# The Principle of Purpose Limitation and Big Data

**Nikolaus Forgó, Stefanie Hänold and Benjamin Schütze**

**Abstract** In recent years, Big Data has become a dominating trend in information technology. As a buzzword, Big Data refers to the analysis of large data sets in order to find new correlations—for example, to find business or political trends or to prevent crime—and to extract valuable information from large quantities of data. As much as Big Data may be useful for better decision-making and risk or cost reduction, it also creates some legal challenges. Especially where personal data is processed in Big Data applications such methods must be reconciled with data protection laws and principles. Those principles need some further analysis and refinement in the light of technical developments. Particularly challenging in that respect is the key principle of "purpose limitation." It provides that personal data must be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes. This may be difficult to achieve in Big Data scenarios. At the time personal data is collected, it may still be unclear for what purpose it will later be used. However, the blunt statement that the data is collected for (any possible) Big Data analytics is not a sufficiently specified purpose. Therefore, this contribution seeks to offer a closer analysis of the principle of purpose limitation in European data protection law in the context of Big Data applications in order to reveal legal obstacles and lawful ways to handle such obstacles.

**Keywords** Big Data · Purpose limitation · Purpose specification · Compatible use · Data protection · General data protection regulation (GDPR) · Data protection directive (DPD)

N. Forgó (✉) · S. Hänold · B. Schütze
Institute for Legal Informatics, Leibniz Universität Hannover, Hannover, Germany
e-mail: forgo@iri.uni-hannover.de



M. Corrales et al. (eds.), *New Technology, Big Data and the Law*,
Perspectives in Law, Business and Innovation, DOI 10.1007/978-981-10-5038-1_2

## Contents



# 1 Introduction

Data, or uninterpreted information, has been collected, stored and processed as long as mankind has existed. Humans have always had a desire to observe and interpret their environment and gather information that would form a solid basis for their decision-making. Yet with the emergence of computers, information technology and digital data processing the game has changed. Since then, the volume of data is growing exponentially and it is expected that by 2020 more than 44 zettabytes (44 Trillion GB) will be generated and approximately 16 zettabytes may be used in the context of Big Data applications.[1] Recent numbers are even more staggering as it is believed that by 2025 the total amount of Data will be as high as 180 zettabytes.[2]

[1] Turner et al. (2014); Cavanillas et al. (2015), p. 3.

[2] Kanellos (2016).

This enormous growth mainly stems from the increasing number of devices generating data, as well as the growing number of built in sensors in each device.[3] More and more devices are connected to the Internet and it is expected that in 2020 nearly 30 billion devices will have an Internet connection.[4] Thus, we find ourselves in an era in which the Internet of Things, i.e., devices communicating with each other, is not a far-fetched dream of the future, but is in the process of happening.

Big Data comes into play when vast amounts of raw data generated by a plethora of different sensors and devices is further stored and processed. It is a challenge for information technology experts to build the pertinent tools to process large quantities of very heterogeneous data, and thus manage this information more effectively. The ability to extract knowledge and value as a result is perceived as a competitive advantage—a future imperative—rather than a luxury. Many organizations, private companies and public institutions alike are expanding their Big Data capabilities and new business models continuously emerge.

However, not only IT professionals are challenged to find solutions for the swelling tide of data. Big Data also poses a considerable number of legal questions and issues of interest for the humanities. Many of them are discussed in research projects such as ABIDA or SoBigData in which the authors of this chapter are (co-) responsible for the legal work package.[5] For example, it is currently legally unclear how far data can be "owned" (in terms of an absolute property right), and if so *who* the owner is.[6] Furthermore, large amounts of data in the hands of one entity raise competition and antitrust law concerns.[7]

One of the most insistent legal challenges of Big Data applications resonates in data protection law, in cases where the data sets processed are to be qualified as personal data. If personal data is processed, then a Big Data provider under the European legal regime has to comply with European data protection law, i.e., the data protection legislation of the European Member States. This legislation was, to some extent, harmonized by the Data Protection Directive (DPD)[8] and will be further modified and reinforced by the European General Data Protection Regulation (GDPR),[9] applicable from May 2018 onwards.

---

[3]Kanellos (2016).

[4]Kanellos (2016).

[5]See http://www.abida.de and http://www.sobigdata.eu/ for further information.

[6]See, e.g., Zech (2012); Grützmacher (2016), pp. 485–495.

[7]See, e.g., Bundeskartellamt, Autorité de la concurrence (2016); Körber (2016), pp. 303–310; pp. 348–356.

[8]European Parliament and the Council (1995) Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

[9]European Parliament and the Council (2016), Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

One of the stable bedrock principles in European data protection law is the principle of purpose limitation. This means in general that processing of personal data in the European Union requires a clearly defined purpose at the time of data collection, and that such data cannot be reused for another purpose that is incompatible with the original purpose. This principle may constrain Big Data applications in Europe because one of the methods to leverage value from Big Data is to use data and further processed datasets for different purposes; and to analyze the data in a way that may not have been envisaged at the time the data was first collected.

This chapter is divided into six parts, which examine the principle of purpose limitation in the context of Big Data applications. Following the introduction, Sect. 2 introduces Big Data technology and delineates the problems commonly associated with the processing of personal information. Section 3 explains the basic legal framework of data protection in Europe and briefly sketches the history and development of the purpose limitation principle. Section 4 then analyses the purpose limitation principle further and outlines its interrelationship with other data protection principles in European law. To conclude, Sects. 5 and 6 focus on the new GDPR and assess whether its interpretation of the principle of purpose limitation and pertinent rules will facilitate Big Data application, in contrast to the current legal situation under the DPD. This may determine whether the law helps to induce economic growth or rather, due to a strict interpretation of the limitation of purpose, hamstrings economic activities involving Big Data.

## 2 Big Data Definition

To understand the context in which the principle of purpose limitation may be relevant, it is useful to explain what Big Data means and, even more importantly, in which (business) environments and value chains it is set. In recent years, the term "Big Data" has been used prolifically. However, until now it remains somewhat obscure what exactly Big Data means and implies. It is not a legal term but rather describes a phenomenon with a multitude of different implications in scientific disciplines, such as economics, technical disciplines, legal and social science, and probably in many further areas of life in the years to come.

Several definitions of the term Big Data have been suggested. The first and best known definition was formulated by *Laney*,[10] who proposed a three-dimensional perspective: the "three Vs" with which he described certain characteristics a Big Data application should have.[11] According to *Laney*, "Big Data is high volume, high **v**elocity, and/or high **v**ariety information assets that require new forms of processing to enable enhanced decision-making, insight discovery and process optimization."[12] "Volume" refers to the amount of data and implies that in Big Data

[10]Laney (2001).

[11]Curry (2015), p. 30.

[12]Laney (2001).

scenarios large amounts of data will be processed. "Variety" on the other hand refers to the range of different types and sources of data. It points to the fact that Big Data infrastructures need to cope with a vast array of different sources as well as a variety of different format and syntactic parameters (e.g., images texts, spreadsheets, database entries) "Velocity" refers to the requirement that, in a Big Data scenario, IT systems need to deal with streams of incoming real time data, for example, real time traffic information or electronic trading. As Big Data applications evolved, further attributes have been suggested of which the most important one is "Veracity." It refers to quality aspects of data, since their accuracy and overall quality may vary greatly. A prediction calculated by Big Data methods may thus be upset by inaccurate raw data.[13]

To understand Big Data and its legal implications it is not necessary to formulate a precise technical or legal definition, but rather to understand the value chains and interdependencies between the entities involved in Big Data ecosystems. To understand the business models and their legal implications one may compartmentalize the data handling into three separate steps, beginning with data acquisition, followed by the actual data processing (i.e., analysis, curation and storage) eventually leading to the use of the results of the Big Data analysis. Every step of such data handling may be associated with certain legal questions and effects.

Data acquisition is the process of gathering and filtering raw data before they are stored and further processed. Data can be gathered from ever increasing sensor networks in the so-called Internet of Things (IoT), acquired on online marketplaces or collected from natural persons in social media or via their smartphones, wearables and other mobile devices. The process of acquisition thus raises questions of data ownership as well as data protection, if personal information is collected. Furthermore, data acquisition raises questions of contractual relations if the data is sold and bought including the rights of the buyer in case of breach of contract following the delivery of defective data, i.e., data that is inaccurate or of lower quality than the parties have agreed upon.

The second phase that follows data acquisition is Big Data sensu stricto because only here data is merged and further processed in order to generate new insights. Although it also involves data curation and storage, more important is the actual analysis of the data by exploring and modeling data in order to highlight and extract information relevant for business or other domain-specific decisions. Merging and combining data to gain new insights may have repercussions in data protection law, as it may be that the envisaged merging and analysis is not compatible with the specified purpose articulated at the time of the collection. Or it may be that non-personal information through combining it with other information becomes personal information, because through such newly extracted data a natural person can be identified. Aside from data protection, the process of data curation and storage may also raise questions of data quality as data must be processed in such a way that it is trustworthy, accessible and in general fits the purpose for which it is cured and stored.

[13]An overview of the different Big Data definitions can be found in Curry (2015), p. 31.

The third phase of a Big Data processing scenario is represented by the usage of the results of the analysis and probably is the most significant phase in a Big Data scenario. Data usage covers a wide range of data driven activities and relies on the access to data and the results of a Big Data analysis. In other words, it is the decision-making process, which is based on the result of the Big Data analysis. This may be a "conscious" decision taken by a natural person, however, Big Data will in the future increasingly result in automated decision-making, where autonomous machines carry out certain tasks without human intervention.

Examples of such machines are robots in autonomous factories that are connected to logistic networks and independently order supplies or manage their repairs and upgrades. Manufacturing and logistics are currently undergoing an industry-wide transformation as part of the so-called "Industry 4.0." The term describes the digitization and interconnection of products, manufacturing facilities, and transport infrastructure for purposes such as supply chain management and maintenance. Industry 4.0 corresponds with Big Data, as a precondition for proper management of the decision-making process is to analyze huge amounts of (real time) data. An even more practical example is the self-driving car or other autonomous vehicles. Driverless cars need to be capable of sensing their environment and navigating without human input. This is only possible through an adequate number of sensors with which the car can detect its surroundings. If the self-driving car is to be embedded in a smart traffic scenario, it must further be capable of receiving live traffic data on congestion, road conditions, etc., to calculate the optimal route or travel speed. In order to navigate in traffic, the self-driving car therefore requires Big Data capabilities. In other words, Big Data is a precondition to operate autonomous vehicles, as the on-board computer has to process large amounts of data in a short period of time to navigate safely and predict potentially dangerous traffic situations and react to unforeseen events.

Events that may occur in connection with data usage raise numerous legal questions. However, the following part of the chapter will focus on the aspects of the protection of personal information and, in particular, the principle of purpose limitation.

## 3 The Development of the Principle of Purpose Limitation

The principle of purpose limitation has served as a key principle and stable element in European data protection law for many years.[14] To understand how it has evolved from the early instruments on human rights and data protection to the most recently enacted GDPR, a brief historical overview is needed. The following section therefore provides a short description of how the concept of purpose limitation came into being, was carved out and redefined.

[14]Article 29 WP, p. 9.

## 3.1 European Convention on Human Rights (ECHR)

The European Convention on Human Rights was adopted in 1950. Article 8 (1) of the Convention incorporates the right to privacy, according to which everyone shall have the right to respect for his private and family life, his home and his correspondence. Article 8 (2) prohibits any interference by a public authority with the exercise of this right unless such interference is in accordance with the law and necessary in a democratic society to satisfy certain public interests listed in Article 8 (2) ECHR.[15] According to Article 8, any interference with the individual's right to privacy requires justification under strictly defined conditions. Such conditions, and the fact that a legal basis is required forms a starting point for the principle of purpose limitation, as without a legal basis, a legitimate purpose, which at the same time sets limits to the interference, cannot be determined.[16]

## 3.2 Council of Europe Resolutions (73) 22 and (74) 29

Two important additional steps that should be mentioned are the Council of Europe (CoE) Resolutions (73) 22[17] and (74) 29,[18] which were elaborated further by later instruments and formulated what have become defining principles of data protection law, inter alia, the principle of purpose limitation. Principle 2 CoE Resolution (73) 22 states that, "information should be appropriate and relevant with regard to the purpose for which it has been stored." Furthermore, principle 5 determines that, "without appropriate authorization, information should not be used for purposes other than those for which it has been stored, nor communicated to third parties." CoE Resolution (74) 29, dealing with the protection of privacy in "electronic data banks" in the public sector, reiterates at first, similar to CoE 73 (22), that the information stored should be "appropriate and relevant to the purpose for which it has been stored".[19] Principle 3 (c) goes on to state "that data stored must not be used for purposes other than those which have been defined unless an exception is explicitly permitted by law, is granted by a competent authority or the rules for the

[15]Article 8 (2) ECHR lists national security, public safety or the economic wellbeing of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

[16]Article 29 WP, p. 7.

[17]Council of Europe Committee of Ministers (1973) Resolution (73) 22 on the protection of privacy of individuals vis-à-vis electronic data banks in the private sector, adopted on 26 Sept 1973.

[18]Council of Europe Committee of Ministers (1973) Resolution (74) 29 on the protection of privacy of individuals vis-à-vis electronic data banks in the public sector, adopted on 20 Sept 1974.

[19]Principle 2 (c).

use of the electronic data bank are amended." In other words, 3 (c) introduces the notion that the purpose of information storage may be changed under certain conditions.

## 3.3 Convention 108

One may say that CoE Resolutions (73) 22 and (74) 29 paved the way for another defining legislative instrument with regard to the principle of purpose limitation: Convention 108 of the Council of Europe.[20] Convention 108 was opened for signatures in January 1981. Article 5 introduces a more elaborate set of data protection principles such as lawfulness, fairness and proportionality. However, three of its five sub clauses refer to key aspects of the principle of purpose limitation. Article 5 (b) determines that personal data undergoing automatic processing shall be "stored for specific and legitimate purposes and not used in a way incompatible with those purposes" (purpose specification). Firstly, this means that it is not permissible to store data for undefined purposes, and it is left to the national legislator, to decide how such purposes must be specified.[21] Secondly it must be emphasized that Article 5 (b) introduces the notion of incompatibility when it determines that the data cannot be used "in a way incompatible" with the specific purposes; this concept has later been incorporated into the Data Protection Directive and General Data Protection Regulation. Article 5 (c) furthermore, addresses the principle of data minimization and determines that personal data must be "adequate, relevant and not excessive in relation to the purpose for which they are stored." In other words, Article 5 (c) connects the principle of data minimization and purpose limitation. Finally, Article 5 (e) interlinks the principle of purpose limitation with anonymization when it determines that "personal information undergoing automatic processing shall be preserved in a form which permits identification of the data subjects for no longer than is required for the purpose for which those data are stored."

Following principle 3 (c) CoE Resolution (74) 29, Article 9 of Convention 108 allows for derogations from Article 5 "when such derogation is provided for by the law of the Party and constitutes a necessary measure in a democratic society in the interests of: protecting State security, public safety, the monetary interests of the State or the suppression of criminal offences; protecting the data subject or the rights and freedoms of others." Furthermore, Article 9 (3) points, by reverse implication, to another important aspect regarding the principle of purpose limitation. This is that for some purposes, the individual's right to privacy may be restricted, namely when automated personal data files are "used for statistics or for

[20]Council of Europe (1981) Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, Strasbourg 28 Jan 1981.

[21]Council of Europe (1981) Explanatory Report to the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, Strasbourg 28 Jan 1981.

scientific research purposes when there is obviously no risk of an infringement of the privacy of the data subject." Admittedly Article 9 (3) is a slightly different case, as it deals with derogations from Article 8 (b–d) of the Convention. Those require additional safeguards for the data subject such as the right of notification, erasure and rectification. However, it would support the argument that changing the purpose of data storage and processing, as long as it is for statistics or scientific research purposes, is less likely to be seen as an infringement of the privacy of the data subject and not incompatible with the specified and legitimate purposes for which personal data has been stored in the first place.

### *3.4 OECD Guidelines*

The OECD Guidelines[22] governing the Protection of Privacy and Transborder Flows of Personal Data, which were adopted in 1980—almost at the same time Convention 108 was signed—have a similar approach to the purpose limitation principle, but are more specific on the exact time at which the purpose must be specified. Paragraph 9 states that the "purposes for which personal data are collected should be specified not later than at the time of data collection." Furthermore, the Guidelines also incorporate the notion of incompatibility when they state that "the subsequent use should be limited to the fulfillment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose." Finally, Paragraph 10 explicitly mentions two exceptions to Article 9, determining that use of personal data for purposes other than those specified in accordance with Paragraph 9 may be admissible "with the consent of the data subject" or "by the authority of law." The 2013 review[23] of the OECD Guidelines left these provisions unchanged.

## 4 The Purpose Limitation Principle Under the Data Protection Directive (DPD) and Its Implications for Big Data Applications

### *4.1 Starting Position*

The Data Protection Directive (DPD) dates back to the year 1995. The DPD was a prominent step to harmonize the data protection rules within the EU. It was the

[22]OECD (1980) Annex to the recommendation of the Council of 23 September 1980: Guidelines governing the protection of privacy and transborder flows of personal data.

[23]OECD (2013) Recommendation of the Council concerning Guidelines governing the Protection of Privacy and Transborder Flows of Personal Data [C(80)58/FINAL, as amended on 11 July 2013 by C(2013)79].

declared aim of the European legislator to remove the obstacles to a free flow of personal data within the EU and at the same time to harmonize the level of protection of the rights and freedoms of individuals with regard to the processing of their personal data.[24] Due to its character as a directive it had to be implemented by the Member States of the European Union into their national legal frameworks.

The European legislator has laid down in Article 6 DPD the basic European data protection law principles, namely the principle of fairness and lawfulness, the purpose limitation principle, the principle of data minimization, the data quality principle and the principle of data security. Referring to the data protection principle of purpose limitation, the Directive determines in Article 6 (1) (b) that personal data must be collected for specified, explicit and legitimate purposes and not be further processed in a manner that is incompatible with these purposes. The specification of the purpose is a core element of the framework established by the Directive. Without specifying the purpose, it is, for instance, not possible to clarify whether the processing is allowed by the applicable data protection regulations. Purpose specification is also necessary in order to determine necessary safeguards for the personal data as well as to fulfill other data protection obligations, such as to inform the data subject of the purposes of the processing of their personal data.[25] In brief, the purpose limitation principle serves two goals. On the one hand, it protects reasonable expectations of data subjects with regard to by whom and how their data shall be processed. On the other hand, it allows data controllers to process data for a new purpose within carefully balanced limits.[26]

## *4.2 Specified, Explicit and Legitimate Purpose (Purpose Specification)*

### 4.2.1 Purpose Must Be Specified

The first building block of the purpose limitation rule is that the controller, when collecting the data, needs to specify the purpose or purposes, which are intended to be served with the collected data.[27] Purpose specification requires an internal assessment and documentation by the data controller who must clearly and specifically identify the purpose of the collection.[28] This step is elementary for consideration of, and compliance with, other data protection requirements. As already mentioned, Article 6 (1) DPD provides further important data protection principles such as the principle of data minimization, which requires that only

[24]Recital 8 Directive 95/46/EC.
[25]Article 29 WP, p. 15.
[26]Article 29 WP, p. 3.
[27]Article 29 WP, p. 15.
[28]Article 29 WP, p. 15.

personal data is processed, which is adequate, relevant and not excessive in relation to the purposes for which the data are collected and/or further processed. Consequently, the data controller must consider carefully, after specifying the purpose, whether the collection and/or processing of the personal data is necessary for the aim he pursues. In order to support transparency and also to improve enforcement of the purpose limitation principle, data subjects must be informed by the data controller of the purpose of the collection, except where they already have it (Article 10 (b) and Article 11 (b) DPD). This shows that there is a connection between transparency and purpose specification. The transparency aspect enhances predictability for data subjects who know what to expect regarding the processing of their personal data by the data controller.[29]

The rule of purpose limitation would not sufficiently protect the data subjects' rights if it was permissible to use vague or very general descriptions of the envisaged purpose of data processing in order to have a broader scope of manoeuvre.[30] In this regard, the Article 29 Working Party has suggested that purported purpose specifications in such terms as "improving user's experience, marketing purposes, IT-security purposes or future research" are invalid.[31] According to the Article 29 Working Party, the required degree of specification depends on the context in which the data is collected and must be determined for every specific case. In some circumstances simple descriptions of the purpose are appropriate, while others require a more detailed specification.[32] This means in effect that, for example, large retail companies selling goods throughout Europe using complex analytic applications to tailor advertisements and offers to their customers will need to specify the purposes in more detail than a local shop, which is collecting only limited information about their customers.[33] If a data controller provides a number of services (e.g., e-mail, photograph upload, social networking functions) it must ensure users are informed about all the different purposes of the envisaged processing activities.[34] Additionally, if a gaming website service is aimed at teenagers, the age of the respective customer must be taken into account. The same is true for website services targeted at elderly people.[35]

In this context, it is also relevant to mention the limitation of purpose by the data subject by giving her informed consent. National courts,[36] as well as data protection agencies of the Member States,[37] have declared vague and/or blanket forms of

[29]Article 29 WP, p. 13.

[30]Article 29 WP, p. 16; Ehmann and Helfrich (1999), p. 113.

[31]Article 29 WP, p. 16.

[32]Article 29 WP, p. 16.

[33]Article 29 WP, p. 51.

[34]Article 29 WP, p. 51.

[35]Article 29 WP, p. 51.

[36]OLG Frankfurt/M., Judgment 17 Dec 2015—6 U 30/15; LG Berlin, Judgement 19 Nov 2013—15 O 402/12; OLG Celle, Judgement 14 Nov 1979—3 U 92/79.

[37]See, e.g., Metschke and Wellbrock (2002), pp. 27–28.

consent in data processing to be invalid. The subject's informed consent is one of the legal grounds that allows the processing of personal data.[38] Since the limits of the consent given by the data subject also constrain the possibilities of the data controller to process the personal data, this also operates as a mechanism for the data subject to stay in control of the purposes for which her personal data are used. For instance, in the medical research field the data subject cannot give valid informed consent if he is not sufficiently aware of the potential ways in which her personal data may be used. It is, in particular, not possible—when obtaining consent—to refer in a general way to future research projects of which the data subject is unable to form any real idea.[39]

### 4.2.2 Purpose Must Be Explicit

Another element of the purpose specification building kit is that the purpose specification must be explicit. This means that the specification of the purpose must be clearly disclosed and explained or expressed in an intelligible form. This must happen no later than the time of the collection of the personal data. This requirement contributes to transparency and predictability, as it allows third parties to understand how the personal data can be used and to identify the limits of the processing of the personal data.[40]

### 4.2.3 Purpose Must Be Legitimate

The purpose for which the data have been collected must be legitimate. This refers in part to the general rules that can be derived from Article 7 and Article 8 DPD, namely that the processing of personal data is prohibited unless there is a legal ground, for example, the consent of the data subject. Moreover, it provides that the purposes must be in accordance with all applicable laws as well as customs, codes of conduct, codes of ethics and contractual arrangements. Finally, the general context and facts of the case may also be considered, for instance, the nature of the relationship between data controller and data subject.[41] Ultimately, the data controller needs to ensure prior to the collection that there is a legal rule allowing the envisaged collection and further envisaged use. Furthermore, they need to take account of other relevant conditions, for example, any civil law obligation they are subject to, or, for instance, in case the data is used in a medical research project,

[38]Article 7 (a) and Article 8 (2) (a) Directive 95/46/EC.

[39]Metschke and Wellbrock (2002), pp. 27–28.

[40]Article 29 WP, p. 17.

[41]Article 29 WP, p. 20.

acknowledged ethical norms such as the Declaration of Helsinki[42] or the International Ethical Guidelines for Epidemiological Studies.[43]

## 4.3 Assessment of Compatibility

The second building block of the purpose limitation principle is the requirement that the collected data must not be further processed in a way incompatible with the purpose for which the data have been originally collected (compatible use). The Directive does not explicitly state what processing steps fall under further processing; it rather distinguishes between the very first processing, which is the collection of data, and all subsequent processing steps such as storage, analysis etc. Any processing steps following the collection of the personal data are to be seen as further processing of personal data, regardless of whether the processing is for the purpose initially specified or for any additional purpose.[44] By providing that further processing is permitted as long as it is not incompatible, it was acknowledged under the Directive that the European legislator intended to give some flexibility with regard to further use of personal data.[45]

In some cases, it is obvious that further processing is compatible, for example, if the data have been collected to specifically achieve the purpose that shall be achieved with the intended further use. In other cases, the decision whether compatibility can be established or not is not that obvious. The compatibility test must be carefully applied as processing of personal data in a way incompatible with the initially determined purposes is unlawful. The data controller cannot legitimize the further processing that is incompatible with the original purpose simply by relying on a legal ground[46] allowing the processing of the personal data.[47]

The Directive explicitly privileges further processing of personal data for historical, statistical or scientific purposes, provided that Member States implement appropriate safeguards (Article 6 (1) (b) DPD). Under the regime of the Directive it is up to the Member States to specify the appropriate safeguards to satisfy this requirement.[48] These safeguards shall preclude that the data will be used to support

[42]WMA General Assembly (2013) WMA Declaration of Helsinki—Ethical Principles for Medical Research Involving Human Subjects.

[43]Council for International Organizations of Medical Sciences (CIOMS), WHO (2008) International Ethical Guidelines for Biomedical Research Involving Human Subjects.

[44]Article 29 WP, p. 21.

[45]Article 29 WP, p. 21.

[46]National implementations of Article 7 and Article 8 Directive 95/46/EC provide legal grounds for processing personal data.

[47]Article 29 WP, p. 3.

[48]Article 29 WP, p. 28.

or justify measures or decisions against any particular individual (Recital 29 DPD).[49] The Article 29 Working Party has interpreted this requirement very broadly: any relevant impact on particular individuals—either negative or positive—shall be avoided.[50] Appropriate safeguards may be for instance early anonymization, or in cases where the purpose of the processing requires the retention of the information in identifiable form other techniques, for instance, pseudonymizing the personal data and keeping the keys coded or encrypted and stored separately.[51] The privileging rule for further processing for historical, statistical or scientific purposes covers a diversity of processing activities ranging from activities supporting public interests—such as medical research—or purely for commercial purposes, for example, market research.[52] In particular, the exemption for statistical purposes is relevant for Big Data applications that try to find correlations and new trends.

Other forms of further use not covered by the privileging rule in Article 6 (1) (b) of the Directive, are as indicated earlier, not precluded per se; this is only so if they qualify as incompatible with the original purpose.[53] Whether a given further use qualifies as compatible or incompatible will need to be assessed on a case-by-case basis.[54] The Article 29 Working Party has analyzed the legal provisions and practices in the Member States to assess the compatibility of further processing and identified key factors to be considered in the compatibility assessment:[55]

(i) The relationship between the purposes for which the personal data have been collected and the purposes of further processing;
(ii) The context in which the personal data have been collected and the reasonable expectations of the data subjects as to their further use;
(iii) The nature of the personal data and the impact of the further processing on the data subjects;
(iv) The safeguards adopted by the controller to ensure fair processing and to prevent any undue impact on the data subjects.[56]

[49]Beyleveld (2004), p. 9.
[50]Article 29 WP, p. 28.
[51]Article 29 WP, pp. 30–32; Metschke and Wellbrock (2002), p. 16.
[52]Article 29 WP, p. 29.
[53]Article 29 WP, p. 21.
[54]Article 29 WP, p. 21.
[55]Article 29 WP, pp. 23–27.
[56]Article 29 WP, p. 40, e.g., example 15: mobile phone locations help inform traffic calming measures, p. 66.

## 4.4 Consequences of the Requirements of the Purpose Limitation Principle Established by the DPD for Big Data Applications

What follows from the above is, firstly, that data controllers cannot simply collect and store any accessible data in order to have the possibility to use such data for a purpose that will be defined in the future.[57] The data processor will need to determine a specific purpose at the latest by the time of data collection. The level of detail required is a matter of degree and depends on the individual circumstances. Vague and blanket specification of the purpose will certainly not suffice.[58] Data controllers can use, for example, the Article 29 Working Party Opinion on purpose limitation[59] or national court decisions[60] for guidance on how to specify the purpose. In general, they should consider that the more the data subject is affected by the envisaged processing of her personal data the more detailed the purpose specification should be.[61] However, the data controller must take into account that there are certain legal fields, for instance, medical research, where the required level of specificity is controversial and where different approaches within the Member States exist.[62]

For the purpose specification data, controllers should determine the types of personal data that are going to be processed, the quantity of the data and also the kind of data they envisage to link with the personal data for the envisaged Big Data application. Possible usage context must be described and—in case the data shall be transferred to third parties—those should be specified, too.[63] This excludes the specification of generic purposes as, for instance, processing of the data for strategy development.[64] Companies need to question the function and the objective of

[57]Werkmeister and Brandt (2016), p. 237.

[58]Article 29 WP, p. 16.

[59]Annex 3 of the Article 29 WP Opinion 03/2013 on purpose limitation gives a number of examples to illustrate purpose specification.

[60]OLG Frankfurt/M., Judgment 17 Dec 2015-6 U 30/15; LG Berlin, Judgement 19 Nov 2013–15 O 402/12; OLG Celle, Judgement 14 Nov 1979—3 U 92/79.

[61]Bretthauer (2016), p. 272; Wolff (2016) margin number 19.

[62]In the UK, broad consent is accepted in some instances (MRC 2011, p. 6). The legal situation in Germany is still unsettled in this regard. German courts (e.g., OLG Celle, Judgement 14 Nov 1979—3 U 92/79) have viewed the use of a broader forms of consent critically in non-medical fields of peronal data processing and it is unsure how this will be translated in medical research. The Data Protection Authorities of the Land Berlin and the Land Hessen seem not to require a consent restricted to a particular research project, but the data subject must be able to gain an idea for what research projects his data will be used for (see Metschke and Wellbrock 2002, p. 27). The working group "Biobanking" published a model broad consent form for biobanks based on recommendations of the National/German Ethics Council (Arbeitskreis Medizinischer Ethikkommissionen in der Bundesrepublik Deutschland e.V. (2013)).

[63]Bretthauer (2016), p. 272; Wolff (2016) margin number 20.

[64]Bretthauer (2016), p. 272.

envisaged Big Data applications in more detail.[65] Open ended Big Data applications where the analysis gives answers to questions that have not been asked before face certain limits here.[66] In order to be in compliance with data protection rules in such cases anonymization of the data may be an option[67] although it also needs to be said that anonymization becomes increasingly difficult in Big Data scenarios due to risks of reidentifiability.

The purpose limitation principle may set another barrier to conduct: Big Data applications as a further processing of the personal data must not be incompatible with the original purpose for which the data was collected. This depends on the individual circumstances. In case the privileging rule for further processing for historical, statistical or scientific purposes does not apply, a compatibility assessment must be conducted. It seems advisable to consider the criteria identified by the Article 29 Working Party as well as the practical examples elaborated on in Annex 3 of the Opinion. The Article 29 Working Party addresses the issue of repurposing data for Big Data analytics.[68] The opinion describes two kinds of further processing relevant for Big Data applications: (1) performing the analysis to detect trends or correlations; or (2) gaining information about certain individuals and making decisions affecting such individuals.[69] The first scenario may pose no further obstacles if appropriate safeguards for the individual's privacy are provided; in the second scenario "free, specific, informed and unambiguous 'opt-in' consent would almost always be required, though, otherwise the further use cannot be considered compatible."[70] For its part, the UK Information Commissioner's Office (ICO) has suggested broadly that a key factor in deciding whether a new purpose is incompatible with the original purpose is whether the further processing can be regarded as fair.[71]

When the purpose for which the data has been collected or further used is fulfilled the data must not be kept for longer in a form, which permits identification of data subjects (Article 6 (1) (e) DPD). Consequently, in such cases, further storage of personal data in order to use them for further analysis that will only be defined in the future is not permitted. However, if appropriate safeguards are provided by the Member States, the personal data can be stored for longer periods for historical, statistical or scientific use (Article 6 (1) (e) DPD. National implementations of this provision facilitate storage of personal data sets that could be used in the future for Big Data analysis which is a supplement to the privileging rule in Article 6 (1) (b) DPD.

[65]Handelsblatt Research Institute (2014), p. 14.

[66]Handelsblatt Research Institute (2014), p. 14; Martini (2014), p. 7; Roßnagel et al. (2016), p. 123.

[67]Raabe and Wagner (2016), p. 437; Handelsblatt Research Institute (2014), p. 14; Martini (2014), p. 15; Dix (2016), p. 60.

[68]Article 29 WP, pp. 46–47.

[69]Article 29 WP, pp. 46–47.

[70]Article 29 WP, p. 46.

[71]Information Commissioner's Office (2014).

# 5 New Developments Regarding the Purpose Limitation Principle Under the General Data Protection Regulation and Its Impact on Big Data Applications

## 5.1 *The General Data Protection Regulation—"A Hybrid of Old and New"*[72]

Continuing differences in the level of data protection between the Member States, which has been evaluated as a danger to the free flow of personal data across the EU, as well as challenges posed by changes in the technological environment and globalization (including the continuously growing scale of collecting and sharing personal data) created impulses for reform, which have led to the enactment of the General Data Protection Regulation.[73] The GDPR shall apply from 25 May 2018 onwards.[74] In contrast to the DPD, the GDPR will be directly applicable in all the Member States. However, it also reserves significant legislative powers to the Member States.[75] The Regulation appears as an "unusual hybrid of old and new."[76] It includes, for example, new rules such as the right to data portability,[77] the "right to be forgotten",[78] and mandatory data breach notifications,[79] and it puts a strong emphasis on privacy by design.[80] But it also reaffirms older principles, such as the requirement of a legal ground to allow processing of personal data, although these also sometimes appear in a new guise.

## 5.2 *Continuation of the Requirement of Purpose Specification and Compatible Use*

The European legislator has placed the purpose limitation principle in Article 5 (1) (b) GDPR. Article 5 of the Regulation also restates other key principles of data protection, such as the principle of data minimization, in Article 5 (1) (c) GDPR, and the principle of fairness and lawfulness, which now includes the explicit requirement of transparency, in Article 5 (1) (a) GDPR. In Article 5 (1) (b) it is

[72]Mayer-Schönberger and Padova (2016), p. 324.
[73]Recitals 5–9 of Regulation (EU) 2016/679; Mayer-Schönberger and Padova (2016), pp. 323–324.
[74]Article 99 (2) Regulation (EU) 2016/679.
[75]Mayer-Schönberger and Padova (2016), p. 325.
[76]Mayer-Schönberger and Padova (2016), p. 324.
[77]Article 20 Regulation (EU) 2016/679.
[78]Article 17 Regulation (EU) 2016/679.
[79]Article 33 Regulation (EU) 2016/679.
[80]Article 25 Regulation (EU) 2016/679.

further stated that personal data shall be "collected for specified, explicit and legitimate 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 shall in accordance with Article 89 (1) GDPR not be considered incompatible with the initial purpose. This reflects the fact that the two earlier identified main building blocks of the purpose limitation principle established by the Data Protection Directive, namely purpose specification and compatible use, still prevail. There is also a provision in Article 5 (1) (b) GDPR privileging further use for statistical purposes or scientific research.

## 5.3 *New Aspects with Regard to Purpose Specification*

Regarding the first building block element—the purpose specification requirement—it is interesting to note that the GDPR recognizes in Recital 33 that "it is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection." In consequence, data subjects shall be able to give consent to certain areas of scientific research if ethical standards are observed.[81] This may resolve to a certain degree the long and intensive debate on the permissibility of a broad consent in the medical field,[82] as it indicates that consent sheets can be formulated in such a way that the consent covers a broader range of research, not only specific research questions. This, in return, may also have an effect upon the question how specifically the purpose must be determined in advance by the researcher collecting and analyzing the personal data.

## 5.4 *Inclusion of the Compatibility Assessment Test into the Legal Text of the GDPR*

Regarding the second building block element—compatible use—the Regulation has become more specific with regard to the question of which secondary uses are to be considered compatible. During the legislative process the purpose limitation principle was heavily debated.[83] While the European Commission had included in its draft a passage, which provided a broad exemption from the requirement of compatibility by allowing further processing by simply identifying a new legal ground

[81]Schaar (2016), pp. 224–225.

[82]See elaborations made in footnote 15.

[83]Werkmeister and Brandt (2016), p. 237.

for the processing,[84] the European Parliament rejected this on grounds of principle.[85] Finally, the European legislative organs found a consensus by retaining the rule that further processing must be compatible (n.b., not identical) with the original purpose. They also agreed to adopt in the legislative text a catalogue of criteria, which are similar to the compatibility test criteria identified by the Article 29 Working Party in its opinion on purpose limitation. Article 6 (4) GDPR provides that when performing the compatibility assessment, the following criteria shall, inter alia,[86] be taken into account:

(i) Any link between the purposes for which the personal data have been collected and the purposes of the intended further processing;
(ii) The context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;
(iii) The nature of the personal data, in particular whether special categories of personal data are processed, pursuant to Article 9, or whether personal data related to criminal convictions and offences are processed, pursuant to Article 10;
(iv) The possible consequences of the intended further processing for data subjects;
(v) The existence of appropriate safeguards, which may include encryption or pseudonymization.

Article 6 (4) of the Regulation also states that the data controller can, in a case where they want to further process the data for a new purpose, obtain the data subject's consent.

Perhaps one of the real achievements of the Regulation is that the compatibility assessment test is now part of the legal text of the Regulation, which may improve attentiveness and enforceability. Legal scholars who have investigated how Big Data applications can comply with the requirements of data protection law often put special emphasis on the last criteria mentioned—the existence of appropriate safeguards—as a key provision to legitimize Big Data applications.[87] This corresponds with the view taken by the Article 29 Working Party, which states that effective anonymization of the data can reduce concerns regarding incompatible processing even in case of relatively sensitive data and where data subjects would not expect their data to be further processed.[88] However, the existence of appropriate safeguards is just one criterion to consider. The more specific and restrictive

[84]European Commission (2012) Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) (COM/2012/011 final—2012/0011 (COD)).

[85]Albrecht (2016), p. 36.

[86]This shows that the criteria catalogue is not excluding other appropriate considerations.

[87]Raabe and Wagner (2016), p. 438; Marnau (2016), p. 432.

[88]Article 29 WP, pp. 66–67.

the context of collection, the more limitations may apply to further use.[89] It also needs to be considered that Big Data processing makes it more and more challenging to achieve and preserve anonymity.[90]

## 5.5 *The New Privileging Rule for Further Processing for Archiving Purposes in the Public Interest, Scientific or Historical Research Purposes or Statistical Purposes*

As mentioned above, Article 5 (1) (b) GDPR, similar to Article 6 (1) (b) DPD, provides that further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes—in accordance with Article 89 (1) GDPR shall not be considered to be incompatible with the initial purposes. Recital 162 GDPR defines statistical purposes as "any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results." The meaning of statistical purposes can be interpreted broadly and does not only cover uses for public interest, but may also include private entities doing research in pursuit for commercial gain.[91] Recital 162 GDPR also states, "The statistical purpose implies that the result of processing is not personal data, but aggregated data and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person." It resembles Recital 29 DPD and it can be clearly followed that statistical analysis used for decision-making that directly affect a particular individual is not covered by the privileging rule.[92]

The use of personal data for scientific research—another privileged purpose—may also be interpreted broadly. Recital 159 GDPR mentions, for example, technological development and demonstration, fundamental research, applied research and also privately funded research. Studies conducted in the public interest in the area of public health are also explicitly referred to in the same recital as scientific research purposes.

The Regulation now refers in Article 5 (1) (b) to Article 89 (1). This article provides that processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall be subject to appropriate safeguards to protect the rights and freedoms of the data subject. Determining those

[89]Article 29 WP, p. 25; The Article 29 Working Party had investigated a considerable number of examples for further processing which is compatible and non-compatible. See Article 29 WP, pp. 51–69.

[90]Dix (2016), pp. 60–61; Sarunski (2016), p. 427; Boehme-Neßler (2016), p. 422; Bretthauer (2016), p. 271.

[91]Mayer-Schönberger and Padova (2016), p. 326.

[92]Mayer-Schönberger and Padova (2016), p. 327.

safeguards will be left to the Member States.[93] Article 89 (1) GDPR imposes some statutory requirements relating to the quality and conditions of the safeguards, which were not explicitly mentioned in the legal text of the Directive. For example, it is provided that the safeguards shall ensure the presence of technical and organizational measures in order to respect the principle of data minimization, for example, pseudonymization or anonymization subject to the circumstance that the purpose pursued with the processing can be fulfilled. In principle, it can be said, that Article 89 (1) GDPR reflects the approach previously suggested, and comprehensively set out, by the Article 29 Working Party in its Opinion on purpose limitation. There it is, inter alia, stated that different scenarios require different safeguards. There are scenarios where anonymized or aggregated data can be used; others require the processing of indirectly identifiable data or directly identifiable data.[94]

## 5.6 *The Waiver of the Requirement of a Legal Basis for the Processing of Personal Data that Qualifies as a Compatible Use*

The Regulation brought another change in respect of the purpose limitation principle that should not be overlooked. Recital 50 GDPR states that if the processing is compatible with the purposes for which personal data were initially collected, no legal basis separate from that which allowed the collection of the personal data is required. This is, at first glance, a surprising shift from the general rule that processing of personal data is prohibited unless covered by existing permissions. The questions arise, however, whether this change in the law is associated with considerable disadvantages for the data subject, and whether it significantly facilitates processing activities on the side of the data controller. Taking a closer look at the new provision on the compatibility assessment in Article 6 (4) of the Regulation, it appears that the interests of the data controller to further process the data and interests of the data subject shall be balanced. The same concept usually applies to a legal ground allowing the processing of personal data where the legitimate interests of the data subjects and data controller concerned are weighed against each other. This is, for example, especially reflected in Article 6 (1) (e) GDPR, which provides that processing of personal data shall be lawful if processing is necessary for the purposes of the legitimate interest of the controller or of 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. Eventually, all legal grounds to be found in Article 6 (1)—and this applies also to the legal grounds established in Article 9 (2) GDPR—are the product of such a weighing of interests of the concerned parties—data subject and data controller—by the European legislator.

[93]Mayer-Schönberger and Padova (2016), p. 327.

[94]Article 29 WP, pp. 27–33.

Article 9 (2) of the Regulation, which concerns special categories of personal data, such as health data or genetic data,[95] provides specific legal grounds for processing such personal data. Due to the sensitive nature of the data and the increased demand for protection on side of the data subject, these legal grounds are generally stricter than those in Article 6 (1) GDPR: for example, consent must be explicit,[96] and a provision comparable to Article 6 (1) (e) GDPR allowing the processing of personal data if it 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 is not existent. One could argue that this specific protection mechanism established by Article 9 GDPR for special categories of personal data may be at risk to become undermined by the omission of the requirement of a legal ground for further processing of personal data that qualifies as a compatible use. However, Article 6 (4) GDPR should be flexible enough to take into account the degree of sensitivity of the data concerned, and to weigh them accordingly. Here the catalogue explicitly mentions that the nature of the data needs to be considered, as well as the possible consequences for the data subject. In order to protect the interests of the data subjects, data controllers that envisage further processing need to make the compatibility assessment in a careful and conscientious way and most probably will have to establish appropriate safeguards to protect the personal data at issue. Nevertheless, one cannot deny that there is a danger of misuse by the data controller through overemphasizing their own interests. In case of further processing for scientific purposes or statistical purposes, it is also interesting to consider that Article 9 (2) (h) GDPR largely resembles the requirements in Article 6 (1) (b) in conjunction with Article 89 (1) GDPR. This, in turn, points toward the conclusion that the data subject is neither placed in a less favorable position in the case of further processing of their personal data for statistical purposes, when the requirements of the privileging rule are fulfilled.

[95]Article 9 (1) Regulation (EU) 2016/679 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.

[96]Article 9 (2) (a) Regulation (EU) 2016/679. For personal data that do not qualify as special categories of personal data in the sense of Article 9 (1) Regulation (EU) 2016/679, Article 6 (1) (a) Regulation (EU) 2016/679 states that processing of such data shall be lawful if the data subject has given consent to the processing of his or her personal data for one or more specific purposes. The standard for explicit consent remains the same as under the Data Protection Directive with the result that, for example, implied consent interpreted out of the data subject's conduct is not enough for an explicit consent in the sense of Article 9 (2) (a) Regulation (EU) 2016/679, but may be a sufficient legal basis for the processing of non-sensitive personal data in the sense of Article 6 Regulation (EU) 2016/679 (see Maldoff 2016).

## 6 Consequences of the Enactment of the GDPR for Big Data Applications and Conclusion

The Regulation retains the purpose limitation principle as one of its basic elements. Consequently, data controllers will, in the future, have to specify the purpose of the collection, which must be clearly and specifically identified. At the same time, in some fields, such as that of scientific and medical research, the European legislator has reacted on the issue of the necessity to reuse personal data and the difficulties to specify all research questions at the time of the collection of the data. Data subjects will be able to give their consent to certain areas of scientific research if ethical standards are complied with.[97]

Further processing of personal data under the Regulation will also need to be compatible with the original purpose for which the data was collected. The requirements regarding the permissibility of a change of purpose have not been loosened. Change of the purpose either needs to be covered by the privileging rule in Article 5 (1) (b) GDPR or pass the compatibility test, which is now explicitly incorporated into the legal text of the Regulation, namely Article 6 (4) GDPR. Further processing of personal data for scientific or statistical purposes shall be deemed in compliance with the purpose limitation principle, subject to appropriate safeguards. The latter, which should exclude or considerably reduce the risk for data subjects, remain (as was the case under the DPD) a matter to be implemented by Member States. A further limitation in such cases is that the processing at issue —including in Big Data scenarios—should not aim to gain information about particular individuals and/or make decisions affecting them. If it is otherwise, the principle of purpose limitation will again apply in full ambit, and the data controller will need to ask for the data subject's consent.

The legal situation under the new GDPR remains somewhat similar to the DPD with regard to the principle that personal data should not be kept in a form which permits identification of data subjects any longer than the purpose of the collection or reuse requires (Article 5 (1) (e) DPD). Again, though, personal data may be stored for longer periods insofar as they will be used solely for privileged purposes, such as statistical purposes or scientific research purposes (assuming appropriate technical and organizational measures to protect the data subject are in place).

Ultimately, then, it appears that the waiver of the requirement of a legal basis for further processing under the GDPR should not have a significant impact on the data subject's interests. These remain protected by the need in such circumstances for the data controller to satisfy the provisions of Article 6 (4), as well as those of Article 5 (1) (b) GDPR in conjunction with Article 89 (1) GDPR.

In short, the legal situation for data controllers wishing to process personal data in Big Data applications has—with regard to the purpose limitation principle—not significantly changed. It will remain a core issue how to specify the purpose of the

[97]Recital 33 Regulation (EU) 2016/679.

collection and further use of the personal data prior to, or at least no later than, the time of collection. As well as under the Directive the purpose specification requirement sets limits to open ended Big Data applications, where the purpose will only be specified after the analysis has commenced.

Big Data applications that involve further processing of personal data for scientific and statistical purposes do not face considerable obstacles if appropriate safeguards for the data subject are maintained. The establishment of such safeguards is, of course, consuming resources on the side of the data controllers. Data controllers wanting to further use personal data for Big Data analysis in order to gain information about particular individuals and/or make decisions affecting them do indeed face larger obstacles to further process the personal data in compliance with the purpose limitation principle. For example, if an organization is aiming to further process the personal data of their customers in order to analyze or predict the personal preferences and behavior of individual customers in order to use such information to base decisions regarding them, then the data controller will be required to obtain the informed consent of those customers.[98]

To conclude, it should also be pointed out that privacy, which the data protection regulations including the purpose limitation principle seek to realize, may not only be seen as a hindering factor for economy and science. European Network and Information Security Agency (ENISA), for example, recently noted that "if privacy principles are not respected, Big Data will fail to meet individual's needs; if privacy enforcement ignores the potentials of Big Data, individuals will not be adequately protected."[99] Involved stakeholders should work together in addressing the challenges and highlight privacy as a core value and a necessity of Big Data. Technology should be used as a support tool to achieve this aim.[100]

**Acknowledgements** This work has been supported by the EU project SoBigData (http://www.sobigdata.eu/) which receives funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 654024 and the German national project ABIDA (http://www.abida.de/) which has been funded by the Bundesminsterium für Bildung und Forschung (BMBF). The authors would like to thank Marc Stauch for his valuable support.

## References

Albrecht JP (2016) The EU's new data protection law—how a directive evolved into a regulation. Comput Law Rev Int 17(2):33–43

Arbeitskreis Medizinischer Ethik-Kommissionen (2013) Mustertext zur Spende, Einlagerung und Nutzung von Biomaterialien sowie zur Erhebung, Verarbeitung und Nutzung von Daten in Biobanken. http://www.med.uni-freiburg.de/Forschung/VerantwortungForschung/mustertext-

[98]Article 29 WP, p. 46.

[99]European Union Agency for Network and Information Security (ENISA) (2005), pp. 17–18.

[100]European Union Agency for Network and Information Security (ENISA) (2005), pp. 17–18.

biobanken-deutsch.doc. Accessed 17 Nov 2016, English Version: http://www.ak-med-ethik-komm.de/index.php?lang=de. Accessed 17 Nov 2016
Article 29 Data Protection Working Party (2013) Opinion 03/2013 on purpose limitation, 00569/13/EN, WP 203
Beyleveld D (2004) An overview of directive 95/46/EC in relation to medical research. In: Beyleveld D et al (eds) The data protection directive and medical research across Europe. Ashgate Publishing Company, Burlington
Boehme-Neßler V (2016) Das Ende der Anonymität – Wie Big Data das Datenschutzrecht verändert. Datenschutz Datensich 40(7):419–423
Bretthauer S (2016) Compliance-by-design-anforderungen bei smart data. Z Datenschutz 6 (2):267–274
Bundeskartellamt, Autorité de la concurrence (2016) Competition law and data. http://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.pdf;jsessionid=9F9A418331598CA75471DEA51872F638.1_cid371?__blob=publicationFile&v=2. Accessed 16 Sept 2016
Cavanillas JM, Curry E, Wahlster W (2015) The big data value opportunity. In: Cavanillas JM, Curry E, Wahlster W (eds) New horizons for a data-driven economy. Springer, Cham
Curry E (2015) The big data value chain: definitions, concepts, and theoretical approaches. In: Cavanillas JM, Curry E, Wahlster W (eds) New horizons for a data-driven economy. Springer, Cham
Dix A (2016) Datenschutz im Zeitalter von Big Data. Wie steht es um den Schutz der Privatsphäre. Stadtforsch Stat 29(1):59–64
European Union Agency for Network and Information Security (ENISA) (2005) Privacy by design in Big Data—an overview of privacy enhancing technologies in the era of Big Data analytics. https://webcache.googleusercontent.com/search?q=cache:bsgvi1hfgTYJ:https://www.enisa.europa.eu/publications/big-data-protection/at_download/fullReport+&cd=2&hl=de&ct=clnk&gl=de&client=firefox-b-ab. Accessed 4 Oct 2016
Ehmann E, Helfrich M (1999) Kurzkommentar zur EG-Datenschutzrichtlinie. Verlag Otto Schmidt, Cologne
Grützmacher M (2016) Dateneigentum – ein Flickenteppich. Comput Recht 32(8):485–495
Handelsblatt Research Institute (2014) Datenschutz und Big Data: Ein Leitfaden für Unternehmen. http://www.umweltdialog.de/de-wAssets/docs/2014-Dokumente-zu-Artikeln/leitfaden_unternehmen.pdf. Accessed 17 Nov 2016
Information Commissioner's Office (2014) Big Data and data protection. https://ico.org.uk/media/1541/big-data-and-data-protection.pdf. Accessed 28 Sept 2016
Kanellos M (2016) 152,000 Smart devices every minute in 2025: IDC outlines the future of smart things. http://www.forbes.com/sites/michaelkanellos/2016/03/03/152000-smart-devices-every-minute-in-2025-idc-outlines-the-future-of-smart-things/#4ec22cb069a7. Accessed 15 Aug 2016
Körber T (2016) "Ist Wissen Marktmacht?" Überlegungen zum Verhältnis von Datenschutz, "Datenmacht" und Kartellrecht – Teil 1. Neue Z Kartellr 4(7):303–310
Laney D (2001) 3D data management: controlling data volume, velocity, and variety. Technical report, META Group, https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf. Accessed 16 Aug 2016
Maldoff G (2016) Top 10 operational impacts of the GDPR: Part 3—consent https://iapp.org/news/a/top-10-operational-impacts-of-the-gdpr-part-3-consent/. Accessed 4 Oct 2016
Marnau N (2016) Anonymisierung, Pseudonymisierung und Transparenz für Big Data. Datenschutz Datensich 40(7):428–433
Martini M (2014) Big Data als Herausforderung für den Persönlichkeitsschutz und das Datenschutzrecht. http://www.uni-speyer.de/files/de/Lehrst%C3%BChle/Martini/PDF%20Dokumente/Typoskripte/BigData-TyposkriptiSd%C2%A738IVUrhG.pdf. Accessed 17 Nov 2016
Mayer-Schönberger V, Padova Y (2016) Regime change? Enabling big data through Europe's new data protection regulation. Columbia Sci Technol Law Rev 17:315–335
Medical Research Council (2011) MRC Policy and Guidance on Sharing of Research Data from Population and Patient Studies. http://www.mrc.ac.uk/publications/browse/mrc-policy-and-

guidance-on-sharing-of-research-data-from-population-and-patient-studies/. Accessed 29 Sept 2016
Metschke R, Wellbrock R (2002) Berliner Beauftragter für Datenschutz und Informationsfreiheit, Hessischer Datenschutzbeauftragter, Datenschutz in Wissenschaft und Forschung. https://datenschutz-berlin.de/attachments/47/Materialien28.pdf?1166527077. Accessed 28 Sept 2016
Raabe O, Wagner M (2016) Verantwortlicher Einsatz von Big Data. Datenschutz Datensich 40 (7):434–439
Roßnagel A et al (2016) Datenschutzrecht 2016 "Smart" genug für die Zukunft. Kassel University Press GmbH, Kassel
Sarunski M (2016) Big Data—Ende der Anonymität? Fragen aus Sicht der Datenschutzaufsichtsbehörde Mecklenburg-Vorpommern. Datenschutz Datensich 40(7):424–427
Schaar K (2016) DS-GVO: Geänderte Vorgaben für die Wissenschaft—Was sind die neuen Rahmenbedingungen und welche Fragen bleiben offen? Z Datenschutz 6(5):224–226
Turner V et al (2014) The digital universe of opportunities: rich data and the increasing value of the Internet of Things. Rep. from IDC EMC. https://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf. Accessed 15 Aug 2016
Werkmeister C, Brandt E (2016) Datenschutzrechtliche Herausforderungen für Big Data. Comput Recht 32(4):233–238
Wolff H (2016) In: Wolff HA, Brink S (eds) Beck'scher Online Kommentar Datenschutzrecht, Prinzipien des Datenschutzrechts. https://beck-online.beck.de/Home. Accessed 17 Nov 2016
Zech H (2012) Information als Schutzgegenstand. Mohr Siebeck Verlag, Tübingen

http://www.springer.com/978-981-10-5037-4

New Technology, Big Data and the Law
Corrales, M.; Fenwick, M.; Forgó, N. (Eds.)
2017, XVI, 330 p. 17 illus., 7 illus. in color., Hardcover
ISBN: 978-981-10-5037-4