp216_en.pdf.

Art. 29-Working Party: Statement on Statement of the WP29 on the impact of the development of big data on the protection of individuals with regard to the processing of their personal data in the EU, WP 221, 16/09/2014. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp221_en.pdf

Art. 29-Working Party: Opinion 01/2016 on the EU – U.S. Privacy Shield draft adequacy decision, WP 238, 13/04/2016. Availabe at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2016/wp238_en.pdf.

Art. 29-Working Party: ANNEX – health data in apps and devices, pp. 1 ff., available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf

Art. 29-Working Party: Guidelines on Consent under Regulation 2016/679, WP 259, 28/11/2017

Art. 29-Working Party: Guidelines on transparency under Regulation 2016/679, WP 260, 28/11/2017

Bergt, Matthias: IP-Adressen: EU-Kommission gibt BGH Nachhilfe in Sachen Grundrechte. 13/09/2015. Available at: http://www.cr-online.de/blog/2015/09/13/ip-adressen-eu-kommission-gibt-bgh-nachhilfe-in-sachen-grundrechte/

Bergt, Matthias: Die Bestimmbarkeit als Grundproblem des Datenschutzrechts – Überblick über den Theorienstreit und Lösungsvorschlag. In ZD 2015, 365 ff.

Boehme-Neßler, Volker: Das Ende der Anonymität – Wie Big Data das Datenschutzrecht verändert. In DuD 2016, 419 ff.

Brink, Stefan; **Eckhardt,** Jens: Wann ist ein Datum ein personenbezogenes Datum? Anwendungsbereich des Datenschutzrechts. In ZD 2015, 205 ff.

Brisch, Klaus: **Pieper,** Fritz: Das Kriterium der „Bestimmbarkeit" bei Big Data-Analyseverfahren: Anonymisierung, Vernunft und rechtliche Absicherung bei Datenübermittlungen. In CR 2015, 724 ff.

Buchner, Benedikt: Grundsätze und Rechtmäßigkeit der Datenverarbeitung unter der DS-GVO. In DuD 2016, 155 ff.

Chassang, Gauthier: The impact of the EU general data protection regulation on scientific research. DOI: 10.3332/ecancer.2017.709. Available at: https://www.researchgate.net/publication/312251732_The_impact_of_the_EU_general_data_protection_regulation_on_scientific_research

Dammann, Ulrich: Erfolge und Defizite der EU-Datenschutzgrundverordnung Erwarteter Fortschritt, Schwächen und überraschende Innovationen. In ZD 2016, 307 ff.

Eckhardt, Jens: IP-Adresse als personenbezogenes Datum – neues Öl ins Feuer. In CR 2011, 339 ff.

Ehmann, Eugen; **Selmayr**, Martin: Datenschutz-Grundverordnung – Kommentar. 1th Edition, Munich, 2017.

El Khoury, Alessandro: Dynamic IP Addresses Can be Personal Data, Sometimes. A Story of Binary Relations and Schrödinger's Cat. In European Journal of Risk Regulation 2017, Vol 8, p. 191

ENISA: Privacy and Data Protection by Design – from policy to engineering, 2014. Available at: https://www.enisa.europa.eu/publications/privacy-and-data-protection-by-design.

ENISA: Privacy by design in big data – An overview of privacy enhancing technologies in the era of big data analytics, 2015. Available at: https://www.enisa.europa.eu/publications/big-data-protection

Esayas, Samson Yoseph: The role of anonymisation and pseudonymisation under the EU data privacy rules: beyond the 'all or nothing' approach. In European Journal of Law and Technology (2015) Vol 6, No 2, 1 ff.

Eskens, Sarah: Profiling the European Citizen in the Internet of Things: How will the General Data Protection Regulation Apply for this Form of Personal Data Processing, and How Should it? University of Amsterdam, Institute for Information Law (IViR) 22/03/2016, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2752010

Forgó, Nikolaus: My health data—your research: some preliminary thoughts on different values in the General Data Protection Regulation. In International Data Privacy Law (2015) Vol. 5, No. 1, 54 ff.

Gola, Peter; **Lepperhoff**, Niels: Reichweite des Haushalts- und Familienprivilegs bei der Datenverarbeitung Aufnahme und Umfang der Ausnahmeregelung in der DS-GVO

Gola, Peter: DS-GVO Datenschutz-Grundverordnung – Kommentar. 1th Edition, Munich, 2017.

El Emam, Khaled; **Álvarez**, Cecilia: A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymisation techniques. In International Data Privacy Law, Vol. 5 2015, 73 ff.

El Emam, Khaled; **Rodgers**, Sam; **Malin**, Bradley: Anonymising and sharing individual patient data. In BMJ, 2015, 350 ff. Available at: http://www.bmj.com/content/350/bmj.h1139

Ernst, Stefan: Die Einwilligung nach der Datenschutzgrundverordnung Anmerkungen zur Definition nach Art. 4 Nr. 11 DS-GVO. In ZD 2017, 110 ff.

Hammer, Volker; **Knopp**, Michael: Datenschutzinstrumente Anonymisierung, Pseudonyme und Verschlüsselung. In DuD 2015, 503

Hansen, Marit: Datenschutz-Folgenabschätzung – gerüstet für Datenschutzvorsorge? In DuD 2016, 587 ff.

Härting, Niko: Datenschutzreform in Europa: Einigung im EU-Parlament: Kritische Anmerkungen. In CR 2013, 715 ff.

Härting, Niko: Datenschutz-Grundverordnung – Das neue Datenschutzrecht in der betrieblichen Praxis. Cologne, 2016.

Helbing, Thomas: Big Data und der datenschutzrechtliche Grundsatz der Zweckbindung. In Kommunikation und Recht 2015, 145 ff.

Hintze, Mike; **El Emam,** Khaled: White Paper on Comparing the Benefits of Pseudonymisation and Anonymisation Under the GDPR. Available at: https://iapp.org/resources/article/comparing-the-benefits-of-pseudonymization-and-anonymization-under-the-gdpr/

Hoeren (ed.): Big Data und Recht. Munich, 2014

Hon, W Kuan; **Hörnle,** Julia; **Millard,** Christopher: Data Protection Jurisdiction and Cloud Computing – When are cloud Users and Providers Subject to EU Data Protection Law?, The Cloud of Unknowing, Part 3. 09/02/2012.

Hon, W Kuan; **Millard**, Christopher; **Walden**, Ian: Who is Responsible of "Personal Data" in Cloud Computing?, The Cloud of Unknowing, Part 2. 21/03/2011. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1794130.

Hullen, Nils: Anonymisierung und Pseudonymisierung in der Datenschutzgrundverordnung. PinG 2015, 210 ff.

Karg, Moritz: Anonymität, Pseudonyme und Personenbezug revisited? In DuD 2015, 520 ff.

Karthauser, Ingemar; **Schmitt**, Florian: Der Niederlassungsbegriff und seine praktischen Auswirkungen Anwendbarkeit des Datenschutzrechts eines Mitgliedstaats auf ausländische EU-Gesellschaften. In ZD 2016, 155 ff.

Kelleher, Denis: In Breyer decision today, Europe's highest court rules on definition of personal data. 19/10/2016. Available at: https://iapp.org/news/a/in-breyer-decision-today-europes-highest-court-rules-on-definition-of-personal-data/

Keppeler, Lutz M.: „Objektive Theorie" des Personenbezugs und „berechtigtes Interesse" als Untergang der Rechtssicherheit? Eine Analyse der Schlussanträge des Generalanwalts in der Rechtssache C-582/14 (Speicherung dynamischer IP-Adressen). In CR 2016, 360 ff.

Knopp, Michael: Pseudonym – Grauzone zwischen Anonymisierung und Personenbezug. In DuD 2015, 527 ff.

Knopp, Michael: Muss die Wirkung von Verschlüsselung neu gedacht werden? In DuD 2015, 542 ff.

Kroschwald, Steffen: Verschlüsseltes Cloud Computing, Auswirkungen der Kryptografie auf den Personenbezug in der Cloud. In ZD 2024, 75 ff.

Kühling, Jürgen; **Martini,** Mario: Die Datenschutz-Grundverordnung: Revolution oder Evolution im europäischen und deutschen Datenschutzrecht? In EuZW 2016, 448 ff.

Kühling, Jürgen; **Buchner**, Benedikt: Datenschutz-Grundverordnung – Kommentar; 1th Edition, Munich, 2017

Lang, Markus: Reform des EU-Datenschutzrechts Einheitliche Regelungen mit hohem Datenschutzniveau geplant. In Kommunikation und Recht 2012, 145 ff.

Laue, Philip: Öffnungsklauseln in der DS-GVO – Öffnung wohin? Geltungsbereich einzelstaatlicher (Sonder-)Regelungen. In 2016, 463 ff.

Leenes, Ronald: Accounability and transparency in Big Data Land, DSC/t Blog, 2016, available at: https://www.tilburguniversity.edu/research/institutes-and-research-groups/data-science-center/blogs/data-sience-blog-ronald-leenes/

Maisch, Michael Marc: Nutzertracking im Internet. In ITRB 2011, 13 ff.

Maldoff, Gabriel: How GDPR changes the rules for research. Available at: https://iapp.org/news/a/how-gdpr-changes-the-rules-for-research/

Maldoff, Gabriel: The Risk-Based Approach in the GDPR: Interpretation and Implementation. Available at: https://iapp.org/resources/article/the-risk-based-approach-in-the-gdpr-interpretation-and-implications/

Mantelero, Alessandro: The EU Proposal for a General Data Protection Regulation and the roots of the 'right to be forgotten'. In Computer Law & Security Review 2013, 229 ff.

Marnau, Ninja: Anonymisierung, Pseudonymisierung und Transparenz für Big Data – Technische Herausforderungen und Regelungen in der Datenschutz-Grundverordnung. In DuD 2016, 428 ff.

Mayer-Schönberger, Viktor; **Padova,** Yann: Regime Change? Enabling Big Data through Europe's New Data Protection Regulation. In: The Columbia Science and Technology Law Review 2016, 315 ff. Available at: http://stlr.org/download/volumes/volume17/SchonbergerPadova.pdf.

Meyerdierks, Per: Sind IP-Adressen personenbezogene Daten?. In MMR 2009, 8 ff.

Mostert, Menno; **Bredenoort**, Annelien L.; **van der Sloot**, Bart ;**van Delden**, Johannes J.M.: From Privacy to Data Protection in the EU: Implications for Big Data Health Research. In European Journal of Health Law Vol. 24, 2017, 43 ff.

Nink, Judith; **Pohle,** Jan: Die Bestimmbarkeit des Personenbezugs - Von der IP-Adresse zum Anwendungsbereich der Datenschutzgesetze. In MMR 2015, 563 ff.

OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. Available at: http://www.oecd.org/sti/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm#memorandum

Ohm, Paul: Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. In UCLA Law Review 2010, Vol. 57 1701 ff.

Ohm, Paul: Sensitive Information. In Southern California Law Review Vol. 88, 2015, 1125 ff.

Oostveen, Manon: Identifiability and the Applicability of Data Protection to Big Data. In International Data Privacy Law 2016, 299 ff.

Paal, Boris; **Pauly,** Daniel A.: Datenschutz-Grundverordnung – Kommentar, 1th Edition, Munich, 2017

Plath, Kai-Uwe (ed.): Kommentar zum BDSG und zur DSGVO sowie den Datenschutzbestimmungen des TMG und TKG. 2nd Edition, Cologne, 2016.

Pordesch, Ulrich; **Steidle**, Roland: Entfernen des Personenbezugs mittels Verschlüsselung durch Cloudnutzer. In DuD 2015, 536

Raabe, Oliver; **Wagner**, Manuela: Verantwortlicher Einsatz von Big Data, Ein Zwischenfazit zur Entwicklung von Leitplanken für die digitale Gesellschaft. In DuD 2016, 434 ff

Roßnagel, Alexander; **Richter,** Philipp; **Nebel**, Maxi: Besserer Internetdatenschutz für Europa Vorschläge zur Spezifizierung der DS-GVO. In ZD 2013, 103

Roßnagel, Alexander; **Richter,** Philipp; **Nebel**, Max: Was Bleibt vom Europäischen Datenschutzrecht? Überlegungen zum Ratsentwurf der DS-GVO. In ZD 2015, 455 ff.

Roßnagel, Alexander; **Scholz,** Philip: Datenschutz durch Anonymität und Pseudonymität – Rechtsfolgen der Verwendung anonymer und pseudonymer Daten. In MMR 2000, 721

Rumbold, John Mark Michael; **Pierscionek,** Barbara: The Effect of the General Data Protection Regulation on Medical Research. In Journal of Medical Internet Research 2017 Feb; 19 (2): e47, doi: 10.2196/jmir.7108

Reid, Alan S: The European Court of Justice case of Breyer. In: Journal of Information Rights, Policy, and Practice 2017, Vol. 2. Available at: https://journals.winchesteruniversitypress.org/index.php/jirpp/article/view/32/14

Schaar, Katrin: Anpassung von Einwilligungserklärungen für wissenschaftliche Forschungsprojekte Die informierte Einwilligung nach der DS-GVO und den Ethikrichtlinien. In ZD 2017, 213 ff.

Schaar, Peter: Privacy By Design. Available at: http://www.bfdi.bund.de/SharedDocs/Publikationen/EN/0610EUPrivacyByDesign.pdf?__blob=publicationFile.

Schantz, Peter: Die Datenschutz-Grundverordnung – Beginn einer neuen Zeitrechnung im Datenschutzrecht. In NJW 2016, 1841 ff.

Specht, Louisa; **Müller-Riemenschneider**, Severin: Dynamische IP-Adressen: Personenbezogene Daten für den Webseitenbetreiber? Aktueller Stand der Diskussion um den Personenbezug. In ZD 2014, 71 ff.

Spindler, Gerald: Persönlichkeitsschutz im Internet – Anforderungen und Grenzen einer Regulierung. In Verhandlungen des 69. Deutschen Juristentages, Band I Gutachten, Munich, 2012.

Spindler, Gerald: Datenschutz- und Persönlichkeitsrechte im Internet - der Rahmen für Forschungsaufgaben und Reformbedarf. In GRUR 2013, 996 ff.

Spindler, Gerald: Datenschutz- und Persönlichkeitsrechte im Internet - Der Rahmen für Forschungsaufgaben und Reformbedarf. In GRUR-Beilage 2014, 101 ff.

Spindler, Gerald; **Schuster,** Fabian (ed.): Recht der elektronischen Medien. 3rd Edition, Munich 2015.

Spindler, Gerald; **Schmechel,** Philipp: Personal Data and Encryption in the European General Data Protection Regulation. In JIPITEC 2016, 163 ff. Available at: https://www.jipitec.eu/issues/jipitec-7-2-2016/4440/spindler_schmechel_gdpr_encryption_jipitec_7_2_2016_163.pdf.

Spindler, Gerald: Big Data und Forschung mit Gesundheitsdaten in der gesetzlichen Krankenversicherung. In Medizinrecht 2016, 691 ff.

Stadler, Thomas: EuGH entscheidet zum Personenbezug von IP-Adressen. 19/10/2016. available at: http://www.internet-law.de/2016/10/eugh-entscheidet-zum-personenbezug-von-ip-adressen.html

Spies, Axel: Cloud Computing: Keine personenbezogenen Daten bei Verschlüsselung. In MMR-Aktuell 2011, 313727

Stiemerling, Oliver; **Hartung,** Jürgen: Datenschutz und Verschlüsselung. In CR 2012, 60 ff.

Richards, Neil M.; **King**, Jonathan H.: Three Paradoxes of Big Data. In 66. Stanford Law Review 2013-2014, 41 ff. Available at: http://heinonline.org/HOL/Page?handle=hein.journals/slro66&collection=journals&id=41&startid=&endid=46

Tamò, Aurelia; **George,** Damian: Oblivion, Erasure and Forgetting in the Digital Age. JIPITEC 2014, 71 ff.

Tene, Omer; **Polonetsky**, Jules: Privacy in the Age of Big Data: A Time for Big Decisions. Stanford Law Review Vol. 64, 63 ff. Available at: http://heinonline.org/HOL/Page?handle=hein.journals/slro64&collection=journals&id=64&startid=&endid=70

Tinnefeld, Marie-Theres; **Buchner**, Benedict; **Petri**, Thomas: Einführung in das Datenschutzrecht. 5th Edition, Munich, 2012

Tikkinen-Piri, Christina; **Rohunen,** Anna; **Markkula,** Jouni: EU General Data Protection Regulation: Changes and implications for personal data collecting companies. In Computer Law & Security Review 2017, doi: 10.1016/j.clsr.2017.05.015

Ursic, Helena; **Custers,** Bart: Legal Barriers and Enablers to Big Data Reuse. In European Data Protection Law Review 2016, 209 ff.

Urgessa, Worku Gedefa: The Protective Capacity of the Criterion of Identifiability under EU Data Protection Law. In European Data Protection Law Review 2016, 521. Available at: http://heinonline.org/HOL/Page?handle=hein.journals/edpl2&collection=journals&id=556&startid=556&endid=570

Van Alsenoy, Brendan: Liability under EU Data Protection Law: From Directive 95/46 to the General Data Protection Regulation. In JIPITEC, Vol. 7, 2016, 271ff

van der Sloot, Bart; **Broeders,** Dennis: **Schrijvers,** Erik (ed.): Exploring the Boundaries of Big Data. The Hague, 2016. Available at: http://www.ivir.nl/publicaties/download/1764.

Yakoubov, Sophia; **Gadepally**, Vijay; **Schear**, Nabil; **Shen**, Emily; **Yerukhimovich**, Arkady: A Survey of Cryptographic Approaches to Securing Big-Data Analytics in the Cloud. Available at: http://www.ieee-hpec.org/2014/cd/index_htm_files/finalpapers/28.pdf

Zarsky, Tal Z: Incompatible: The GDPR in the Age of Big Data. In Seton Hall Law Review 2017, 995 ff.

Ziegenhorn, Gero; **von Heckel**, Katharina: Datenverarbeitung durch Private nach der europäischen Datenschutzreform – Auswirkungen der Datenschutz-Grundverordnung auf die materielle Rechtmäßigkeit der Verarbeitung personenbezogener Daten. In NVwZ 2016, 1585

Zuiderveen Borgesius, Frederik J.: Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation. In Computer Law & Security Review 2016, 256 ff.