Tilburg Law School, LL.M. Law and Technology

The legal protection against inferences drawn by AI under the GDPR

July 2020

Author:
Celin Fischer
Snr. 2045976
Anr. 779616

First Supervisor:
Dr. Nadezhda Purtova LLM MSc
Second Supervisor:
Dr. Esther Keymolen

Table of Contents

Abbreviations

Article 29 Working Party	Art. 29 WP
Artificial Intelligence	AI
Artificial Neural Network	ANN
Court of Justice of the European Union	CJEU
Data Protection Directive	DPD
Etcetera	Etc.
European Data Protection Board	EDPB
Example given	E.g.
Machine Learning	ML

1. Introduction

1.1. Problem statement

Decisions that were once based on human intelligence are increasingly being replaced and supplemented by artificial intelligence (AI).[1] AI can be used in various areas, e.g. in human resources for recruitment, in the financial sector for credit scoring and in healthcare for tumor detection. Unlike humans, AI can evaluate large amounts of data, search for links in these datasets and make inferences based on the findings. The use of AI promises to offer great benefits and is therefore being implemented across a variety of fields. However, its implementation is accompanied by many challenges. The data an algorithm is trained with and also the data serving as input are influencing the performance, the outputs and the inferences drawn tremendously.[2] This is being referred to as the "garbage in, garbage out" principle.[3] The concern that inferences drawn about individuals are biased, discriminating or in any other way inaccurate, are serious, as they have the possibility to not only influence a single decision or prediction made about an individual, but these inferences can be fed back into the AI system and influence any future decision or prediction.[4] Inferences are drawn, for instance, to determine someone's credit risk. The credit score calculated on a variety of data is itself the inference drawn.[5]

---

1 Puaschander, J. and Feierabend, D. 'Artificial Intelligence in the healthcare sector' [2019] 2(4) International Journal of Multidisciplinary Research < https://ssrn.com/abstract=3469423>, page 1; Ferretti, A. et al. 'Machine Learning in Medicine: Opening the New Data Protection Black Box' [2018] 4(3) European Data Protection Law Review https://doi.org/10.21552/edpl/2018/3/10, page 320.

2 Independent High-Level Expert Group On Artificial Intelligence, 'A Definition of AI: Main capabilities and disciplines' [2019] Set up by the European Commission https://ec.europa.eu/futurium/en/ai-alliance-consultation, page 5; Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 684; Danks, D. and London, A. 'Algorithmic Bias in Autonomous Systems' [2017] Twenty-Sixth International Joint Conference on Artificial Intelligence <DOI: 10.24963/ijcai.2017/654>, page 2.

3 Mittelstadt, B. et al. 'The Ethics of Algorithms: Mapping the Debate' [2017] 3(2) Big Data & Society <https://ssrn.com/abstract=2909885>, page 5; Hand, D. et al. Principles of Data Mining (A Bradford Book The MIT Press 2001) 32; Raso, F. et al. 'Artificial Intelligence & Human Rights: Opportunities & Risks' [2018] Bergmann Klein Center Research Publication No 2018-6 https://ssrn.com/abstract=3259344, page 15; Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 683.

4 See Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018); Raso, F. et al. 'Artificial Intelligence & Human Rights: Opportunities & Risks' [2018] Bergmann Klein Center Research Publication No 2018-6 https://ssrn.com/abstract=3259344, page 7; Danks, D. and London, A. 'Algorithmic Bias in Autonomous Systems' [2017] Twenty-Sixth International Joint Conference on Artificial Intelligence <DOI: 10.24963/ijcai.2017/654>, page 4692.

5 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 8; Kamp, M. et al.

While the data the credit score is based on can be reasonable, such as credit histories, current loans, and employment status, data such as the zip code one lives in, the gender or skin color, seem unreasonable for anyone or anything, making a decision in regard to a credit application.

The concern is that individuals will be put into categories that are for one, inaccurate, and secondly, hard to break out of. This can lead to individuals being discriminated and social inequality being amplified, by categorizing individuals based on inferences, that cannot be evaluated for accuracy or are not subject to checks for up-to-datedness.[6] For this reason, inferences can implicate individuals in negative ways and individuals need to be protected from those implications.

AI systems are often trained with personal data, that were either once provided by the data subject or observed, possibly combined with ordinary data about the data subject.[7] The notion of personal data is the trigger for the application of the European Union General Data Protection Regulation (GDPR)[8] that took effect on May 2018, by which data protection law was further unified in the EU.[9] For the material scope of the GDPR to apply to inferences drawn by AI, they must be considered as personal data that are processed in accordance with the definitions provided under Art. 4 GDPR.[10] However, despite the importance of the concept of personal data, the boundaries of what can constitute such are unclear. [11] Even though, European doctrine and jurisprudence generally have adopted a broad and technology-neutral definition of the concept of

---

'Profiling of Customers and Consumers – Customer Loyalty Programmes and Scoring Practices' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 207.

[6] See also: Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018) 1.

[7] Raso, F. et al. 'Artificial Intelligence & Human Rights: Opportunities & Risks' [2018] Bergmann Klein Center Research Publication No 2018-6 https://ssrn.com/abstract=3259344, page 18.

[8] 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 (General Data Protection Regulation), and repealing Directive 95/46/EC (henceforth "the data protection Directive), published in Official Journal of the European Union L, 119, 4 May 2016.

[9] Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 1, 2.

[10] Art. 2(1), Art. 4(1), (2) GDPR; Edwards, L. 'Data Protection: Enter the General Data Protection Regulation' [2018] Forthcoming in L Edwards ed Law, Policy and the Internet (Hart Publishing, 2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3182454>, page 5.

[11] Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 1, 2.

personal data, but whether inferences drawn can constitute personal data, is disputed.[12] Therefore, in order to assess whether the GDPR can provide protection for data subjects over their inferred data, the four elements of the personal data definition under Art. 4(1) GDPR have to be individually interpreted, to establish whether drawn inferences can fit the definition. If only one, of the four elements, is not fulfilled, inferences will not constitute personal data.[13] If inferences would constitute personal data, data subjects would be granted several rights under the GDPR and controllers would face many obligations.[14] Not only would data controllers be obliged to provide the information under Articles 13 and 14 GDPR about the processing of personal data, but they would have to comply with the seven data protection principles on which the GDPR is built upon.[15] The obligations posed on the data controllers aim to ensure that they comply with each data protection principle when processing personal data.[16] Therefore, the question arises whether the GDPR, with its current framework of rights and obligations, provides sufficient protection for individuals in regard to inferences drawn by AI?

1.2. Literature Review

Regarding the state of the art of the literature, there is extensive and growing literature in regard to AI technologies and their possible implications for individuals and society. There are scholars taking a human rights-based approach to the implementation of AI in society[17] and other scholars focusing more on the ethical concerns in general[18]. Many legal scholars which identify the implications posed by AI turn to privacy and data

---

[12] Ibid 2.
[13] Ibid 3.
[14] Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review https://ssrn.com/abstract=3248829, page 5; Humerick, Matthew 'Taking AI Personally: How the EU must learn to balance the interests of personal data privacy and artificial intelligence', page 402.
[15] Article 5 GDPR; Article 29 Working Party, Opinion 8/2014 on the on Recent Developments on the Internet of Things, Adopted on 16 September 2014, 14/EN WP 223, < https://www.pdpjournals.com/docs/88440.pdf>, page 16; Schreurs, Wim et al. 'Cogitas, Ergo Sum. The Rolfe of Data Protection Law and Non-discrimination Law in Group Profiling in the Private Sector' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 243.
[16] Hoofnagle, C. et al. 'The European Union general data protection regulation: what it is and what it means' [2019] 28(1) Information & Communications Technology Law, https://doi.org/10.1080/13600834.2019.1573501, page 92.
[17] For instance, Raso, F. et al. 'Artificial Intelligence & Human Rights: Opportunities & Risks' [2018] Bergmann Klein Center Research Publication No 2018-6 https://ssrn.com/abstract=3259344.
[18] Mittelstadt, B. 'Principles Alone cannot Guarantee Ethical AI' [2019] Nature Machine Intelligence, < https://ssrn.com/abstract=3391293>; Mittelstadt, B. et al. 'The Ethics of Algorithms: Mapping the Debate' [2017] 3(2) Big Data & Society https://ssrn.com/abstract=2909885.

protection in order to seek effective regulation or in order to demonstrate whether current frameworks are sufficient or failing to address the implications.[19] Three areas of scholarship have caught my eye specifically, when reviewing what has been written in regard to regulating AI under European data protection law. For one, this is the extensive discussion around the so-called 'right to explanation' which may or may not be entailed in the GDPR and which may or may not be effective in regulating automated decision-making and profiling.[20] Furthermore, the article of Wachter et al. on 'the right to reasonable inferences', which doesn't specifically state, that the right to explanation is not existent, but rather doubts its effectiveness and proposes a different right to be included into the framework of the GDPR.[21] And thirdly, the 'Right to Legibility of Automated Decision-Making' by Malgieri and Comandé, which also does not dismiss the existence of a right to explanation but proposes a different concept of the right.[22]

In regard, to the material scope of the GDPR and more specifically its definition of personal data, Purtova[23] and Dalla Corte[24] have both published articles with detailed

---

[19] Among many other: Humerick, M. 'Taking AI Personally: How the EU must learn to balance the interests of personal data privacy and artificial intelligence' [2018] 34(4) Santa Clara High Technology Law Journal < https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=1633&context=chtlj>; Mitrou, L. 'Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) "Artificial Intelligence-Proof"?' [2019] University of the Aegean Dpt. of Information and Communication Systems Engineering; Athens University of Economics and Business - Department of Informatics < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914>; Zarsky, T. 'Incompatible: The GDPR in the Age of Big Data' [2017] 47(4) Seton Hall Law Review (2017) < https://ssrn.com/abstract=3022646>.

[20] Supporters of the right: Goodman, B. and Flaxman, S. 'European Union regulations on algorithmic decision-making and a "right to explanation"' [2017] 38(3) AI Magazine, < https://arxiv.org/ct?url=https%3A%2F%2Fdx.doi.org%2F10.1609%2Faimag.v38i3.2741&v=f2c797e9>; Selbst, A. and Powles J. 'Meaningful Information and the Right to Explanation' [2017] 7(4) International Data Privacy Law https://ssrn.com/abstract=3039125; Kaminski, M. 'The Right to Explanation, Explained' [2018] 35(1) Berkeley Technology Law Journal (2019) < https://ssrn.com/abstract=3196985>. Doubters of the right: Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855; Wachter, S et al. 'Why a Right to Explanation of Automated Decision-Making Does not Exist in the General Data Protection Regulation' [2017] International Data Privacy Law < https://ssrn.com/abstract=2903469>.

[21] Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review https://ssrn.com/abstract=3248829.

[22] Malgieri, G. and Comandé, G. 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' [2017] 7(4) International Data Privacy Law, < https://ssrn.com/abstract=3088976>.

[23] Purtova, N. 'The law of everything. Broad concept of personal data and future of EU data protection law' [2018] 10(1) Law, Innovation and Technology (2018) < https://ssrn.com/abstract=3036355>.

[24] Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>.

interpretations of the definition. The issues of inferences drawn by AI, have been specifically addressed by Hildebrandt and Koops.[25] Also the discussion of applying the provisions of the GDPR on the data subject rights to the case of inferences drawn by AI has been addressed by scholars[26] and discussing inferences drawn by AI under the definition of personal data under the GDPR or the former Data Protection Directive has been approached by Wachter et al.[27]

Despite scholars proposing the GDPR, and former DPD, as regulatory frameworks for addressing implications posed by AI technologies and to some extent those posed by inferences drawn, none of the literature addresses inferences under the GDPR in detail in regard to whether the category of inferred data is possibly entailed by the material scope of the GDPR, specifically if all elements of the personal data definition can be fulfilled by inferences drawn, in order to apply the provisions on data subject rights to the inferences drawn or those intended to being drawn. Wachter et al. do this briefly, by portraying the relevant case law of the CJEU and the guidance documents from the Art. 29 WP, however, quickly concluding that the GDPR in its current state is unable to mitigate the implications posed for individuals by inferences, rather than focusing on how the GDPR, in its current state, may be interpreted in order to cover inferences sufficiently.

Therefore, this thesis aims at closing the gap in the literature by focusing on the current state of the art of the GDPR's legislative text. For one, this will be done by defining the category of inferred data under the definition of personal data under the GDPR and secondly, by applying the provisions on the data subjects' rights on inferences

---

[25] Hildebrandt M. and Koops, E.J. 'The challenges of ambient law and legal protection in the profiling areas' [2010] 73(3) Moden Law Review, < https://pure.uvt.nl/ws/portalfiles/portal/1248058/Koops_The_Challenges_of_Ambient_Law_100712.pdf.

[26] Hallinan, D. and Borgesius, F. 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' [2020] 10(1) International Data Privacy Law (2020) https://doi.org/10.1093/idpl/ipz025; Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855; Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review https://ssrn.com/abstract=3248829.; Blanke, J. 'Protection for 'Inferences Drawn: A Comparison between the General Data Protection Rule and the California Consumer Privacy Act' [2020] < https://ssrn.com/abstract=3518164>; Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018).

[27] Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review https://ssrn.com/abstract=3248829.

drawn, while focusing on the characteristics of inferred data, which differ from those of provided and observed data, leading to hurdles in applying the GDPR on inferences.[28] Acknowledging the hurdles, this thesis will not jump to conclude that the GDPR is inefficient to mitigate the risks, but focusing on how it must be or could be interpreted in order to be sufficient, rather than relying on proposing new rights to being added to the framework of the GDPR.

1.3. Main research question and sub-questions

Addressing the identified gap in the literature the following main research question, which this thesis aims to answer, has been raised:

To what extent can the EU General Data Protection Regulation address the implications posed to individuals of inferences drawn by artificial intelligence?

In order to answer the main research question, the following sub-questions have been formulated:

1. *How does artificial intelligence draw inferences and what are the implications posed for individuals if left unaddressed?*
2. *Can inferences be defined as personal data under Art. 4(1) GDPR and does the GDPR apply to inferences?*
3. *Are the data subject rights able to sufficiently address inferences drawn in order to mitigate the implications posed for individuals?*

1.4. Methodology

In order to create a sound overview and understanding of AI technology and how inferences get drawn by them, the second chapter will consult computer science literature. Computer science sources were especially used for explaining the various definitions of the sub-fields of AI that are currently widely discussed, such as machine learning, deep learning, artificial neural networks and data mining. Furthermore, computer science literature was used for explaining the functioning of AI technologies, such as how input data can be assessed by the algorithms in order to create certain outputs. This was important in order to demonstrate how AI technologies draw inferences, as this understanding is necessary to evaluate the different steps in the following analytical chapters. However, the technical descriptions are more functioning as an overview of the

[28] See below under 4.2. “Specific issues relating to challenging inferences”.

technology for readers of other sciences, rather than aiming at providing technologically precise descriptions in terms of computer science. Since this field is inherently complex, the claim, to make accurate and comprehensive statements about AI from the perspective of computer science, is not raised. The literature review was based on a small collection of books on AI and machine learning technology, retrieved from academic databases.

This research led to defining inferences, which is based on a variety of legal and social sciences literature, in order to demonstrate the relevant attempts that have been made in scholarship to define inferences drawn by AI. The sources have been found by using the term 'inferences' in the search function of various academic databases.

Furthermore, literature from the social science community was used to present the societal and ethical problems and concerns that accompany the technology of artificial intelligence and are currently part of many scholarly discussions. Essentially, this literature was used to portray the issues accompanying AI in general and further to analyze which issues concern inferences drawn specifically, in the second chapter. The necessary literature was retrieved through the use of the search functions on various academic databases. The chosen sources were primarily selected for their topicality and scientific relevance, and, where appropriate, further narrowed down based on their significance for the main research question. The second chapter was solely based on secondary literature research and sources.

The third and fourth chapter relied more on doctrinal legal analysis, as they aim to define and interpret the relevant provisions and terms of the GDPR. The legislative texts of the GDPR and the Data Protection Directive are next to case law important sources. This black-letter law research is accompanied by legal scholarship. The necessary secondary legal literature was retrieved on academic databases by using the search function. It mostly consists of European scholars, as the GDPR is the European Union's data protection framework. However, there are important scholars from outside of the EU and their literature has been included, when their work was relevant for the research question and necessary in order to portray the extent of the scholarly discussion. The primary literature was found by actively searching for the relevant case law and legislative frameworks. Next to the legal scholarship, official documents of institutions of the European Union were essential for these chapters. They have been retrieved by searching the websites of the European Data Protection Board and the European Data

Protection Supervisor, for relevant guidance documents. Non-academic sources have been scarcely used as examples of the implementation of AI technologies or to demonstrate the implications that the implementation can cause for individuals and society. These were chosen based on an internet search.

1.5. Structure

This thesis will start off by demonstrating the technological background of how inferences get drawn, by introducing the various sub-fields of AI and the state of the art of the technology in the second chapter. After explaining the training phase and the phase of drawing inferences in AI systems, this chapter defines inferences and portrays the characteristics of the category of inferred data. Building on the technological knowledge, this chapter will illustrate some of the implications that inferences drawn by AI can pose for individuals.

The third chapter evolves around the material scope of the GDPR under Art. 2(1) GDPR. After an introduction into the different elements of the "personal data" definition under Art. 4(1) GDPR, inferences drawn will be analyzed under each element of the definition individually. Following this, it will be further analyzed whether inferences can constitute a special category of data under Art. 9(1) GDPR. The chapter will conclude by answering the second sub-question, whether or not, the GDPR is applicable to inferences drawn by AI.

As defining inferences drawn as personal data triggers the GDPR's material scope, the fourth chapter, aims to apply provisions of the GDPR to scenarios in which inferences get drawn. After highlighting the hurdles data subjects are faced with, due to the characteristics of inferred data as portrayed in the second chapter, this chapter applies Articles 13, 14, 15 and 20 GDPR in order to tackle the issue of individuals' unawareness of inferences being drawn about them. This analysis is followed, by the application of the data subject rights under Articles 16-22 GDPR, in order to evaluate their possibility of mitigating the implications raised by inferences drawn.

2. Inferences drawn by Artificial Intelligence: Applications and implications

2.1. Introduction

This chapter provides the basis that is necessary for the later in-depth analysis of inferences drawn under the GDPR. Firstly, the different fields that have established under the term of AI, such as machine learning and data mining, will be explained in order to gain an understanding of how these systems work and draw inferences. This will be done, by distinguishing between the training of AI and the subsequent drawing of inferences. Secondly, the notion of inferences will be defined, by referring to the different definition's existent in scholarship and how drawing inferences relates to the practice of profiling. At last, the implications for individuals that can arise when drawing inferences or using drawn inferences further, will be portrayed.

2.2. Artificial Intelligence and its types

2.2.1. Artificial Intelligence (AI) in general

AI is a universal field which encompasses a variety of subfields, ranging from general fields, such as learning, to specific ones, e.g. diagnosing diseases or automatic driving.[29] Due to the various fields that AI is implemented, no single definition for "artificial intelligence" has been established.[30] The attempted definitions though share a variety of necessary conditions a technology has to meet, to characterize as AI.[31] For this thesis AI is defined as consisting of sets of algorithms, usually coming as part of a software, which can perform a certain task, that would have been otherwise performed by a human, by making use of different techniques.[32]

---

[29] Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 1.

[30] See four approaches to defining AI in: Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 2.

[31] Siapka, A. 'The Ethical and Legal Challenges of Artificial Intelligence: The EU response to biased and discriminatory AI' [2018] Panteion University of Athens <https://ssrn.com/abstract=3408773> accessed 13 January 2020, page 12.

[32] Kiseleva, A. 'AI as a Medical Device: Is It Enough to Ensure Performance Transparency and Accountability in Healthcare?' [2019] European Pharmaceutical Law Review (1/2020) <https://ssrn.com/abstract=3504829> accessed 13 January 2020, page 2. See also the definition provided by Independent High-Level Expert Group On Artificial Intelligence, 'A Definition of AI: Main capabilities and disciplines' [2019] Set up by the European Commission https://ec.europa.eu/futurium/en/ai-alliance-consultation.

Before any AI system can be used for specific tasks, for instance drawing inferences in order to make predictions about individuals, it has to be trained.[33] This will be discussed below for each AI field individually.

2.2.2. Machine Learning (ML)

If the AI algorithms have the ability to learn from data, we enter the field of machine learning (ML).[34] Machine learning works by getting trained with a collection of paired inputs and outputs, in order to learn the function of predicting outputs from new, unpaired, inputs.[35] ML can be characterized as having the ability to learn, supervised or unsupervised, how to perform a certain task.[36] In supervised learning the algorithm is presented with example input-output pairs, input data is labeled with the desired outputs, from which the ML algorithm is supposed to learn how to develop a pattern to get from the input to the output.[37] In unsupervised machine learning the input data is unlabeled and the algorithm must find its own pattern to get from the input to an output.[38] Based on those identified patterns it builds models.[39] After the training phase, the ML algorithm can then use the created models to analyze new input data, without supervision.[40] So, the data that ML algorithms are trained with can either be input-output pairs or just input data.[41] In both cases the ML algorithms develop decision-making rules for handling new input data, with the distinction that one is supervised by a human and the other works autonomously.[42] ML algorithms own the capability to improve their own performance on

---

33 Anrig, B. et al. 'The Role of Algorithms in Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010), 70.
34 Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 37, 41.
35 Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 693.
36 Ferretti, A. et al. 'Machine Learning in Medicine: Opening the New Data Protection Black Box' [2018] 4(3) European Data Protection Law Review https://doi.org/10.21552/edpl/2018/3/10, page 320.
37 Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 695; Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 43.
38 Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 43; Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 694.
39 Selbst, A. and Powles J. 'Meaningful Information and the Right to Explanation' [2017] 7(4) International Data Privacy Law <https://ssrn.com/abstract=3039125> accessed 27 January 2020, page 14.
40 Ibid 14.
41 Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 677; Anrig, B. et al. 'The Role of Algorithms in Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010), 78.
42 Mittelstadt, B. et al. 'The Ethics of Algorithms: Mapping the Debate' [2017] 3(2) Big Data & Society <https://ssrn.com/abstract=2909885>, page 6.

future tasks through the use in real-world practice and by making experiences and observations about the world.[43]

2.2.3. Artificial Neural Networks (ANN)

Among others, there are artificial neural networks, which is a framework for how ML algorithms can work.[44] Neural networks can be defined as logical networks with the complexity needed to classify information in a similar way the human brain would.[45] They are structured as an interconnected group of neurons, transmitting signals between the neutrons through connections.[46] So called, feed-forward networks are organized in layers and each layer performs a different kind of transformation on the input data.[47] An untrained neural network is trained by providing the first layer with training data.[48] Based on the task, the neurons categorize the input data.[49] New input data gets passed along each layer of the neural network until the final output is determined.[50] The signals travel from the first layer, which can be considered as the input, to the last layer, the final output.[51] This framework makes it possible for ML algorithms to find patterns in major amounts of data.[52] The results of the output layer are compared with the known results of the training data. According to the error value, the algorithm can adapt his framework based

---

[43] Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 693; Kiseleva, A. 'AI as a Medical Device: Is It Enough to Ensure Performance Transparency and Accountability in Healthcare?' [2019] European Pharmaceutical Law Review (1/2020) <https://ssrn.com/abstract=3504829> accessed 13 January 2020, page 1; U.S. Food and Drug Administration (FDA), 'Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)', https://www.fda.gov/media/122535/download, page 2.

[44] Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 727; Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 54.

[45] Anrig, B. et al. 'The Role of Algorithms in Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010), 77; Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 727.

[46] Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 727; Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 53.

[47] Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 54; Anrig, B. et al. 'The Role of Algorithms in Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010), 77.

[48] Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 729.

[49] Anrig, B. et al. 'The Role of Algorithms in Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010), 78.

[50] Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 729.

[51] Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 54; Anrig, B. et al. 'The Role of Algorithms in Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010), 77.

[52] Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 54.

on how correct or incorrect the value is. This is called re-evaluation of the error value which is an important step in training because this is how an algorithm can improve its performance.[53]

2.2.4. Data mining techniques

Another field of AI is data mining. "Data Mining is the analysis of (large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner."[54] The task of data mining is the extraction of patterns and knowledge from large data sets.[55] Thus, in order to use data mining algorithms, a data set is needed, on which certain tasks are to be performed, such as finding links between the data.[56] Data mining algorithms can perform a variety of tasks, such as detecting anomalies in data, searching for links between data, clustering and classification.[57] In data mining the algorithms involved also consist of the ability to learn, supervised or unsupervised, in order to build models, but they follow different goals than those of ML.[58] While ML algorithms aim to form predictions based on what they have learned in the training phase from the input data, data mining algorithms focus on already existing datasets, to search them for unknown patterns and relationships.[59]

2.3. Drawing inferences

One thing the beforementioned AI techniques have in common, is that after the algorithms have been trained, they are able and ready to draw inferences.[60] The drawing of inferences can be described as the step in between of training the algorithms and

---

53 Russel, S. and Norvig, P., Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010) 733.
54 Hand, D. et al. Principles of Data Mining (A Bradford Book The MIT Press 2001) 6.
55 Hand, D. et al. Principles of Data Mining (A Bradford Book The MIT Press 2001) 6; Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 108.
56 Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 109.
57 Canhoto, A. and Backhouse, J. 'General Description of the Process of Behavioural Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 51; Custers, B. The Power of Knowledge Ethical, Legal and Technological Aspects of Data Mining and Group Profiling in Epidemiology (Wolf Legal Publishers 2004) 19; Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 109.
58 Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 110.
59 Bloch, D., Machine Learning: Models and Algorithms, Quantitative Analytics (2018) 51; Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 680.
60 Anrig, B. et al. 'The Role of Algorithms in Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010), 78; Goodfellow, I. et al. Deep Learning (MIT Press 2016) 104.

applying the drawn inferences to individuals, in order to make predictions or decisions. Inferences can be drawn in algorithmic decision-making processes, in big data analysis, in other machine-learning techniques and in profiling.[61] Similar to how the human brain gains and uses its knowledge, algorithms can infer things about new data based on its training.[62] Artificial neural networks, for instance, equal a huge database, that are able to retain the training and quickly apply it to new input data.[63] This is inference. "Inference refers to the process of using a trained machine learning algorithm to make a prediction."[64] Wachter, et al. define inferences "as information relating to an identified or identifiable natural person created through deduction or reasoning rather than mere observation or collection from the data subject."[65] Hallinan and Borgesius use the term 'opinions' instead of inferences and define them as "an assertation about an entity, built on the back of facts about that entity subjected to some interpretative framework to produce new, probable facts."[66]

Edwards and Veale describe the process of drawing inferences by comparing supervised and unsupervised machine learning algorithms. According to them, supervised learning algorithms take a set of variables, for instance a set of certain characteristics, and a correct label is added to this set, which could be for instance a medical diagnosis.[67] They refer to this as the ground truth.[68] "The aim of supervised learning is to accurately predict this ground truth from the input in variables in cases where we only have the

---

61 Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 13.
62 Anrig, B. et al. 'The Role of Algorithms in Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010), 78.
63 Ibid 78.
64 Paul DeBeasi, 'Training versus inference' [2019] Gartner Blog Network, https://blogs.gartner.com/paul-debeasi/2019/02/14/training-versus-inference/, last accessed 15 June 2020.
65 Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 22.
66 Hallinan, D. and Borgesius, F. 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' [2020] 10(1) International Data Privacy Law (2020) https://doi.org/10.1093/idpl/ipz025, page 6.
67 Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855, page 25.
68 Ibid 25.

latter."[69] In contrast, unsupervised learning algorithms are not based on the ground truth, but rather they draw inferences based on new insights.[70]

2.3.1. Categorizing inferred data

The Art. 29 WP[71] and several scholars[72] distinguish between four categories of data: Provided, observed, derived and inferred. Provided data is collected or created by direct actions of the individual himself, which he is aware of, for instance, by filling out an application or a questionnaire.[73] Observed data is created by observing actions of the individual, which the individual is not necessarily aware of, for instance, cookies or tracking of location data via an application.[74] Derived data is created of existing data on the individual, for instance, the average purchase of a customer visiting an online shop.[75]

Inferred data is developed by processing other data and can be defined as the product of an analytical process which aims to find correlations in datasets to, for example, make predictions, without direct or indirect action of the individual.[76] On the one hand,

---

[69] Ibid 25.
[70] Ibid 25.
[71] Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 8.
[72] Abrams, M. 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, <https://ssrn.com/abstract=2510927>, page 1; Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 23; Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018) 3.
[73] Abrams, M. 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, <https://ssrn.com/abstract=2510927>, pages 5, 6; Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 8.
[74] Abrams, M. 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, <https://ssrn.com/abstract=2510927>, pages 5, 6; Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 8.
[75] Abrams, M. 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, <https://ssrn.com/abstract=2510927>, pages 5, 7.
[76] Information Commissioner`s Office (ICO) (2017), 'Big data, artificial intelligence, machine learning and data protection', https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf accessed 20 January 2020, page 13; Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>; Abrams, M. 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, <https://ssrn.com/abstract=2510927>, page 8; Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt,

inferred data is the output of the processing of other data and on the other, inferred data is the input for creating new data, outputs, like predictions.[77] This leads to more data being increasingly developed, with the level of individual awareness and involvement of the data subjects being low.[78] In general, inferred data is created without the affected individual being aware of its creation, due to the key characteristic of inferred data, it being not provided, directly or indirectly, by the individual.[79]

2.3.2. Inferring profiles from individuals

Data mining can lead to the construction of profiles about individuals or groups of individuals.[80] This process is referred to as profiling.[81] According to Hildebrandt profiling can be defined as follows:

> *"The process of 'discovering' correlations between data in databases that can be used to identify and represent a human or nonhuman subject (individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category"*[82].

Mendoza and Bygrave define profiling as

> *"the process of (i) inferring a set of characteristics about an individual person or group of persons (i.e. the process of creating a profile), and/or (ii)*

Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018) 2.

77 Abrams, M. 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, <https://ssrn.com/abstract=2510927>, page 8; See also Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018) 2.

78 Abrams, M. 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, <https://ssrn.com/abstract=2510927>, page 8; Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 17; Moerel, L. and Wolk, A. v. d. 'Big data analytics under the EU General Data Protection Regulation', [2017], https://ssrn.com/abstract=3006570, page 28.

79 Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018) 2; Abrams, M. 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, <https://ssrn.com/abstract=2510927>, page 3.

80 Hildebrandt, M. 'Defining Profiling: A New Type of Knowledge?' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 18.

81 Ibid 18.

82 Ibid 19.

treating that person or group (or other persons/groups) in light of these characteristics (i.e., the process of applying a profile)"[83].

Profiling is also legally defined in the GDPR as

"any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location or movements;"[84].

Profiling can be used to infer certain information about an individual, such as risks, habits and other characteristics.[85] Two relevant steps are distinguished here, first, the construction of the profile, that can be inferred from the individual's data and/or the pre-existing dataset, and secondly, the application of the profile to the individual, e.g. using the profile to make a decision or prediction about the individual.[86]

A well-known example of where profiling is used, and inferences are drawn, is that of credit scoring.[87] The credit score is generally used to estimate whether an individual qualifies for a loan and if so, it also influences the duration of the loan, the rate and the credit limit.[88] The credit score itself is the inference drawn.[89] This is in line with

---

[83] Mendoza, I. and Bygrave, L. 'The Right not to be subject to automated decisions based on profiling' [2017] 20 University of Oslo Faculty of Law Legal Studied Research Paper Series No. 2017-20, page 1.
[84] Article 4(4) GDPR.
[85] Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855, page 32; Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 9; Jaquet-Chiffelle, D.-O. 'Defining Profiling: A New Type of Knowledge?' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 35.
[86] Hildebrandt, M. 'Defining Profiling: A New Type of Knowledge?' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 19; Jaquet-Chiffelle, D.-O. 'Defining Profiling: A New Type of Knowledge?' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 41.
[87] See for an extensive discussion on profiling in credit scoring: Kamp, M. et al. 'Profiling of Customers and Consumers – Customer Loyalty Programmes and Scoring Practices' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 205 following.
[88] Kamp, M. et al. 'Profiling of Customers and Consumers – Customer Loyalty Programmes and Scoring Practices' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 206.
[89] Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 8; Kamp, M. et al. 'Profiling of Customers and Consumers – Customer Loyalty Programmes and Scoring Practices' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 207.

the argumentation of Custers, that "profiles are not regarded as knowledge, but rather as (new) data, namely as inferred data"[90]. Profiling techniques allow inferences to be drawn from either the data of the individual itself, or from data of other persons.[91]

2.4. Challenges of inferences drawn and their implications for individuals

AI promises to solve a wide range of problems and improve many facets of our lives and while it may provide many opportunities, these are accompanied by various challenges.[92] The use of AI can lead to neutral, non-discriminatory and accurate decisions, outcomes or predictions. At the same time though, it can lead to outcomes being based on bias, discriminatory factors and inaccurate data, resulting in individuals being treated unfair or even discriminated.[93] This section aims to illustrate the importance of addressing inferences drawn by illustrating a selection of commonly discussed challenges relating to AI in general, before analyzing how these challenges specifically relate to the inferences drawn and how, left unaddressed, they can implicate individuals.

2.4.1. An overview of specific challenges that accompany AI

As defined above, any AI system is based on algorithms and these algorithms work with massive amounts of data.[94] Data can come from various sources and the performance of an algorithm is impacted by the quality of the provided data, as any algorithm is only as good as the data it works with.[95] The data an algorithm is trained with and also the data serving as input are influencing the performance, the outputs and the

---

[90] Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018) 1.
[91] Ibid 1.
[92] European Commission, 'Artificial Intelligence – A European Perspective' [2018] Joint Research Center EUR 29425 EN,< https://publications.jrc.ec.europa.eu/repository/bitstream/JRC113826/ai-flagship-report-online.pdf>, page 16.
[93] Borgesius, F. 'Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence' [2020] The International Journal of Human Rights, < https://ssrn.com/abstract=3561441>, page 2.
[94] See above under 2.2.1. Artificial Intelligence (AI) in general.
[95] European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 351; Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 680; Roger, A. and Price, W. 'Privacy and Accountability in Black-Box Medicine' [2016] 23 Mich. Telecomm & Tech. L. Rev. 1 (2016) https://ssrn.com/abstract=2758121, pages, 3, 9; Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 4.

inferences drawn tremendously.[96] In computer science, legal and social sciences scholarship this is being referred to as the "garbage in, garbage out" principle.[97] Training data can be of low quality, for instance, when it includes bias, inaccurate data or when it takes into account discriminatory factors.[98] Training data can be biased if the dataset is not balanced or inclusive enough in order to produce neutral outcomes, as it cannot generalize with such a dataset, also referred to as uncertainty bias[99].[100] This can result in the outcomes, decisions or predictions, to be in favor of individuals or groups over others.[101] Simply put, you provide the AI with biased input, the AI provides you with biased output.[102] The same applies for when the input data is inaccurate, you won't receive accurate results by the AI.[103]

The quality of the training data is not the only challenge. As humans are the developers of the algorithms in AI, they decide the factors and values and the weight

---

[96] Independent High-Level Expert Group On Artificial Intelligence, 'A Definition of AI: Main capabilities and disciplines' [2019] Set up by the European Commission https://ec.europa.eu/futurium/en/ai-alliance-consultation, page 5; Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 684; Danks, D. and London, A. 'Algorithmic Bias in Autonomous Systems' [2017] Twenty-Sixth International Joint Conference on Artificial Intelligence <DOI: 10.24963/ijcai.2017/654>, page 2.

[97] Mittelstadt, B. et al. 'The Ethics of Algorithms: Mapping the Debate' [2017] 3(2) Big Data & Society <https://ssrn.com/abstract=2909885>, page 5; Hand, D. et al. Principles of Data Mining (A Bradford Book The MIT Press 2001) 32; Raso, F. et al. 'Artificial Intelligence & Human Rights: Opportunities & Risks' [2018] Bergmann Klein Center Research Publication No 2018-6 https://ssrn.com/abstract=3259344, page 15; Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 683.

[98] Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 684; Danks, D. and London, A. 'Algorithmic Bias in Autonomous Systems' [2017] Twenty-Sixth International Joint Conference on Artificial Intelligence <DOI: 10.24963/ijcai.2017/654>, page 2; Roger, A. and Price, W. 'Privacy and Accountability in Black-Box Medicine' [2016] 23 Mich. Telecomm & Tech. L. Rev. 1 (2016) https://ssrn.com/abstract=2758121, page 14.

[99] Goodman, B. and Flaxman, S. 'European Union regulations on algorithmic decision-making and a "right to explanation"' [2017] 38(3) AI Magazine, < https://arxiv.org/ct?url=https%3A%2F%2Fdx.doi.org%2F10.1609%2Faimag.v38i3.2741&v=f2c797e9>, page 54.

[100] Independent High-Level Expert Group On Artificial Intelligence, 'A Definition of AI: Main capabilities and disciplines' [2019] Set up by the European Commission https://ec.europa.eu/futurium/en/ai-alliance-consultation, page 5.

[101] Ibid 5.

[102] Danks, D. and London, A. 'Algorithmic Bias in Autonomous Systems' [2017] Twenty-Sixth International Joint Conference on Artificial Intelligence <DOI: 10.24963/ijcai.2017/654>, page 4692.

[103] See Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 681; Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 4.

ascribed to each of them, which the algorithms consider when creating the outputs.[104] Due to the fact that through this practice the biases and values of the developers can be embedded into each step of the AI development process, a big concern is that AI systems will reflect existing bias and discriminatory rules that persist in society in their outcomes.[105]

Assessing the training data, to detect discriminatory factors, bias or inaccuracies, evaluating the values set for each factor to be taken into account and tracing the inner process, starting with the AI receiving a task to be performed, its drawing upon the training to perform its task and to finally produce an output, are possible solutions to prevent biased or inaccurate outcomes.[106] Tracing the inner process of the AI system however has proven to be a challenge. The notion of "black-box"[107] has evolved in scholarship in regard to AI, referring to the challenge of understanding and following the reasons for a certain decision made by AI.[108] This is due to opaque algorithms which transform the input data into output data, without humans being able to explain or understand the inner process of how an outcome was developed.[109] For instance, we are provided with what the best option for a patient might be in terms of treating a disease,

---

[104] Citron, D. and Pasquale, F. 'The Scored Society: Due Process for Automated Predictions' [2014] 89 Washington Law Review, p 1- U of Maryland Legal Studies Research Paper No 2014-8 https://ssrn.com/abstract=2376209, page 5.

[105] Citron, D. and Pasquale, F. 'The Scored Society: Due Process for Automated Predictions' [2014] 89 Washington Law Review, p 1- U of Maryland Legal Studies Research Paper No 2014-8 https://ssrn.com/abstract=2376209, page 14; European Commission, 'Artificial Intelligence – A European Perspective' [2018] Joint Research Center EUR 29425 EN,< https://publications.jrc.ec.europa.eu/repository/bitstream/JRC113826/ai-flagship-report-online.pdf>, page 23; Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 674; Raso, F. et al. 'Artificial Intelligence & Human Rights: Opportunities & Risks' [2018] Bergmann Klein Center Research Publication No 2018-6 https://ssrn.com/abstract=3259344, page 5.

[106] European Commission, 'Artificial Intelligence – A European Perspective' [2018] Joint Research Center EUR 29425 EN,< https://publications.jrc.ec.europa.eu/repository/bitstream/JRC113826/ai-flagship-report-online.pdf>, page 23; Danks, D. and London, A. 'Algorithmic Bias in Autonomous Systems' [2017] Twenty-Sixth International Joint Conference on Artificial Intelligence <DOI: 10.24963/ijcai.2017/654>, page 4691.

[107] Mittelstadt, B. et al. 'The Ethics of Algorithms: Mapping the Debate' [2017] 3(2) Big Data & Society <https://ssrn.com/abstract=2909885>, page 6; Citron, D. and Pasquale, F. 'The Scored Society: Due Process for Automated Predictions' [2014] 89 Washington Law Review, p 1- U of Maryland Legal Studies Research Paper No 2014-8 https://ssrn.com/abstract=2376209, page 6.

[108] Independent High-Level Expert Group On Artificial Intelligence, 'A Definition of AI: Main capabilities and disciplines' [2019] Set up by the European Commission https://ec.europa.eu/futurium/en/ai-alliance-consultation, page 5.

[109] Mittelstadt, B. et al. 'The Ethics of Algorithms: Mapping the Debate' [2017] 3(2) Big Data & Society <https://ssrn.com/abstract=2909885>, page 3; Independent High-Level Expert Group On Artificial Intelligence, 'A Definition of AI: Main capabilities and disciplines' [2019] Set up by the European Commission https://ec.europa.eu/futurium/en/ai-alliance-consultation, page 5.

but what stays hidden and without an explanation is how specific patient characteristics relate to the outcome.[110] Though not all ML algorithms can be referred to as black-boxes, many of them nowadays do display black-box characteristics.[111]

2.4.2. Assessing these challenges in the case of drawing inferences

Depending on the task the AI aims to fulfil, inferences drawn can either be the necessary step in between of providing the AI with data and it coming to an output in the form of a prediction or decision, for instance. Or the inference drawn can constitute the output of the AI, for instance, when AI is supplementing human decision making, leaving, making the actual decision based on the inference drawn, to humans (human in the loop[112]).

Anrig et al. raise the concern that in black-box algorithms, especially those using neural networks, the information learned from the input data stays hidden and it can therefore not function as evidence for a certain output delivered by AI.[113] Taken the scenario that the inference drawn is not the end result of the process performed by the AI, but rather the in between step of the decision-making process, this can mean that the inference drawn stays hidden, as it is inside the black box. This is a concern in evaluating whether the output is inaccurate or biased. It is reasonable to assume that it is possible to evaluate the training data sets and the input data, whether they are discriminating against certain groups or individuals, whether there is bias eminent or whether the data is plain inaccurate. Furthermore, the process of developing the AI and its algorithms can be revised to gain insights, whether accurate weights have been assigned to certain factors and whether there are any values embedded which may lead to an inaccurate performance of the AI. However, if the AI works as a black box, the drawn inference, which constitutes simply said, new information which the AI learned from existing data, cannot be checked for inaccuracies and it cannot be checked for whether it portrays the right values. Without

---

110 Roger, A. and Price, W. 'Privacy and Accountability in Black-Box Medicine' [2016] 23 Mich. Telecomm & Tech. L. Rev. 1 (2016) https://ssrn.com/abstract=2758121, page 7; Danks, D. and London, A. 'Algorithmic Bias in Autonomous Systems' [2017] Twenty-Sixth International Joint Conference on Artificial Intelligence <DOI: 10.24963/ijcai.2017/654>, page 1.

111 European Commission, 'Artificial Intelligence – A European Perspective' [2018] Joint Research Center EUR 29425 EN,< https://publications.jrc.ec.europa.eu/repository/bitstream/JRC113826/ai-flagship-report-online.pdf>, page 23.

112 Danks, D. and London, A. 'Algorithmic Bias in Autonomous Systems' [2017] Twenty-Sixth International Joint Conference on Artificial Intelligence <DOI: 10.24963/ijcai.2017/654>, page 4691.

113 Anrig, B. et al. 'The Role of Algorithms in Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010), 66, 78.

being able to revise the inferences drawn by AI on which its outcomes are based on, the outcomes cannot be proven to be entirely accurate, non-discriminating and without bias.

The inability to check drawn inferences for their accuracy is not just an issue which can lead to one final output, for instance a decision on credibility, being inaccurate or unreliable. According to Custers, profiles constitute inferences and so he portrays another issue by explaining that "profiles are not only an end result or an end product but can also be reused as ingredients for further data analytics"[114]. As inferences drawn, either as the in between step or as the final output, constitute new data on an individual,[115] this data can be fed back into the AI for making other future decisions regarding the individual.[116] In the end, this can result to not only one AI outcome being negatively impacted by inaccurate inferences drawn, but all future decisions based on that inference or which take the inferred data into account, can be negatively impacted.[117] "In this way, profiling processes may function as amplifiers, amplifying bias and inaccuracies via positive feedback loops, that further entrench consequences for data subjects."[118] The creation of this new data, the inferences drawn, wouldn't have existed without the technology, and the above has shown that they can harm individuals, by paving the way for discrimination, by perpetuating bias or by creating inaccurate data about individuals.[119]

---

[114] Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018) 1.
[115] See above under 2.3.1. Categorizing inferred data.
[116] See Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018).
[117] See also Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 681; Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 4.
[118] Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018) 1.
[119] Privacy International, 'Data is Power: Profiling and Automated Decision-Making in GDPR' [2017] <https://privacyinternational.org/sites/default/files/2018-04/Data%20Is%20Power-Profiling%20and%20Automated%20Decision-Making%20in%20GDPR.pdf>, page 7; Roger, A. and Price, W. 'Privacy and Accountability in Black-Box Medicine' [2016] 23 Mich. Telecomm & Tech. L. Rev. 1 (2016) https://ssrn.com/abstract=2758121, page 26.

2.4.3. Left unaddressed: Possible implications for individuals

Will AI drawn inferences exacerbate discrimination in society and of already discriminated individuals and groups? Will prejudices persist and accompany individuals throughout their life due to AI drawn inferences? What are the implications of individuals getting judged based on inferences drawn by AI? The concern is that individuals will be put into categories that are for one inaccurate and secondly, hard to break out of. This can lead to individuals being discriminated and social inequality being amplified, by categorizing individuals based on inferences, that cannot be evaluated for accuracy or are not subject to checks for up-to-datedness.[120] Individuals might therefore get judged, not based on what "they've done, or will do in the future, but because inferences or correlations drawn by algorithms suggest they may behave in ways that make them poor credit or insurance risks"[121].

Inferences are essentially being drawn about individuals in order to use the newly gained inferred data to base decisions or predictions on them.[122] However, the above mentioned concerns in regard to accuracy, bias and discrimination leave a chilling effect, when taking into account that inferences drawn can have quite the harmful potential for individuals.[123] Kamarinou et al. demonstrate the concern by making the example of an individual belonging to a specific credit risk group due to their residence within a certain zip code area.[124] Inferences in this specific case are not only based on the individuals own data but also on characteristics shared with other members of the "specific credit risk group".[125] Van der Sloot and Borgesius also take up on this example to illustrate that data accuracy constitutes a big concern in terms of profiling, while making the example that

---

120 See also: Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018) 1.
121 Ramirez, E. 'Privacy Challenges in the Era of Big Data: A View from the Lifeguard's Chair 3' https://www.ftc.gov/sites/default/files/documents/public_statements/privacy-challenges-big-data-view-lifeguard's-chair/130819bigdataaspen.pdf.
122 See above under 2.3. Drawing inferences.
123 See also Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 27; Privacy International, 'Data is Power: Profiling and Automated Decision-Making in GDPR' [2017] <https://privacyinternational.org/sites/default/files/2018-04/Data%20Is%20Power-Profiling%20and%20Automated%20Decision-Making%20in%20GDPR.pdf>, page 7.
124 Kamarinou, D. et al. 'Machine Learning with Personal Data' [2016] 247 Queen Mary University of London, School of Law, Legal Studies Research Paper, < https://ssrn.com/abstract=2865811>, pages 10, 11.
125 Ibid 10, 11.

"not everybody who lives in a poor town is a credit risk"[126]. Numerous examples exist where the use of AI has provided biased and discriminating outcomes because it was trained with an unbalanced dataset, for instance, women being discriminated in employment decisions performed by AI because of their gender[127], black men being discriminated by associating them to gangs because of their skin color, age and gender[128], individuals with disabilities being discriminated by not showing them housing-related advertisements[129]. Inferences drawn by AI can lead to discrimination in various scenarios and this might not just stay a single occurrence, but if those inferences get fed back into the analytical process of the AI, every single outcome in regard to the individual can be discriminating.

2.5. Conclusion

While AI promises great benefits the challenges that are accompanied by it drawing inferences about individuals cannot be overlooked. Inferred data has one of its characteristics that individuals in general are not aware of its creation or its use for decision or predictions made about them. As this alone is already worrisome, the possible perpetuation of bias, the possible reliance on discriminatory factors and its possible inaccuracy plays into it as well. The implications for individuals raised by these issues accompanying inferences, call for protection of affected individuals. This chapter aimed to pave the way for the next steps of applying the GDPR to inferences drawn by AI, that is by evaluating whether inferences drawn can constitute personal data under the GDPR, and if they can, how the legal framework of the GDPR may or may not be sufficient in order to address the demonstrated implications. To achieve this, this chapter has not only illustrated what is understood by inferences drawn or inferred data, but also how they are created, how they might be used and the various concerns that arise.

---

126 Sloot, Bart van der and Borgesius, Frederik 'Google and Personal Data Protection' [2012] 22 n A. Lopez-Tarruella (Ed.), Google and the Law. Empirical Approaches to Legal Aspects of Knowledge-Economy Business Models. Series: Information Technology and Law Series, Vol. 22 VIII, T.M.C. Asser Press (Springer 2012), < https://ssrn.com/abstract=2146968>, page 16.

127 Barocas, S. and Selbst, A. 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>, page 682.

128 See Amnesty International UK, 'Trapped in the Gangs Matrix' [2018] < https://www.amnesty.org.uk/trapped-gangs-matrix>, last accessed 15.05.2020.

129 See Marks, M. 'Algorithmic Disability Discrimination' [2019] I. Glenn Cohen et al., Eds., Title TBD Cambridge University Press, Forthcoming, < https://ssrn.com/abstract=3338209>, page 3.

3. Establishing the applicability of the GDPR to inferences drawn by AI

3.1. Introduction

The previous showed that there are various implications raised for individuals and society by the drawing of inferences by AI. This chapter will therefore evaluate whether the material scope of the GDPR is fit to cover inferences drawn in order to assess in chapter four whether the provisions of the GDPR are able to address the raised implications sufficiently.

The material scope of the GDPR is triggered according to Art. 2(1) GDPR by "the processing of personal data". Having portrayed how inferences get created, what they are defined as and how the AI technology behind it works, this chapter focuses on subsuming inferences drawn under the GDPR. This chapter starts off by defining the four elements of the definition of 'personal data' under Art. 4(1) GDPR, by going into the relevant case law of the CJEU, the opinions of the Art. 29 WP and legal scholarship, in order to subsequently subsume inferences drawn under this definition. Following this analysis, it will be assessed whether inferences drawn can constitute a special category of personal data under Art. 9(1) GDPR

3.2. The material scope of the GDPR: Processing of personal data

For the drawing of inferences to fall within the scope of the GDPR, the inferences have to constitute personal data which are being processed, Art. 2(1) GDPR. The triggering definitions of 'personal data' and 'processing' are contained in Art. 4(1), (2) GDPR.[130]

Processing is defined under Art. 4(2) GDPR as

> *"any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaption or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction;"*[131].

[130] Hoofnagle, C. et al. 'The European Union general data protection regulation: what it is and what it means' [2019] 28(1) Information & Communications Technology Law, https://doi.org/10.1080/13600834.2019.1573501, page 72.
[131] Article 4(2) GDPR.

A wide range of processing operations are covered under the GDPR, which generally "includes any including any AI/ machine learning operation performed on personal data"[132].[133]

What constitutes 'personal data' under the GDPR is legally defined in Art. 4(1) GDPR,

> *"(1) 'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;"*[134].

The GDPR replacing the Data Protection Directive (DPD) has not affected the definition of personal data, as Art. 4(1) GDPR is similar to the former Art. 2(a) DPD.[135] Thus, the published materials by the Article 29 Working Party (Art. 29 WP) in regard to the definition of 'personal data' in the former DPD and the case law of the Court of Justice of the European Union (CJEU) remain to be relevant also for the GDPR's definition of 'personal data' in Art. 4(1).[136]

---

[132] Mitrou, L. 'Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) "Artificial Intelligence-Proof"?' [2019] University of the Aegean Dpt. of Information and Communication Systems Engineering; Athens University of Economics and Business - Department of Informatics < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914>, page 31.

[133] Edwards, L. 'Data Protection: Enter the General Data Protection Regulation' [2018] Forthcoming in L Edwards ed Law, Policy and the Internet (Hart Publishing, 2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3182454>, page 8; Hoofnagle, C. et al. 'The European Union general data protection regulation: what it is and what it means' [2019] 28(1) Information & Communications Technology Law, https://doi.org/10.1080/13600834.2019.1573501, page 72.

[134] Article 4(1) GDPR.

[135] Opinion of Advocate General Kokott, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:582, par. 3

[136] With the implementation of the GDPR on May 25, 2018, replacing the Data Protection Directive (DPD), the Art. 29 WP has ceased to exist and has been replaced by the European Data Protection Board (EDPB). The former Art. 29 WP published an Opinion on the concept of personal data in 2007 in regard to the definition of 'personal data' as it was defined under the DPD, which is still the key source for interpreting the definition of 'personal data' also under the GDPR, see: Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Though it is reasonable to believe that the EDPB will publish additional materials in the future that regard the definition of 'personal data' which may be relevant for the following analysis.

Four cumulative "building blocks"[137] are distinguished in the definition which establish whether data constitutes "personal data", (1) *any information*, (2) *relating to*, (3) *identified or identifiable* and (4) *natural person*, each of which will be looked at separately when subsuming inferences drawn under the definition.[138] If data is unable to fulfil one of the four elements, it is not personal.[139]

The CJEU has mentioned on several occasions that personal data covered by the DPD (a statement also applicable to the GDPR) is varied.[140] As the Advocate General in *Google Spain* put it: "The concept of personal data is given a wide definition in the Directive, this wide definition has been applied by the Article 29 Working Party and it has been confirmed by the Court."[141]

3.2.1. "Any information"

The concept of information is constructed broadly under the GDPR.[142] The Art. 29 WP indicates that this phrasing illustrates the legislator's intention to "design a broad concept of personal data"[143], which would lead to a wide interpretation. Mentioned examples state that this element includes both "the presence of a certain substance in one's blood" [144] as objective information and also the fact that someone is not expected to die soon for insurance purposes, as subjective information.[145] Objective information describes information which can be subject to revision, as they have a factual basis.[146]

---

137 Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 6.
138 Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology, page 3.
139 Ibid 3.
140 CJEU, Case C-553/07, College van burgemeester en wethouders van Rotterdam v. M.E.E. Rijkeboer [2009] ECLI:EU:C:2009:293, par. 59; CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, par. 33.
141 AG Opinion, Case C-131/12, Google Spain SL, Google Inc. v. Agencia Espanola de Protección de Datos (AEPD), Mario Costeja González, [2013] ECLI:EU:C:2013:424, par. 71.
142 Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology, page 3.
143 Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 6.
144 Ibid 6.
145 Ibid 6.
146 See also, Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 57.

Subjective information, on the other hand, in general does not have a factual basis, but rather constitutes assumptions, opinions or assessments.[147]

Furthermore, the Art. 29 WP is of the opinion that for information to be considered "personal data" it does not need to be true or proven.[148] In regard to the content of the information, the word "any" implies that the information can be of any sort of information.[149] Information can be contained in any form, be it numerical or graphical or stored in a computer memory by means of binary code, for instance.[150]

The CJEU as well adopted a broad interpretation of 'any information' and includes both objective and subjective information.[151] This becomes clear in *Nowak* where the CJEU concluded, "The use of the expression 'any information' in the definition of the concept of 'personal data', within Article 2(a) of Directive 95/46, reflects the aim of the EU legislator to assign a wide scope to that concept, which is not restricted to information that is sensitive or private, but potentially encompasses all kinds of information, not only objective but also subjective, in the form of opinions and assessments, provided that it 'relates' to the data subject."[152].

While it is obvious, that the term 'any information' shall include a broad range of things, neither the CJEU nor the Art. 29 WP actually define what 'information' is. This

147 See, CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, par. 34; Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 6. See also, Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 57.

148 Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 6.

149 Ibid 6.

150 Ibid 7.

151 CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, par. 34.

152 Ibid [34] (Nowak).

has led to various scholars concentrating on what information is under data protection law.[153][154]

3.2.2. "Relating to"

"Relating to" describes the necessary relationship or link between the information and the natural person.[155] The way in which information can relate to a certain natural person are threefold: Information can relate to a person by virtue of its content, its processing purpose or of its result.[156] The content element is described as information given about a certain person, for instance the results of an medical analysis are about a certain patient.[157] Information is linked to a natural person, when information is used with the purpose "to evaluate, treat in a certain way or influence the status or behavior of an individual"[158]. The result element leads to information being linked to a person in cases where the use of information is "likely to have an impact on a certain person's rights and interests"[159]. It shall be sufficient that as a result of the processing of such information, the person is treated differently than others, which does not require "that the potential result be a major impact"[160].[161] The three ways of linking information to a natural person are alternative.[162] While the GDPR does not specifically propose the threefold approach

---

[153] Gellert on the notions of data and information, "Whereas data protection is based on the vernacular idea of information as the communication of knowledge, machine learning is predicated on a different premise, namely the creation of knowledge. This leads to different concepts of data and information. (…) These definitional differences lead to conclusions on the GDPR's potential for algorithmic regulation. (…), it might simply be inadequate in order to regulate a technology which is based upon a fundamentally different logic: the production of knowledge." Gellert, R. 'Data Protection and notions of information: a conceptual exploration' [2018] Working Paper, last updated 06.11.2018 < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3284493>, pages 2, 3.

[154] See also, Bygrave, L. 'Information Concepts in Law: Generic Dreams and Definitional Daylight' [2015] 35(1) Oxford Journal of Legal Studies (2015) <https://doi.org/10.1093/ojls/gqu011>.

[155] Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology, page 3.

[156] Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 2; Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, pages 10, 11.

[157] Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 10.

[158] Ibid 10.

[159] Ibid 11.

[160] Ibid 11.

[161] Ibid 11.

[162] Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 3; Article 29 Data Protection Working Party, Opinion 4/2007

to creating the link between information and a natural person, this approach has also been taken by the CJEU in its *Nowak*[163] judgment.[164]

Following the threefold approach to establish the element of 'relating to' for information, both the Art. 29 WP and the CJEU have chosen a broad interpretation of the element. According to Dalla Corte, the many links of how information can relate to a natural person entails the potential that, "depending on each individual processing instance, all data can thus potentially become personal"[165]. The claim that the definition of personal data under the GDPR is becoming increasingly broader, encompassing ever more data as personal data, is not knew and has been subject to discussion in legal scholarship.[166]

3.2.3. "Identified and identifiable"

In order to trigger the applicability of the GDPR to information relating to a natural person, that person needs to be identified or identifiable. The Art. 29 WP explains the meaning as follows: "In general terms, a natural person can be considered as "identified" when, within a group of persons, he or she is "distinguished" from all other members of the group."[167] If it is possible to distinguish the person from all members of the group (singling out), but this has not occurred yet, then the person is identifiable, meaning for this element to be fulfilled, the person doesn't need to be identified already, as long as the possibility of doing so exists.[168] Identification of a person can be

---

on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, pages 10, 11.

163 CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, par. 35.

164 Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 4.

165 Ibid 2.

166 See Purtova, N. 'The law of everything. Broad concept of personal data and future of EU data protection law' [2018] 10(1) Law, Innovation and Technology (2018) < https://ssrn.com/abstract=3036355>; Graef, I. et al. 'Feedback to the Commission's Proposal on a framework for the free flow of non-personal data' [2018] < https://ssrn.com/abstract=3106791>, page 2.

167 Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 12.

168 Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 4; Borgesius, F. 'Case Note: Breyer Case of the Court of Justice of the European Union: IP addresses and the personal data definition' [2017] 3(1) European Data Protection Law Review (2017), < https://ssrn.com/abstract=2933781>, page 15.

accomplished through so-called identifiers, pieces of information, that link to a certain person, either directly (e.g. a name) or indirectly (e.g. an IP address).[169] Whether the information that relates to a person is able to identify that exact individual, is something that needs to be considered separately for each individual case, based on the circumstances.[170] It is not required that the name of a person alone, or just the social security number, are able to identify a certain person, rather whether these pieces of information, combined with other information, can identify a certain person.[171]

In cases, where the purpose of the processing implies the identification of an individual, means likely reasonable to be used for identification of the person can be assumed to exist.[172] For instance, data processing's in medical diagnosis or treatment, the patient necessarily needs to be identified.

In terms of identifiability, the CJEU adopted a broad approach to the identifiability element in *Breyer*.[173] As the definition of 'personal data' under the GDPR requires the identifiability directly or indirectly, the CJEU states in *Breyer* that the legislators intent with the term 'indirectly' was, that it is not required that the information alone allows for the person to be identified.[174] Meaning, additional information can be used in order to identify an individual.[175] According to Recital 26 GDPR "all the means reasonably likely

---

Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 12.

169 Art. 4(1) GDPR; Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 3; Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, pages 12, 13.

170 Graef, I. et al. 'Feedback to the Commission's Proposal on a framework for the free flow of non-personal data' [2018] < https://ssrn.com/abstract=3106791>, page 2; Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 13.

171 Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 4; Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 13.

172 Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 16.

173 CJEU, Case C-582/14, Patrick Breyer v. Bundesrepublik Deutschland [2016] ECLI:EU:C:2016:779.

174 Ibid [41] (Breyer).

175 Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 2.

to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly"[176] need to be taken into account. In light of Recital 26 of the DPD (similarly phrased in Recital 26 of the GDPR), the Court concludes that in light of "means likely reasonably" to identify a person, it is not necessary "that all the information enabling the identification of the data subject must be in the hands of one person"[177], but may be spread among several persons.[178] Similar to the examples made by the Art. 29 WP, according to the CJEU, an indirect identifier, can be an assigned identification number of the individual,[179] and a direct identifier is the name of a person, as several times confirmed by the Court.[180]

3.2.4. "Natural person"

The fourth element, that of "natural person", is simply analyzed as affording protection under the GDPR only to human beings. Personal data is information relating to living individuals.[181] This element is left to the Member States to determine, as in general it does not raise any particular problems.[182]

3.3. Inferences drawn within the material scope of the GDPR

For the material scope of the GDPR to be triggered, drawn inferences must constitute personal data which are processed under the GDPR. All four of the beforementioned definitional elements of personal data have to be fulfilled.

3.3.1. Inferences drawn as 'any information'

As 'information' is neither defined by the Art. 29 WP nor by CJEU case law, it is reasonable to assume that 'any information' shall "potentially encompass all kinds of

---

176 Recital 26 GDPR.
177 CJEU, Case C-582/14, Patrick Breyer v. Bundesrepublik Deutschland [2016] ECLI:EU:C:2016:779, par. 43.
178 See also, AG Opinion, Case C-582/14, Patrick Breyer v. Bundesrepublik Deutschland [2016] ECLI:EU:C:2016:339, par. 68.
179 CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, par. 29; Opinion of Advocate General Kokott, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:582, par. 28.
180 CJEU, Case C-524/06, Heinz Huber v. Bundesrepublik Deutschland [2008] ECLI:EU:C:2008:724, par. 43; CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, par. 29.
181 Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 22.
182 Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 3.

information, not only objective but also subject, in the form of opinions and assessments"[183].

As stated above it is hard to prove or test whether inferences drawn by AI are accurate and not laced with bias, due to algorithms often working as a black box. This though does not seem to hinder inferences being able to fall under the element of "any information", as the working party says that "data protection rules already envisage the possibility that information is incorrect"[184]. This line of reasoning also complies with the view of the CJEU in *Nowak,* as assessments fall within the scope of 'any information' and inference is just the result of such an assessment of data. Hallinan and Borgesius refer to inferences as opinions[185], which the CJEU explicitly named as being able to comply with the element of 'information[186].[187]

3.3.2. Inferences as 'relating to'

To establish whether there is a link between the inference, the information, and a natural person, it is important to recall that inferences can either be drawn as part of an analytical process, a necessary step to take, in order to get a result, such as a decision or prediction about a person, or the drawn inference can constitute the desired result itself.[188]

3.3.2.1. Inferences and the threefold approach of 'relating to'

According to Wachter et al. the link of information to a natural person by virtue of its result, "is key to the legal status of inferences"[189] as personal data. Arguing that inferences drawn relate to a natural person by virtue of its result is the most obvious one of the three, as the drawing of inferences by AI has been put in the spotlight because of

---

183 CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, par. 46.

184 Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 6.

185 Hallinan, D. and Borgesius, F. 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' [2020] 10(1) International Data Privacy Law (2020) https://doi.org/10.1093/idpl/ipz025, page 2.

186 CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, par. 46 and following.

187 Also, Korff enumerates 'opinions', intentions and predictions to be included under 'any information', Korff, D. 'Data Protection laws in the EU: The difficulties in meeting the challenges posed by global social and technical developments' [2010] Working Paper 2, European Commission Directorate-General Justice, Freedom and Security, 20 January 2010 < https://dx.doi.org/10.2139/ssrn.1638949>, page 41.

188 See above under 2.3.1. Categorizing inferred data.

189 Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 25.

its impacts it can have on individuals.[190] However, Wachter et. al. make it seem that inferences will most likely relate to a natural person by virtue of its result. Inferences though have the ability to relate to a natural person also by virtue of its content or its purpose. Referring to the example made by the Art. 29 WP, that the result of a medical analysis is clearly about a natural person by virtue of its content, inferences drawn by AI in the medical field, which infer, for instance, a certain health status of a natural person constitutes information about that person.[191] According to the Art. 29 WP the information can be linked to a person by virtue of its purpose, for instance when it is used to treat the person in a certain way or influences the status of the person.[192] When applying the three alternative elements the Art. 29 WP and the CJEU in *Nowak*[193] propose to determine whether information relates to a person, it becomes clear that inferences can fulfil the element of relating to a person. Which of the three elements establishes the link, between the inference and a natural person, has to be evaluated for each individual separately.[194]

3.3.2.2. Inferences as part of legal analysis

In *YS and others* the CJEU had to conclude whether a legal analysis can constitute personal data within data protection law.[195] It ruled in accordance with the AG's opinion, that while the legal analysis "may contain personal data, it does not in itself constitute such data within the meaning of Art. 2(a) of Directive 05/46"[196]. How does this case relate to the status of inferences as personal data in terms of 'relating to'?

According to Wachter et. al., "a legal analysis is comparable to an analysis of personal data where new data is derived or inferred"[197].

---

190 See above under 2.4.3. Left unaddressed: Possible implications for individuals.
191 See for the example, Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, pages 9, 10.
192 Ibid 10.
193 CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, 40.
194 See also, Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 11.
195 CJEU, Joined Cases C-141/12 and C-372/12, YS v. Minister voor Immigratie, Integratie en Asiel, and Minister voor Immigratie, Integrie en Asiel v. M, S [2014] ECLI:EU:C:2014:2081.
196 Ibid [39] (YS and others).
197 Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 31.

Applying this argument to the AG's opinion, the following would present itself. The AG uses an example of a person's weight which can be expressed objectively in kilos or subjectively in the terms obese or underweight, to distinguish between what personal data is and what the legal analysis constitutes.[198] Both forms of expressing a person's weight shall constitute personal data, while "the steps of reasoning by the which the conclusion is reached that a person is 'underweight' or 'obese' are not facts, any more than legal analysis is"[199]. The inference drawn from an analysis of data, as its reached conclusion or output, can according to this interpretation constitute personal data. For instance, when it is inferred that due to all the data the AI has on one person, it infers that the person is obese. According to the threefold approach, the link between the information and the person would be created here by virtue of its content. Clearly, when the inference is a result from an analysis it can constitute personal data. However, in regard to the inferences that are only drawn as part of the process of the analysis and that don't constitute the result of the analysis, the AG and CJEU in *Ys and others* are not so clear on the status. Wachter et al. interpret this as meaning that inferences, that are only drawn as part of the analysis process, don't constitute personal data.[200]

This interpretation seems to narrow. The inference cannot be compared to the legal analysis itself. "For example, a person's address is personal data but an analysis of his domicile for legal purposes is not."[201] Instead of interpreting this example as narrow as Wachter did, meaning inferences drawn in the process of analyzing where a person lives and thereby excluding inferences from constituting personal data, the example can be interpreted as meaning, that the analysis by which inferences are created does not constitute personal data, but inferences that are drawn in the process can constitute 'facts' about a person. Simply said, this would mean that the process of searching through a database and trying to find patterns or correlations to link certain data pieces together, is a process which is not protected under data protection law. However, if this process infers

198 Opinion of Advocate General Sharpston, Joined Cases C-141/12 and C-372/12, YS v. Minister voor Immigratie, Integratie en Asiel, and Minister voor Immigratie, Integrie en Asiel v. M, S [2013] ECLI:EU:C:2013:838, par. 57.
199 Ibid [58] (AG, YS and others).
200 Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 32.
201 Opinion of Advocate General Sharpston, Joined Cases C-141/12 and C-372/12, YS v. Minister voor Immigratie, Integratie en Asiel, and Minister voor Immigratie, Integrie en Asiel v. M, S [2013] ECLI:EU:C:2013:838, par. 56.

data about a person, e.g. his address or him being underweight, this data might as well constitute personal data if it fulfils all four elements. The AG also explicitly states that "the possibility that assessments and opinions may sometimes fall to be classified as data"[202] is not excluded and therefore the opinion should not be interpreted to narrow in regard to inferences.

Furthermore, in the later *Nowak* case the CJEU lists "assessments" as being encompassed by the expression 'any information', as long as it also 'relates to' a person.[203] The CJEU is satisfied with the criterion of 'relate to', when the information is linked to an individual "by reason of its content, purpose or effect"[204]. The CJEU in *Nowak* is in line with how the Art. 29 WP defines the element and inferences can relate to a natural person. The CJEU in *Nowak* therefore "effectively reversed"[205] the narrow interpretation of "relating to" adopted in *YS and others*.

3.3.3. Inferences and the element of 'identified/identifiable'

The drawn inference is required to identify a certain person or at least be able to identify a person, meaning the inference about a person has to be able to distinguish this person from members of the same group, taking all means reasonably likely used for identification into account.[206] According to the CJEU in *Breyer*, means are considered not likely reasonably to be used if they are illegal or practically impossible.[207]

What is crucial for the element of 'identified or identifiable' is the state of the art of the technology at the time of the processing, the drawing of the inference, in terms of "means likely reasonably to be used"[208] to identify a person, and the identification possibilities to come with the further development of the technology for the time the data

---

[202] Ibid [57] (AG, YS and others).
[203] CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, par. 34.
[204] Ibid [35] (Nowak).
[205] Purtova, N. 'The law of everything. Broad concept of personal data and future of EU data protection law' [2018] 10(1) Law, Innovation and Technology (2018) < https://ssrn.com/abstract=3036355>, page 67.
[206] Recital 26 GDPR.
[207]CJEU, Case C-582/14, Patrick Breyer v. Bundesrepublik Deutschland [2016] ECLI:EU:C:2016:779, par. 46; Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 7.
[208] Recital 26 GDPR.

will be processed.[209] This aims to keep the identifiability test a dynamic one, which takes into account the expected technological developments.[210]

> *"(…) The same dataset may not obviously be personally identifiable at the stage of processing, or from the perspective of the controller, given the tools and data available to him, but become, or appear to have been all along, identifiable from the perspective of another person or once the circumstances change."*[211]

This dynamic characteristic of the personal data definition, combined with its broad interpretation of all its elements, has led to the concern, that everything will become personal data.[212][213]

In the case of drawing inferences on a natural person, especially in cases of being able to predict the behavior of a person or making a decision on the person due to the inferences that have been drawn about him, controllers intentionally and purposefully have to identify the person in order to make use of the inferences drawn. The purpose of drawing inferences is often related to the offering of tailored advertisements, goods and services, medical treatments to a specific person, or calculating credit scores or insurance rates for a specific person, where identification of the natural person is possible and, in many cases, desired. According to the Art. 29 WP,

---

209 Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 15.

210 Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 15; Korff, D. 'Data Protection laws in the EU: The difficulties in meeting the challenges posed by global social and technical developments' [2010] Working Paper 2, European Commission Directorate-General Justice, Freedom and Security, 20 January 2010 < https://dx.doi.org/10.2139/ssrn.1638949>, page 46; Graef, I. et al. 'Feedback to the Commission's Proposal on a framework for the free flow of non-personal data' [2018] < https://ssrn.com/abstract=3106791>, page 2.

211 Purtova, N. 'The law of everything. Broad concept of personal data and future of EU data protection law' [2018] 10(1) Law, Innovation and Technology (2018) < https://ssrn.com/abstract=3036355>, page 47.

212 For further discussion of the topic see: Purtova, N. 'The law of everything. Broad concept of personal data and future of EU data protection law' [2018] 10(1) Law, Innovation and Technology (2018) < https://ssrn.com/abstract=3036355>.

213 Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 1; Schwartz, P. and Solove, D. 'Reconciling Personal Information in the US and EU' [2013] 102 California Law Review 877 (2014); UC Berkeley Public Law Research Paper No. 2271442; GWU Legal Studies Research Paper No. 2013-77; GWU Law School Public Law Research Paper No. 2013-77 https://ssrn.com/abstract=2271442, page 892.

"In these cases, where the purpose of the processing implies the identification of individuals, it can be assumed that the controller or any other person involved have or will have the means "likely reasonably to be used" to identify the data subject."[214]

Whether the identifiability criterion is given, has to be determined, nevertheless, for each case individually, but in general it will be given.

3.3.4. Inferences and the element of 'natural person'

Only inferences drawn relating to living persons can constitute personal data. In the case that inferences are drawn about deceased persons, which are then used to make predictions or decisions impacting living persons, it has to be closely evaluated whether the personal data indeed relates to a living person. For instance, a patient has died of a disease, and the AI evaluates all the patient data and draws inferences, which contain information on how certain patient characteristics related to his death. While this drawn inference relates to a deceased person, it would not constitute personal data, for that individual. However, if this drawn inference is taken and used for another patient, who is still alive, for treatment purposes, who has similar characteristics as the deceased, it can constitute personal data, if the other elements of the definition are fulfilled as well. This portrays, that for every person the assessment of whether a drawn inference constitutes personal data, has to be evaluated independently.[215]

3.4. Inferences drawn as a special category of personal data, Art. 9(1) GDPR

Having established, that inferences can constitute personal data under Art. 4(1) GDPR, this section will assess whether inferences can also constitute a special category of data under Art. 9(1) GDPR. Art. 9(1) GDPR aims to provide certain categories of personal data with extra protection, also referred to as "sensitive data"[216]:

---

[214] Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 16.

[215] Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 2; Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 11.

[216] Article 29 Data Protection Working Party, 'Advice paper on special categories of data ("sensitive data")', 20 April 2011, https://ec.europa.eu/justice/article-29/documentation/other-document/files/2011/2011_04_20_letter_artwp_mme_le_bail_directive_9546ec_annex1_en.pdf., page 4.

"[…] 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 […]".

The Art. 29 WP clearly states that "the term 'revealing' (…) is to be understood that not only data which by its nature contains sensitive information is covered by this provision, but also data from which sensitive information with regard to an individual can be concluded"[217]. Furthermore, in another guidelines the Art. 29 WP also states that special category data can be inferred from profiling activity.[218] As an example the Art. 29 WP states that someone's state of health may be inferred from "the record of their food shopping combined with data in the quality and energy content of foods"[219]. The possibility, that inferences could constitute a special category of personal data under Art. 9(1) GDPR was therefore seen and acknowledged by the Art. 29 WP. Malgieri and Comandé have raised the question in this regard, what "the appropriate degree of sophistication"[220] is, when assessing whether data can be used to, for instance, infer a health status and came to the conclusion that this requires a flexible case-by-case approach.[221] They propose a new category, that of quasi-health data, "which are indirectly related to health and which are nearly as sensitive as health data, (…) and not immediately revealing of health status information", such as data which allows the inference of individuals health conditions.[222] Inferences, which can be in itself a special category of personal data, would not fall under the newly proposed "quasi"-category. In my opinion, the proposed solution does not offer an answer to the issue that we are faced with, that almost any data can be used, e.g. by inference, analysis or combination, to reveal sensitive attributes. Creating a new category of quasi-data, with a different level of protection, just seems to shift the issue rather than answering it, because in the end the same question

---

217 Ibid 6.
218 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 15.
219 Ibid 15.
220 Malgieri, G. and Comandé, G. 'Sensitive-by-distance: quasi-health data in the algorithmic era' [2017] 26(3) Information & Communication Technology Law, https://doi.org/10.1080/13600834.2017.1335468, page 234.
221 Ibid 234.
222 Ibid 236.

remains; how do we then determine what is 'ordinary' personal data, where does 'quasi-data' begin and where does 'special category data' begin? By drawing inferences, meaning creating new data from existing data, the possibility arises, that sensitive data is being inferred from 'ordinary', non-sensitive data.[223] Malgieri and Comandé propose to put this ordinary data in the quasi-sensitive data category, because sensitive personal data may be inferred from it. However, if all data from which possibly sensitive data can be inferred from, is put in a category, we stand where we stand now, just that except for all data to either constitute personal data or ordinary data, it will fall under the category of quasi-sensitive data.

A different discussion currently focuses on the question whether inferred sensitive data should be regulated under the higher protection of Art. 9(1) GDPR, if a controller had not intended to infer sensitive data, for instance, by processing shopping receipts and behavior[224].[225] Social media networks are allegedly already able to infer a person's sexual orientation, political opinions or whether someone is suicidal.[226] Presuming, that drawing these sensitive inferences was not intended, should these inferences nevertheless be subject to the stricter provision of Art. 9(1) GDPR? Edwards[227] and other scholars[228]

---

[223] Mitrou, L. 'Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) "Artificial Intelligence-Proof"?' [2019] University of the Aegean Dpt. of Information and Communication Systems Engineering; Athens University of Economics and Business - Department of Informatics < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914>, page 22; Edwards, L. 'Data Protection: Enter the General Data Protection Regulation' [2018] Forthcoming in L Edwards ed Law, Policy and the Internet (Hart Publishing, 2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3182454>, page 18; Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 73.

[224] Zarsky, T. 'Incompatible: The GDPR in the Age of Big Data' [2017] 47(4) Seton Hall Law Review (2017) < https://ssrn.com/abstract=3022646>, page 1013.

[225] Edwards, L. 'Data Protection: Enter the General Data Protection Regulation' [2018] Forthcoming in L Edwards ed Law, Policy and the Internet (Hart Publishing, 2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3182454>, page 18; Article 29 Working Party, Annex – health data in apps and devices to the Advice paper on special categories of data ("sensitive data"), April 2011, < https://ec.europa.eu/justice/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf>, page 4.

[226] See for more information, Marks, M. 'Artificial Intelligence Based Suicide Prediction' [2019] 18(3) Yale Journal of Health, Policy, Law and Ethics, 21(3) Yale Journal of Law and Technology, https://ssrn.com/abstract=3324874.

[227] Edwards, L. 'Data Protection: Enter the General Data Protection Regulation' [2018] Forthcoming in L Edwards ed Law, Policy and the Internet (Hart Publishing, 2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3182454>, page 19.

[228] See Zarsky, T. 'Incompatible: The GDPR in the Age of Big Data' [2017] 47(4) Seton Hall Law Review (2017) < https://ssrn.com/abstract=3022646>, page 1013.
Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 73.

argue if these sensitive inferences were to fall under Art. 9(1) GDPR, "given the increasing power of ML algorithms, this might lead us to the inexorable conclusion that not only is 'everything personal data' per Purtova but perhaps even, 'everything is special category data'"[229]. According to Zarsky, the distinction between ordinary and special personal data should be abandoned as a whole, as it does not serve any purpose anymore in the time of big data, if all data fall under a special category.[230] "At the end of the day, there will be no real special treatment for these special categories as this stricter standard will be applied across the board."[231] Ultimately this would lead to all data being treated the same, data which is considered to be the most private can be processed just as regularly as other personal data.[232] How this concern is best addressed is not part of this thesis' aim, but what can be taken from this discussion is that, inferences obviously entail the possibility of inferring sensitive data from various ordinary data or sources, which causes for a lot of distress on how efficient the distinction of personal data and special category data in times of big data and AI still proves to be.[233]

3.5. Conclusion

This chapter illustrated the four elements of the definition of 'personal data' under the GDPR, portraying the relevant case law, the opinions of the Art. 29 WP and relevant discussions by scholars. It was found that AI drawn inferences can constitute personal data within the meaning of the GDPR under Art. 4(1). However, even though, any of the four elements of the definition can be fulfilled by drawing inferences, as the definition, with all its elements, is a broad and dynamic one, each inference, when considering its state of being personal data, needs to be considered within the concrete processing

---

229 Edwards, L. 'Data Protection: Enter the General Data Protection Regulation' [2018] Forthcoming in L Edwards ed Law, Policy and the Internet (Hart Publishing, 2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3182454>, page 19.

230 Zarsky, T. 'Incompatible: The GDPR in the Age of Big Data' [2017] 47(4) Seton Hall Law Review (2017) < https://ssrn.com/abstract=3022646>, page 1014; See also: Malgieri, G. and Comandé, G. 'Sensitive-by-distance: quasi-health data in the algorithmic era' [2017] 26(3) Information & Communication Technology Law, https://doi.org/10.1080/13600834.2017.1335468, page 234.

231 Zarsky, T. 'Incompatible: The GDPR in the Age of Big Data' [2017] 47(4) Seton Hall Law Review (2017) < https://ssrn.com/abstract=3022646>, page 1015.

232 Ibid 1015.

233 See also, Article 29 Working Party, Annex – health data in apps and devices to the Advice paper on special categories of data ("sensitive data"), April 2011, < https://ec.europa.eu/justice/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf>, page 3.

instance.[234] After having come to this finding it was furthermore assessed that inferences drawn by AI can also constitute a special category of personal data pursuant to Art. 9(1) GDPR, which provides these inferences with additional protection under the GDPR. In both analyses it was shown that special attention needs to be paid to the types of data being processed, throughout the entire process of the AI drawing its inferences. Even though the input data may constitute neither personal data nor a special category of personal data, the inferences drawn can fall within the scope of the definitions. As the material scope of the GDPR can include the inferences drawn by AI, the next step is to analyze to which extent the provisions of the GDPR actually apply to inferences drawn in order to mitigate the implications they possibly pose for individuals and society.

[234] Dalla Corte, L. 'Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law' [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909>, page 11.

4. The possibility of mitigating the implications of inferences drawn by AI under the GDPR

4.1. Introduction

The previous chapters have for one, shown the implications inferences drawn by AI can pose for individuals, and secondly, that inferences indeed can be considered as personal data under Art. 4(1) GDPR, leading to its applicability under Art. 2(1) GDPR. This chapter aims to assess whether the provisions on data subject rights under the GDPR are sufficient to mitigate the implications posed by inferences. Starting, this chapter will portray the main issues of awareness, inaccuracy and treating inferences under Art. 22 GDPR that arise in regard to applying the provisions of the GDPR to the data category of inferred data, mainly due to its inherent differences to other data categories, as in provided or observed data. For one, this chapter will analyze the information rights of Art. 13 and 14 and the right to access under Art. 15 GDPR and under which it will be addressed to what extent information can help data subjects become aware of the drawing of inferences in the first place, their existence in general and their subsequent use. Following this, the data subject rights of Articles 16-21 GDPR will be assessed in regard to how data subjects can exercise them to gain control and oversight of inferences drawn, especially in the case of inaccurate inferences. At last, Art. 22 GDPR will be discussed, in how it may apply to AI practices where inferences get drawn and whether data subjects gain any protection specifically towards their inferences under this provision.

4.2. Specific issues relating to challenging inferences

As inferred data is not personal data which was provided by the data subject to a controller, but rather new data created by a controller on the basis of other data, also referred to as input data, several challenges arise when applying the data subject rights under the GDPR.[235] The input data can be collected from various sources and can be personal and non-personal data.[236] Some may be data, which was once provided by the data subject, directly or indirectly, either to the controller, who draws the inferences, or to a third party, but the input data may also be data which comes from other sources. This

[235] See Article 29 Working Party, Guidelines on the right to data portability, WP 242 rev.01, 05 April 2017, https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233, page 10; Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 24.

[236] See above under 2.3. Drawing inferences.

leads to the level of awareness of individuals, that inferences are being drawn about them, the existence of inferred data and the subsequent use to make decisions concerning the individual, being relatively low to non-existent.[237] Without awareness, individuals cannot exercise their data subject rights.[238] Awareness has to be achieved before it can be assessed whether the data subject rights, e.g. the right to rectify[239], the right to erasure[240], the right to object[241], can apply to inferred data. Awareness might be achieved through the right of information[242], the right of access[243] and the right of data portability[244], but because inferred data is not provided but rather newly created, there are also challenges in applying these provisions.

If these challenges can be overcome, individuals being aware of inferences and applying the GDPR provisions to inferred data, one issue remains. For instance, Art. 16 GDPR that stipulates the right to rectify, talks about "inaccurate personal data", and Art. 18 GDPR stipulates the right to restrict processing, when "the accuracy of personal data is contested by the data subject". In essence, these provisions and the remaining data subject rights aim to give data subjects control over their data, being able to correct incorrect data, complete incomplete data and rectify inaccurate data. Thus, data subjects need to gain information under Articles 13-15 GDPR to the extent, that they are not just made aware of the existence and use of inferences drawn about them, but also, they need to gain the possibility to ascertain the accuracy of those inferences.

Art. 22 GDPR might be the solution, as it addresses automated decision making and profiling, and under Articles 13(2)(f), 14(2)(g), 15(1)(h) of the GDPR "meaningful information about the logic involved" has to be provided to the data subjects. If this

---

237 Abrams, M. 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, <https://ssrn.com/abstract=2510927>, page 8; Hof, S. v. d. and Prins, C. 'Personalisation and its Influence on Identities, Behavior and Social Value' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 116; Kindt, E. 'Biometric Profiling: Opportunities and Risks' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 143; Moerel, L. and Wolk, A. v. d. 'Big data analytics under the EU General Data Protection Regulation', [2017], https://ssrn.com/abstract=3006570, page 28.
238 European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 349; Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 51.
239 Art. 16 GDPR.
240 Art. 17 GDPR.
241 Art. 21 GDPR.
242 Articles 13, 14 GDPR.
243 Art. 15 GDPR.
244 Art. 20 GDPR.

includes information about the inferences drawn in the process of automated decision-making and profiling will be assessed.

4.3. Raising awareness of inferences among data subjects

"Awareness among and control by individuals are key to ensuring rights enforcement."[245] The principle of transparency in Art. 5(1)(a) GDPR requires personal data to be processed "in a transparent manner in relation to the data subject".[246] It provides the obligation for controllers to keep data subjects informed according to Art. 13 and 14 GDPR about how their personal data is used, to enable individuals to exercise their data subject rights according to Articles 15 to 22 GDPR.[247] The transparency rights don't just aim to make individuals aware of their data being processed, but also aim to give them effective control over how their data is being processed.[248]

4.3.1. The right to be informed, Art. 13 and 14 GDPR

Articles 13 and 14 GDPR provide the right to information and describe what kinds of information a controller has to provide to a data subject, when the personal data was either collected from the data subject or from a third party, while Art. 12 GDPR describes how the information must be provided.[249]

Art. 13 GDPR enumerates the information which is to be provided to a data subject "where personal data relating to a data subject are collected from the data subject"[250]. According to Wachter et al. to challenge inferences that have been drawn, Art. 13 GDPR cannot help data subjects.[251] This seems reasonable as Art. 13 GDPR specifically talks about data collected from the data subject. It portrays the main difference between the categories of provided and observed data and inferred data, as inferred data is neither collected through direct actions of the data subject, e.g. registering

---

245 European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 349.
246 Art. 5(1)(a) GDPR.
247 Art. 12 GDPR; Articles 13 and 14 GDPR; Articles 15 to 22 GDPR.
248 Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855, page 41; Article 29 Working Party, Opinion 8/2014 on the on Recent Developments on the Internet of Things, Adopted on 16 September 2014, 14/EN WP 223, < https://www.pdpjournals.com/docs/88440.pdf>, page 17.
249 Art. 12, 13, 14 GDPR.
250 Art. 13(1) GDPR.
251 Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 52.

on a website, nor by indirect actions, browsing on websites, but rather new created by actions of someone else, which may be based on provided and observed data of the data subject.[252] After inferences have been drawn, Art. 13 GDPR therefore is not applicable.

Nevertheless, Art. 13 GDPR may be helpful in order to raise awareness prior to inferences being drawn. As inferences are essentially based on input data, of which at least some will have been provided by the data subject, Art. 13 GDPR can help inform data subjects of the practice that the controller intends to draw inferences about him.[253] Informing data subjects in a transparent manner of the purposes of processing[254], including the existence of automated decision-making, including profiling[255], would enable data subjects to know whether inferences would be drawn in the first place and give them control over their personal data, by informing them also of the data subject rights[256] and the recipients of their personal data[257]. While data subjects may not be able to prevent inferences from being drawn about them altogether, they at least would be informed about the practice and what rights they have in that relation.

Another right to information can be found in Art. 14 GDPR that covers informing the data subject about the processing of personal data "where personal data have not been obtained from the data subject"[258]. Different scenarios are covered by this provision.

For one, this can include the controller processing inferred data about the data subject, that he obtained from a third party, for further purposes, e.g. making a decision about the data subject. According to Art. 14(3)(a) the controller has to provide the information enumerated in par. 1 and 2 "within a reasonable period after obtaining the personal data". Essential is here, when the controller has received the inferred data on the data subject from the third party, because then he is obligated to notify the data subject about, for instance, the purposes of processing the inferred data (par. (1)(c)), the categories of data in which the inferred data falls (par. 1(d)), the existence of certain data subject rights (par. 2(c)) and importantly, from which source the controller has received

252 Abrams, M. 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, <https://ssrn.com/abstract=2510927>, pages 6, 7, 8.
253 Art. 13(1)(c), (2)(f) GDPR.
254 Art. 13(1)(c) GDPR.
255 Art. 13(2)(f) GDPR.
256 Art. 13(2)(b) GDPR.
257 Art. 13(1)(e) GDPR.
258 Art. 14(1) GDPR.

the inferred data (par. 2(f)).[259] Not only will the data subject be made aware of the existence of inferred data about him, but also, he will know from where the inferred data comes, which is essential if he wants to prevent inferences from being drawn about him in the first place. Furthermore, he will know what the controller will use the inferred data for and whether they get transferred to another party, which is essential in order for a data subject to gain oversight of the data flows of inferences about him. Knowing the purposes for what his inferred data are used for, will also help him evaluate, to some extent at least, how decisions, for instance in regard to insurances or credit loans, may be impacted by this data.

The second scenario covered by Art. 14 GDPR, is when the controller himself creates the inferred data. The moment in which the AI presents the inference drawn is the moment the controller 'obtained' this data, which according to Art. 14(3)(a) GDPR triggers the notification requirements of par. 1 and 2, leading to the same information required to be provided to the data subject as in the first scenario.

Therefore, Art. 14 GDPR entails the potential for data subjects to gain awareness of the existence of inferences, of the responsible party that drew the inferences, what category of data the inferences fall into, for what purposes the inferences are being further processed and who else receives the inferred data. This helps data subjects to a great extent, to gain oversight of all data flows regarding inferences drawn on them, enabling them to gain control over their inferred data.[260] In theory, this would give data subjects the information needed in order to exercise their data subject rights according to Articles 15-22 GDPR.

4.3.2. The right of access, Art. 15 GDPR

Art. 15 GDPR sets out the right of the data subject to "obtain from the controller confirmation as to whether or not personal data concerning him or her are being processed,

[259] See also Article 29 Working Party, Annex – health data in apps and devices to the Advice paper on special categories of data ("sensitive data"), April 2011, < https://ec.europa.eu/justice/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf>, page 6.
[260] See also Article 29 Working Party, Opinion 8/2014 on the on Recent Developments on the Internet of Things, Adopted on 16 September 2014, 14/EN WP 223, < https://www.pdpjournals.com/docs/88440.pdf>, page 6.

and, where that is the case, access to the personal data"[261] and certain information.[262] This right does not depend on the fact from whom the personal data, e.g. inferred data, was collected or obtained from, it provides for data subjects to gain information from a controller, of which they suspect that he has either created inferences about them or has received inferences and is processing them further. The controller has to provide the data subject "with a copy of the personal data being processed"[263], in an intelligible, understandable and easily accessible form[264]. Content-wise the information to be provided to the data subject is similar to Art. 13 and 14 GDPR.[265] Similar as described under Art. 14 GDPR, the data subject can request information on the purposes of processing[266] of his personal data of a specific controller, the categories of personal data[267] concerned, any possible recipients of his personal data[268] and the source of where the controller has obtained his personal data[269]. The data subject could for one, make use of Art. 15 GDPR, if he suspects that inferences have been drawn about him, either of a controller he believes is drawing or aims to draw inferences in the first place and disclosing those to other parties. Secondly, he can use the right of access to receive information on whether a controller uses inferred data, to further process those, for instance, for making decisions about the data subject. This poses as an additional mechanism for the data subject to become aware of the existence of inferred data about him, about controllers aiming to process their data to draw inferences and for what purposes these inferences are further used for and who receives the inferences.

Wachter et al. argue that Art. 15 GDPR provides only ex-post knowledge on a processing activity[270], which Edwards and Veale don't seem to disagree with, "it seems

---

[261] Art. 15(1) GDPR.
[262] Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855, page 51.
[263] European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 218.
[264] Art. 12(1) GDPR.
[265] See text of Art. 13, 14 and 15 GDPR.
[266] Art. 15(1)(a) GDPR.
[267] Art. 15(1)(b) GDPR.
[268] Art. 15(1)(c) GDPR.
[269] Art. 15(1)(g) GDPR.
[270] Wachter, Sandra et al. 'Counterfactual explanations without opening the black-box: Automated decisions and the GDPR' [2017] 31(2) Harvard Journal of Law & Technology (2018), https://ssrn.com/abstract=3063289, page 870.

access comes after processing"[271]. However, in my opinion it is necessary to be clear about what the relevant processing event is. Two different processing events should be distinguished here, that are relevant for gaining awareness and oversight in regard to inferences. The event of processing ordinary data and provided or observed personal data from which ultimately inferences are drawn from and the event of using the drawn inference further to make a decision or prediction. In the first event, Art. 15 GDPR can act as an ex-ante mechanism, when the aim of the data subject is to become aware of whether a controller plans to draw inferences on him. In the second event, the inferences have been drawn and the data subject seeks information on how those are further processed by the controller; therefore, acts as an ex-post mechanism. This distinction of when is the access ex-ante and when it is ex-post can only be narrowed down to a concrete processing event, in this case here, the drawing of inferences and whether access is requested prior or after it has occurred. The focus on the distinction of whether the right entails an ex-ante or ex-post mechanism does not provide a benefit to the discussion of how data subjects can gain awareness and oversight of inferences being drawn about them.

4.3.3. The right to data portability, Art. 20 GDPR

Art. 20 GDPR provides data subjects with the right to "receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format"[272]. "The new right to data portability aims to empower data subjects regarding their own personal data, as it facilitates their ability to move, copy or transmit personal easily from one IT environment to another."[273] The idea is that while the data subjects receive their personal data, they could be made aware of the existence of inferred data about them, which would not only raise awareness but also provide a solution for checking inferred data on their accuracy.[274] The idea though

[271] Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855, page 52.
[272] Art. 20(1) GDPR.
[273] Article 29 Working Party, Guidelines on the right to data portability, WP 242 rev.01, 05 April 2017, https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233, page 4.
[274] European Data Protection Supervisor, Opinion 7/2015, Meeting the challenges of big data, A call for transparency, user control, data protection by design and accountability, 19 November 2015, < https://edps.europa.eu/sites/edp/files/publication/15-11-19_big_data_en.pdf>, page 13; Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855, page 74.

is not feasible. Art. 20 GDPR, just as Art. 13 GDPR, only applies to data which the data subjects "has provided to a controller"[275]. The Art. 29 WP addresses this issue in its guidance on data portability by stating that

> *"even though such data may be part of a profile kept by a data controller and are inferred or derived from the analysis of data provided by the data subject, these data will typically not be considered as 'provided by the data subject' and thus will not be within scope of this new right"*[276].

Edwards and Veale, who propose that Art. 20 GDPR might offer a solution for individuals to gain control over inferences being drawn about them, do emphasize that there is no consensus on whether inferred data can be covered by Art. 20 GDPR and that in order for data subjects to gain greater control over inferences drawn under Art. 20 GDPR, many problems arise and from a practical point of view it may not offer the best solution for data subjects.[277] I have to agree with the later. As discussed above, data subjects already have a number of rights under Articles 13-15 GDPR to gain information on inferences being drawn about them and I don't see the need to interpret another right, in the way it would also provide data subjects with information. I believe the need for information is adequately fulfilled by the other provisions.

4.3.4. Data subjects sufficiently aware of their inferred data?

To summarize this section, Art. 13 GDPR entails the possibility to provide ex-ante information for data subjects, before inferences get drawn based on personal data provided by them or observed of them. Art. 14 GDPR provides ex-post information for data subjects, after inferences have been either drawn by a third-party and transferred to a controller who further processes these, or inferences that have been drawn by the controller himself. Art. 15 GDPR can address both the scenario, to gain access to certain information before inferences have been drawn about a data subject and to gain access to information after inferences have been drawn, thus providing an ex-ante and ex-post mechanism. Art. 20 GDPR though does not provide any additional benefit for individuals

---

[275] Art. 20(1) GDPR; See also Hoofnagle, C. et al. 'The European Union general data protection regulation: what it is and what it means' [2019] 28(1) Information & Communications Technology Law, https://doi.org/10.1080/13600834.2019.1573501, page 89.

[276] Article 29 Working Party, Guidelines on the right to data portability, WP 242 rev.01, 05 April 2017, https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233, page 10.

[277] Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855, page 73.

to gain awareness on inferences about them, that the other three rights have not already covered. Articles 13, 14 and 15 GDPR each entail the potential to inform data subjects sufficiently in order to facilitate them to exercise their data subject rights under Articles 16-21 GDPR effectively. However, the above demonstrations are quite optimistic. The following section will assess the individual data subject rights of Articles 16-21 GDPR in regard to inferences, and it will be demonstrated how detailed provided information would practically be needed, in order for data subjects to exercise their rights effectively.

4.4. Proving the inaccuracy of inferences drawn by data subjects with the general data subject rights

After having become aware of inferences being drawn, proving their inaccuracy can be challenging, but is a prerequisite for various data subject rights.[278] The data accuracy principle, laid down in Art. 5(1)(d) GDPR, requires controllers to take reasonable steps in order to assure that inaccurate personal data are erased or rectified or to assure that personal data are accurate and up to date, in the context of the purpose of the data processing.[279] While controllers should apply the data accuracy principle at all stages of a processing process, the variety of sources that input data comes from, from which the inferred data is created of, challenges the principle.[280] As portrayed in the second chapter, the drawing of inferences that are inaccurate, can occur for various reasons, such as unbalanced datasets, biased input data or biased design choices of the AI system and algorithms which take discriminating factors into account.[281] For such reasons, it is essential that data subjects who fear that inaccurate inferences are being drawn about them, to have a way to either rectify or erase that data or to be able to object to or restrict the processing of that personal data, so it won't implicate them and their lives in a negative way.

---

278 See, for instance, Art. 16 GDPR ("the rectification of inaccurate personal data"); Art. 18(1)(a) GDPR ("the accuracy of the personal data is contested").

279 Art. 5(1)(d) GDPR; See also European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 127.

280 European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 356; Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, pages 12, 17.

281 See above under 2.4. "Challenges of inferences drawn and their implications for individuals".

4.4.1. Right to Rectification, Article 16 GDPR

Art. 16 GDPR provides data subjects with the right to rectify inaccurate personal data and to complete incomplete personal data.[282] Data subjects wanting to challenge the accuracy of inferences drawn might be able to do so with Art. 16 GDPR.[283]

In order for a data subject to obtain the rectification of inaccurate inferences from the controller, he has to show that the inferences are indeed inaccurate. "Rectification implicitly relies upon the notion of verification, meaning that a record can demonstrably be shown to be invalid and thus corrected by the data subject."[284] This becomes problematic in cases where inferences are drawn by AI and the inferences constitute probabilistic assumptions and cannot be verified.[285] As explained above, in terms of the element "any information" a distinction exists between objective and subjective information.[286] If the data used for drawing the inference or the inference itself relies on verifiable data, e.g. the customer is pregnant[287] or the individual is male[288], then the data subject should be capable of providing evidence which contradict the inference drawn or the data used to draw the inference.[289] The inference would be objective information, as it has a verifiable, factual basis, which can be proven to be accurate or inaccurate.[290] However, making use of the right to rectify in terms of where the inference constitutes subjective information, is more of a challenge, maybe even impossible.[291] Some inferences have a predictive or assuming character, such as inferring that someone is a

---

[282] Art. 16 GDPR.
[283] Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 17.
[284] Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019] (2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 57.
[285] Ibid 57.
[286] See above under 3.2.1. "Any information".
[287] Example, where a company send advertisement for pregnancy goods to women, they inferred to be pregnant at the time, see: Kashmir Hill, 'How Target Figured out a teen girl was pregnant before her father did' [2012], https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#1fc1e24c6668.
[288] Example where a woman was denied access to the female locker room because she carried the title 'Dr.' which the algorithm associated with men only.
[289] Similar, Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 57.
[290] See also above under 3.2.1. "Any information".
[291] For an explanation of what subjective information is, see above under 3.2.1. "Any information".

member of a gang[292], that someone is a risky driver or likely to be in car accidents in the future, also referred to as non-verifiable data, where it is unclear how an individual could prove the data to be inaccurate.[293] In these cases, exercising the right to rectify, even when all necessary information as described above under Articles 13 – 15 GDPR has been provided to the data subject, may be impossible.

4.4.2. Rights to Erasure and Restriction, Articles 17, 18 GDPR

Art. 17 GDPR entails the data subject right "to obtain from the controller the erasure of personal data concerning him or her without undue delay"[294], while also obligating the controller to erase the data subject's personal data if one of the enumerated grounds apply.[295]

According to the Art. 29 WP "the right to rectification and erasure apply both to the 'input personal data' (the personal data used to create the profile) and the 'output data' (the profile itself or score assigned to the person)"[296]. Assessing this statement under what has been defined as inferences in chapter two[297], both the profile and the credit score can be defined as inferences, thus the rights in Art. 16 and 17 GDPR would be applicable to inferences, if they constitute the output of processing. Unfitting, Edwards and Veale rely on the Art. 29 WP guidance on the right to portability, when assessing whether a data subject has the right to erase his inferences drawn by AI.[298] According to them, the Art. 29 WP states in that guidance, that inferences belong to the system that drew them, rather than belonging to the data subject, based on which they make the false assumption that this also applies for the right to erasure. However, the Art. 29 WP only addresses inferred

---

292 For example, the Gangs Matrix, see: https://www.amnesty.org.uk/london-trident-gangs-matrix-metropolitan-police.

293 Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 57. See also: CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994, par. 34; Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, page 6.

294 Art. 17(1) GDPR.

295 Art. 17(1) GDPR.

296 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 18.

297 See above under 2.3. Drawing inferences.

298 Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855, page 69.

data in the context of whether they constitute 'provided data' by the data subject, as it is required under Art. 20 GDPR, which in that context is correct.[299] As, Art. 17 GDPR does not distinguish between the categories of personal data but rather applies to all personal data of the data subject, the analogy of the guidance on Art. 20 GDPR does not fit in regard to Art. 17 GDPR. Therefore, Art. 17 GDPR is applicable to inferences drawn, as assessed above. However, there is no general right of the data subject to obtain the erasure of his personal data, thus, a data subject cannot prevent the drawing or use of inferences altogether, if none of the enumerations apply to his situation. This narrows the impact Art. 17 GDPR can have on inferences quite extensively.

Art. 18(1)(a) GDPR provides the data subject with the right to restrict the processing of their personal data for the period of time where "the accuracy of the personal data is contested by the data subject, for a period enabling the controller to verify the accuracy of the personal data"[300]. According to the Art. 29 WP "the right to restrict processing (Art. 18) will apply to any stage of the profiling process"[301]. So, even if the profile as the output is not defined as the inference drawn, inferences would be covered by Art. 18 GDPR, as it is at least one step of the profiling process.[302]

In general, all three rights therefore entail the possibility of applying to inferences drawn, while their actual impact might not be as encompassing as one would hope.

4.4.3. Right to object, Art. 21 GDPR

The data subject does not have a general right to object to the processing of personal data, but rather the right in Art. 21 GDPR relies "on grounds relating to his or her particular situation"[303] which then has to be balanced with the legitimate rights of the controllers for processing the data. Even though, according to Recital 69 of the GDPR the controller has to "demonstrate that its compelling legitimate interest overrides the interests or the fundamental rights and freedoms of the data subject"[304], meaning the

---

299 Article 29 Working Party, Guidelines on the right to data portability, WP 242 rev.01, 05 April 2017, https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233, pages 10,11, 17.
300 Art. 18(1)(a) GDPR.
301 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 18.
302 See above under 2.3.2. Inferring profiles from individuals.
303 Art. 21(1) GDPR; See also, European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 229.
304 Recital 69 GDPR.

burden of proof lies with the controller, the data subject first will have to provide certain reasons for why he believes he has grounds for objection for which he needs to receive the necessary information on his inferred data.[305] If the results of the balancing is in favor of the data subject, then then controller can no longer legally process the data subject's personal data, e.g. inferred data.[306] The data subject can then request the erasure of the inferred data pursuant to Art. 17 GDPR.[307] Processing activities performed prior to the objection, however, remain legitimate.[308]

4.4.4. General information provided enough to facilitate the data subject rights?

All four of the above-mentioned rights can apply to inferences drawn. However, essential for data subjects being able to exercise their data subject rights is, whether the information they have been provided with under Articles 13 and 14 or the information they have gotten access to under Art. 15 GDPR, is sufficient enough in enabling them to effectively make use of the rights. In general, Articles 13-15 GDPR should provide the data subject with the needed information in order to enforce data subject rights in cases where he finds inferences drawn about him to be inaccurate. This possibility though depends on how informative and detailed the information of the controller will actually be. For each right it is essential that the data subject gets the information specific enough in order to determine whether inferred data about him is accurate or not. The European Agency for Fundamental Rights and the Council of Europe specifically state that

> *"accessing his or her personal data will help a data subject to determine whether or not the data are accurate. It is, therefore, essential that the data subject is informed, in an intelligible form, not only of the actual personal data that are being processed, but also the categories under which these personal data are processed, such as name, IP address, geolocation coordinates, credit card number, etc."*[309]

---

[305] Art. 21(1) GDPR, Recital 69 GDPR; See also, European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 230.
[306] European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 231.
[307] Art. 17(1)(c) GDPR. See also, Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 62.
[308] European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 231.
[309] Ibid 218.

However, the data subject will not only need to know the drawn inference itself, for instance, high insurance risk or low credibility, because while he can ascertain for himself whether he finds this to be accurate or not, he will not know what led to this inference. He will need to know, which data about him, which factors, have resulted in leading to this inference, for instance, his eating and fitness behavior combined together concluded that he is or will be obese or his driving behavior combined with his history of accidents concluded that he is a risky driver, both leading to the inference that he constitutes a high insurance risk. According to Art. 12(2) GDPR a controller therefore must provide information to that detail, that the data subject is put in the position of sufficiently exercising his data subject rights under Articles 15-22 GDPR, as this is important for two reasons.

Only when he receives information to this detail, he has the possibility for one, to prove that he indeed does put effort into leading a healthy lifestyle, facilitating him to exercise his data subject rights. Secondly, this also gives the data subject, the possibility to change certain behavior, for instance, his driving, in order not to be categorized as a high insurance risk. Individuals need to be given the possibility to change their behavior in order to change the inferences that will be drawn about them so they can positively better the implications inferences have on decisions being made about them. Otherwise, 'old' inferences can keep determining and affecting every new inference to be drawn and every future decision being made, when drawn inferences keep getting fed back into the algorithm. For instance, when insurance risks are re-evaluated every ten years in order to decide on adequate insurance rates, when these are based on historic inferences drawn, rather than taking into account the new adapted positive behavior of the last ten years adequately, individuals will have a hard time getting out of a category, that they have once been put in.

However, the provision on Art. 16 GDPR has shown, that proving the inaccuracy of inferences, that can be evaluated as constituting subjective information, might just be a hurdle, hard to overcome. Even if data subjects had been provided with all information necessary according to Articles 13 – 15 GDPR, proving one will not have a number of car accidents in the future, would be a challenge.

The next section will assess what impact Art. 22 GDPR can have on inferences drawn in automated decision-making and profiling.

4.5. Art. 22 GDPR's impact on inferences drawn

Art. 22(1) GDPR entails the right of data subjects "not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects (…) or similarly significantly affects him or her"[310]. Art. 22(2) GDPR provides for exemptions from the prohibition[311], while Art. 22(3) GDPR requires the controller to implement "suitable measures to safeguard the data subject's rights and freedoms and legitimate interests"[312], by mentioning, among others, the right to contest the decision. Art. 22(4) GDPR puts down specific limitations for processing special categories of personal data under Art. 9(1) GDPR. Articles 13(2)(f), 14(2)(g) and 15(1)(h) GDPR require data controllers to provide specific information or access to information in regard to automated decision-making under Art. 22 GDPR. In order to assess to what extent Art. 22 GDPR impacts the drawing of inferences by AI, this section will start off by explaining the cases Art. 22(1) GDPR applies to, followed by a discussion on what information is required under Articles 13(2)(f), 14(2)(g) and 15(1)(h) GDPR. At last the existence of a 'right to explanation' and the impact it could have on inferences will be shortly portrayed.

4.5.1. Application of Art. 22 GDPR

Starting with Art. 22(1) GDPR, the Art. 29 WP[313] specifies that the data subject is not required to actively seek an objection to such a decision, but the right rather constitutes a general prohibition, which is in line with the interpretation done by Mendoza and Bygrave[314].[315] It aims to protect individuals who are subject to automated decisions against negative consequences.[316] In light of how Art. 22 GDPR can impact inferences drawn, the elements of "solely based on automated processing" and "legal or similar

---

310 Art. 22(1) GDPR.
311 Art. 22(2) GDPR: "Paragraph 1 shall not apply if the decision: (a) is necessary for entering into, or performance of, a contract between the data subject and a controller; (b) is authorized by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject`s rights and freedoms and legitimate interests; or (c) is based on the data subject's explicit consent."
312 Art. 22(3) GDPR.
313 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 15.
314 Mendoza, I. and Bygrave, L. 'The Right not to be subject to automated decisions based on profiling' [2017] 20 University of Oslo Faculty of Law Legal Studied Research Paper Series No. 2017-20, page 9.
315 Also, the view of other scholars, Brkan, M. 'Do Algorithms rule the world? Algorithmic decision-making in the framework of the GDPR and beyond' [2017] Proceedings of the 16th International Conference on Artificial Intelligence and Law, < https://ssrn.com/abstract=3124901>, page 8;
316 European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 233.

significant effects" will be further elaborated. Inferences are likely to have been drawn in the process of any automated decision making.[317] Whether intended or not.

"Solely automated decision-making is the ability to make a decision by technological means without human involvement."[318] It has to be assessed how strictly the solely automated processing is to be understood, as this determines whether Art. 22 GDPR is applicable to a processing situation or not.[319] "In real life, much of automated decision-making supplements human judgment, and these systems appear to escape the prohibition."[320] Is the possibility, that a human can impact, for instance, change the decision made by the automated process, enough, in order to escape Art. 22(1) GDPR?[321] For instance, in credit scoring or in insurance rate evaluations, the decisions made by the automated systems often only assist the human manager in his decision.[322] According to the Art. 29 WP, if a human reviews the decision made before he approves or makes the final decision, it falls outside the scope of Art. 22(1) GDPR.[323] However, the human needs to have actual influence on the decision, meaning he must be able to actually change the decision made by the automated process in order to escape the scope of Art. 22(1) GDPR.[324] Whether the human then actually takes a different decision than the advised

---

317 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 8; Agreeing see Wachter, S. and Mittelstadt, B. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review <https://ssrn.com/abstract=3248829>, page 78.

318 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 8

319 Wiedemann, K. 'Automated Processing of Personal Data for the Evaluation of Personality Traits: Legal and Ethical Issues' [2018] Max Planck Institute for Innovation and Competition Research Paper No. 18-04, https://ssrn.com/abstract=3102933, page 22.

320 Hoofnagle, C. et al. 'The European Union general data protection regulation: what it is and what it means' [2019] 28(1) Information & Communications Technology Law, https://doi.org/10.1080/13600834.2019.1573501, page 91.

321 Brkan, M. 'Do Algorithms rule the world? Algorithmic decision-making in the framework of the GDPR and beyond' [2017] Proceedings of the 16th International Conference on Artificial Intelligence and Law, < https://ssrn.com/abstract=3124901>, page 9; See also Borgesius, F. 'Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence' [2020] The International Journal of Human Rights, < https://ssrn.com/abstract=3561441>, page 17.

322 Hoofnagle, C. et al. 'The European Union general data protection regulation: what it is and what it means' [2019] 28(1) Information & Communications Technology Law, https://doi.org/10.1080/13600834.2019.1573501, page 91.

323 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 20.

324 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018,

one from the automated system, is irrelevant, as the controller can prove, that there is a human affecting the decision and he escapes Art. 22 GDPR.[325] This leads to the conclusion, that many practically automated decisions will not be covered by Art. 22(1) GDPR in the end, leading to a narrow applicability of the prohibition.[326] However, only inferences drawn in decision making processes which occur completely automated, may be impacted by Art. 22 GDPR. As inferences are generally drawn in profiling scenarios[327], these are also covered, when the profiling process is solely relying on automated processing.[328]

Only inferences drawn, that led to decision which produces legal effects or similarly significantly affects the data subject, are possibly treated under Art. 22 GDPR. Legal effects might be the denial of social benefits, citizenship or cancellations of a contract.[329] A similar significant impact on the life of the data subject can be assumed when the decision relates to, for example, creditworthiness, employee recruitment or reliability analysis.[330] Especially in these cases inferences are generally drawn in the process of making a decision, e.g. credit scores are a well-known example of where inferences get drawn.[331]

Having set the scope of which inferences can be impacted by Art. 22 GDPR by determining the decisions covered, the next part will portray the "special" requirements that are to be complied with, when the scope of Art. 22 GDPR is triggered, and the exemptions of Art. 22(2) GDPR do not apply.

<http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 21; Wiedemann, K. 'Automated Processing of Personal Data for the Evaluation of Personality Traits: Legal and Ethical Issues' [2018] Max Planck Institute for Innovation and Competition Research Paper No. 18-04, https://ssrn.com/abstract=3102933, page 22.

325 Brkan, M. 'Do Algorithms rule the world? Algorithmic decision-making in the framework of the GDPR and beyond' [2017] Proceedings of the 16th International Conference on Artificial Intelligence and Law, < https://ssrn.com/abstract=3124901>, page 9.

326 Edwards, L. and Veale, M. 'Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?' [2018] 16(3) IEEE Security & Privacy, < https://ssrn.com/abstract=3052831>, page 3.

327 See above under 2.3.2. Inferring profiles from individuals.

328 See Art. 22(1) GDPR 'including profiling'.

329 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 21.

330 Recital 71 of the GDPR; European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018), page 233.

331 Kamp, M. et al. 'Profiling of Customers and Consumers – Customer Loyalty Programmes and Scoring Practices' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010) 207.

4.5.2. Information rights in regard to Art. 22 GDPR

Articles 13(2)(f), 14(2)(g) and 15(1)(h) GDPR each require controllers to provide data subjects with information or access to information in regard to

> *"the existence of automated decision-making, including profiling, referred to in Articles 22 (1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject"*[332].

One part especially needs further elaboration in regard to how Art. 22 GDPR impacts inferences drawn: Does meaningful information about the logic include information about the inferences drawn in the process of making a decision?

The GDPR itself does not provide further guidance of what "meaningful information" would constitute. Selbst and Powles interpret the element from the data subjects point of view, as they argue, "Articles 13-15 relate to the rights of the data subject"[333] and therefore, the information provided under those articles should help data subjects to exercise their rights guaranteed under the GDPR.[334] "(…), If an individual receives an explanation of an automated decision, she needs to understand the decision well enough to determine whether she has an actionable discrimination claim."[335] I agree with this observation. "Meaningful information" needs to be assessed in light of Article 12 GDPR. The information rights under Art. 13 and 14 and the access to information of Art. 15 GDPR shall enable data subjects to exercise their data subject rights.[336] The transparency principle in Art. 5(1)(a) GDPR therefore, requires that the information be 'meaningful' in regard to the data subject.[337] "The information about the logic involved needs to enable the data subject to express his or her point of view and to contest the automated decision."[338][339] This information goes beyond what the controller is required

---

[332] Articles 13(2)(f), 14(2)(g) and 15(1)(h) GDPR.
[333] Selbst, A. and Powles J. 'Meaningful Information and the Right to Explanation' [2017] 7(4) International Data Privacy Law https://ssrn.com/abstract=3039125, page 7.
[334] Ibid 8.
[335] Ibid 8.
[336] Art. 12(3) GDPR.
[337] Selbst, A. and Powles J. 'Meaningful Information and the Right to Explanation' [2017] 7(4) International Data Privacy Law <https://ssrn.com/abstract=3039125>, page 8.
[338] Brkan, M. 'Do Algorithms rule the world? Algorithmic decision-making in the framework of the GDPR and beyond' [2017] Proceedings of the 16th International Conference on Artificial Intelligence and Law, < https://ssrn.com/abstract=3124901>, page 15.
[339] See also Mendoza, I. and Bygrave, L. 'The Right not to be subject to automated decisions based on profiling' [2017] 20 University of Oslo Faculty of Law Legal Studied Research Paper Series No. 2017-20, pages 16, 17.

to inform about in regard to the other data subject rights, in the sense that, not only does the controller have to inform about the purposes of processing certain data, of its source and whom it will be disclosed to, but in regard to Art. 22 GDPR he actually has to provide inside knowledge on the process, at least to some extent.[340] For a data subject to understand the logic of a decision making process, the following information could be required: information about the input or training data of the algorithm and information about the factors that are taken into account for making an decision and their ascribed weights.[341]

The Art. 29 WP states that "meaningful information about the logic involved" does not require a complex explanation of the algorithms involved.[342] This seems reasonable, as an explanation of the algorithm would probably not be very meaningful to the general data subject. The information that is to be provided should put the data subject in a position of understanding the reasons for a certain decision.[343] In credit or insurance decision making, this can include information, which was provided by the data subject, information about the data subject in public records, for instance about insolvency or fraud, and information about the account conduct, for instance in regard to payments.[344] If followed the guidelines of the Art. 29 WP inferences drawn would not necessarily be included in the information provided to the data subject. However, it would make more sense providing individuals with the inferences that were drawn about them in the process, rather than disclosing the technological processes of the automated system, which also face the barrier of being balanced with IP rights and trade secrecy.[345] Knowing what the inferences are that have been drawn about a data subject, would rather put them in a position of being able to contest a decision or expressing his view on the matter (Art. 22(3) GDPR), as data subjects could assess whether they find this information to be accurate or not. Receiving information only on the input data which was provided by the data subject

340 Brkan, M. 'Do Algorithms rule the world? Algorithmic decision-making in the framework of the GDPR and beyond' [2017] Proceedings of the 16th International Conference on Artificial Intelligence and Law, < https://ssrn.com/abstract=3124901>, page 15.
341 Ibid 15.
342 Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 25.
343 Ibid 25.
344 Ibid 25.
345 See Recital 63 GDPR.

and a technological description of what is done with this data, seems less sufficient to help a data subject, assessing the accuracy of a decision made.

In the context of what information needs to be provided to the data subject, there is an extensive discussion in legal scholarship[346], on whether the Articles 13(2)(f), 14(2)(g), 15(1)(h) GDPR and 22 GDPR constitute a 'right to explanation', as Recital 71 of the GDPR "to obtain an explanation of the decision reached"[347] as a suitable safeguard under Art. 22(3) GDPR. However, there is not much consensus in regard to the details, for instance, what the explanation would exactly entail and what role Recital 71 of the GDPR plays. Having an explanation of the reached decision may help individuals gain knowledge on drawn inferences, on which the decision relies. Goodman and Flaxman expect of such an explanation to gain insight into how certain input data relates to the output, the decision.[348] As an example, they state that the explanation should enable data subjects being able to answer questions such as, "Is the model more or less likely to recommend a loan if the applicant is a minority? Which features play the largest role in prediction?"[349]. However, explanations seem unlikely to be as detailed. While detailed explanations would enable data subjects, to also check the accuracy of the inferences being drawn in the process, the characteristics of AI will make the occurrence of such an explanation unlikely.[350] As many AI systems currently already work with algorithms which are characterized as black-boxes, where even for developers it is impossible to explain how the system came from the input data to a certain output, knowing which

---

346 Doubters of the effectiveness and proposal of other rights or mechanisms: Edwards, L. and Veale, M. 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' [2017] 16 Duke Law & Technology Review 18 (2017), https://ssrn.com/abstract=2972855; Wachter, S. et al. 'Why a Right to Explanation of Automated Decision-Making Does Not Exists in the General Data Protection Regulation' [2017] International Data Privacy Law < https://ssrn.com/abstract=2903469>; Malgieri, G. and Comandé, G. 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' [2017] 7(4) International Data Privacy Law, < https://ssrn.com/abstract=3088976>. Supporters of the right, see Goodman, B. and Flaxman, S. 'European Union regulations on algorithmic decision-making and a "right to explanation"' [2017] 38(3) AI Magazine, < https://arxiv.org/ct?url=https%3A%2F%2Fdx.doi.org%2F10.1609%2Faimag.v38i3.2741&v=f2c797e9>; Selbst, A. and Powles J. 'Meaningful Information and the Right to Explanation' [2017] 7(4) International Data Privacy Law https://ssrn.com/abstract=3039125.

347 Recital 71 GDPR.

348 Goodman, B. and Flaxman, S. 'European Union regulations on algorithmic decision-making and a "right to explanation"' [2017] 38(3) AI Magazine, < https://arxiv.org/ct?url=https%3A%2F%2Fdx.doi.org%2F10.1609%2Faimag.v38i3.2741&v=f2c797e9>, pages 55, 56.

349 Ibid 55, 56.

350 See above under 2.4.1. An overview of specific challenges that accompany AI.

inferences are drawn in the process, is even more hidden in the black-box. "With unsupervised models, it may not be possible to trace the AI's learning processed or to explain its decisions, due to a lack of data labels and relationships."[351] Nevertheless, the aim of this thesis is not, whether such a right exists, but to what extent the provisions, in their current state, can mitigate the implications posed by inferences. Therefore, this section will focus on whether the mentioned articles apply to the case of inferences drawn and not how a possible right to explanation would look like, as this discussion would excess the limits set for this thesis.

For now, the conclusion is found that inferences drawn could be part of the 'meaningful information about the logic involved' to be provided under Articles 13(2)(f), 14(2)(g) and 15(1)(h) GDPR in regard to Art. 22 GDPR. However, time will tell if this will be the case in practice, when the GDPR is applied on automated decisions.

4.5.3. Introduction of appropriate procedures and measures, Recital 71 of the GDPR

Recital 71 of the GDPR recommends controllers to introduce "appropriate mathematical or statistical procedures for the profiling" for fair and transparent processing and to "implement technical and organizational measures" to ensure that inaccuracies and errors are minimized. As discussed above under Art. 16 GDPR data subjects face the challenge of being able to prove or assess the inaccuracy of inferences drawn about them.[352] To what extent can this provision of Recital 71 be relevant for data subjects? The recommendation to introduce certain procedures is directed at the controller which conducts profiling. The Art. 29 WP provides a list of exemplary procedures and measures controllers could take into account, for instance, "regular quality assurance checks", "testing the algorithms used and developed by machine learning systems to prove that they are actually performing as intended, and not producing discriminatory, erroneous or unjustified results", "independent third party auditing" or an "ethical review board to assess the potential harms and benefits to society".[353]

---

[351] Humerick, M. 'Taking AI Personally: How the EU must learn to balance the interests of personal data privacy and artificial intelligence' [2018] 34(4) Santa Clara High Technology Law Journal < https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=1633&context=chtlj>, page 412.
[352] See above under 4.4.1. "Right to Rectification, Art. 16 GDPR".
[353] Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, <http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826>, page 32.

This might be beneficial for the profiling results in general, as auditing entails the possibility for controllers to correct occurring errors in the machine learning system and therefore, improve the overall quality of the machine learning outputs.[354] However, the extent to which this sort of auditing is actually helpful in regard to the challenges of AI in general and inferences has been discussed above[355], and the black-box characteristic of some AI systems remains a hurdle. If inferences are drawn, not as the output but as the necessary step in-between the input data and the aspired output, e.g. a decision, the inferences drawn cannot be audited for their accuracy.[356] Furthermore, these inferences can be fed back into the system for future decisions to be made about an individual, therefore remaining in the system.[357] As even for the developers of the systems, they constitute a black-box, it does not seem reasonable to assume that the recommendations of Recital 71 will have a large impact on the current challenges posed by AI systems, if not the systems themselves become more transparent.

I do recognize, that auditing entails the possibility to determine the accuracy of those systems to a certain degree, for instance, by testing, when the system is given known input-output pairs, how accurate its results are and whether certain factors or weights taken into account by the system have to be adapted. However, an important concern raised by Wiedemann is "how close to the "truth" the results of profiling have to be in order to fulfil the requirements of Recital 71, and how many false positives or false negatives are acceptable"[358]; one I fully agree with.[359][360]

If the black-box barrier is not existent for a system or can be overcome to a large extent, having third-party audits or ethical review boards to assess the impacts of machine

---

354 Malgieri, G. and Comandé, G. 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' [2017] 7(4) International Data Privacy Law, < https://ssrn.com/abstract=3088976>, pages 248, 249.
355 See above under 2.4.1. "An overview of specific challenges that accompany AI".
356 See above under 2.4.1. "An overview of specific challenges that accompany AI".
357 See above under 2.4.2. "Assessing these challenges in the case of drawing inferences".
358 Wiedemann, K. 'Automated Processing of Personal Data for the Evaluation of Personality Traits: Legal and Ethical Issues' [2018] Max Planck Institute for Innovation and Competition Research Paper No. 18-04, https://ssrn.com/abstract=3102933, page 14.
359 On a similar note, see: Korff, D. 'Data Protection laws in the EU: The difficulties in meeting the challenges posed by global social and technical developments' [2010] Working Paper 2, European Commission Directorate-General Justice, Freedom and Security, 20 January 2010 < https://dx.doi.org/10.2139/ssrn.1638949>, page 52.
360 For an interesting paper on auditing algorithms, see: Mittelstadt, B. 'Auditing for Transparency in Content Personalization Systems' [2016] 10(12) International Journal of Communication, https://www.researchgate.net/publication/309136069_Auditing_for_Transparency_in_Content_Personalization_Systems.

learning results, would not only help controllers improve the systems they use, but could also benefit the data subjects wanting to asses or challenge the accuracy of inferences drawn about them, if they would get access to the results of said audits or reviews. Either to help data subjects make up their own mind about whether an inference drawn about them might be appropriate and accurate or to support their inaccuracy claim for inferences drawn. While discussing the extent to what Recital 71 can benefit data subjects, its non-binding nature has to be kept in mind and how reasonable it is to assume, that controllers will have third parties reviewing their algorithms.

4.5.4. Assessing the impact of Art. 22 GDPR on inferences drawn in the process of making a decision

Art. 22 GDPR specifically addressed the processing of personal data in automated decision-making processes and profiling in which in general inferences are drawn. It has been assessed that only inferences which are drawn in the process of a decision or profiling process which relies on a completely automated process can be impacted by Art. 22 GDPR and that many processes will escape the scope of Art. 22 GDPR due to humans in the loop, which have the ability to change the decision made by the automated process, regardless of whether the human will actually do so or aims to do so. This leads to many inferences drawn in decision-making processes and profiling not being covered by Art. 22(1) GDPR in the first place, leading to the "special" information requirements of Articles 13(2)(f), 14(2)(g) and 15(1)(h) GDPR not being applicable.

Processes that do fall within the scope of Art. 22(1) GDPR however entail the possibility of addressing inferences drawn in light of the information requirements of Articles 13(2)(f), 14(2)(g) and 15(1)(h) GDPR by covering them within the scope of "meaningful information about the logic involved" of the processing. Whether this will be the case and what effect this will have on inferences drawn, though, will only be seen once the GDPR is actively applied to such processing scenarios.[361] Another question that arises in regard to Art. 22 GDPR is, when a processing is prohibited under Art. 22(1) GDPR and no exemption applies, the decision-making process itself might be prohibited, however the inferences to be drawn will not be.

[361] See also, Borgesius, F. 'Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence' [2020] The International Journal of Human Rights, < https://ssrn.com/abstract=3561441>, page 18.

Another issue in the application of Art. 22 GDPR poses the fourth paragraph, which prohibits decision based on special categories of personal data.[362] In chapter three it was shown, that in many cases it is possible to infer sensitive data from ordinary input data.[363] It is unclear, how this would be taken into account by Art. 22 GDPR. Does the term 'based on' in Art. 22(4) GDPR only refer to the input data, or also data that was created in the process of making a decision?[364]

This section has shown that many details of how Art. 22 GDPR is applicable to processing activities remains unclear and the treatment of inferences under the provision can only be assumed so far.

4.6. Conclusion

Articles 13, 14 and 15 GDPR are the essential provisions when wanting to address inferences drawn under the GDPR, as for one they are important in raising awareness about inferences drawn and their level of detail in of the provided information determines how effective data subjects will be in exercising their data subject rights under Articles 16 to 22 GDPR. Articles 13, 14 and 15 GDPR each entail the potential for data subjects to be sufficiently made aware of controller's intentions of drawing inferences about them, of the existence of inferred data about them and of the purposes the inferences are intended to be used for. However, whether the information provided under Art. 13-15 GDPR is sufficient in order to exercise data subject rights against inaccurate inferences, depends on how strict and broad the information obligations for controllers will be interpreted as. The provisions definitely entail the possibility to facilitate data subjects to exercise their rights against inaccurate inferences; to what extent, depends on future interpretations of the requirements under Articles 13-15 GDPR.

To what extent, the provision of Art. 22 GDPR will impact inferences drawn in the process of making an automated decision similarly depends on the interpretation of its specific information requirements under Articles 13(2)(f), 14(2)(g) and 15(1)(h) GDPR. Only with enough detail, will the provisions of the GDPR be sufficient to mitigate the risks, inaccurate inferences can pose for individuals.

---

362 Art. 22(4) GDPR.
363 See above under 3.4. Inferences drawn as a special category of personal data, Art. 9(1) GDPR.
364 Similar so, see Hoofnagle, C. et al. 'The European Union general data protection regulation: what it is and what it means' [2019] 28(1) Information & Communications Technology Law, https://doi.org/10.1080/13600834.2019.1573501, page 92.

5. Conclusion

The thesis focused on the found gap in literature which evolved around the hurdles of applying the provisions of the GDPR on inferences drawn by AI, which are caused by the characteristics, the category of inferred data portrays. Especially analyzing the legislative text of the GDPR in its current state in regard to inferences drawn, without concentrating on finding solutions beyond the text, by proposing new rights to be interpreted into or added to the GDPR, constitutes an approach which had not been taken so far. Therefore, the main research question was formulated, which this thesis aimed to answer:

To what extent can the EU General Data Protection Regulation address the implications posed to individuals of inferences drawn by artificial intelligence?

The following will demonstrate the steps taken in order to answer the main research question, concluding with the found answer, before addressing the implications of the findings. In order to address the issues raised three sub-questions have been formulated.

The first sub-question aimed at clarifying the technological background of inferences drawn by AI and to demonstrate the implications posed thereof for individuals: *How does artificial intelligence draw inferences and what are the implications posed for individuals if left unaddressed?* The second chapter was devoted to this question and it was portrayed that AI systems have to be trained before they are able to draw inferences. Once drawn, the inferences constitute new data, which is categorized as inferred data, next to the other three categories of provided, observed and derived data. Depending on the task the AI system aims to fulfil, a drawn inference can either constitute the step in between of providing the AI with input data and it coming to an output, e.g. in the form of a decision, or the drawn inference can constitute the output itself, for instance, when AI is supplementing human decision making, leaving, making the actual decision based on the inference drawn, to humans. Some issues raised in regard to inferences apply to AI systems in general, as in bias built in the design of the AI by the developers itself or inaccuracy of the outputs, due to inaccurate, biased or unbalanced input and training datasets. However, there are concerns raised specifically in regard to inferences drawn. Due to the black-box characteristics of some AI systems, when the inference constitutes the middle step between the input data and the output, rather than itself constituting the aspired output, it can stay completely hidden inside the AI system, without anyone

knowing which inference was drawn about an individual. Checking inferences in this scenario for their accuracy seems impossible, even for the AI developers themselves. Furthermore, the drawn inferences, whether as a middle step or the output, can be fed back into the AI system, for future use, for instance, for future decision to be made about an individual. This entails the possibility, that once an individual has been categorized, he may be trapped in this category, without the possibility of escaping it. As inferences can therefore entail the possibility to impact the lives of individuals for a long term and can impact their lives in a variety of decisions and predictions to be made about them, making sure that the inferences drawn about one are accurate, is essential. For individuals to tackle these issues themselves, one characteristic of inferred data constitutes a major hurdle: Individuals are generally not aware that inferences are drawn about them or that inferred data about them exists, as neither active nor inactive participation in the creation of this inferred data is required from the individual. Overall, the second chapter portrayed the necessary background information on inferences drawn and how harmful they could implicate individuals.

The second sub-question triggered the material scope of the GDPR by asking '*Can inferences be defined as personal data under Art. 4(1) GDPR and does the GDPR apply to inferences?'*. It was interpreted that inferences can fulfil all four elements required for data to constitute personal data under Art. 4(1) GDPR, which is partly due to the broad scope the CJEU and the Art. 29 WP ascribe to the definition. Important conclusions found during the interpretation were that, for one, even if inferences are inaccurate, they can constitute 'any information', as the element itself does not require information to be accurate. Secondly, inferences are able to relate to a natural person either by result, content or purpose. Thirdly, it was analyzed, that while inferences can be the output of a legal analysis, those themselves are not comparable to the process of a legal analysis, and therefore the uncertainty some authors might have experienced in applying the CJEU *YS and others* judgment to inferences, should be resolved. Fourthly, the dynamic characteristic of the personal data definition should not be underestimated, as it needs to be assessed in each individual case, whether the drawn inferences can fulfill all four elements of the definition. Thus, a general claim that inferences constitute personal data should be avoided and rather a statement that the definition provides for the possibility to treat inferences as personal data, should be chosen. Furthermore, the chapter

demonstrated the risks, of inferring sensitive personal data from ordinary input data, highlighting the possible risk of everything becoming a special category of data, due to the increasing power of AI technologies.

The third sub-question was formulated in order to apply provisions of the GDPR to the case of inferences drawn by AI to ascertain if and how this could mitigate the implications of inferences raised in chapter two: *Are the data subject rights able to sufficiently address inferences drawn in order to mitigate the implications posed for individuals?* The fourth chapter was applied the GDPR under three questions to be resolved. First, can the information requirements under Art. 13, 14, 15 GDPR raise awareness in individuals regarding inferences drawn about them? Second, are the provisions of the GDPR on the data subject rights able to put data subjects in the position to prevent inaccurate inferences to be drawn, from existing and to be further used? Third, how does Art. 22 GDPR on automated decision-making and profiling impact inferences drawn by AI? All questions were answered positively under one condition: It depends on how detailed the information, controllers are obligated to provide to the data subjects, under Art. 13, 14 and 15 GDPR will be. The information must be detailed and encompassing to that extent that a data subject is made aware of who aims to draw inferences, who drew existing inferences, who is processing drawn inferences for what purposes and what the inferences state about the individual. Only by having this oversight, a data subject can sufficiently exercise all his data subject rights under Art. 16-21 GDPR to tackle inaccurate inferences and therefore, mitigate the risks they pose for them. The impact Art. 22 GDPR can have on inferences drawn will to a large extent depend on what is to be understood under 'meaningful information about the logic involved', whether this also includes information about inferences drawn or not. However, it is to be expected that most cases where inferences get drawn will escape the scope of Art. 22 GDPR by not complying with the element of being 'solely' automated.

This thesis did not intend to add to the discussion of whether rights, such as the "right to explanation", "right to reasonable inferences" and "right to legibility" exist, or what form they would take, if they are existent or to be included in the GDPR. The scope of this thesis was limited to exploring the text of the GDPR in its current state, focusing on interpreting specific individual provisions in light of inferences drawn. Furthermore, the author acknowledges, that the above analysis of applying the GDPR to inferences

drawn and its possible subsequent impact it can have on controllers, aiming to draw or use inferences, and individuals, subject to the drawing of inferences or use thereof, is done without taking into account some important barriers. Intellectual property rights and trade secrets, among others, will likely play an important role in regard to the impact, the GDPR will have on the challenges posed by inferences.

In light of the above said, the main research question must be answered in concluding that the extent to which the GDPR will be able to address the implications posed by inferences drawn by AI for individuals, depends to a large extent on how broad and encompassing the information required under Articles 13-15 GDPR will be interpreted. The GDPR entails the capacity to address and mitigate the implications; it is now up to jurisprudence and practice to interpret the provisions of the GDPR in a sufficient way. The found answer is accompanied by the conclusion, that the GDPR provides the right basis for tackling issues of inferences drawn by AI, without the need for new rights to being implemented into the framework, nor looking at a different framework other than the GDPR entirely.

I will conclude with a statement which perfectly fits the point this thesis aimed to prove: "The GDPR can be a toothless or a powerful mechanism to protect data subjects dependent upon its eventual legal interpretation: the wording of the regulation allows either to be true."[365]

[365] Mittelstadt, B. et al. 'The Ethics of Algorithms: Mapping the Debate' [2017] 3(2) Big Data & Society <https://ssrn.com/abstract=2909885>, page 14.

Bibliography

Primary Sources

Law

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

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, OJ 1995 L 281

Case Law

CJEU, Joined Cases C-141/12 and C-372/12, YS v. Minister voor Immigratie, Integratie en Asiel, and Minister voor Immigratie, Integrie en Asiel v. M, S [2014] ECLI:EU:C:2014:2081.

CJEU, Case C-524/06, Heinz Huber v. Bundesrepublik Deutschland [2008] ECLI:EU:C:2008:724.

CJEU, Case C-582/14, Patrick Breyer v. Bundesrepublik Deutschland [2016] ECLI:EU:C:2016:779.

CJEU, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:994.

CJEU, Case C-553/07, College van burgemeester en wethouders van Rotterdam v. M.E.E. Rijkeboer [2009] ECLI:EU:C:2009:293.

Opinion of Advocate General Kokott, Case C-434/16, Peter Nowak v. Data Protection Commissioner [2017] ECLI:EU:C:2017:582.

AG Opinion, Case C-582/14, Patrick Breyer v. Bundesrepublik Deutschland [2016] ECLI:EU:C:2016:339.

AG Opinion, Case C-131/12, Google Spain SL, Google Inc. v. Agencia Espanola de Protección de Datos (AEPD), Mario Costeja González, [2013] ECLI:EU:C:2013:424.

Opinion of Advocate General Sharpston, Joined Cases C-141/12 and C-372/12, YS v. Minister voor Immigratie, Integratie en Asiel, and Minister voor Immigratie, Integrie en Asiel v. M, S [2013] ECLI:EU:C:2013:838.

Secondary Sources

Abrams, Martin 'The Origins of Personal Data and its Implications for Governance' [2014] The Information Accountability Foundation, https://ssrn.com/abstract=2510927.

Amnesty International UK, ‚Trapped in the Gangs Matrix' [2018], < https://www.amnesty.org.uk/trapped-gangs-matrix>, last accessed 15.05.2020.

Anrig, Bernhard/ Browne, Will/ Gasson, Mark 'The Role of Algorithms in Profiling' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010).

Article 29 Working Party, Guidelines on the right to data portability, WP 242 rev.01, 05 April 2017, https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233.

Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679', (17/EN, WP251rev.01), 6 February 2018, http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826.

Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007 (WP 136), < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>.

Article 29 Data Protection Working Party, 'Advice paper on special categories of data ("sensitive data")', 20 April 2011, https://ec.europa.eu/justice/article-29/documentation/other-document/files/2011/2011_04_20_letter_artwp_mme_le_bail_directive_9546ec_annex1_en.pdf.

Article 29 Working Party, Annex – health data in apps and devices to the Advice paper on special categories of data ("sensitive data"), April 2011, < https://ec.europa.eu/justice/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf>.

Article 29 Working Party, Opinion 8/2014 on the on Recent Developments on the Internet of Things, Adopted on 16 September 2014, 14/EN WP 223, < https://www.pdpjournals.com/docs/88440.pdf>.

Barocas, Solon and Selbst, Andrew D., 'Big Data's Disparate Impact' [2014] 104 California Law Review 671 (2016), < https://ssrn.com/abstract=2477899>.

Borgesius, Frederik Zuidverveen 'Case Note: Breyer Case of the Court of Justice of the European Union: IP addresses and the personal data definition' [2017] 3(1) European Data Protection Law Review (2017), < https://ssrn.com/abstract=2933781>.

Borgesius, Frederik Zuidverveen ‘Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence’ [2020] The International Journal of Human Rights, < https://ssrn.com/abstract=3561441>.

Bloch, Daniel Alexandre, Machine Learning: Models and Algorithms, Quantitative Analytics (2018) < https://ssrn.com/abstract=3307566>.

Brkan, Maja ‘Do Algorithms rule the world? Algorithmic decision-making in the framework of the GDPR and beyond’ [2017] Proceedings of the 16th International Conference on Artificial Intelligence and Law, < https://ssrn.com/abstract=3124901>.

Bygrave, Lee A. ‘Information Concepts in Law: Generic Dreams and Definitional Daylight’ [2015] 35(1) Oxford Journal of Legal Studies (2015) <https://doi.org/10.1093/ojls/gqu011> accessed 09 June 2020.

Canhoto, Ana and Backhouse, James ‘General Description of the Process of Behavioural Profiling’ in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010).

Citron, Danielle Keats and Pasquale, Frank 'The Scored Society: Due Process for Automated Predictions' [2014] 89 Washington Law Review, p 1- U of Maryland Legal Studies Research Paper No 2014-8 <https://ssrn.com/abstract=2376209> accessed 27 January 2020.

Custers, B. 'Profiling As Inferred Data Amplifier Effects and Positive Feedback Loops' in Bayamlıoğlu, Emre/ Baraluic, Irina/ Janssens, Liisa and Hildebrandt, Mireille (eds), Being Profiled: Cogitas Ergo Sum 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018), <https://ssrn.com/abstract=3466857> accessed 27 January 2020.

Custers, Bart, The Power of Knowledge: Ethical, Legal and Technological Aspects of Data Mining and Group Profiling in Epidemiology (Wolf Legal Publishers 2004), https://ssrn.com/abstract=3186639.

Dalla Corte, Lorenzo ‘Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law’ [2019] 10(1) European Journals of Law and Technology < http://ejlt.org/article/view/672/909> accessed 20 January 2020.

Danks, David and London, Alex John 'Algorithmic Bias in Autonomous Systems' [2017] Twenty-Sixth International Joint Conference on Artificial Intelligence <DOI: 10.24963/ijcai.2017/654> accessed 27 January 2020.

DeBeasi, Paul ‘Training versus inference’ [2019] Gartner Blog Network, https://blogs.gartner.com/paul-debeasi/2019/02/14/training-versus-inference/, last accessed 15 June 2020.

Edwards, Lilian, ‘Data Protection: Enter the General Data Protection Regulation’ [2018] Forthcoming in L Edwards ed Law, Policy and the Internet (Hart

Publishing, 2018) < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3182454> accessed 09 June 2020.

Edwards, Lilian and Veale, Michael 'Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For' [2017] 16 Duke Law & Technology Review 18 https://ssrn.com/abstract=2972855.

Edwards, Lilian and Veale, Michael 'Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?' [2018] 16(3) IEEE Security & Privacy, < https://ssrn.com/abstract=3052831>.

European union agency for fundamental rights and Council of Europe, Handbook on European data protection law (2018 edn., Publications Office of the European Union 2018).

European Commission, 'Artificial Intelligence – A European Perspective' [2018] Joint Research Center EUR 29425 EN, < https://publications.jrc.ec.europa.eu/repository/bitstream/JRC113826/ai-flagship-report-online.pdf>.

European Data Protection Supervisor, Opinion 7/2015, Meeting the challenges of big data, A call for transparency, user control, data protection by design and accountability, 19 November 2015, < https://edps.europa.eu/sites/edp/files/publication/15-11-19_big_data_en.pdf>.

Ferretti, Agata et al. 'Machine Learning in Medicine: Opening the New Data Protection Black Box' [2018] 4(3) European Data Protection Law Review https://doi.org/10.21552/edpl/2018/3/10.

Gellert, Raphael 'Data Protection and notions of information: a conceptual exploration' [2018] Working Paper, last updated 06.11.2018 < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3284493>.

Goodfellow, Ian et al. Deep Learning (MIT Press 2016), 104.

Goodman, Bryce and Flaxman, Seth 'European Union regulations on algorithmic decision-making and a "right to explanation"' [2017] 38(3) AI Magazine, < https://arxiv.org/ct?url=https%3A%2F%2Fdx.doi.org%2F10.1609%2Faimag.v38i3.2741&v=f2c797e9>.

Graef, Inge/ Gellert, Raphael/ Purtova, Nadya/ Husovec, Martin 'Feedback to the Commission's Proposal on a framework for the free flow of non-personal data' [2018] < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3106791>.

Hallinan, Dara and Borgesius, Frederik Zuiderveen 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' [2020] 10(1) International Data Privacy Law (2020) < https://doi.org/10.1093/idpl/ipz025> accessed June 2020.

Hand, David et al. Principles of Data Mining (A Bradford Book The MIT Press 2001).

Hildebrandt, Mireille 'Defining Profiling: A New Type of Knowledge?' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010).

Hof, Simone van der and Prins, Corien 'Personalisation and its Influence on Identities, Behavior and Social Value' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010).

Hoofnagle, Chris Jay/ Sloot, Bart van der/ Borgesius, Frderik Zuiderveen 'The European Union general data protection regulation: what it is and what it means' [2019] 28(1) Information & Communications Technology Law, https://doi.org/10.1080/13600834.2019.1573501.

Humerick, Matthew 'Taking AI Personally: How the EU must learn to balance the interests of personal data privacy and artificial intelligence' [2018] 34(4) Santa Clara High Technology Law Journal < https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=1633&context=chtlj>.

Independent High-Level Expert Group On Artificial Intelligence, 'A Definition of AI: Main capabilities and disciplines' [2019] Set up by the European Commission https://ec.europa.eu/futurium/en/ai-alliance-consultation.

Information Commissioner`s Office (ICO) (2017), 'Big data, artificial intelligence, machine learning and data protection', https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf accessed 20 January 2020, last accessed 15 June 2020.

Jaquet-Chiffelle, David-Olivier 'Defining Profiling: A New Type of Knowledge?' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010).

Kamarinou, Dimitra et al. 'Machine Learning with Personal Data' [2016] 247 Queen Mary University of London, School of Law, Legal Studies Research Paper < https://ssrn.com/abstract=2865811>.

Kamp, Meike/ Körffer, Barabara/ Meints, Martin 'Profiling of Customers and Consumers – Customer Loyalty Programmes and Scoring Practices' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010).

Kindt, Els 'Biometric Profiling: Opportunities and Risks' in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010).

Kiseleva, Anastasiya 'AI as a Medical Device: Is It Enough to Ensure Performance Transparency and Accountability in Healthcare?' [2019] European

Pharmaceutical Law Review (1/2020), https://ssrn.com/abstract=3504829, accessed 13 January 2020.

Korff, Douwe 'Data Protection laws in the EU: The difficulties in meeting the challenges posed by global social and technical developments' [2010] Working Paper 2, European Commission Directorate-General Justice, Freedom and Security, 20 January 2010 < https://dx.doi.org/10.2139/ssrn.1638949> accessed June 2020.

Malgieri, Gianclaudio and Comandé, Giovanni 'Sensitive-by-distance: quasi-health data in the algorithmic era' [2017] 26(3) Information & Communication Technology Law, https://doi.org/10.1080/13600834.2017.1335468.

Malgieri, Gianclaudio and Comandé, Giovanni 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' [2017] 7(4) International Data Privacy Law, < https://ssrn.com/abstract=3088976>.

Marks, Mason 'Algorithmic Disability Discrimination' [2019] I. Glenn Cohen et al., Eds., Title TBD Cambridge University Press, Forthcoming, < https://ssrn.com/abstract=3338209>.

Marks, Mason 'Artificial Intelligence Based Suicide Prediction' [2019] 18(3) Yale Journal of Health, Policy, Law and Ethics, 21(3) Yale Journal of Law and Technology, https://ssrn.com/abstract=3324874.

Mendoza, Isak and Bygrave, Lee A. 'The Right not to be subject to automated decisions based on profiling' [2017] 20 University of Oslo Faculty of Law Legal Studied Research Paper Series No. 2017-20.

Mitrou, Lilian, 'Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) "Artificial Intelligence-Proof"?' [2019] University of the Aegean Dpt. of Information and Communication Systems Engineering; Athens University of Economics and Business - Department of Informatics < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914> accessed 09 June 2020.

Mittelstadt, Brent et al. 'The Ethics of Algorithms: Mapping the Debate' [2017] 3(2) Big Data & Society <https://ssrn.com/abstract=2909885> accessed 27 January 2020.

Moerel, Lokke and Wolk, Alex van der 'Big data analytics under the EU General Data Protection Regulation', [2017], https://ssrn.com/abstract=3006570.

Privacy International, 'Data is Power: Profiling and Automated Decision-Making in GDPR' [2017] <https://privacyinternational.org/sites/default/files/2018-04/Data%20Is%20Power-Profiling%20and%20Automated%20Decision-Making%20in%20GDPR.pdf.

Puaschander, Julia and Feierabend, Dieter ‘Artificial Intelligence in the Healthcare Sector’ [2019] 2(4) International Journal of Multidisciplinary Research <https://ssrn.com/abstract=3469423>.

Purtova, Nadezhda ‘The law of everything. Broad concept of personal data and future of EU data protection law’ [2018] 10(1) Law, Innovation and Technology (2018) < https://ssrn.com/abstract=3036355> accessed January 2020.

Ramirez, Edith ‘Privacy Challenges in the Era of Big Data: A View from the Lifeguard’s Chair 3’ https://www.ftc.gov/sites/default/files/documents/public_statements/privacy-challenges-big-data-view-lifeguard’s-chair/130819bigdataaspen.pdf.

Raso, Filippo A. et al. 'Artificial Intelligence & Human Rights: Opportunities & Risks' [2018] Bergmann Klein Center Research Publication No 2018-6 https://ssrn.com/abstract=3259344.

Roger, Allan and Price, W. Nicholson 'Privacy and Accountability in Black-Box Medicine' [2016] 23 Mich. Telecomm & Tech. L. Rev. 1 (2016) <https://ssrn.com/abstract=2758121> accessed 27 January 2020.

Russel, Stuart and Norvig, Peter, Artificial Intelligence, A modern approach (3rd edition., Pearson Education, Inc. 2010).

Schreurs, Wim et al. ‘Cogitas, Ergo Sum. The Rolfe of Data Protection Law and Non-discrimination Law in Group Profiling in the Private Sector’ in Hildebrandt, Mireille and Gutwirth, Serge (eds.), Profiling the European Citizen (Springer 2010).

Schwartz, Paul M and Solove, Daniel J. ‘Reconciling Personal Information in the US and EU’ [2013] 102 California Law Review 877 (2014); UC Berkeley Public Law Research Paper No. 2271442; GWU Legal Studies Research Paper No. 2013-77; GWU Law School Public Law Research Paper No. 2013-77 https://ssrn.com/abstract=2271442 accessed April 2020.

Selbst, Andrew D. and Powles, Julia, 'Meaningful Information and the Right to Explanation' [2017] 7(4) International Data Privacy Law 233 <https://ssrn.com/abstract=3039125> accessed 27 January 2020.

Siapka, Anastasia 'The Ethical and Legal Challenges of Artificial Intelligence: The EU response to biased and discriminatory AI' [2018] Panteion University of Athens <https://ssrn.com/abstract=3408773> accessed 13 January 2020.

Sloot, Bart van der and Borgesius, Frederik ‘Google and Personal Data Protection’ [2012] 22 n A. Lopez-Tarruella (Ed.), Google and the Law. Empirical Approaches to Legal Aspects of Knowledge-Economy Business Models. Series: Information Technology and Law Series, Vol. 22 VIII, T.M.C. Asser Press (Springer 2012), < https://ssrn.com/abstract=2146968>.

U.S. Food and Drug Administration (FDA), 'Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)', https://www.fda.gov/media/122535/download.

Wachter, Sandra and Mittelstadt, Brent 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2019](2) Columbia Business Law Review https://ssrn.com/abstract=3248829.

Wachter, Sandra et al. 'Counterfactual explanations without opening the black-box: Automated decisions and the GDPR' [2017] 31(2) Harvard Journal of Law & Technology (2018), https://ssrn.com/abstract=3063289.

Wachter, Sandra/ Mittelstadt, Brent/ Floridi, Luciano 'Why a Right to Explanation of Automated Decision-Making Does Not Exists in the General Data Protection Regulation' [2017] International Data Privacy Law < https://ssrn.com/abstract=2903469>.

Wiedemann, Klaus 'Automated Processing of Personal Data for the Evaluation of Personality Traits: Legal and Ethical Issues' [2018] Max Planck Institute for Innovation and Competition Research Paper No. 18-04, https://ssrn.com/abstract=3102933.

Zarsky, Tal 'Incompatible: The GDPR in the Age of Big Data' [2017] 47(4) Seton Hall Law Review (2017) < https://ssrn.com/abstract=3022646> accessed June 2020.