DO ALGORITHMS RULE THE WORLD? ALGORITHMIC DECISION-MAKING AND DATA PROTECTION IN THE FRAMEWORK OF THE GDPR AND BEYOND

Dr. Maja Brkan
Assistant Professor
Faculty of Law, Maastricht University, The Netherlands
maja.brkan@maastrichtuniversity.nl

1 INTRODUCTION

'This is beautiful,' acclaimed Fan Hui, the multiple European Go champion, when he saw Google's AlphaGo making an unconventional move during a Go game.[1] AlphaGo eventually won the ancient game, much more complex than chess, and proved that machines could outperform humans…again.[2] With the rise of intelligent machines, the input of big data and more or less complex algorithms, autonomous artificially intelligent agents are becoming more and more powerful and enabling numerous decisions to be taken entirely automatically. The rapid increase of the use of algorithms, processing large amounts of data, in financial, banking and insurance services, in medical services, in public administration and on the stock markets, offers infinite possibilities of invention of new machine-learning algorithms or configuration of existing ones, such as decision trees or neural networks. The use of these algorithms boosts companies' speed and arguably (!) also the accuracy of decision-making; at the same time, algorithmic decision-making can lead to biased decisions, in particular when sensitive data such as race, ethnic origin or sexual orientation are at stake.

Artificial Intelligence (AI) is expected to be the major trigger for the 'fourth industrial revolution' that is predicted to change the way our society functions and how humans relate to each other, to alter the job market and job demands as well as the entre industries which will take the path of digitalisation. The European and global society is witnessing an exponential technological advancement in the field of Big Data (BD) and Artificial Intelligence. In the recent years, technical developments in the field of robotics and appertaining software have seen an undreamed-of progress, ranging from humanoid, autonomous and care robots (such as Paro therapeutic robot[3]), autonomous vehicles, care robots, robot nannies and toys, robotic assistants to AI agents used for predictive policing or medical diagnosis and more. Other examples of deployment of AI are numerous, such as AI-supported voice-generating features in smartphones such as Siri; personal assistants such as Alexa; voice, facial and pattern recognition; automated profiling which enables companies to send targeted advertising to their consumers; finally, media are increasingly reporting about quantum computing.[4] The robots help paralysed people to walk and, in 2016, the first autonomous robotic surgery took place.[5] The latest trends in robotic

[1] See Cade Metz, 'The Sadness and Beauty of Watching Google's AI Play Go <https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/> accessed 21 November 2016.

[2] For example, an AI outperformed a human also in playing chess (AI Deep Blue) and Jeopardy (AI Watson).

[3] <http://www.parorobots.com/> accessed 11 July 2017.

[4] Adams, R. L., '10 Powerful Examples Of Artificial Intelligence In Use Today', Forbes, 10 January 2017; <https://www.forbes.com/sites/robertadams/2017/01/10/10-powerful-examples-of-artificial-intelligence-in-use-today/#2aba3ab6420d> accessed 10 July 2017.

[5] K.G. Orphanides, 'Robot carries out first autonomous soft tissue surgery' <http://www.wired.co.uk/article/autonomous-robot-surgeon> accessed 25 July 2017.

innovation and industry have enormous business potentials as well as ethical and legal limitations. BD & AI are at the top of the EU's agenda for digitalising European economy through Digital Single Market[6] and the EU institutions are pioneering in the establishment of clear legal and ethical guidelines for the AI. The European Parliament, with its recently adopted Resolution on Civil Law Rules on Robotics,[7] seeks to formulate legal and ethical standards for robots. Moreover, the EU has already built a solid legal framework for data protection any cybersecurity that could be applied to BD & AI, including the General Data Protection Regulation, Network Security Directive[8] and the proposed ePrivacy Regulation.[9]

However, the crucial role that the EU plays in this field raises many unanswered questions. While European companies exponentially use BD & AI in their business models, this approach not only needs to be embedded into a clear and concise EU legal framework providing for privacy,[10] transparency[11] and accountability,[12] but also has to respect ethical rules.[13] Transformation into a true 'BD & AI society' with broader use of Big Data mining and AI agents is only possible if the users trust these agents. This links back to legal, technological, economic and ethical queries: if technology is designed in a way to generate trust and enable compliance with legal requirements of transparency and accountability, while respecting high ethical demands, it leads to its increased use by businesses and higher competitiveness on the EU digital single market.

Against this backdrop, the purpose of this paper is to analyse the rules of the EU General Data Protection Regulation (GDPR)[14] on automated decision making in the age of Big Data and to explore how to ensure transparency of such decisions, in particular those taken with the help of algorithms. The paper thus analyses the rules of the GDPR and the Directive on Data Protection in Criminal Matters[15] on automated individual decision-making; the relevant provisions of the EU legislation regulating data protection are taken under the loop and their consequences for data subjects are examined. The paper further elaborates on the necessity of safeguards in automated decision-making, such as providing data subject with an explanation of an automated decision, guaranteeing algorithmic transparency and determining accountability in automated decision-

[6] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A Digital Single Market Strategy for Europe (COM/2015/192 final).

[7] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)).

[8] Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union (OJ L 194, 19.7.2016, p. 1).

[9] Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications) (COM/2017/010 final).

[10] Ryan Calo, 'Peeping HALs: Making Sense of Artificial Intelligence and Privacy' (2010) 2 *European Journal of Legal Studies* 3, http://www.ejls.eu/6/83UK.htm.

[11] Bryce Goodman, Seth Flaxman, 'European Union regulations on algorithmic decision-making and a "right to explanation"', *ICML Workshop on Human Interpretability in Machine Learning*, New York (2016).

[12] Andrea Bertolini, 'Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules' (2013) 5 *LIT* 2, 214–247.

[13] Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, Luciano Floridi, 'The ethics of algorithms: Mapping the debate' (2016) 3 *Big Data & Society* 1, 1–21.

[14] 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 L 119, 4.5.2016, p. 1.

[15] Directive (EU) 2016/680 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 by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA, L 119, 4.5.2016, p. 89.

making. Obstacles to algorithmic transparency are discussed, paying particular attention to technical obstacles, obstacles due to intellectual property and those relating to state secrets and other confidential information. Before concluding remarks, the paper puts forward arguments as to how the provisions of the GDPR will be relevant globally for businesses established in the US and other parts of the world.

2 AUTOMATED DECISION-MAKING

Automated decision-making could be defined as taking a decision without human intervention; according to the GDPR, 'automated individual decision-making' is 'a decision based solely on automated processing'.[16] The human can of course feed the system with data – although even this can be an automatic procedure – and interpret the decision once it is taken. If the automated decision-making does not have any binding effect on data subjects and does not deprive them of their legitimate rights, such decision-making is of a low impact. However, when a decision is binding for individuals and affects their rights, by deciding for example that a client should be awarded credit, tax return or to be employed, the law has to provide sufficient safeguards to protect this individual.[17]

Automated decision-making seems to encompass a multitude of decision types, ranging from displaying search results, profiling,[18] high-frequency trading,[19] decisions on granting of a loan by a bank, administrative decisions[20] (such as deciding which company to check for tax purposes) and to a certain extent even judicial decisions.[21] The notion of automated decision-making is not a unitary concept, comprising only a particular type of decisions. Rather, it is broad, multifaceted and prone to be divided into several sub-categories. Before analysing the provisions of the GDPR and the Directive on Data Protection in Criminal Matters, it is important to distinguish between procedural and substantive automated decision-making; algorithmic and non-algorithmic automated decision-making; and rule-based as opposed to law-based decisions.

Procedural/substantive. The procedural/substantive divide does not refer to taking procedural or substantive decisions; it rather means that automated decisions will have to be adopted in a way that guarantees procedural and substantive fairness and accurateness. The requirement of procedural fairness requires that all decisions relating to the same or comparable facts are taken

[16] Article 22(1) of the General Data Protection Regulation.

[17] See further on efficiency and fairness in automated decision-making Tal Zarsky, 'The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making' (2016) 41 Science, Technology, & Human Values 118-132.

[18] More on profiling see Mireille Hildebrandt, Serge Gutwirth (Eds.), Profiling the European Citizen. Cross-Disciplinary Perspectives (Springer 2008).

[19] Frank Pasquale, The Black Box Society. The Secret Algorithms That Control Money and Information (Harvard University Press 2015); Jacob Loveless et al., 'Online Algorithms in High-frequency Trading. The challenges faced by competing HFT algorithms' (2013) 11 acmqueue 1.

[20] Melissa Perry, 'iDecide: Administrative Decision-Making in the Digital World' (2017, forthcoming) Australian Law Journal, courtesy of the author.

[21] See Trevor Bench-Capon, Thomas F. Gordon, 'Tools for Rapid Prototyping of Legal Cased-Based Reasoning' (2015) ULCS-15-005, University of Liverpool, United Kingdom; Giovanni Sartor, Luther Branting (Eds.), *Judicial Applications of Artificial Intelligence* (Springer, 1998); Angèle Christin, Alex Rosenblat, Danah Boyd, 'Courts and Predictive Algorithms' (2015) Data & Civil Rights: A New Era of Policing and Justice <http://www.law.nyu.edu/sites/default/files/upload_documents/Angele%20Christin.pdf> accessed 16 January 2017.

according to the same automated procedure.[22] This procedural fairness is closely linked with substantive fairness since it would lead to the result that the same cases would have the same outcome. However, decisions also have to be substantively fair, meaning that they should not be discriminatory in any way, especially not decisions taken on the basis of algorithms.[23]

Algorithmic/non-algorithmic. Algorithmic decision-making is automated decision-making with the support of algorithms. There is no common definition of the notion of algorithm across literature. However, it has to be specified that, in automated decision-making, we are dealing with computer algorithms that can be defined as 'a set of steps to accomplish a task that is described precisely enough that a computer can run it'.[24] Many – if not most – automated decisions nowadays are taken with a support of algorithms. With the increasing use of big data and more and more complex decision-making, algorithmic intervention has become almost indispensable.

Rule-based/law-based automated decisions. In fact, both 'rule-based' and 'law-based' decisions are taken on the basis of rules, but the source of the rule for both types of decisions is different. For rule-based decisions the rule is mostly an outcome of a business decision, for example profiling for the purposes of targeted advertising (e.g. a company sending an advertisement about vacation in Bali to all people searching for vacation in Asia). The law-based decisions are based on a legal rule that is binding on everyone. An example of a rule that is prone to automated decision, is a rule prescribing that everyone who exceeds the speed limit gets a fine. Unless the law-based rule is very clear and precise, automated decisions based on law have to face a challenge of law's open texture and notions requiring interpretation. Autonomous decision-making presupposes that the rules needed to be applied are not prone to interpretation and do not leave to the decision-maker much or any discretion in taking the decision.

3 GDPR'S TAKE ON AUTOMATED INDIVIDUAL DECISION-MAKING

This section contains an in-depth analysis of the GDPR provision on automated decision-making, explaining the circumstances in which such decision-making is possible according to the GDPR and under which conditions. It does not come as a surprise that the GDPR, in its Article 22, regulates automated individual decision-making, including profiling. According to the first paragraph of this provision,

> '[t]he data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.'

This provision continues the legacy of the Data Protection Directive,[25] more precisely its Article 15, according to which the data subject equally had the right not to be subject to a decision producing legal effects or significantly affecting him and which is based solely on automated processing of data. While the wording of the provision did not undergo substantial changes with the adoption of the GDPR, the practical importance of the provision increased with augmented

[22] Every decision in administrative procedure will fall under procedural administrative law – this means that the procedural laws will have to be amended to give some new procedural rules for automated decision-making.

[23] On algorithmic discrimination see for example Bryce Goodman, 'Discrimination, Data Sanitisation and Auditing in the European Union's General Data Protection Regulation' (2016) European Data Protection Law Review 493.

[24] Thomas H. Coormen, *Algorithms Unlocked* (MIT Press 2013) 1.

[25] 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 L 281, 23.11.1995, p. 31.

use of automated decision-making. The Data Protection Directive also contained some examples of automated individual decisions, namely decisions 'to evaluate certain personal aspects relating to him, such as his performance at work, creditworthiness, reliability, conduct, etc'. These examples demonstrate that the provision of the Data Protection Directive seemed to focus mostly on instances of profiling based on automated processing of data, not including other types of automated decision-making involving processing of personal data.

Against this backdrop, it is interesting to observe how Article 22 GDPR developed throughout the legislative procedure leading to the adoption of the GDPR as it shows the evolution from focusing specifically on profiling to a more general formulation using a broader notion of automated individual decision-making. Differently from the final GDPR, in the initial Commission's proposal, this article, titled 'Measures based on profiling',[26] regulated profiling based on automated processing and not generally automated decision-making as the provision in the final GDPR does. Moreover, the initial provision contained a separate paragraph on the obligation to inform the data subject about the existence of automated processing and 'the envisaged effects of such processing on the data subject'.[27] Differently from the current provision, the scope of application of this provision thus seems to be more limited since it only applied for profiling; moreover, the obligation to inform the data subject about such processing was moved to Articles 13 and 14 under the general obligation that needs to be provided to the data subject. While in the first reading in the European Parliament, the provision kept the focus on profiling and added the right to human intervention regarding profiling, the paragraph on informing the data subject about the envisaged effects of profiling was deleted.[28] In the first reading in the Council, the provision then took the shape of the current Article 22 GDPR, not restricting the scope of the article merely on profiling, but rather including profiling into a more general category of individual automated decision-making.[29] Nevertheless, as mentioned in the GDPR proposal,[30] this provision still takes into consideration the Recommendation on profiling issued by the Council of Europe.[31]

Regardless a broader formulation of the GDPR, it is questionable to what extent the scope of application of Article 22 GDPR covers decision-making that is indeed broader than decisions based on profiling.[32] The data subject has a right not to be subject to a decision based exclusively on

[26] Article 20 of GDPR proposal, COM(2012) 11 final.

[27] Ibid, Article 20(4).

[28] Article 20 (Profiling), European Parliament legislative resolution of 12 March 2014 on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) (COM(2012)0011 – C7-0025/2012 – 2012/0011(COD)) (Ordinary legislative procedure: first reading).

[29] Position of the Council at first reading with a view to the adoption of a Regulation of the European Parliament and of the Council 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) - Adopted by the Council on 8 April 2016.

[30] Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) COM(2012) 11 final, p. 9.

[31] Recommendation CM/Rec(2010)13 of the Committee of Ministers to member states on the protection of individuals with regard to automatic processing of personal data in the context of profiling (Adopted by the Committee of Ministers on 23 November 2010 at the 1099th meeting of the Ministers' Deputies).

[32] Isak Mendoza, Lee A. Bygrave, 'The Right not to be Subject to Automated Decisions based on Profiling', University of Oslo Faculty of Law Legal Studies Research Paper Series No. 20/2017, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2964855> accessed 11 July 2017, p. 7, rightly point out that the legislative process leading to the adoption of the GDPR leads to the result that it does not give the right to the data subject to object to all profiling, but only to certain types of decisions arising from profiling.

automated processing, and profiling is a type of processing mostly leading to such decisions. According to the GDPR, profiling means processing of personal data in a way to use it to 'evaluate certain personal aspects relating to a natural person', such as for example 'to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements' (Article 4). It is difficult to imagine examples where person's personal data *not* leading to profiling would lead to an automated decision. A potential example would be automated application of tax rules in order to determine how much tax return a tax resident would get. However, would that decision again not be based on her personal tax profile? There are automated decisions and predictions that do not involve profiling, such as high-frequency trading or predictions of outcomes of judicial decisions, but they do not involve processing of personal data and would thus not fall within the ambit of Article 22 GDPR.

Article 22 GDPR reflects, on the one hand, European scepticism towards biases and potentially false decisions that can be taken by automated means if they are not verified by humans. On the other hand, this provision, by giving certain guarantees to the data subject, notably the right to human intervention, addresses concerns around the lack of ability of data subjects to influence decisions which are to an increasing extent taken by automated means.[33] On the first impression, the general negative stance towards such automated decisions comes across as a forceful fortress for strongly protecting individuals and potentially even hampering the future development of AI in decision making. However, on a more comprehensive level of evaluation, it can be argued that this provision, containing numerous limitations and exceptions, looks rather like a Swiss cheese with giant holes in it.

3.1 Automated decisions barred by the GDPR and the Directive on Data Protection in Criminal Matters

In order for the data subject to have the right not to be subject to automated decision-making, the decision itself needs to fulfil certain requirements laid down by Article 22(1). However, before delving deeper into these conditions, it is crucial to analyse the nature of the 'right' of the data subject not to be subjected to automated decision-making.

The 'right' of the data subject not to be subject to automated decision-making can be understood either as a right that the data subject has to actively exercise or as a 'passive' right that the controllers taking an automated decision have to observe themselves without an active claim from the data subject. If the 'right' from Article 22(1) GDPR is constructed in a former way, the exercise of the right would depend on data subject's free will and her choice. Not choosing to exercise this right would, on a proper construction, lead to the result that automated decisions having the characteristics described in Article 22(1) could be lawfully taken. That would, for example, lead to the possibility to take a fully automated decision having legal consequences for the data subject, without providing her with necessary safeguards from paragraph 3 of that provision. Legal consequences of such a decision, taken as a result of a failure of the data subject to exercise her right, could therefore be rather detrimental for the data subject. On the other hand, data subjects' choice to exercise this right would have unclear legal consequences. For example, would exercise of this right be translated into the right to object, barring such automated decision altogether? Would it be understood as a request for human intervention?

[33] Isak Mendoza, Lee A. Bygrave, 'The Right not to be Subject to Automated Decisions based on Profiling', University of Oslo Faculty of Law Legal Studies Research Paper Series No. 20/2017, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2964855> accessed 11 July 2017.

Interpreting Article 22(1) as giving data subject a right that she has to actively exercise could therefore lead to detrimental effects for her and run contrary to the purpose of this provision which aims to protect data subject against a general possibility to subject her to automated decision-making. A systematic interpretation of Article 22 implies that only automated decisions fulfilling the requirements of paragraph 2 and allowing for safeguards from paragraph 3 of this provision are authorised by the GDPR. Therefore, as Mendoza and Bygrave correctly claim, it is more appropriate to construct the data subjects' 'right' as a prohibition of fully automated decision-making that the data controllers have to comply with.[34] Such interpretation of Article 22(1) aligns this provision to Article 11 of the Directive on Data Protection in Criminal Matters which gives the Member States a clear obligation to prohibit automated decisions having certain characteristics.

Constructing data subjects' 'right' as a general prohibition of certain types of automated decisions also sheds a different light on conditions from Article 22(1); on the basis of this reading, a decision having the following characteristics is prohibited by this provision: (1) the decision has to be individual (the same condition is imposed by the Directive), (2) it needs to be based solely on automated processing (the same goes for the Directive) and (3) it needs to have legal or significant effects on the data subject (the Directive contains an additional requirement of 'adverse' legal effects).

From that perspective, the first condition has to be understood prohibiting *individual* automated decisions, that is, decisions relating only to a particular natural[35] person, a single data subject. Individual decisions can be binding on an individual (such as a decision on loan application, credit card application, welfare and financial decisions, granting a visa, choosing taxpayer for audit) or non-binding (such as profiling, e.g. sending targeted advertisements to an air traveller on the basis of her profile). In line with the general scope of application *ratione personae* of the GDPR which covers the protection of natural persons (Article 1(1)) and hence regulates only the protection of individuals and not groups, the textual interpretation of Article 22 GDPR equally seems to exclude *collective* decisions affecting several natural persons or a group of those linked together either by virtue of their common characteristics, their belonging to a group or their living in a particular area.[36] The same reasoning can be put forward regarding Article 11 of the Directive on Data Protection in Criminal Matters which equally prohibits only individual decisions. An illustration of a collective decision in criminal matters is, for example, a machine-based decision taken by the police to increase police monitoring in a certain geographical area, affecting all data subjects residing in this area. A collective decision in non-criminal matters would, for instance, be a decision on dynamic pricing, selling certain product for a certain price to a category of data subjects belonging to a certain income bracket.

In their current wording, neither Article 11 of the Directive on Data Protection in Criminal Matters, nor Article 22 GDPR, read together with their respective Articles 1, would *ratione personae* not cover (and in consequence not prohibit) such a collective decision. Both the GDPR and the Directive seem to follow the logic that the underlying rationale for data protection of groups of data subjects differs from the rationale for data protection of an individual data subject. An argument that is sometimes put forward in this regard is that a collective decision is not necessarily linked to personal data of a particular individual, but can be easily based on anonymised data which

[34] Mendoza, Bygrave, 9.

[35] According to its Article 1(1), the GDPR applies only for natural and not legal persons.

[36] For a collective data protection aspect in the age of Big Data analytics see Alessandro Mantelero, 'Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection' (2016) 32 Computer Law & Security Review 2, 238-255.

would render EU data protection legislation inapplicable.[37] However, anonymisation of data is not sufficient as long as the data subject remains *identifiable*.[38] With an increasing importance and use of big data, re-identification of an individual appertaining to a certain group is significantly facilitated. Moreover, big data greatly enables group-related automated decision-making. Classifying data subjects into a specific category (man/woman, low/high income) enables collective decisions pertaining to this group. Excluding collective automated decisions from the scope of application of GDPR would not only create an enormous imbalance in how individual and collective automated decisions are treated, but could also open the door to circumvent the prohibition of individual automated decisions by adopting collective decisions whenever possible.[39] Therefore, in the light of the high level of protection of data subject, it is submitted that collective automated decisions should be covered by the scope of application of Article 22 GDPR and Article 11 of the Directive on Data Protection in Criminal Matters. A possible way to include such decisions into the scope of application of these two legal instruments is to consider the decision regarding a group as actually being a bundle of individual decisions. A purposeful (teleological) interpretation of Article 22 GDPR and Article 11 of the Directive on Data Protection in Criminal Matters, coupled with the need to guarantee the data subject a high level of safeguarding her fundamental right to data protection could lead the Court of Justice of the EU to adopt this interpretative stance.

Secondly, the GDPR and the Directive on Data Protection in Criminal Matters do not allow for a decision to be based *solely* on automated processing. Whether a decision is fully automated or not depends, in the first place, on whether human intervention is technically possible in the process of decision-making. For example, if the price of a product sold online is determined on the basis of data subject's income and the price is shown automatically on the website, without a human being involved in the process of determining the price, such a decision is surely based solely on automated processing. However, if the process allows for human involvement, it is to be verified whether the mere possibility that a human has the power to change a decision automatically renders this decision not to be based solely on automated processing. In other words, if a human merely rubberstamps an automated decision without verifying its correctness, can it be assumed that such a decision was not taken by fully automated means?

The answer to this question should be in the negative. Such a formalistic interpretation, involving the human only as a necessary part of procedure but ultimately leaving the decision power to the machine, would not ensure a sufficiently high enough level of data protection of the data subject. In order for the decision *not* to be based solely on automated processing, the human judgment needs to be such as to verify the machine-generated decision and the human should assess[40] the substance of the decision and not be involved merely as another (empty) procedural step. In other words, in order to escape the prohibition from Article 22 GDPR or Article 11 of the Directive on Data Protection in Criminal Matters, the human as to use the machine only as decision support, whereas the final decision is taken by the human.

Third, the GDPR and the Directive on Data Protection in Criminal Matters prevent only decision-making, including profiling, which produces legal effects (in case of Directive, 'adverse'

[37] According to Recital 26 GDPR and Recital 21 of the Directive on Data Protection in Criminal Matters, the 'principles of data protection should … not apply to anonymous information'.

[38] On identifiability, see Worku Gedefa Urgessa, 'The Protective Capacity of the Criterion of 'Identifiability' under EU Data Protection Law' (2016) 2 European Data Protection Law Review 4, 521 – 531.

[39] Given that a cluster of multiple data subjects does not necessarily constitute a group of data subject with the same or similar characteristics, this might not always be possible.

[40] Mendoza and Bygrave, 10, point out that the human has to 'actively assess the result of the processing prior to its formalisation as a decision'.

legal effects) for the individual or significantly affects the individual. Even though neither of the two legal instruments defines the notion of legal effects, it can be assumed that a decision having legal effects is a binding decision that impacts legal position or legal interests of data subject. For example, a decision of a tax authority on tax return of a particular data subject, calculated on the basis of her income, is a decision having legal effects relating to this data subject within the meaning of the GDPR. A decision taken by a police to interrogate a data subject or to seize her mobile device, taken on the basis of her personal data, is a decision having adverse legal effects on that person within the meaning of the said Directive. While it seems relatively straightforward to determine which decision would have legal effects on an individual, it is less clear what kind of decision making or profiling 'significantly affects' such an individual. GDPR gives examples of a refusal of an online credit application or the use of automated decision-making in e-recruiting practices. These are instances where the data subject acts as an applicant for credit card, insurance contract with a certain premium or a job position. However, establishing significant effect on data subject with regard to profiling seems less straightforward. For example, when does sending advertisements by Google and Facebook 'significantly affect' an individual? Given different potential impacts that such targeted advertising can have on data subject, it is close to impossible to clearly answer it in the affirmative or negative. For example, if the data subject ignores such targeted advertising and does not follow up on it, it is rather difficult to argue that the advertising 'significantly affects' this data subject. To the contrary, if a person systematically shapes his/her purchasing decisions on the basis of such targeted advertising, the significant effect would be more easily established. This of course raises the question whether, for the criterion of significant effect to be fulfilled, a causal link between the profiling and the action of the data subject would need to be required. Requirement of existence of a causal link would, on the one hand, ensure that only limited instances of targeted advertising would have significant effect on the data subject; on the other hand, the requirement of such a causal link would render the analysis of significant effect extremely complicated. An alternative test to establish significant effect in such cases would be to take as a benchmark an average consumer rather than the actual consumer on which the advertising was targeted.

3.2 AUTOMATED DECISIONS AUTHORISED BY THE GDPR AND THE DIRECTIVE ON DATA PROTECTION IN CRIMINAL MATTERS

The GDPR in Article 22(2) and the Directive on Data Protection in Criminal Matters in Article 11 expressly authorise certain types of automated decisions. According to the GDPR, the prohibition from paragraph 1 of that provision does 'not apply if the decision: (a) is necessary for entering into, or performance of, a contract between the data subject and a data controller; (b) is authorised by Union or Member State law to which the controller is subject …; or (c) is based on the data subject's explicit consent.' The Directive authorises only decisions 'authorised by Union or Member State law to which the controller is subject'.

The first possibility of automated decisions allowed by the GDPR are those that are *necessary* to enter into or perform a contract between data subject and data controller. If the meaning of this provision is to be constructed on the basis of a very strict textual interpretation, it is questionable whether it would ever open the door for automated decisions. For example, it can be argued that the conclusion of the insurance or loan contract necessitates an assessment of risk – but does this risk necessarily need to be assessed by automated means? Prices of flights are often determined through 'dynamic pricing', taking into account the profile of the potential buyer – but is such an automated determination of price really necessary for conclusion of performance of this purchase contract? Therefore, it is submitted that the 'necessity' requirement will have to be understood more as an 'enabling' requirement for the conclusion of a contract. If the automated assessment of a credit risk enables conclusion of a contract on the basis of which the data subject receives a credit card, such an assessment enabled the conclusion of this contract. Sometimes these contracts are termed 'algorithmic contracts'[41] and are ever more frequent in online trading, Amazon being the most used example.

Secondly, automated decisions and profiling are allowed if they are authorised by Union or Member State law that provide for sufficient safeguards to protect data subject's rights, freedoms and legitimate interests. An example of Union legislation potentially allowing for an automated decision with sufficient safeguards is the new PNR Directive.[42] While the Directive in principle does not allow for an automated decision 'that produces an adverse legal effect on a person or significantly affects a person' (Article 7(6)), it does provide for the possibility of automated matching or identification of persons who should be further examined by the competent authorities in view of potential involvement in terrorism, provided that such matching is individually reviewed by non-automated means.[43]

An example of a Member State law regulating automated decision-making is the recently adopted German law implementing the GDPR[44] which expressly allows for automated decisions in the field of insurance. On the one hand, an automated decision is allowed if it is taken in the framework of performance of an insurance contract and the request of the person in question was approved. As it is clarified by the explanations of this German law, this provision allows for such an automated decision in tortious (and not contractual!) relationship between the insurance company of the person who caused damage and the person who suffered damage, under the condition that the latter wins with her claim.[45] On the other hand, German law also allows for an automated decision about insurance services of a private health insurance when the decision is based on binding rules on remuneration for medical treatment.[46] Moreover, the German administrative law also allows for automated adoption of administrative acts in the framework of fully automated administrative procedure.[47]

Third, the GDPR also allows for automated decision-making if such a decision is based on the explicit consent of the data subject. In certain cases, notably with regard to decisions based on profiling, where the data subject has to give his consent online, it can be problematic whether the consent obtained online was indeed explicit or not. Profiling is often done without data subject even knowing about it[48] and if the data subject did not give explicit consent for profiling, she also did not consent with a decision taken on the basis of such profiling. For example, an explicit

[41] More on this type of contracts see Lauren Henry Scholz, 'Algorithmic Contracts' (2017) Stanford Technology Law Review, available at <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=274770> accessed 10 November 2016.

[42] Directive (EU) 2016/681 of the European Parliament and of the Council of 27 April 2016 on the use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime (OJ L 119, 4.5.2016, p. 132).

[43] Article 6(5) of the PNR Directive; compare also paragraph 2 of this provision, and Mendoza, Bygrave, 6.

[44] See § 37 (Automatisierte Entscheidungen im Einzelfall einschließlich Profiling) of the Gesetz zur Anpassung des Datenschutzrechts an die Verordnung (EU) 2016/679 und zur Umsetzung der Richtlinie (EU) 2016/680 (Datenschutz-Anpassungs- und -Umsetzungsgesetz EU – DSAnpUG-EU).

[45] See ibid., p. 106: explanations to § 37 (Automatisierte Entscheidungen im Einzelfall einschließlich Profiling)

[46] Ibid.

[47] See § 35a of Verwaltungsverfahrensgesetz (VwVfG) and explanation to § 37 of DSAnpUG-EU above.

[48] Article 29 Working Party, 'Advice paper on essential elements of a definition and a provision on profiling within the EU General Data Protection Regulation', adopted on 13 May 2013, <http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2013/20130513_advice-paper-on-profiling_en.pdf> accessed 11 November 2016.

consent to cookies should not necessarily mean consent to an automated decision based on such profiling. While the GDPR allows for profiling itself, provided that the GDPR requirements are respected,[49] the decisions based on profiling should conform to certain safeguards.[50]

3.3 Safeguards in automated decision-making: From revealing the logic behind the decision to accountability

3.3.1 Safeguards in Article 22 GDPR

In all cases when an automated decision is allowed, the data subject has to be provided with appropriate safeguards in order to prevent wrong or discriminatory decision or a decision that does not respect data subject's rights and interests. Whenever automated decision is authorised on the basis of Union or Member State law, this law also has to provide for 'suitable measures to safeguard the data subject's rights and freedoms and legitimate interests' (Article 22(2)(b)). In other two examples – conclusion of a contract and explicit consent – the GDPR equally requires such safeguards, but clarifies which minimum measures should be provided for: the data subject shall have at least (1) the right to obtain human intervention on the part of the controller, (2) to express his or her point of view and (3) to contest the decision.

The data subject always has a right to obtain human intervention, meaning that she has the right that the fully automated decision becomes non-automated through human intervention. For example, in the insurance contract the risk assessment is made by automated means, but the human assesses the results and takes the final decision. Sometimes it might be difficult to exercise this right in practice. For example, if the data subject concludes an online contract with dynamic pricing, how can she request human intervention if the website does not provide for that possibility? Furthermore, the data subject also has the right to express her point of view, albeit the GDPR does not clarify what the legal consequence should be if such an opinion is expressed. And finally, the data subject has the right to contest the decision. In practice that means that the procedure becomes adversarial and, in the light of this, it is questionable who should decide about such an objection of the data subject. If, for example, data subject gave his explicit consent to automated assessment of her credit rating and then objects to such a decision, would this objection need to be dealt with by the bank official handling the file, by another employee within this organisation or by an independent body?

3.3.2 The existence of the right to explanation?

In case of automated decisions involving personal data of the data subject, the GDPR obliges the controller to provide the data subject with 'meaningful information about the logic involved'

[49] See Recital 72 GDPR according to which '[p]rofiling is subject to the rules of this Regulation governing the processing of personal data, such as the legal grounds for processing or data protection principles'.

[50] Countries which specifically allow for profiling mostly require additional safeguards in this regard. Italy can be used as an example of a country which specifically allows for profiling, but the data subject has to be notified prior to processing of data aimed at profiling: See Guidelines on online profiling issued by *Garante per la protezione dei dati personali*; for a summary see <http://blogs.dlapiper.com/iptitaly/?p=56970> accessed 26 May 2017. Moreover, some countries, such as the Netherlands, even allow for ethnic profiling, which may be problematic both from data protection and non-discrimination perspective: For more on this issue see Simone Vromen, 'Ethnic profiling in the Netherlands and England and Wales: Compliance with international and European standards', Public Interest Litigation Project (PILP-NJCM) / Utrecht University, <https://pilpnjcm.nl/wp-content/uploads/2015/06/Research-project-B-FINAL-version.pdf> accessed 16 May 2017.

in such decision-making, regardless of whether personal data is collected from the data subject[51] or not (notification duties of the controller)[52] Moreover, within the framework of the right to access, the GDPR provides for a similar right of the data subject to receive not only information on the existence of automated decision-making, but also 'meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject'.[53] These provisions fit well within the broader framework of GDPR's quest for a high level of transparency which requires that the processing of personal data should be transparent to natural persons whose personal data are 'collected, used, consulted or otherwise processed'.[54] The principle of transparency of data processing, epitomised in Article 5(1)(a) GDPR, requires not only that the information to the data subject is 'concise, easily accessible and easy to understand',[55] but also that the data subject is informed 'of the existence of the processing operation and its purposes'.[56] Given the circumstance that the transparency within the GDPR relates to the particular individual and not to the society at large, it can be understood as 'individual transparency' as it, in principle, gives the data subject rights of access, explanation and understanding the reasons behind a decision in case of automated processing. EDPS correctly points out that it is not up to individuals to seek disclosure of such logic, but that the organisations have to proactively seek for such transparency.[57]

This quest for transparency, however, raises several questions: what exactly needs to be revealed to the data subject? Does revealing meaningful logic mean that the data subject has the right to explanation of the automated decision? If yes, how detailed does the explanation have to be? It is to be noted that the explicit right to explanation is not mentioned either in Article 22 GDPR or in Articles 13 and 14 on notification duties, giving the data subject the right to obtain meaningful information about the logic involved. The only instance where the right of explanation is mentioned in the GDPR is its Recital 71, according to which processing under Article 22:

> "should be subject to suitable safeguards, which should include … the right to obtain human intervention, to express his or her point of view, *to obtain an explanation of the decision reached after such assessment* and to challenge the decision."[58]

There is a vigorous discussion in the academic literature whether such a right to explanation should indeed be given to the data subject. Goodman and Flaxman ignited the by inferring such right from the requirement to give the data subject meaningful information about the logic involved (Articles 13 and 14).[59] Wachter et al. claim that the GDPR only requires an *ex ante* explanation of

[51] Article 13(2)(f) GDPR.

[52] Article 14(2)(g) GDPR.

[53] Article 15(1)(h) GDPR.

[54] Recital 39 GDPR.

[55] Recital 58 GDPR.

[56] Recital 60 GDPR.

[57] 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' <https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2015/15-11-19_Big_Data_EN.pdf> accessed 15 November 2016.

[58] Emphasis added.

[59] Bryce Goodman, Seth Flaxman, 'European Union regulations on algorithmic decision-making and a "right to explanation"' <https://arxiv.org/pdf/1606.08813v3.pdf> accessed 18 July 2017.

how the system functions and not *ex post* explanation of the reasons behind the decision.[60] Edwards and Veale accept the possibility of the right to explanation, but point out practical difficulties of its exercise from the perspective of machine learning algorithms.[61] Mendoza and Bygrave equally put forward arguments in favour of the right to explanation.[62]

It is submitted that the provisions of the GDPR should be interpreted in such a way as to give the data subject such a right to explanation and that the CoJ should follow this approach when deciding on this issue. 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.[63] This information should go beyond the information that needs to be offered to the data subject in all cases of data processing, such as the identity of a controller or the purposes for which personal data is processed.[64] Therefore, we submit that the meaningful information about the logic involved would ideally comprise: (a) information about the data that served as the input for automated decision, (b) information about the list of factors that influenced the decision, (c) information on the relative importance of factors that influenced the decision, and (d) a reasonable explanation about *why* a certain decision was taken (textual information). In reality and given numerous obstacles for (b) and (c) as elaborated further in this paper, the right to explanation would probably encompass 'only' textual information explaining crucial reasons for decisions. Several arguments in favour of such a right to explanation can be put forward.

First, the methodological approach that should be used in this regard is to interpret several GDPR provisions together. In the light of that, the provisions of the GDPR, more precisely, Article 22, read in the light of Recital 71, in combination with Articles 13(2)(f), 14(2)(g) and 15(1)(h) GDPR, should be interpreted in a way that they give the data subject the right to an *ex post* explanation of the automated decision. Such methodological grouping of different data protection provisions in order to create a certain right of data subject is not unusual in the case law of the CoJ. For example, in *Google Spain*, the Court relied on the combination of the right of access and the right to object from Directive 95/46[65] in order to judicially construct the right to erasure (popularly described as 'the right to be forgotten').[66] Contrary to the approach of Wachter et al. who analyse these provisions separately, it is submitted here that the CoJ, when interpreting the provisions of the GDPR, should read them together if it seeks to construct the right to explanation.

Second, in the absence of the data subject's right to explanation, her right to contest the decision taken by automated means would be entirely ineffective.[67] If the data subject wants to substantively contest such a decision, she needs to obtain information at least about the data that was used as an input for automated decision and a reasonable explanation of grounds for the decision. The right

[60] For a view that the right to explanation of an automated decision does not exist in the GDPR, see Sandra Wachter, Brent Mittelstadt, Luciano Floridi, 'Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation' available on <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2903469> accessed 18 May 2017.

[61] Lilian Edwards, Michael Veale, 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for', <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972855>.

[62] Mendoza, Bygrave, 16.

[63] Article 22(3) GDPR.

[64] See Articles 13 and 14 GDPR.

[65] 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 [1995] OJ L 281/31.

[66] More precisely, the CoJ relied on Article 12(b) and subparagraph (a) of the first paragraph of Article 14 of Directive 95/46; Case C-131/12 *Google Spain and Google* ECLI:EU:C:2014:317.

[67] See also Mendoza, Bygrave, 16.

to contest is a right relating to substance of the decision and it would be an empty shell if the data subject was faced merely with a final decision without any explanation relating to it.

Third, opponents to the right to explanation base themselves on the argument that the right to explanation does not expressly appear in the text of the GDPR and that Recital 71 GDPR, containing that right, is not legally binding.[68] While the legislative procedure shows that the right to explanation was indeed omitted from the text of Article 22,[69] the circumstance that this right still appears in the recital demonstrates that the legislator did not want to do away with it entirely. Putting the right to explanation into the recital was a compromise solution, born from a disagreement on whether this right should be enshrined in the GDPR or not.[70] Drafting legislation in such a way also demonstrates that the legislator left the final decision on the existence of this right to the CoJ which is, as it has been repeatedly demonstrated in the recent case law, rather purposeful and activist when interpreting data protection legislation. Dismissing the possibility of the existence of the right to explanation altogether because recitals are not legally binding is too formalistic, in particular in the light of the CoJ's case law which regularly uses recitals as an interpretative aid.[71] A closer look into the case law of this court reveals that, indeed, 'the preamble to a European Union act has no binding legal force' and cannot be used to derogate from the actual provisions or 'for interpreting those provisions in a manner clearly contrary to their wording'.[72] It thus stems from the case law that recitals cannot be used for a *contra legem* interpretation of EU law provisions. However, relying on an approach of interpreting various GDPR provisions together as suggested above and the use of Recital 71 to strengthen the interpretation supporting the existence of the right to explanation would not lead to a *contra legem* interpretation. Rather, it would serve as a means to resolve ambiguity resulting from common reading of relevant GDPR provisions.[73] Given an activist stance and efforts of the CoJ to ensure a high level of data protection, it is possible that the Court will opt for an interpretation giving the data subject the right to explanation.

Fourth, it is also important to note the subtleties of the wording of relevant provisions. A careful reading of Article 22(3) reveals that the safeguards from this provision (human intervention, expression of point of view, contesting) are not necessarily the only possible safeguards. This provision namely requires that the data subject is guaranteed *at least* those safeguards. Adding an additional safeguard – the right to explanation from Recital 71 – through judicial interpretation would thus clearly not amount to an interpretation derogating from this provision or being contrary to its wording. Moreover, it has also been claimed that Article 13(2)(f) and Article 14(2)(g) GDPR omit to refer directly to Article 22(3) GDPR on safeguards, but only make reference to Article 22(1) and (4).[74] The same can be claimed for Article 15(1)(h). However, the wording of the former provisions requires that the data subject is provided with meaningful information about the logic involved in automated decision *at least* in cases of Article 22(1) and (4). It seems that the wording

[68] Wachter, Mittelstadt, Floridi, 4.

[69] Wachter, Mittelstadt, Floridi, 9.

[70] Compare Edwards and Veale, 33.

[71] See, for example, Case C-283/16 *M.S.* ECLI:EU:C:2017:104, paras 34-35; Case C-436/16 *Leventis and Vafias* ECLI:EU:C:2017:497, para 33; Case C-578/16 PPU *C. K., H. F., A. S.* ECLI:EU:C:2017:127, para 43; Case C-111/17 PPU *OL* ECLI:EU:C:2017:436, para 40.

[72] Case C-308/97 *Manfredi* ECLI:EU:C:1998:566, para 30; Case C-136/04 *Deutsches Milch-Kontor* ECLI:EU:C:2005:716, para 32; Case C-134/08 *Tyson Parketthandel* ECLI:EU:C:2009:229, para 16; Case C-7/11 *Caronna* ECLI:EU:C:2012:396, para 40; Case C-345/13 *Karen Millen Fashions* ECLI:EU:C:2014:2013, para 31.

[73] The role of recitals to resolve ambiguity in legislative provisions see Tadas Klimas, Jurate Vaičiukaitė, 'The Law of Recitals in European Community Legislation' (2008) 15 ILSA Journal of International & Comparative Law, 26

[74] Wachter, Mittelstadt, Floridi, 13.

was explicitly left open not to entirely preclude a possiblity of judicial interpretation giving the data subject the right to explanation.

Fifth, even if a general right to explanation is not recognised, this right should exist at least in case automated decision is based on sensitive data. In principle, automated decisions should not be based on special categories of personal data (Article 22(4)), except if data subject gives explicit consent to processing for specified purposes or if such processing is necessary to safeguard an important public interest (Article 9(2)(a) and (g)). Article 22(4) remains rather vague when it comes to safeguards, stating that 'suitable measures to safeguard the data subject's rights and freedoms' should be in place. However, all the GDPR provisions requiring that the data subject should be familiarised with the logic involved (Articles 13(2)(f), 14(2)(g) and 15(1)(h)), expressly require that this should be guaranteed in case automated decision is based on sensitive data. What needs to be equally taken into account is not only the textual interpretation of these provisions, but also the purpose of high level of data protection when it comes to sensitive data. If the inclusion of sensitive data leads to a biased decision, the data subject should be able to understand the reasons behind such a decision. Her rights would not be sufficiently safeguarded if she would only receive a general information about the functioning of the system.

In order to comply with the GDPR requirement that the logic behind the decision must be explained to the data subject, it is not enough to ensure merely what Kroll et al. term 'procedural regularity'. Such procedural regularity ensures only that the decisions are based on the same decision policy, that the policy was determined before knowing the inputs and that the outcomes can be reproduced.[75] It therefore addresses only aggregate procedural regularity of all cases, safeguarding that all cases are decided upon the same rules. However, the concept of procedural regularity does not answer the question of *why* the algorithm reached a certain decision with a certain dataset as an input. The transparency required by the GDPR is of a different kind: the data subject has to understand reasons behind the decision.

The individual transparency relating to *non-algorithmic* automated decisions will not pose particular problems regarding the explanation of the logic behind the decision. For example, if a camera detecting the speed of the driver communicates to the public authorities that the speed limit was exceeded, the issuing of speeding ticket follows automatically. The logic behind the decision as well as the rule on which the decision is based can be easily explained to the data subject: speeding ticket is issued if the speed limit is exceeded (taking into account the correction factor if applicable).

Differently, automated decision-making based on *algorithms* faces numerous complications when it comes to the explanation of the reasons underlying a decision. As the technology advances and the use of algorithms for decision-making is exponentially growing, both the legal regulation and academic work call for more transparent algorithmic decision-making, often described with the buzzword 'algorithmic transparency'. The basic quest of proponents of algorithmic transparency is to reveal the logic behind the algorithm that adopts a certain decision. While some commentators consider that it is near to impossible to explain an algorithm because even its developers cannot exactly pinpoint the reasons why a particular decision was taken,[76] others take a more optimistic

[75] Joshua A. Kroll et al., 'Accountable Algorithms' (2017, forthcoming) 165 University of Pennsylvania Law Review, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765268>, 1, 18.

[76] See for example <https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/> and <https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/>.

approach[77] and even propose technical solutions[78] that would lead to a higher algorithmic transparency. Algorithmic transparency is deemed to cover different transparency degrees from revealing a source code to an explanation of its functioning. For the purposes of this paper, we believe that the algorithmic transparency, legally speaking, should encompass transparency of the process of algorithmic decision-making to the extent that this is necessary to ensure the respect of rights under the GDPR, notably the information to the data subject about meaningful information about the logic involved. Technically speaking, the degree to which the functioning of the algorithm is revealed might be different for different decisions.

Furthermore, algorithmic transparency is a necessary tool for the prevention of discrimination in algorithmic decision-making. The decisions taken by algorithms are sometimes discriminatory, even without the algorithm being programmed to discriminate. For example, the decision can be discriminatory because the data on which the decision is based is discriminatory in itself. This can arise, in particular, when sensitive data such as race or gender is involved in decision-making. In this regard, Zarsky points out that 'a skewed and biased data set may cause outcomes of the algorithm process that discriminate against protected groups'.[79] The decision may be biased also if the algorithm is trained on biased data.[80] However, other authors claim that, in order to avoid algorithmic discrimination, it is necessary to use sensitive personal data in the process of building of decision-making models.[81] In any event, when the algorithm takes the decision, sensitive personal data such as race should not be required as 'input variables' relevant for decision-making.[82] Apart from discrimination due to biased entry datasets, algorithmic decision can be discriminatory also due to biased programming of the algorithm, leading to discriminatory decisions. In any event, algorithmic transparency would help understand the reasons behind biased decisions and would therefore be the first step towards preventing such algorithmic discrimination.

Finally, apart from providing the data subject with reasons for the decision concerning him/her and apart from the prevention of discrimination, algorithmic transparency plays also an important role in determining the distribution of accountability in automated decision-making (sometimes also termed 'algorithmic accountability'). Logically speaking, algorithmic transparency is a predisposition for algorithmic accountability. Who should be responsible if an algorithm makes a mistake or takes a discriminatory decision – the developers, the user or even the autonomous agent itself? Who is accountable if a search engine uses an algorithm that favours a particular political party instead of being politically neutral; or if it displays results regarding certain companies above the others and aims to run competitors out of business? While the doctrine does not specifically address the issue of *algorithmic* accountability, inspiration could be drawn from the literature examining the accountability of *robots* when they act as autonomous agents. The proposals for accountability (and hence liability) range from strict liability of the developer and accountability of

[77] Joshua A. Kroll et al., 'Accountable Algorithms' (2017, forthcoming) 165 University of Pennsylvania Law Review, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765268>, 1, 15-16.

[78] Anupam Datta, Shayak Sen and Yair Zick, 'Algorithmic Transparency via Quantitative Input Influence' <https://www.andrew.cmu.edu/user/danupam/datta-sen-zick-oakland16.pdf> accessed 10 December 2016.

[79] Tal Zarsky, 'The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making' (2016) 41 Science, Technology, & Human Values 126.

[80] Bryce Goodman, 'Discrimination, Data Sanitisation and Auditing in the European Union's General Data Protection Regulation' (2016) 2(4) European Data Protection Law Review 498.

[81] Indrė Žliobaitė, Bart Custers, 'Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models' (2016) 24 Artif Intell Law 183–201.

[82] Ibid.

the user to accountability of the autonomous agent itself.[83] While the full-fledged analysis of accountability issues goes beyond the scope of this paper, it is submitted that, on the one hand, the accountability should be shared between the developer and the user of the algorithm, provided that the algorithm has self-learning features and, on the other hand, that both strict liability as well as full accountability of autonomous agents should be rejected. A forceful argument against strict liability for actions of an algorithm stems from the circumstance that it is highly unlikely to consider an algorithm as a dangerous product.[84] Furthermore, we argue that exclusive accountability of autonomous agents could lead to the avoidance of accountability altogether and should therefore not be endorsed. This is true in particular with regard to potential criminal liability which is impossible to be attributed to an algorithm. With regard to civil liability, it is admittedly feasible to establish an insurance fund from which damages could be paid for responsibility of the autonomous agent. Nevertheless, in civil law, accountability would in principle require either intent or negligence on the part of the autonomous agent which would be obviously impossible to be established. Therefore, accountability of autonomous agents themselves might be a tool to dispense the developers or users of accountability they should be attributed.

3.4 OBSTACLES TO ALGORITHMIC TRANSPARENCY

It is submitted that there are several obstacles that stand in a way of providing data subject with a meaningful explanation of logic behind algorithmic decisions.[85] Among the obstacles to algorithmic transparency are the following: (1) technical obstacles, (2) intellectual property obstacles and (3) state secrets and other confidential information of state authorities.

3.4.1 Technical obstacles

The amount of technical obstacles standing in a way of explaining algorithmic-based autonomous decisions depends on the complexity of an algorithm. Many authors claim that it is nearly impossible to explain the logic behind an algorithm taking a decision. The reasons for a decision on the basis of a simple decision tree could perhaps still be explained. Decision trees, which are a form of reasoning or decision-support that use a graph similar to a tree to reach a conclusion, were the main decision-making tools until 1980s and are still used nowadays.[86] The way a decision tree reaches comes to a conclusion is 'by performing a sequence of tests'.[87] In a simple decision tree, each 'decision node' represents a turning point, a question to which different answers are possible; at the end of these turning points are 'leaf nodes' which represent different possible answers to the initial question contained in the 'root node'.[88] Even though decision trees can be learning algorithms, the method of learning they use is induction.[89] This makes them take linear decisions that are easy to analyse *ex post* and that enable algorithmic transparency. For example, a

[83] Erica Palmerini, Andrea Bertolini, 'Liability and Risk Management in Robotics' in Reiner Schulze, Dirk Staudenmayer (ed), *Digital Revolution: Challenges for Contract Law in Practice* (Nomos 2016) 225 – 260. For liability of autonomous agents see also Jaap Hage, 'Should Autonomous Agents be Liable for What They Do?', available on <http://www.jaaphage.nl/pdf/LiableAutonomousAgents.pdf> accessed 18 May 2017.

[84] See, per analogy regarding robots, Erica Palmerini and Andrea Bertolini, 'Liability and Risk Management in Robotics' in Reiner Schulze, Dirk Staudenmayer (ed), *Digital Revolution: Challenges for Contract Law in Practice* (Nomos 2016) 240.

[85] Burell distinguishes between three types of opacity of algorithms: corporate or state secrecy; technical illiteracy; and opacity arising from characteristics of machine learning; see Jenna Burrell, 'How the machine 'thinks': Understanding opacity in machine learning algorithms' (2016) Big Data & Society 1-12.

[86] Stuart J. Russell, Peter Norvig (eds), *Artificial Intelligence. A Modern Approach*, 3rd ed. (Pearson 2010) 638.

[87] Ibid., 698.

[88] For a mathematical explanation of decision trees, see ibid. 698 et seq.

[89] On 'decision tree induction' see ibid., 697.

simple decision tree could provide decision assistance with determination whether a person who applied for a job fulfils formal requirements. The decision nodes could contain questions such as 'Has this person completed education of a required level?', 'Does this person have required work experience?' and 'Does this person have a recommendation from his/her previous work?' Depending on whether the answers to these questions are affirmative or negative, leaf nodes would provide affirmative or negative answers to the initial question of whether the formal requirements for a job are fulfilled.

However, if the algorithm used for decision-making is a neural network, prone to very fast learning,[90] it will be close to impossible to explain the reasons behind its decision.[91] Neural networks function fundamentally differently from simple algorithms. Neural networks are built on the model of human brain where different 'nodes' connect with each other in a network.[92] Due to this ability of interconnection and due to the fact that it mirrors human brain, neural network has an ability to learn while processing data. According to Russell and Norvig, 'neural networks remain one of the most popular and effective forms of learning system'.[93] Because of their ability to learn, it is almost impossible to guarantee their transparency and to identify the factors that influence a decision taken by a neural network. However, Datta et al. developed a system called Quantitative Input Influence (QII) that could explain autonomously-made decisions.[94] The idea behind QII is that the degree of influence from input data to output data could be measured.[95] It seems that, in order to reach the transparency of an algorithm, another algorithm would need to be developed to clarify which factors were taken into account and what was their weight.[96]

3.4.2 IP-related obstacles

To a certain extent, intellectual property rights can also represent obstacles for algorithmic transparency. We submit that this is not the case as far as patent and copyright are concerned. Differently, trade secrets or confidential information can stand in a way of algorithmic transparency.

Even though the European Patent Convention (EPO) allows for patenting "computer-implemented inventions",[97] it is not enough that the software is inventive, but it also has to allow

[90] Masnick claims that the faster the machine learns, the more difficult it is to understand the reasons behind its decisions; Mike Masnick, 'Activists Cheer On EU's 'Right To An Explanation' For Algorithmic Decisions, But How Will It Work When There's Nothing To Explain?' <https://www.techdirt.com/articles/20160708/11040034922/activists-cheer-eus-right-to-explanation-algorithmic-decisions-how-will-it-work-when-theres-nothing-to-explain.shtml> accessed 10 January 2016.

[91] Metz points out that "[d]eep neural nets depend on vast amounts of data, and they generate complex algorithms that can be opaque even to those who put these systems in place." See Cade Metz, 'Artificial Intelligence Is Setting Up the Internet for a Huge Clash With Europe' <https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/> accessed 10 January 2016. Compare also Bryce Goodman, Seth Flaxman, 'European Union regulations on algorithmic decision-making and a "right to explanation"' <https://arxiv.org/abs/1606.08813v3> accessed 1 September 2016.

[92] For a mathematical explanation see Stuart J. Russell, Peter Norvig (eds), *Artificial Intelligence. A Modern Approach*, 3rd ed. (Pearson 2010) 728 et seq.

[93] Ibid., 728.

[94] Datta et al., 1.

[95] *Ibid.*

[96] Compare 'Artificial Intelligence, Robotics, Privacy and Data Protection' Room document for the 38th International Conference of Data Protection and Privacy Commissioners, October 2016.

[97] See Article Art. 52(2)(c) of the European Patent Convention in combination with Guidelines for Examination, point 3.6 Programs for computers, available at <https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_ii_3_6.htm> accessed 20 December 2016.

for an industrial application.[98] Along this line of reasoning, the EPO does not allow to patent a computer algorithm, as the "programmer must have had technical considerations beyond 'merely' finding a computer algorithm".[99] However, even if the algorithm was subject to a patent, this would still not create an obstacle for algorithmic transparency as having a patent would oblige the patent holder to disclose the composition and the modalities of functioning of an algorithm.

Copyright protection of a computer software leads to a similar result: while both the TRIPS agreement [100] and the WIPO[101] copyright treaty[102] allow for copyright protection of software from the moment of its creation, it is not entirely clear whether algorithms themselves can be a subject matter of copyright protection.[103] It seems that the EU does not allow for such protection and this approach is in line with some other non-European jurisdictions.[104] In any event, it is worth mentioning that the EU computer programmes directive[105] allows the user of a computer program "to observe, study or test the functioning of the program in order to determine the ideas and principles which underlie any element of the program". In practice this means that a user of a computer programme is allowed to determine the functioning of the algorithm and, if it is technically possible, to reveal the importance of particular factors involved in the algorithmic decision-making. That would imply that even if the software and/or algorithm are protected by a copyright, such protection could not stand in a way of algorithmic transparency.

However, an IP right that does stand in the way of algorithmic transparency is a trade secret (confidential or undisclosed information).[106] According to TRIPS, trade secret allows natural and legal persons to prevent that the information is disclosed to or used by others in a way that goes against honest commercial practices if that information is secret, has commercial value and steps have been taken to keep it secret.[107] In its recently adopted EU Trade Secrets Directive[108] which follows this definition, the EU however provides for an exception that allows for suspension of a trade secret "for the purpose of protecting a legitimate interest recognised by Union or national law".[109] Explaining an algorithmic decision to a data subject could fall under this exception as it is provided by the GDPR and seeks to protect a legitimate interest.

[98] See Article 52(1) of the European Patent Convention.

[99] See Opinion of EPO G 0003/08 (Programs for computers) of 12.5.2010, ECLI:EP:BA:2010:G000308.20100512, point 13.5.

[100] Agreement on Trade-Related Aspects of Intellectual Property Rights. See its Article 10(1), according to which "[c]omputer programs, whether in source or object code, shall be protected as literary works".

[101] World Intellectual Property Organization.

[102] See its Article 4, according to which "[c]omputer programs are protected as literary works within the meaning of […] the Berne Convention."

[103] For US scholarship on this issue (as well as patents on algorithms) see Richard H. Stern, 'On Defining the Concept of Infringement of Intellectual Property Rights in Algorithms and Other Abstract Computer-Related Ideas' (1995) 23 AIPLA Quarterly Journal 401; John Swinson, 'Copyright or Patent or Both: An Algorithmic Approach to Computer Software Protection' (1991) 5 Harvard Journal of Law & Technology 145.

[104] For example, in Japan, copyright protection of algorithms is not allowed; see Dennis S. Karjala, 'Japanese Courts Interpret the 'Algorithm' Limitation on the Copyright Protection of Programs' (1991) 31 Jurimetrics Journal 233.

[105] Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs, OJ L 111, 5.5.2009, p. 16.

[106] More precisely on balancing between trade secrets and personal data, see Gianclaudio Malgieri, 'Trade Secrets v Personal Data: a possible solution for balancing rights' (2016) 6(2) International Data Privacy Law 102-116.

[107] See Article 39(2) of TRIPS agreement.

[108] Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, OJ L 157, 15.6.2016, p. 1.

[109] See Article 5(d) of the Trade Secrets Directive.

Even if that exception is not applicable, we argue that algorithmic transparency does not necessarily need to run against trade secrets. In order to provide for such transparency within the GDPR, the source code or even the way the algorithm operates does not need to be disclosed. The 'logic behind the decision' from GDPR points to the (argumentative) tool that was deployed by an algorithm without the necessity to fully disclose that tool. Per analogy with computational journalism Diakopoulos points out that algorithmic transparency would command merely "the disclosure of certain key pieces of information, including aggregate results and benchmarks".[110]

3.4.3 State secrets and other confidential information

The biggest obstacles to algorithmic transparency are state secrets or other information held by public authorities that cannot be revealed to the public. It is in the interest of the state and those authorities that they do not reveal exactly why a certain decision was taken.[111] For example, a tax authority will not reveal the algorithm that chooses taxpayers whose tax (ir)regularity needs to be checked. Likewise, customs authorities will not reveal the pattern-match system that chooses which company needs to undergo customs check. Equally, police authorities will not disclose the rule behind the choice of neighbourhood or persons to monitor, for example for the purposes of prevention of terrorism or drug trafficking.

The question of whether the algorithm should be revealed to the data subject in these situations depends on the balancing of privacy with competing interests. It is obvious that state secret or confidential information will stand in a way of algorithmic transparency, but the latter is not an ultimate value in our society that would need to always prevail over other interests. In such cases, revealing the logic behind the automated decision will depend on the proportionality analysis in each particular case.

4 THE RELEVANCE OF THE GDPR FOR US AND GLOBAL BUSINESSES[112]

4.1 ARTICLE 3(2)(A) GDPR

The provisions of the GDPR analysed above are relevant not only for companies based in the EU, but also for businesses outside Europe. The second and the third paragraph of Article 3 of GDPR deal with a situation where a controller does not have an establishment in the Union. According to Article 3(2)(a) of GDPR, such a controller has to comply with the rules established in the regulation if his activities relate to the offering of goods or services to data subjects in the Union.[113] It is not entirely clear what 'offering of goods or services' entails, but from reading Recital 20 of GDPR it becomes clear that it should be 'apparent that the controller is *envisaging* the offering' of goods/services to European data subjects.[114] Through this recital, it is also clarified that, whereas mere access to a website or e-mail address are not sufficient for the GDPR to apply, other criteria, such as the mention of Member State's currency or offering of goods/services in a language of this Member State could point to controller's intention to offer goods/services to European data

[110] Nicholas Diakopoulos, 'Accountability in Algorithmic Decision Making' (2016) 59 Communications of the ACM, 56, 58-59.

[111] For example, in the Dutch tax authority 95% of decisions taken are automatic and the customs authorities also rely heavily on automated decision-making. I am grateful to Prof. dr. ir. Sjir Nijssen for this insight.

[112] A previous version of this part of the paper has been published in Brkan, M., 'Data Protection and Conflict-of-laws: A Challenging Relationship' (2016) 2 European Data Protection Law Review 3, 324-341.

[113] See Article 3(2)(a) of the GDPR.

[114] Emphasis added.

subjects.[115] In practice this means that many online stores based in the US and not having a subsidiary in the Union will have to comply with the European data protection legislation when they offer goods or services online to European data subjects.[116]

The criterion of envisaging doing business in a particular Member State of the European Union can be compared with criteria for 'orientating' business activities to a Member State, as developed by the CJEU in *Pammer and Hotel Alpenhof,*[117] *Mühlleitner*[118] and *Emrek*.[119] Just as in *Pammer*, the accessibility of the website, e-mail address or other contact details is insufficient to fulfil the criteria of 'orientating'.[120] However, the criteria developed in *Pammer* are much more detailed than those in Article 3(2)(a) of the GDPR and the related Recital 20. The *Pammer* judgment offers a non-exhaustive list of criteria, among which are, other than the use of language (of a particular Member State) and possibility to make reservation in that language (also used in Recital 20 GDPR) also the international nature of the activity, mention of telephone numbers with an international code, use of a top-level domain name other than that of the state of establishment.[121] Moreover, the criteria from Article 3(2)(a) GDPR and Recital 20 can also be compared with the CJEU's reasoning in *Google Spain and Google*, which requires that the subsidiary of a search engine 'orientates' its activity towards the Member State where it promotes and sells advertising space.[122]

Therefore, with the GDPR, the question of applicable law will move more on a global level in that the European authorities and third-country companies will have to, prior to raising the issue of compliance with the GDPR, address the question whether EU law applies or not. From the mere text of Article 3(2)(a) it seems that the applicability of GDPR will extend far over EU borders[123] and is therefore controversial[124] for two principal reasons.

On the one hand, the third-country controller only has to *offer* goods or services within the Union for the GDPR to apply. In view of Svantesson, this means that "this provision seems likely to bring all providers of Internet services…under the scope of EU Regulation as soon as they interact with data subjects…in the European Union".[125] It has been argued that the GDPR changed the connecting element from 'country of origin' to 'country of destination'.[126] For GDPR to apply, it is not important whether the controller offering goods or services has any territorial connection with the EU and it is not important whether the data subject in the EU actually buys goods or

[115] See Recital 20 of the GDPR.

[116] As a comparison, on extra-territorial application of Data Protection Directive, see Article 29 Data Protection Working Party, *Working document on determining the international application of EU data protection law to personal data processing on the Internet by non-EU based web sites* (2002), 5035/01/EN/Final, WP 56, p. 4. See also Christopher Kuner, 'Internet Jurisdiction and Data Protection Law: An International Legal Analysis (Part 1)' (2010) 18 International Journal of Law and Information Technology, 178.

[117] Case C-585/08 *Pammer and Hotel Alpenhof* [2010] ECR I-12527.

[118] Case C-190/11 *Mühlleitner* [2012] not yet published in ECR.

[119] Case C-218/12 *Emrek* [2013] not yet published in ECR.

[120] *Pammer and Hotel Alpenhof*, para 94.

[121] *Pammer and Hotel Alpenhof*, para 93.

[122] Case C-131/12 *Google Spain and Google* [2014] not yet published in ECR, para 60.

[123] James Castro-Edwards, 'The Proposed European Data Protection Regulation' (2013) Journal of Internet Law, 6, points out that non-EU businesses "will need to be mindful of the potential 'long arm' of the Regulation and the potential heavy sanctions for failing to comply".

[124] Jacob M. Victor, The EU General Data Protection Regulation: Toward a Property Regime for Protecting Data Privacy (2013) 123 *The Yale Law Journal*, 514, stresses that the GDPR is "controversial for its potential applicability to any corporation that processes the personal data of EU citizens (including U.S. corporations)".

[125] Dan Jerker B. Svantesson, *Extraterritoriality in Data Privacy Law* (Ex Tuto 2013), 107.

[126] Dennis Holmes, 'Debating the "Where" of Online Jurisdiction' The Privacy Advisor <https://iapp.org/news/a/debating-the-where-of-online-jurisdiction/> accessed 11 April 2016.

services.[127] In practice this will probably mean that third-country controllers will build different websites for different countries (or at least a special website for EU customers) with in-built Privacy-By-Design settings complying with GDPR. For smaller businesses which cannot afford such settings, it will be much more difficult to comply with GDPR, in particular in a stage where they merely offer goods or services to data subjects in the EU.

On the other hand, the criterion from Article 3(2)(a) GDPR only requires that data subjects are 'in the Union'. There is no requirement that data subjects have to reside (permanently or temporarily) in the Union, but merely that they are present on its territory. It is questionable though how this phrase should be interpreted. If a friend from Japan pays a visit to the EU and orders a travel guide online to be delivered to her hotel during her stay in Europe, does the GDPR apply? Svantesson understands this criterion as requiring residence in the EU,[128] but such an interpretation is not directly supported by the text of the article.

A very extensive reading of this provision could even lead to an interpretation according to which the Union legislation on data protection would apply even if a European data subject buys goods or receives services while being physically in the territory of a third state and not online.[129] Such an interpretation would however lead to a too extensive extraterritorial application of Union legislation on the territory of a third state and can therefore not be upheld.[130] It would also go against the wording of Article 3(2)(a) GDPR.

Therefore, if the conditions for territorial applicability on the basis of this provision are fulfilled, the business with a seat outside the EU will have to respect the requirements related to automated decision-making stipulated in Article 22 and related articles of the GDPR.

4.2 ARTICLES 3(2)(B) AND 3(3) GDPR

The GDPR will apply also if activities of the controller not established in the Union relate to the monitoring of the behaviour of data subjects in the Union (Article 3(2)(b)). It is not clear how broad this article should be interpreted. On a more realistic interpretation, this provision covers monitoring of behaviour by companies established in third countries (such as Google or Facebook), in order to use the gathered information for commercial purposes, such as targeted advertising. Such an interpretation seems to be confirmed by Recital 21 of the GDPR where 'monitoring' is explained as an activity where "individuals are tracked on the Internet including potential subsequent use of data processing techniques which consist of profiling an individual, particularly in order to take decisions concerning her or him or for analysing or predicting her or his personal preferences, behaviours and attitudes". Therefore, the third-country companies that profile[131] data subjects in the Union will be able to perform their marketing activities and target

[127] Alexander Dix, 'The Commission's Data Protection Reform After Snowden's Summer' (2013) 5 *Intereconomics*, 269, points out that many US companies have accepted this rule.

[128] Dan Jerker B. Svantesson, *Extraterritoriality in Data Privacy Law* (Ex Tuto 2013), 107.

[129] Compare *ibid.*

[130] Compare also Lokke Moerel, 'The long arm of EU data protection law: Does the Data Protection Directive apply to processing of personal data of EU citizens by websites worldwide?' (2011) 1 International Data Privacy Law, at 44; who points out that 'an unbridled expansion of applicability of EU data protection laws to processing of data on EU citizens wherever in the world should be prevented'.

[131] The notion of 'profiling' is defined in Article 4 GDPR as "any form of automated processing of personal data consisting of using those data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements".

their advertising, but only under the legal regime and requirements of the GDPR.[132] It therefore seems that the requirements from Article 22 regarding profiling will become relevant also for foreign undertakings through this jurisdictional basis. If the Court of Justice of the EU recognises the right to explanation to the data subject for automated decision-making, this means that US and other global businesses will have to provide such a right to explanation.

On a (much) more expansionist interpretation, it could also be argued that even the NSA, when processing data of Union citizens or obtained from Union authorities, has to respect Union law. Although this is, admittedly, an extremely broad interpretation of this article, the text of the article does not seem to limit such an application *ratione personae* of this article (the Recital 21, however, does). One could stretch this interpretation even further and ask a question whether this would also mean that the US authorities have to observe Union law when a Union citizen travels to the US and gives his fingerprints on the US border. Such an interpretation, however, seems to be rather far-reaching, in particular because it would lead to a broad extraterritorial application of the EU data protection legislation. It should therefore not be accepted. In any event, if this jurisdictional basis is triggered, the US national authorities will have to respect the provisions of the Directive on Data Protection in Criminal Matters.

Finally, the third paragraph of Article 3 of the General Data Protection Regulation is, again, comparable to the rule set out in the current Article 4(1)(b) of the Data Protection Directive, since both legal instruments provide for the applicability of, respectively, Union and Member State's law, in case where the national law of a Member State 'applies by virtue of public international law'. This paragraph is not often applicable in practice and does not raise questions of conflict-of-laws.

5 CONCLUSION

This paper analyses rules of the GDPR and the Directive on Data Protection in Criminal Matters regarding automated decision making. It is established that, while these rules clearly give the right to the data subject not to be subjected to a fully automated decision, including profiling, the exceptions to this right hollow it out to the extent that the exceptions themselves become a rule. Within the EU, the most important exception is the possibility to allow automated decision-making either by Member State or Union law, allowed by both, the GDPR and the Directive. Outside the EU, the most important exception will probably be the one allowing automated decisions in case of a conclusion of a contract as many EU citizens conclude contracts with businesses in US and other countries, notably when using social media.

The paper further argues that the data subject should, from the perspective of a high transparency quest in the EU data protection legislation, have the right for an explanation of automated decision and be given 'meaningful information about the logic involved' in such decision. Despite voices expressed against such a right to explanation,[133] it is submitted that a purposeful interpretation of GDPR provisions, coupled with a broader quest of a high level of transparency of data processing, should lead to establishment of such a right for a data subject. Nevertheless, it is the Court of Justice of the EU that will have a final say on this matter when interpreting the relevant provisions of the GDPR.

[132] See Olivier Proust, 'Getting to know the GDPR, Part 5: Your big data analytics and profiling activities may be seriously curtailed' *Privacy, Security and Information Blog* <http://privacylawblog.fieldfisher.com/2015/getting-to-know-the-gdpr-part-5-your-big-data-analytics-and-profiling-activities-may-be-seriously-curtailed/> accessed 11 April 2016.

[133] Sandra Wachter, Brent Mittelstadt, Luciano Floridi, 'Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation' available on <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2903469> accessed 18 May 2017.

Furthermore, eliminating the obstacles to algorithmic transparency would not only enable the data subject to gain an insight on the reasons for which it was taken, but it would also be important in two other respects.[134] On the one hand, it would help to eliminate the discriminatory biases in the decision-making process itself. Discrimination in algorithmic decision-making can arise because the datasets with which an algorithm operates might be biased and therefore the algorithm learns to be biased itself. However, some authors argue that removing biased data from the automated process can nevertheless lead to biased results.[135] Recognising and being able to determine the bias is an important step in removing it and making the substantive decision fairer.[136] On the other hand, closely linked to algorithmic transparency is accountability. The paper argues that the accountability should be either distributed between or attributed to the developer and the user of the algorithm. The developer would bear the responsibility for issues relating to the way an algorithm was designed, whereas the user, feeding the algorithm with data, could be responsible for the algorithmic learning on the basis of this data. The paper however warns against strict liability of both developer and user for the decisions of an algorithm since an algorithm should not be considered a dangerous product. Moreover, the paper equally rejects the possibility of accountability of the autonomous agent itself as this approach could lead to avoidance of accountability altogether. Here, again, the question of applicability of GDPR rules to US and other global businesses arises. What if the user of the algorithm is based in the EU and the developer outside the EU? Could GDPR apply to such developer? In theory it could be argued that this is the case, but the accountability might be more difficult to be enforced in practice.

The paper also analyses different types of obstacles to algorithmic transparency and hence to the right of data subject to an explanation: technical obstacles, intellectual property obstacles and state secrets and other confidential information of state authorities. While the IP-related obstacles are considered to be merely paper tigers, state secrets can and will, in many cases justifiably, continue to stand in the way of algorithmic transparency. As transparency is not the ultimate societal value, it will have to be balanced against the state and public interests, notably in cases where public security is at stake. Finally, even though the technical obstacles represent the biggest hurdle towards a high level of algorithmic transparency, the scientists are gradually developing the means to overcome such obstacles. Nevertheless, autonomous decisions taken with the help of neural networks might remain 'black boxes', mostly inexplicable to data subjects.

Finally, the paper puts forward arguments justifying why the GDPR provisions on automated decision-making can, in certain circumstances, be applicable also to undertakings based outside the EU. This will be the case notably for profiling of EU data subjects by such undertakings, and in case of conclusion of contracts with EU data subjects. While the arguments put forward in this paper might show legal relevance of GDPR provisions for such businesses, it is yet to be seen whether these provisions will be effectively applied in practice. The biggest challenge in this regard will be enforcement to ensure compliance with GDPR; the beginning of application of this legal instrument in May 2018 will show how effective ensuring of such compliance will be. In any event,

[134] On discrimination related to algorithmic decision-making, see Sue Newell, Marco Marabelli, 'Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'' (2015) 24 Journal of Strategic Information Systems, 3–14.

[135] See, for example, Indrė Žliobaitė, Bart Custers, 'Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models' (2016) 24 Artif Intell Law 183–201; Toon Calders, Indrė Žliobaitė, 'Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures' in Bart Custers et al. (eds), *Discrimination and Privacy in the Information Society. Data Mining and Profiling in Large Databases* (Springer 2013) 43-57.

[136] More on algorithmic discrimination, see Bryce Goodman, 'Discrimination, Data Sanitisation and Auditing in the European Union's General Data Protection Regulation' (2016) 2(4) European Data Protection Law Review 493-506.

the provisions of the GDPR seek to contain decision-making based exclusively on the basis of algorithms and thus aim to prevent the algorithms from ‘ruling the world’.

BIBLIOGRAPHY

Article 29 Data Protection Working Party, Working document on determining the international application of EU data protection law to personal data processing on the Internet by non-EU based web sites (2002), 5035/01/EN/Final, WP 56, p. 4.

Auditing in the European Union's General Data Protection Regulation' (2016) 2 European Data Protection Law Review 493-506.

Bench-Capon, Thomas F. Gordon, 'Tools for Rapid Prototyping of Legal Cased-Based Reasoning' (2015) ULCS-15-005, University of Liverpool, United Kingdom; Giovanni Sartor, Luther Branting (Eds.), Judicial Applications of Artificial Intelligence (Springer, 1998).

Bertolini, 'Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules' (2013) 5 LIT 2, 214–247.

Brkan, M., 'Data Protection and Conflict-of-laws: A Challenging Relationship' (2016) 2 European Data Protection Law Review 3, 324-341.

Burrell, 'How the machine 'thinks': Understanding opacity in machine learning algorithms' (2016) Big Data & Society 1-12.

Bertolini and Palmerini, 'Regulating Robotics- a challenge for Europe', in: European Parliament Directorate General For Internal Policies Policy Department C: Citizens' Rights And Constitutional Affairs: Workshop on upcoming issues of EU law. Strasbourg: Policy Department C: Citizens' Rights And Constitutional Affairs (2014).

Calders, Indrė Žliobaitė, 'Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures' in Bart Custers et al. (eds), Discrimination and Privacy in the Information Society. Data Mining and Profiling in Large Databases (Springer 2013) 43-57.

Calo, 'Peeping HALs: Making Sense of Artificial Intelligence and Privacy' (2010) 2 European Journal of Legal Studies 3, http://www.ejls.eu/6/83UK.htm.

Castro-Edwards, 'The Proposed European Data Protection Regulation' (2013) Journal of Internet Law, 3-8.

Christin, Rosenblat, Boyd, 'Courts and Predictive Algorithms' (2015) Data & Civil Rights: A New Era of Policing and Justice <http://www.law.nyu.edu/sites/default/files/upload_documents/Angele%20Christin.pdf> accessed 16 January 2017.

Coormen, Algorithms Unlocked (MIT Press 2013).

Datta, Shayak Sen and Yair Zick, 'Algorithmic Transparency via Quantitative Input Influence' <https://www.andrew.cmu.edu/user/danupam/datta-sen-zick-oakland16.pdf>.

Dix, 'The Commission's Data Protection Reform After Snowden's Summer' (2013) 5 Intereconomics, 269.

de Bruin, 'Autonomous Intelligent Cars on the European Intersection of Liability and Privacy - Regulatory Challenges and the Road Ahead' (2016) 7 European Journal of Risk Regulation 3, 485-501.

Diakopoulos, 'Accountability in Algorithmic Decision Making' (2016) 59 Communications of the ACM, 56, 58-59.

Edwards, Veale, 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for', <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972855>.

Goodman, 'Discrimination, Data Sanitisation and Auditing in the European Union's General Data Protection Regulation' (2016) European Data Protection Law Review 493.

Goodman, Seth Flaxman, 'European Union regulations on algorithmic decision-making and a "right to explanation"' <https://arxiv.org/abs/1606.08813v3> [ICML Workshop on Human Interpretability in Machine Learning, New York (2016)].

Gurney, 'Sue My Car Not Me: Products Liability And Accidents Involving Autonomous Vehicles', (2013) University of Illinois Journal of Law 2, 247-277.

Holmes, 'Debating the "Where" of Online Jurisdiction' The Privacy Advisor <https://iapp.org/news/a/debating-the-where-of-online-jurisdiction/> accessed 30 August 2017.

Karjala, 'Japanese Courts Interpret the 'Algorithm' Limitation on the Copyright Protection of Programs' (1991) 31 Jurimetrics Journal 233.

Klimas, Vaičiukaitė, 'The Law of Recitals in European Community Legislation' (2008) 15 ILSA Journal of International & Comparative Law.

Kroll et al., 'Accountable Algorithms' (2017, forthcoming) 165 University of Pennsylvania Law Review, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765268>, 1, 18.

Kuner, 'Internet Jurisdiction and Data Protection Law: An International Legal Analysis (Part 1)' (2010) 18 International Journal of Law and Information Technology, 176-193.

Lohmann, 'Liability Issues Concerning Self-Driving Vehicles', (2016) 7 European Journal of Risk Regulation, 335-340.

Mantelero, 'Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection' (2016) 32 Computer Law & Security Review 2, 238-255.

Marchant and Lindor, 'The Coming Collision between Autonomous Vehicles and the Liability System', (2012) 52 Santa Clara Law Review 4, 1321-1340.

Masnick, 'Activists Cheer On EU's 'Right To An Explanation' For Algorithmic Decisions, But How Will It Work When There's Nothing To Explain?' <https://www.techdirt.com/articles/20160708/11040034922/activists-cheer-eus-right-to-explanation-algorithmic-decisions-how-will-it-work-when-theres-nothing-to-explain.shtml>.

Mendoza, Bygrave, 'The Right not to be Subject to Automated Decisions based on Profiling', University of Oslo Faculty of Law Legal Studies Research Paper Series No. 20/2017, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2964855>

Metz, 'Artificial Intelligence Is Setting Up the Internet for a Huge Clash With Europe' <https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/>.

Metz, 'The Sadness and Beauty of Watching Google's AI Play Go <https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/>.

Mittelstadt, Allo, Taddeo, Wachter, Floridi, 'The ethics of algorithms: Mapping the debate' (2016) 3 Big Data & Society 1, 1–21.

Moerel, 'The long arm of EU data protection law: Does the Data Protection Directive apply to processing of personal data of EU citizens by websites worldwide?' (2011) 1 International Data Privacy Law, 28-46.

Opinion of EPO G 0003/08 (Programs for computers) of 12.5.2010, ECLI:EP:BA:2010:G000308.20100512, point 13.5.

Orphanides, 'Robot carries out first autonomous soft tissue surgery' <http://www.wired.co.uk/article/autonomous-robot-surgeon>.

Palmerini, Bertolini, 'Liability and Risk Management in Robotics' in Reiner Schulze, Dirk Staudenmayer (ed), *Digital Revolution: Challenges for Contract Law in Practice* (Nomos 2016) 225 – 260.

Pasquale, The Black Box Society. The Secret Algorithms That Control Money and Information (Harvard University Press 2015); Jacob Loveless et al., 'Online Algorithms in High-frequency Trading. The challenges faced by competing HFT algorithms' (2013) 11 acmqueue 1.

Perry, 'iDecide: Administrative Decision-Making in the Digital World' (2017, forthcoming) Australian Law Journal.

Pillath, 'Briefing January 2016: Automated Vehicles in the EU', (2016) European Parliamentary Research Service, PE 573.902, available at <http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/573902/EPRS_BRI(2016)573902_EN.pdf >.

Proust, 'Getting to know the GDPR, Part 5: Your big data analytics and profiling activities may be seriously curtailed' Privacy, Security and Information Blog <http://privacylawblog.fieldfisher.com/2015/getting-to-know-the-gdpr-part-5-your-big-data-analytics-and-profiling-activities-may-be-seriously-curtailed/> accessed 30 August 2017.

Russell, Norvig (eds), *Artificial Intelligence. A Modern Approach*, 3rd ed. (Pearson 2010).

Schellekens, 'Self-driving cars and the chilling effect of liability law', (2015) 31 Computer Law & Security Review 4, 506-517.

Scholz, 'Algorithmic Contracts' (2017) Stanford Technology Law Review, available at <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=274770> accessed 10 November 2016.

Shah, Warwick, Lucivero, Schellekens, ´Self-Driving Cars´, in Frederico Azzari et al., (2014) Guidelines on Regulating Robotics, RoboLaw, available at <http://www.robolaw.eu/RoboLaw_files/documents/robolaw_d6.2_guidelinesregulatingrobotics_20140922.pdf>.

Stern, 'On Defining the Concept of Infringement of Intellectual Property Rights in Algorithms and Other Abstract Computer-Related Ideas' (1995) 23 AIPLA Quarterly Journal 401.

Sue Newell, Marco Marabelli, 'Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'' (2015) 24 Journal of Strategic Information Systems, 3–14.

Svantesson, Extraterritoriality in Data Privacy Law (Ex Tuto 2013).

Swinson, 'Copyright or Patent or Both: An Algorithmic Approach to Computer Software Protection' (1991) 5 Harvard Journal of Law & Technology 145.

Victor, The EU General Data Protection Regulation: Toward a Property Regime for Protecting Data Privacy (2013) 123 The Yale Law Journal, 513.

Vromen, 'Ethnic profiling in the Netherlands and England and Wales: Compliance with international and European standards', Public Interest Litigation Project (PILP-NJCM) / Utrecht University, <https://pilpnjcm.nl/wp-content/uploads/2015/06/Research-project-B-FINAL-version.pdf>.

Wachter, Mittelstadt, Floridi, 'Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation' available on <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2903469>.

Zarsky, 'The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making' (2016) 41 Science, Technology, & Human Values 118-132.

Žliobaitė, Bart Custers, 'Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models' (2016) 24 Artif Intell Law, 183–201