17

Artificial Intelligence as a Challenge for Data Protection Law

And Vice Versa

*Boris P. Paal**

I. INTRODUCTION

Artificial Intelligence (AI) as an area of research within the field of computer science concerns itself with the functioning of autonomous systems and, as such, not only affects almost all areas of modern life in the age of digitisation but has also – and for good reasons – become a focal point within both academic and political discourse.[1] AI scenarios are mainly driven and determined by the availability and evaluation of data. In other words, AI goes hand in hand with what may be referred to as an enormous 'appetite for data'. Thus, the accumulation of relevant (personal or non-personal) data regularly constitutes a key factor for AI-related issues. The collected personal data may then be used to create (personality) profiles as well as to make predictions and recommendations with regard to individualised services and offers. In addition, non-personal data may be used for the analysis and maintenance of products. The applications and business models based on the collection of data are employed in both the private and public sector. The current and potential fields of application for AI are as diverse and numerous as the reactions thereto, ranging from optimism to serious concerns – oftentimes referring to a potential 'reign of the machines'. However, there is a general consensus regarding the fact that the development and use of AI technologies will have significant impact on the state, society, and economy. For instance, the use of such applications may greatly influence the protection of personal rights and privacy, because the development of AI technologies regularly requires the collection of personal data and the processing thereof. This chapter will focus on and examine provisions concerning the handling of personal data as set out in the European Union's General Data Protection Regulation (GDPR)[2] which entered into force on 24 May 2016 and has been applicable since 25 May 2018.

The prerequisites and applications of AI on one hand and the regulatory requirements stipulated by the GDPR on the other, give rise to a number of complicated, multi-sided tensions and conflicts. While the development of AI is highly dependent on the access to large amounts

* Transcript of a presentation held at the Conference Global Perspectives on Responsible AI 2020 in Freiburg on June 26, 2020. The presentation form was maintained for the most parts. Fundamental considerations of this paper are also published in B Paal, 'Spannungsverhältnis von KI und Datenschutzrecht' in M Kaulartz and T Braegelmann (eds), *Rechtshandbuch Artificial Intelligence und Machine Learning* (2020) 427–444.

[1] On defining AI see for example J Kaplan, *Artificial Intelligence* (2016) 1 *et seq*.

[2] 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), OJ 2016 L 119/1.

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

of data (i.e. big data), this access is subject to substantial limitations imposed by the data protection law regime. These restrictions mainly apply to scenarios concerning personal (instead of non-personal) data and primarily stem from the GDPR's preventive prohibition subject to authorisation[3] and its general principles relating to the processing of personal data.[4] One of the most fundamental problems which arises in connection with big data is referred to as 'small privacy'. This term alludes to the inherent conflict between two objectives pursued by data protection law, the comprehensive protection of privacy and personal rights on the one hand and the facilitation of an effective and competitive data economy on the other. The tension arising from this conflict is further illustrated by Article 1 GDPR, according to which the Regulation contains provisions to promote both the protection of natural persons with regard to the processing of personal data and the free movement of such data. An instrument intended to facilitate an appropriate balance between the protection of personal data and seemingly contradictory economic interests may be seen in the users' data sovereignty.[5]

At this point, it should be noted that the GDPR does not (or, if at all, only marginally) address the implications of AI for data protection law. Thus, in order to be applied to individual cases and to specific issues arising in connection with AI, the general provisions of the GDPR need to be construed. This may oftentimes lead to substantial legal uncertainties, especially when considering the vague wording, unclear exemptions, and considerable administrative discretion provided by the GDPR. The aforementioned uncertainties may not only impede innovation but may also give rise to a number of issues concerning the (legal) accountability for AI, for instance, in connection with the so-called black-box-phenomenon[6] regularly encountered when dealing with self-learning AI systems (i.e. deep or machine learning).

II. AI AND PRINCIPLES RELATING TO THE PROCESSING OF DATA

The development and use of AI may potentially conflict with almost all principles concerning the processing of data as enshrined in the GDPR. In fact, the paradigms of data processing in an AI-context are very difficult, if not impossible, to reconcile with the traditional principles of data protection. The complex and multi-layered legal issues resulting from this contradiction are first and foremost attributable to the fact that AI scenarios were not (sufficiently) taken into account during the drafting of the GDPR. This raises the question of whether and to what extent AI scenarios can be adequately addressed and dealt with under the existing legal regime by utilising the available technical framework and by interpreting the relevant provisions accordingly. Where the utilisation of such measures and, consequently, the application of the law and the compliance[7] with the principles of data protection is not possible, it has to be assessed whether there are any other options to adapt or to amend the existing legal framework.[8]

The aforementioned data protection issues have their roots in the general principles of data protection. Hence, in order to fully comprehend the (binding) provisions that a 'controller' in the sense of the GDPR must observe when processing data, it is necessary to take a closer look at these principles. This is especially important considering the very prominent role of the legal

[3] Cf. GDPR, Article 6(1).
[4] Cf. GDPR, Article 5.
[5] On data sovereignty see for example PL Krüger, 'Datensouveränität und Digitalisierung' (2016) *ZRP* 190.
[6] On the 'black box-phenomenon' see for example F Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information* (2015).
[7] For this see the following Section III.
[8] For this see the following Section IV.

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

framework in nearly all AI scenarios. The addressee of the principles relating to the processing of personal data laid down in Article 5(1) GDPR, is responsible for the adherence thereto and must, as required by the principle of accountability, be able to provide evidence for its compliance therewith.[9] The obligations set out in Article 5(1) GDPR range from the lawfulness, fairness, and transparency of data processing as well as the adherence to and compatibility with privileged purposes (purpose limitation) to the principle of data minimisation, accuracy, storage limitation, as well as integrity and confidentiality.[10] Beyond the scope of the present analysis in this chapter lie questions concerning conflicts of law and the lawfulness of data transfer in non-EU Member States, although these constellations are likely to become increasingly important in legal practice especially in light of the growing importance of so-called cloud-solutions[11].

1. *Transparency*

In accordance with Article 5(1)(a) alt. 3 GDPR, personal data must be 'processed [...] in a transparent manner in relation to the data subject'. These transparency requirements are of particular importance for matters relating to AI. As set out in Recital 39 of the GDPR, the principle of 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.[12]

These transparency requirements are specified by the provisions contained in Articles 12-15 GDPR, which stipulate the controllers' obligation to provide information and to grant access to personal data. They are further accompanied by the obligation to implement appropriate technical and organisational measures.[13] Moreover, Article 12(1) sentence 1 GDPR requires the controller to provide the data subject with any information and communication 'in a concise, transparent, intelligible and easily accessible form, using clear and plain language'. Especially with regard to issues relating to AI, the implementation of these requirements is very likely to present responsible controllers with a very complex and onerous task.

In an AI scenario, it will often be difficult to state and substantiate the specific purposes for any given data analysis in advance. Controllers may also face enormous difficulty when tasked with presenting the effects that such an analysis could have on the individual data subject in a sufficiently transparent manner. In fact, the very nature of self-learning AI which operates with unknown (or even inexplicable) variables seems to oppose any attempt to present and provide any transparent information.[14] In addition, the aforementioned 'black-box-phenomena' may

[9] Cf. GDPR, Article 5(2).

[10] GDPR, Article 5(1)(a)–(f).

[11] On GDPR and the cloud see J Krystelik, 'With GDPR, Preparation Is Everything' (2017) *Computer Fraud & Security* 5 (7).

[12] See also Article 29 Data Protection Working Party, 'Guidelines on Transparency under Regulation 2016/679' WP 260 rev.01.

[13] Cf. GDPR, Recital 78. For this see the following Section III 3.

[14] L Mitrou, 'Data Protection, Artificial Intelligence and Cognitive Services: Is the GDPR "Artificial Intelligence-Proof"?' (2018) Tech Report commissioned by Microsoft, 58 https://ssrn.com/abstract=3386914 (hereafter Mitrou, 'Data Protection').

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

occur if, for instance, artificial neural networks on so-called hidden layers[15] restrict or even prohibit the traceability of the respective software-processes. Thus, on the one hand, it may be difficult to break down the complex and complicated AI analyses and data collection processes into 'concise, transparent, intelligible and easily accessible' terms that the affected data subject can understand. On the other hand, the lack of transparency is an inherent feature and characteristic of self-learning, autonomous AI technologies.[16] Furthermore, these restrictions on transparency also come into play when considering potential justifications for the processing of data. This is particularly relevant where the justification is based on the data subject's consent as this (also) requires an informed indication of the subject's agreement.[17]

However, according to the principles of the GDPR, even controllers who use systems of AI and, thus, carry out extensive analyses of huge amounts of data of different origins, should have the (realistic) possibility to process data in a manner which allows them to adequately inform the subject about the nature and origin of the processed data. Further difficulties are likely to arise in situations where personal data are generated in the course of analyses or as a result of combinations of originally non-personal data. Because, in this case, the legally relevant collection of data is to be found in the analysis, it is difficult if not impossible to pinpoint the data's initial origin and source. In such constellations, it should, thus, be assumed that the responsible controller is permitted to merely provide general information, for instance by naming the source of the data stock or the systems utilised to process the data in addition to the means used for their collection. In this context, it also has to be emphasised that the obligation to inform the data subject as set out in Article 14(5)(b) GDPR may be waived if the provision of such information would be disproportionally onerous. The applicability of this waiver must be determined by balancing the controller's efforts required for the provision of information with the data subject's right and interest to be informed. The outcome of this (case-by-case) balancing process in big-data-situations – not only in the context of AI – will largely depend on the effects that the data analysis and processing have on the subject's fundamental rights, as well as on the nature and degree of risks that arise in connection thereto. For the purposes of such an assessment, the principle of transparency should extend beyond the actual data processing procedures to include the underlying technical systematics and the decision-making systems employed by the (responsible) controller.

2. *Automated Decisions/Right to Explanation*

Article 22 GDPR is intended to protect the individual from being made subject to decisions based solely on an automated assessment and evaluation of the subject's personal profile, because this would risk degrading the individual to a mere object of computer-assisted programs. Against this background, the GDPR imposes additional obligations to provide information in situations where the responsible controller utilises automated decision-making procedures in Articles 13(2)(f), 14(2)(g), and 15(1)(h) GDPR. Pursuant to these provisions, the controller has to provide 'meaningful information about the logic involved' in the data processing.

[15] On artificial neuronal networks see for example Y LeCun, Y Bengio and G Hinton, 'Deep Learning' (2017) *Nature Deep Review* 436 (437); T Sejnowski, *The Deep Learning Revolution* (2018) 37 *et seq*.

[16] A Deeks, 'The Judicial Demand for Explainable Artificial Intelligence' (2019) *Colombia Law Review* 1892 (1833 *et seq*.).

[17] Cf. Recital 32. For this see the following Section II 6(a).

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

This obligation may be called into question[18] when considering the aforementioned difficulties that controllers may face when tasked with providing information about complex and potentially inexplicable (autonomous) AI processes and the results based thereon. In these scenarios, the controller should merely have to provide (and the subject should merely be entitled to) general information on the functioning of the specific AI technology, whereas a right to a substantiated explanation should be rejected. In accordance with Article 35(3)(a) GDPR, an evaluation of personal data which is based on automated processing requires a data protection impact assessment. It should also be emphasised that the use of AI as such is not restricted as of today. Instead, the restrictions apply solely to decision-making processes based on the use of AI.

3. *Purpose Limitation/Change of Purpose*

Pursuant to the principle of purpose limitation as set out in Article 5(1)(b) GDPR, the purposes for processing and collection of (personal) data must be specified and made available to the data subject in a concise and intelligible way.[19] This principle also applies to any further processing of data. The requirement of a pre-defined purpose limitation generally opposes the basic concept of AI, according to which AI should develop independently (or possibly within a certain pre-defined framework) and should be used for purposes not defined in advance.[20] Against this backdrop, the prescription of purpose limitations threatens to impede the (unhindered) development and potentials of AI technologies.[21] Thus, the limitation of legitimate purposes of data processing may lead to a considerable restriction of technological AI potentials.[22] In situations in which AI can (and frequently even should) lead to unforeseen and possibly unforeseeable applications and results, it can, therefore, be very challenging to find an appropriate equilibrium between the principle of purpose limitation and the innovation of AI technologies. In many AI scenarios, it is virtually impossible to predict what the algorithm will learn. Furthermore, the purpose in the sense of Article 5(1)(b) GDPR may change in the course of the (autonomous) development of self-learning AI, especially as the relevant objectives of the data processing may not be known at the time of data collection. Moreover, it is reasonable to be concerned about a distortion of the results (freely) generated by AI tools as potentially induced by data protection law, if such technologies are only granted restricted (or no) access to certain data sources.

[18] In favour of such a right to explanation B Goodman and S Flaxman, 'European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"' (2017) 38(3) *AI Magazine* 50, 55 *et seq*.; in contrast S Wachter, B Mittelstadt, and L Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the GDPR' (2017) 7 (2) *IDPL* 76.; cf. also M Temme, 'Algorithms and Transparency in View of the New GDPR' (2017) 3(4) *EDPL* 473, 481 *et seq*.; *L Edwards* and *M Veale*, 'Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For' (2017) 16 *DLTR* 18; critical of the GDPR's significance in principle for AI methods also R van den Hoven van Genderen, 'Privacy and Data Protection in the Age of Pervasive Technologies in AI and Robotics' (2017) 3(3) *EDPL* 338, 346 *et seq*.; on the ethical dimension and the efforts to supplement Convention No 108 of the Council of Europe with corresponding transparency provisions, see Committee of Experts on Internet Intermediaries, *Algorithms and Human Rights: Study on the Human Rights Dimensions of Automated Data Processing Techniques* (DGI (2017)12) 13 *et seq*. in particular Algorithms and Possible Regulatory Implications.

[19] See the above comments on transparency Sub II 1.

[20] Mitrou, 'Data Protection' (n 14) 20; N Purtova, 'The Law of Everything. Broad Concept of Personal Data and Future of EU Data Protection Law' (2018) 10(1) *Law, Innovation and Technology* 40, 56 (hereafter Purtova, 'The Law of Everything').

[21] N Wallace, and D Castro, *The Impact of the EU's New Data Protection Regulation on AI*, 14 (Centre for Data Innovation Policy Brief, 2018) https://euagenda.eu/upload/publications/untitled-140069-ea.pdf (hereafter Wallace and Castro, 'Data Protection Regulation').

[22] Norwegian Data Protection Authority, *Artificial Intelligence and Privacy*, 18 (Datatilsynet Report, 2018) www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf.

There is, thus, a notable risk of conflicts between the interests and objectives of the individual on the one hand and public welfare on the other. In order to avoid such conflicts, it is crucial to explicitly list the application and use of AI as one of the purposes for the collection of data. Data controllers should, therefore, seek to identify, document, and specify the purposes of future data processing at an early stage. Where these measures are not taken, the requirements for a permissible change of purpose follow from Article 6(4) GDPR.

Article 6(4) GDPR, which addresses purpose changes, lists a number of criteria for the evaluation of the compatibility of such changes in situations where the data processing is carried out for purposes other than the ones for which the data has been originally collected. This creates a direct link to the principle of purpose limitation as laid down in Article 5(1)(b) GDPR. It should further be emphasised that the compatibility of a change of purpose with the original purpose does not affect the cumulative prerequisites for the lawfulness of the processing in question. Because Article 6(4) GDPR itself does not constitute a legal basis for the processing of data for other purposes, recourse must be taken to Article 6(1) subpara. (1) GDPR which requires the existence of a legal justification also for other purposes. In consequence, the controller is responsible to ensure that the data processing for the new purpose is compatible with the original purpose and based upon a legal justification in the sense of Article 6(1) subpara. (1) GDPR. In many cases, relevant personal data will not have been collected for the purposes of training or applying AI technology.[23] In addition, controllers may sometimes have the hope or expectation to subsequently use the collected data for other purposes, for instance in exploratory data analyses. If one were to pursue a more restrictive line of interpretation regarding the change of purposes by applying the standard of Article 6(4) GDPR, it would be impossible to use AI with a sufficient degree of legal certainty. Especially, situations, in which data is generated in different contexts and subsequently combined or used for (new) purposes, are particularly prone to conflict.[24] In fact, this scenario demonstrates the far-reaching implications of and issues arising in connection with the principle of purpose limitation and AI scenarios: if the purpose for the processing of data cannot (yet) be determined, the assessment of its necessity becomes largely meaningless. Where the purpose limitation remains vague and unspecified, substantial effects of this limitation remain unlikely.

4. Data Minimisation/Storage Limitation

Pursuant to the principle of data minimisation,[25] personal data must be adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed. The principle of data minimisation is specified by the requirement of storage limitation (as will be elaborated in the following) and the provisions concerning data protection through the implementation of technical measures and data protection 'by design and by default'.[26] Similarly to the principles described above, the principle of data minimisation oftentimes directly contradicts the general concept of AI technologies which is based on and requires the collection of large

[23] On the consequences of the prohibition on repurposing data see Wallace and Castro, 'Data Protection Regulation' (n 21) 14.

[24] M Butterworth, 'The ICO and Artificial Intelligence: The Role of Fairness in the GDPR Framework' (2018) 34(2) *Computer Law & Security Review: The International Journal of Technology Law and Practice* 257, 260 (hereafter Butterworth, 'GDPR Framework').

[25] Cf. GDPR, Article 5(1)(c).

[26] GDPR, Article 25.

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

amounts of data.[27] Given the very nature of AI applications, it is exceedingly difficult to make any kind of prediction regarding the type and amount of data necessary in constellations which have yet to be determined by the application itself. In addition, the notion of precautionary protection of fundamental rights by way of data avoidance openly conflicts with the high demand for data in any given AI scenario.[28]

The principle of storage limitation[29] prescribes that where personal data is stored, the identification of the data subject is only permissible for as long as this is necessary for the processing purposes. This principle also poses considerable difficulties in AI constellations, because the deletion or restriction of personal data after the fulfilment of their purpose can significantly impede both the development and use of AI technologies. According to Recital 39 sentences 8 and 10 of the GDPR, the period for which personal data is stored must be limited to a strict minimum. The controller should further establish time limits for the data's erasure or their periodic review. Correspondingly, Article 17 GDPR contains the data subject's right to demand the immediate erasure of any data concerning him or her under certain conditions.[30]

5. *Accuracy/Integrity and Confidentiality*

Another principle of data protection law which may be affected in AI scenarios is the principle of accuracy as set out in Article 5(1)(d) GDPR. This principle is intended to ensure that the collected (personal) data accurately depicts reality so that the affected data subjects will not suffer any disadvantages resulting from the use of inaccurate data. In situations in which the procedure and systems used for the processing of data present themselves as a 'black box' to both data subject and controller, it can be very difficult to detect inaccurate information and to restore their accuracy.[31] However, situations concerning the accuracy of data require a distinction between data input and output; as the latter is a result of data-processing analyses and processes – also and in particular in situations involving AI – it will regularly constitute a (mere) prognosis.

Pursuant to Article 5(1)(f) GDPR, personal data must be processed in a manner that ensures their appropriate security. The controller is thereby required to take adequate measures to ensure the data's protection against unauthorised or unlawful processing and against accidental loss, destruction, or damage.

6. *Lawfulness/Fairness*

The lawfulness of data processing[32] requires a legal basis authorising the processing of data as the normative concept of data protection law envisages a prohibition subject to authorisation. In order to be deemed lawful in the sense of Article 5(1)(a) GDPR, the processing must fulfil at least one of the prerequisites enumerated in Article 6(1) GDPR. In this context, Article 6(1) subpara. 1(b) GDPR permits the processing of data if it is necessary for the performance of a contract which the data subject is party to or for the implementation of pre-contractual measures.

[27] Butterworth, 'GDPR Framework' (n 24) 260.
[28] T Zarsky, 'Incompatible: The GDPR in the Age of Big Data' (2017) 47 *Seton Hall Law Review* 995, 1005 *et seq*.
[29] GDPR, Article 5(1)(e).
[30] On Article 17 and the implications for AI technologies see M Humerick, 'Taking AI Personally: How the EU Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence' (2018) 34 *Santa Clara High Tech L.J.* 393, 407 *et seq*.
[31] On AI and the accuracy principle see Butterworth, 'GDPR Framework' (n 24) 257, 260 *et seq*.; Mitrou, 'Data Protection' (n 14) 51 *et seq*.
[32] GDPR, Article 5(1)(a) alt. 1 and 2.

However, in scenarios involving AI, such pre-contractual constellations will not arise regularly. Similarly, AI scenarios are very unlikely to fall within the scope of any of the other authorisations listed in Article 6(1) subpara. (1) GDPR which include the existence of a legal obligation, the protection of vital interests, or the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.[33]

In contrast, the authorisations set out in Article 6(1) subpara. (1)(a) and (f) GDPR, which are based on the data subjects' consent[34] and the balancing of interests[35] are of great practical importance for the development and use of AI applications. Especially in connection with authorisations relying on consent, attention must be paid to the data subject's right to withdraw his or her consent[36] and to provisions regulating the processing of special categories of personal data.[37]

a. Consent

The most prominent justification for the processing of data is the subject's consent.[38] The requirements for consent can be derived from a conjunction of the provisions stipulated in Article 4(11), Article 6(1) subpara. 1(a), Article 7, and Article 8 GDPR as well as from the general principles of data protection law. The processing of data is only lawful to the extent that consent has been given, meaning that the data subject must give his or her consent for one or more specific purpose(s).[39] Thus, the scope of the justification is determined by the extent of consent. It should also be pointed out that abstract purposes such as 'advertisement' or 'IT-security' are insufficient.[40] This will also apply in the context of AI. Furthermore, Article 4(11) defines consent as the 'freely given' and 'informed' indication of the subject's declaration of intent. The requirement of an 'informed' decision corresponds directly with the previously elaborated principle of transparency[41] which is also laid down in Article 7(2) GDPR. In an AI scenario, this requirement gives rise to further tension between the controllers' obligation to provide adequate information on the one hand and the information's comprehensibility for the average data subject on the other.

The requirement of 'specific' and 'informed' consent may also pose significant challenges where the controller neither knows nor is able to foresee how and for which purposes the personal data will be processed by self-learning and autonomous AI systems. In principle, the practicability of a consent-based justification may be called into question, particularly when considering the voluntary element of such consent in situations lacking any viable alternatives or scenarios of market dominance. In this regard, it may be said that the requirements of a justification based on consent are more fictional than practicable, especially in view of the ubiquity of data-related consent agreements: 'no one has ever read a privacy notice who was not paid to do so.'[42]

b. Withdrawal of Consent

In addition to the fulfilment of the requirements for a consent-based justification, the technical and legal implementation of the withdrawal of consent as set out in Article 7(3) sentence 1 GDPR

[33] GDPR, Article 6(1) subpara (1)(c)–(e).
[34] GDPR, Article 6(1) subpara (1)(a).
[35] GDPR, Article 6(1) subpara (1)(f).
[36] GDPR, Article 7(3).
[37] GDPR, Article 9.
[38] GDPR, Article 6(1) subpara (1)(a).
[39] For information on earmarking see Section II 3.
[40] Cf. Article 29 Data Protection Working Party, 'Guidelines on Consent under Regulation 2016/679' WP 259 rev. 01, 10.
[41] For transparency see the Section II 1.
[42] Butterworth, 'GDPR Framework' (n 24) 257, 262 *et seq*.

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

is also highly problematic. According to this provision, the data subject has the right to withdraw his or her consent at any time and without having to adhere to any formal requirements. After the consent has been effectively withdrawn, the justification for the processing of data in the sense of Article 6(1) subpara. 1(a) GDPR ceases to exist. In consequence, any further processing of data will only be lawful if, as a substitute, another ground for justification were to apply.[43] Furthermore, a distinction must be made between the right to withdraw in the aforementioned sense, the right to object to unconsented processing of data as regulated by Article 21 GDPR, and, finally, a generally permissible time limitation. As a consequence of the withdrawal of consent, the controller is required to erase the relevant personal data. In cases involving the use of AI, especially scenarios in which certain data is used to train an AI application, it is doubtful whether (and if so, to what extent) the imposition of an obligation to delete is even practicable.[44]

c. Balancing of Interests

The justification based on a balancing of interests allows the processing of personal data in cases where there cumulatively exists (i) a legitimate interest pursued by the controller or by a third party and, (ii) where the processing is necessary to safeguard these legitimate interests, and (iii) where these interests are not overridden by the interests or fundamental rights and freedoms of the data subject who requires the protection of his or her data. The vague wording of this provision is likely to give rise to complications, which do not only apply in the context of AI. For instance, the GDPR does not provide any specific points of reference regarding the general admissibility of and the specific requirements for the processing of data in connection with the balancing of interests within the meaning of Article 6(1) subpara. 1(f) GDPR.

Thus, the task to specify the requirements of the abovementioned balancing process is mostly assigned to academic discourse, courts, and public authorities. However, such an interpretation of the GDPR must, in any case, comply with and adhere to the objective of a consistent standard of (data) protection throughout the EU.[45] It is, therefore, subject to the requirement of a harmonised interpretation of the law which, in turn, is intended to guarantee equal data processing conditions for all market participants in the EU.[46] In addition, by establishing codes of conduct designed to contribute to the appropriate application of the GDPR, Member States are encouraged to provide legal certainty by stating which (industry-specific) interests can be classified as legitimate in the sense of Article 6(1) GDPR. Finally, the European Data Protection Board may, pursuant to Article 70(1)(e) GDPR, further ensure the consistent application of the Regulation's provisions by issuing guidelines, recommendations, and best practices, particularly regarding the practical implementation of the aforementioned balancing process.

d. Special Categories of Personal Data

Article 9 GDPR establishes a separate regulatory regime for special categories of personal data and prohibits the processing of these types of data. These include, for instance, genetic and biometric data, or data concerning health, unless their processing falls under one of the exemptions listed in Article 9(2) GDPR. In accordance with Article 22(4) GDPR, automated decisions, including profiling, must not be based on sensitive data unless these exemptions

[43] It has to be taken into account that it could present itself as contradictory behaviour if, in the case of the omission of consent, an alternative legal justification is applied.
[44] Wallace and Castro, 'Data Protection Regulation' (n 21) 12 *et seq*.
[45] GDPR, Recital 13.
[46] GDPR, Recitals 9 and 10.

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

apply. Furthermore, the processing of large amounts of sensitive data, as referred to in Article 35 (3)(b) GDPR, requires an obligatory data protection impact assessment. Overall, the use and application of AI impose new challenges for the protection of sensitive data. The accumulation of personal data in conjunction with improved methods of analysis and (re-)combination will certainly increase the likelihood of cases affecting potentially sensitive data within the meaning of Article 9 and Recital 51 of the GDPR. Consequently, an increasing amount of data may fall under the prohibition of Article 9(1) GDPR. It is, therefore, necessary to closely follow new trends and developments in the technical field, including but not limited to AI, in order to correctly determine the scope of application of Article 9 GDPR. These findings leave controllers with considerable (legal) uncertainties regarding their obligations.

In light of the new possibilities for a fast and effective AI-based evaluation of increasingly large amounts of data (i.e. big data), the question arises whether metadata, source data, or any other types of information which, by themselves, generally do not allow the average observer to draw any conclusions as to the categories mentioned in Article 9(1) GDPR, nevertheless fall under this provision. If so, one may consider adding the application of AI technology to the list of potential exemptions under Article 9(2) GDPR. In this context, however, regard must be paid to the principle of purpose limitation as previously mentioned.

7. *Intermediate Conclusion*

Given its rather broad, oftentimes undefined and vague legal terminology, the GDPR, in many respects, allows for a flexible application of the law. However, this flexibility goes hand in hand with various (legal) uncertainties. These uncertainties are further perpetuated by the GDPR's notable and worrisome lack of reference to and regulation of AI-specific constellations. As shown above, these constellations are particularly prone to come into conflict with the general principles of data protection as set out in Article 5(1) GDPR and as specified and reiterated in a number of other provisions. In this context, the principles of data minimisation and storage limitation are particularly problematic. Other conflicts, especially involving the GDPR's principles of purpose limitation and transparency, may arise when considering the rather complex and ambiguous purposes and structures for the processing of data as well as the open-ended explorative analyses frequently observed in AI-scenarios. This particularly applies to subsequent changes of purpose.[47] It must also be emphasised that the requirement of transparency serves as a regulatory instrument to ensure the lawfulness of data processing and to detect tendencies of dominance[48] or, rather, the abuse thereof. However, legal uncertainties entail considerable risks and burdens for controllers implementing AI technologies which are amplified and intensified by the GDPR's new and much stricter sanctions regime.[49] Finally, it has to be pointed out that these conflicts by no means only apply to known concerns of data protection law, but rather constitute the starting point for new fundamental questions in this field.

III. COMPLIANCE STRATEGIES (*DE LEGE LATA*)

Based on these findings, it is necessary to examine potential strategies to comply with the provisions of the GDPR and to establish a workable and resilient framework which is capable of fostering the future development and application of AI technologies under the given legal framework. It should

[47] Cf. GDPR, Article 6(4).
[48] For this see the following Section IV 3.
[49] Wallace and Castro, 'Data Protection Regulation' (n 21) 18 *et seq*.

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

also be emphasised that the enactment of the GDPR has fundamentally increased the requirements for compliance with data protection law. This development was further accompanied by substantially higher sanctions for the infringement of data protection law.[50] In addition to potential sanctions, any infringement of data protection law may also give rise to private damage claims pursuant to Article 82 GDPR which cover both material and non-material damage suffered by the data subject. The legally compliant implementation of AI may further be impeded by the interplay and collision of different or conflicting data protection guarantees. Such guarantees can, for instance, be based on data protection law itself, on other personal rights, or on economic and public interests and objectives. In an AI context, this will become particularly relevant in connection with the balancing of interests required by Article 6(1) subpara. 1(f) GDPR.

Article 25 GDPR contains the decisive normative starting point for data protection compliance, in other words the requirement that data protection-friendly technical designs and default settings must be used. However, the rather vague wording of this provision (again) calls for an interpretation as well as specification of its content. The obligation of the responsible controller to implement appropriate technical and organisational measures is essential in terms of data protection compliance. Overall, the GDPR pursues a risk-based approach.[51] From a technical and organisational point of view, it is, thus, necessary to ask how the protection of personal data can be achieved by way of a data protection management system and other measures, for instance through anonymisation and pseudonymisation. The starting point of these considerations is the connection between the data in question and an individual (personal reference), which is decisive for the opening of the substantive scope of application of the GDPR.[52]

1. *Personal Reference*

The existence of such a personal reference is a necessary prerequisite for the application of the GDPR. From a factual point of view, as set out by Article 2(1) GDPR, the GDPR applies in cases of a 'wholly or partially automated processing of personal data and for non-automated processing of personal data stored or to be stored in a file system'. Therefore, it must be asked whether, in a given case and under specific circumstances, personal data is being processed.[53] According to the legal definition stipulated in Article 4(1) GDPR, personal data is

> any information relating to an identified or identifiable natural person [...]; 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 one or more specific characteristics expressing the physical, physiological, genetic, psychological, economic, cultural or social identity of that natural person.

According to the pertinent case law of the European Court of Justice (ECJ), it is sufficient for the responsible data controller to have legal means at his or her disposal to make the data of the third party available (so-called absolute personal reference); this can also encompass detours via state authorities.[54]

[50] Cf. Article 83 GDPR: up to € 20 million or 4% of worldwide turnover.
[51] Cf. GDPR Articles 30(5), 33(1), 35, 36, and 37(1).
[52] Already critical about the old legal situation before the GDPR regarding the (legal) uncertainties regarding personal references and anonymisation J Kühling and M Klar, 'Unsicherheitsfaktor Datenschutzrecht – Das Beispiel des Personenbezugs und der Anonymität' (2013) *NJW* 3611.
[53] On the expanding scope of personal data under the GDPR see Purtova, 'The Law of Everything' (n 20) 40, 43 *et seq*.
[54] CJEU, C-582/14, *Patrick Breyer v Bundesrepublik Deutschland* (19 October 2016), paras 47 *et seq*.

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

The application of the GDPR – and thus the application of its strict regulatory regime – could be avoided by way of, for instance, the data's anonymisation. Article 3(2) of Regulation No. 2018/1807 concerning the free movement of non-personal data states that, in the event of personal and non-personal data[55] being inseparable, both sets of rules (regarding personal and non-personal data) must, in principle, be applied. However, in many cases, it will not be easy to determine with any (legal) certainty whether and to what extent data records may also contain personal data. Hence, in order to remain on the 'safe side' regarding the compliance with the current data protection regime, controllers may feel the need to always (also) adhere to the provisions and requirements of the GDPR even in cases where its application may be unnecessary. This approach may result in considerable (and needless) expenditures in terms of personnel, material, and financial resources.

a. Anonymisation

In contrast to personal information in the aforementioned sense, the GDPR does not apply to anonymous information because they are, by their nature, the very opposite of personal. Recital 26 of the GDPR states: 'The data protection principles should therefore not apply to anonymous information, i.e. information which does not relate to an identified or identifiable natural person, or personal data which has been anonymised in such a way that the data subject cannot or can no longer be identified.'

The Regulation, therefore, does not address the processing of such anonymous data, including data for statistical or research purposes. It follows from the aforementioned Recital that, when it comes to the identifiability of an individual person, the technological capabilities and developments available at the time of the processing must be taken into account. However, when it comes to the technical specifications with regard to the actual anonymisation process, the GDPR, with good reason, does not stipulate a specific procedure to follow. This lack of a prescribed procedure not least benefits innovation and development of new technologies and the concept of technological neutrality.[56] The relevant time of evaluation is always the time of the processing in question.

This is further not changed by mere reference to the fact that almost all anonymised data may be restored by means of advanced sample formations, because such an objection is far too broad and, thus, certainly falls short of the mark.[57] Nevertheless, it should also be noted that, with respect to data relating to location, an anonymisation is considered virtually impossible. Thus, Article 9 GDPR bears particular significance when it comes to the inclusion of location data in the relevant applications. In any case, the issue of de-anonymisation, for which especially the available data stock, background knowledge, and specific evaluation purposes have to be considered, remains highly problematic. According to Recital 26 of the GDPR, in order to identify means likely to be used for the identification of an individual, all objective factors such as costs, time, available technologies, and technological developments, should be considered. In this context, the continuously more advanced big data analysis techniques tend to lead to an ever further reaching re-identification of persons in a constantly growing data pool. In addition, the change of the underlying technological framework and the conditions thereof may (over time) result in the 'erosion' of the former anonymisation and subsequently uncover or expose a

[55] Regulation EU 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free movement of non-personal data in the European Union (2018) OJ L 303, 59.

[56] Due to legal uncertainties companies might be deferred from using such data, Wallace and Castro, 'Data Protection Regulation' (n 21) 15.

[57] On the discussion see Purtova, 'The Law of Everything' (n 20) 40, 42 *et seq*.

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

personal reference of the respective data. Naturally, the consequential (legal) uncertainties may pose a considerable risk and problem to users and other affected parties, especially with respect to issues of practical manageability and incentive. In this context, the opinion on Anonymisation Techniques of the Article 29 Data Protection Working Party (particularly relating to the robustness of randomisation and generalisation) may give helpful indications, but will certainly not be the solution to all potential issues arising in connection thereto.[58] Thus, the findings and developments mentioned earlier give rise to well-founded doubts as to whether the comprehensive anonymisation of data can be successfully achieved under the current framework conditions (e.g. technological progress and available data volumes).

b. Pseudonymisation

According to the legal definition provided in Article 4(5) GDPR, pseudonymisation means the processing of personal data in such a way that personal data cannot (any longer) be assigned to a specific data subject without the use of additional information. Although the GDPR does not expressly permit or privilege the processing of personal data in the event of a pseudonymisation, there are a number of substantial incentives to carry out such a pseudonymisation: in the case of a pseudonymisation, the balancing of interests within the meaning of Article 6(1) subpara. 1(f) GDPR is more likely to sway in favour of the processor. Furthermore, in the case of data protection violations pursuant to Article 34(3)(a) GDPR, the obligation to notify the data subject does not apply in cases of encryption as a sub-category of pseudonymisation. In addition, the procedure may decrease the need for further technical and organisational protection and may, in the event of a previously mentioned change of purpose, be included as a factor in the balancing process as required by Article 6(4)(e) GDPR. Pseudonymisation, therefore, has the potential to withdraw the processing of certain data from the scope of the GDPR and to avoid the application of the Regulation's strict requirements.

c. Synthetic Data

Another possibility to avoid a personal reference and, thus, the application of the GDPR is the production and use of synthesised data which constitute a mere virtual representation of the original set of data. The legal classification of synthetic data is directly linked to the existence or producibility of a personal reference. As a result, the lack of a personal reference allows synthetic data to be equated to anonymous data. In this context and in connection with all related questions, the decisive issue is, again, the possibility of a re-identification of the data subject(s). Another potential problem that must be taken into account relates to eventual repercussions on the data subjects of the (underlying) original data set from which the synthetic data were generated. For instance, processing operations subject to the provisions of the GDPR may hereby arise due to the predictability of sensitive characteristics resulting from a combination of multiple data sets.

2. *Responsibility*

The question of who is responsible for the compliance with the requirements of data protection and to whom data subjects can turn in order to exercise their rights is of great importance.[59] Article 4(7) GDPR defines the data controller as 'the natural or legal person, public authority,

[58] See Article 29 Data Protection Working Party, 'Opinion 5/2014 on Anonymisation Techniques' WP 216.
[59] In detail see Mitrou, 'Data Protection' (n 14) 60 *et seq*.

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

agency or other body which alone or jointly with others determines the purposes and means of the processing of personal data'. In practice, an essential (distinguishing) characteristic of a data controller within the sense of the GDPR is, thus, the authority to make a decision about the purposes and means of data processing. In a number of recent rulings, the ECJ has further elaborated the criterion of responsibility by specifying the nature and extent of a controller's decision concerning the purpose and means of processing personal data: Facebook Fanpage,[60] Jehovah´s Witnesses,[61] and Fashion ID/Facebook Like Button.[62] According to the (previous) case law of the ECJ, those involved in the processing of data do not necessarily have to bear an equal amount of responsibility. Instead, the criterion of responsibility is met if the participants engage in the data processing at different stages and to varying extents, provided that each participant pursues its own purposes for the processing.

3. *Privacy by Default/Privacy by Design*

Article 25 GDPR contains provisions concerning the protection of data by way of (technology) design and data protection-friendly default settings.[63] The first paragraph of the provision stipulates the principles for privacy by design, that is, the obligation to design technology in a manner that facilitates and enables effective data protection (in particular to safeguard the implementation of data-protection principles such as data minimisation). In its scope, the provision is limited to an enumeration of various criteria to be taken into account by the controller with regards to the determination of appropriate measures and their respective durations. The provision does not further specify any concrete measures to be taken by the responsible controller – with the exception of pseudonymisation as discussed earlier. In addition, Article 25(2) GDPR sets out the principle of privacy by default, in other words, the controller's obligation to select data protection-friendly default settings to ensure that only data required for the specific purpose are processed. Finally, in order to demonstrate compliance with the requirements of Article 25(1) and (2) GDPR, Article 25(3) allows the use of an approved certification procedure in accordance with Article 42 GDPR. The challenges previously described typically arise in cases where the GDPR's transparency requirements coincide with complex AI issues, which – by themselves – already present difficulties for the parties concerned. Against this background, certification procedures, data protection seals, and test marks in the sense of Article 42 GDPR could represent valuable instruments on the way to data protection compliance.

4. *Data Protection Impact Assessment*

The data protection impact assessment pursuant to Article 35 GDPR addresses particularly high-risk data processing operations with regard to the rights and freedoms of natural persons. The provision requires the controller to carry out a preventive review of the potential consequences

[60] ECJ, Case C-210/16 Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH (*Facebook Fanpage Case*, 5 June 2018).

[61] ECJ, Case C-25/17 *Tietosuojavaltuutettu v Jehovan todistajat*-uskonnollinen yhdyskunta (*Jehovah's Witnesses Case*, 10 July 2018).

[62] ECJ, Case C-40/17 *Fashion ID GmbH & Co. KG v Verbraucherzentrale NRW eV* (29 July 2019).

[63] For detail on privacy by default and privacy by design see L Bygrave, 'Minding the Machine v2.0: The EU General Data Protection Regulation and Automated Decision Making' in K Yeung and M Lodge (eds), *Algorithmic Regulation* (2019) 9 *et seq*. https://ssrn.com/abstract=3329868.

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

of any processing operations likely to result in such a high risk and to subsequently select and implement the appropriate risk-minimising remedial measures. The obligation to carry out a data protection impact assessment serves the purpose to ensure the protection of personal data and, thus, the compliance with the provisions of the Regulation.[64] At the same time, Articles 35 (4) and (5) require the responsible supervisory authority to establish and make public a list of the kind of the specific processing operations which require such an impact assessment[65] and of those operations which do not require an assessment.[66] These lists are intended to ensure legal certainty for and equal treatment of responsible controllers and to facilitate transparency of all parties concerned. By including processing operations, for which a data protection impact assessment must be carried out, in a list,[67] the supervisory authority can also positively establish an obligation to carry out a data protection impact assessment.

Furthermore, it should be noted that Article 35(1) GDPR explicitly requires the conduction of a data protection impact assessment 'in particular' where 'new technologies' are used. Naturally, this provision is of particular relevance in cases where large amounts of data are processed using 'new' AI systems and technologies and might indicate that the use of such applications may automatically trigger the need for a comprehensive and onerous impact assessment. The GDPR does not explicitly provide any examples of technologies or areas of technology which qualify as 'new'. However, a new technology is likely to pose a high risk to the rights and freedoms of natural persons if it enables the execution of large-scale processing operations which allow the processing of large quantities of personal data at a regional, national, or supranational level and which may involve data relating to a large number of individuals and data of a particularly sensitive nature. Thus, developments such as Smart Car, Smart Health, Big Data, and Tracking procedures as well as new security and monitoring technologies are likely to fall within the scope of Article 35(1) GDPR, hence requiring controllers using and offering such technologies to conduct a data protection impact assessment in accordance with Article 35 GDPR.[68] In the context of AI systems, it remains highly doubtful whether such an obligation would even be feasible, given the fact that self-learning programs develop continuously and more or less unpredictably.

5. *Self-Regulation*

In order to specify, construe, and interpret the large number of indeterminate legal provisions of the GDPR, it is also necessary to give consideration to the elements of self-regulation or co-regulation.[69] Article 40 GDPR gives associations and other bodies the possibility to draw up, amend, or extend rules of conduct which clarify the application of the GDPR. Thus, pertinent rules of conduct can be developed (e.g. by means of best practices) which can subsequently be approved by the responsible supervisory authorities[70] or given general binding force by the EU Commission. In addition, certification procedures, data protection seals, and test marks[71] could also serve as valuable instruments when it comes to the compliance with data protection law.

[64] Purtova, 'The Law of Everything' (n 20) 77.
[65] GDPR, Article 35(4).
[66] GDPR, Article 35(5).
[67] Cf. GDPR Article 35(4).
[68] Mitrou, 'Data Protection' (n 14) 65 *et seq*.
[69] Cf. GDPR Articles 40–43.
[70] GDPR Article 40(5).
[71] Cf. GDPR Article 42(1).

On the basis of Article 42 GDPR, data controllers may also voluntarily seek certification of their data processing operations by the responsible supervisory authority or an accredited body within the meaning of Article 43 GDPR. Recital 100 of the GDPR emphasises that the associated certification procedures, data protection seals, and marks are intended to increase transparency and improve compliance with the GDPR's requirements.

IV. LEGAL POLICY PERSPECTIVES (*DE LEGE FERENDA*)

In view of the earlier points, it becomes evident that the use and implementation of AI-based technologies necessitates a thorough review of the current data protection framework. Such a review may indicate the need for the modification, amendment, or development of the GDPR's current regime. From a legal policy perspective, legislative initiatives should hereby be the main point of focus.

1. *Substantive-Law Considerations*

From a substantive-law point of view, one key element of the GDPR that merits a closer examination is the personal reference as a prerequisite for the GDPR's application. This is not least due to the structural narrowness of the personal reference in its current definition as well as its frequent lack of adequate relevance. Presently, the personal reference as a connecting criterion only insufficiently reflects the existing multiple rationalities of data processing constellations and lacks the capability to take into account the specific characteristics of each case-by-case context. In fact, the one-size-fits-all-approach of the GDPR does not appropriately distinguish between different risk situations, which means that – due to the ubiquitous relevance of personal data – there exists the risk of an excessive application of the law. Among others, this certainly applies to issues relating to the use of AI as presently discussed. With this in mind, it is both necessary and important to create sector-specific regulations for AI constellations, for instance regarding the permissibility of data processing and the specific requirements thereof. Furthermore, the ubiquity of data processing operations in the present age of digitalisation frequently calls into question the general concept of data protection in its current state. It is, therefore, necessary to (at least partially) move away from the current approach, in other words, the prohibition subject to permission in favour of a more general clause. Such a provision should differentiate between different data protection requirements according to specific risks that specific situations are likely to pose. Such a stringent risk-based approach would have the advantage of facilitating the weighing and balancing of the interests of all affected parties as well as appropriately taking into account their respective purposes for protection. In addition, the readjustment of the objectives that the GDPR serves to protect may help to realise an adequate protection of an individual's personality and privacy rights whilst also incentivising the development and use of AI applications. In this context, the overarching objective should always be to reassess the balance of interests pursued by data subjects, responsible processors, third parties, and the public welfare in general.

Another issue that ought to be addressed relates to the granting of access to data and the corresponding rights of usage. This further encompasses questions as to the law of obligations in a data law context, data ownership, and data economics. Finally, due consideration should be given to whether the existing legal framework should be supplemented by specific provisions governing the use of AI. These provisions should not least be capable of overcoming the currently existing tensions resulting from the bi-dimensional, two-person relationship between

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

controller and data subject. This could necessitate an amendment of data protection law with regard to AI in order to move away from an approach based solely on the individual and to appropriately take into account the challenges that may arise in connection with the quantity, heterogeneity, inter-connectivity, and dynamism of the data involved. Such an amendment should be accompanied by more systematic protective measures. A valuable contribution could hereby be made by technical design and standardisation requirements. In addition, all of these measures must be safeguarded and supported by way of an adequate and effective supervisory and judicial protection.

2. *Conflicts between Data Protection Jurisdictions*

Due to the cross-border ubiquity of data (processing) and the outstanding importance of AI-related issues, efforts must be made to achieve a higher degree of legal harmonisation. Ideally, such a development could result in the establishment of an overarching supra- or transnational legal framework, containing an independent regulatory regime suited to the characteristics of AI. Such a regime would also have to take into account the challenges resulting from the interplay of multi-level legal systems as well as the conflicts arising between different data protection legal regimes. For instance, conflicts may arise when the harm-based approach of US data protection law, which is focused on effects and impairments, the Chinese system, which allows for far-reaching data processing and surveillance (e.g. a Social Credit System), and the GDPR approach, which is based on a preventive prohibition subject to permission, collide. Assuming that a worldwide harmonisation of the law is hardly a realistic option in the foreseeable future, it is important to aim for an appropriate balance within one's 'own' data (protection) regime.

3. *Private Power*

In connection with the transnationalisation of the legal framework for data protection and the conflicts between different regulatory regimes, regard must also be paid to the influence exerted by increasingly powerful private (market) players. This, naturally, raises questions as to the appropriate treatment and, potentially, the adequate containment of private power, the latter of which stems from considerations regarding the prevention of a concentration of power and the sanctioning of the abuse of a dominant market position. However, the GDPR itself does not directly stipulate any specific protective measures governing the containment of private power. Legal instruments capable of addressing the aforementioned issues must, therefore, be found outside of the data protection law body. For this purpose, recourse is frequently taken to the (unional or national) competition law, because it expressly governs questions relating to the abuse of market power by private undertakings and, in addition, provides a reliable system and regulatory framework to address such issues. In this regard, the German Federal Cartel Office (FCO) served as a pioneer when it initiated proceedings against Facebook for the alleged abuse of a dominant market position through the use of general terms and conditions contrary to data protection law, specifically the merging of user data from various sources.[72] In any case, the role and power of private individuals as an influential force in the field of data protection should certainly not be underestimated. In fact, by establishing new technological standards and,

[72] BKartA, *Facebook Inc. i.a. Case – The use of abusive business terms pursuant to Section 19 (1) GWB* (B6–22/16, 6 February 2019).

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

thereby, elevating their processing paradigms and business models to a *de facto* legal power, they have the potential to act as substitute legislators.

V. SUMMARY AND OUTLOOK

There is an inherent conflict of objectives between the maximisation of data protection and the necessity to make use of (large quantities of) data, which transcends the realms of AI-related constellations. On the one hand, the availability and usability of personal data bears considerable potential for innovation. On the other hand, the possibilities and limitations of data processing for the development and use of AI are (above all) determined by the requirements of the GDPR. In consequence, the permissibility of the processing of personal data must be assessed in accordance and adherence with the powers to collect, store, and process data as granted by the GDPR. The law of data protection thereby imposes strict limits on the processing of personal data without justification or sufficient information of the data subject. These limitations have a particular bearing on issues relating to AI, as it is frequently impossible to make a comprehensive *ex ante* determination of the scope of the processing operations conducted by a self-learning, autonomous system. This is not least due to the fact that such systems may only gain new information and possibilities for application – potentially relating to special categories of personal data – after the processing operation has already started. In addition, the processing of such large amounts of personal data is oftentimes likely to result in a significant interference with the data subjects' fundamental rights. All of these considerations certainly give rise to doubts as to whether a complete anonymisation of data is even a viable possibility under the given framework conditions (i.e. technological progress and available data volumes).

In order to combat these shortcomings of the current data protection framework, the establishment of a separate legal basis governing the permissibility of processing operations using AI-based applications should be considered. Such a separate provision would have to be designed in a predictable, sector-specific manner and would need to adhere to the principle of reasonableness, thus also ensuring the adequate protection of fundamental rights and the rule of law. The GDPR's *de lege lata* approach to the processing of personal data, in other words, the comprehensive prohibition subject to permission leaves controllers – as previously elaborated – in a state of considerable legal uncertainty. As of now, controllers are left with no choice but to seek the users' consent (whereby the requirements of informing the data subject and the need for their voluntary agreement apply restrictively) and/or to balance the interests involved on a case-by-case basis. These input limits not only burden controllers immensely, but are also likely to ultimately limit output significantly, especially in an AI context. In fact, the main principles of the applicable data protection law, (i.e. the principles of transparency, limitation, reservation of permission, and purpose limitation), appear to be in direct conflict with the functioning and underlying mechanisms of AI applications which were, evidently, not considered during the drafting of the GDPR's legal regime. In practice, this is especially problematic considering that the GDPR has significantly increased the sanctions imposed for violations of data protection law.

Multidimensional border dissolutions occur and do mainly affect the levels of technology and law, territories, and protection dimensions: on the one hand, these border dissolutions may promote innovation, but at the same time they threaten to erode the structures of efficient law enforcement. The previously mentioned tensions between the GDPR and the basic concepts underlying AI also raise the fundamental question of whether traditional data protection principles in the age of digitalisation, especially with regard to AI, Big Data, the Internet of Things, social media, blockchain, and other applications, are in need of a review. Among others,

https://doi.org/10.1017/9781009207898.023 Published online by Cambridge University Press

the instrument of consent as a justification in AI constellations, which are typically characterised by unpredictability, and limited explainability, must be called into question. In any case, the legal tools for the protection of privacy need to be readjusted in the context of AI. This also and especially applies to the data protection law regime. Against this background, legislative options for action at national, unional, and international level should be examined. In this context, the protection of legal interests through technology design will be just as important as interdisciplinary cooperation and research.

Overall, (legal) data policy is a central industrial policy challenge that needs to be addressed – not only for AI constellations. Legal uncertainties may cause strategies of evasion and circumvention, which in turn (can) trigger locational disadvantages and enforcement deficits, bureaucratic burdens, and erosion with respect to legal compliance. Thus, AI-specific readjustments of data protection law should – where necessary – prevent imminent disadvantages in terms of location and competition and ensure that technology and law are open to innovation and development. Both new approaches to the interaction between data protection law and AI should be examined and existing frameworks retained (and, where appropriate, further developed). By these means, a modern data and information usage right may be established which does not result in a 'technology restriction right' but rather gives rise to new development opportunities. The legal questions raised and addressed in this article concern not only isolated technical issues but also the social and economic order, social and individual life, research, and science. In this sense, the existing legal framework (the European approach) should be further enhanced/developed to make it an attractive alternative to the approaches taken in the US and China, while the current model of individual protection should be maintained, distinguishing it from the other data protection regimes. With the ongoing GDPR evaluation, it is an opportune time for such an initiative. However, such an initiative requires the cooperation of all actors (users and developers, data protection authorities and bodies, policy and legislation, science, and civil society) in order to reconcile data protection with the openness of technology and law for necessary developments.