Algorithms and the GDPR: An analysis of article 22

Sandra Barbosa[*]
Sara Félix[**]

Abstract: The ever more current use of automated decisions, in the most various fields, has strike society's attention to the (lack) of protection given to data subjects when these decisions come to fruition. Since it is not a question of if they occur, but more so, of when and how, the General Data Protection Regulation, on its article 22, attempted to provide a framework for those decisions, aiming to put the data subject's rights, freedoms and legitimate interests at their forefront. The question that remains, and that we intend to answer, is if the use of automated decision-making is hindering that aimed protection so highly that it must be withdrawn, despite the supposed benefits they might bring to the entities making use of them.

Keywords: *Algorithms; Automated decision-making; Profiling; Bias; General Data Protection Regulation.*

Resumo: A utilização cada vez mais recorrente de decisões automatizadas, nas mais variadas áreas, tem despertado a atenção da sociedade para a (falta) de proteção dos titulares dos dados quando estas decisões são postas em prática. Uma vez que não se trata de uma questão de se podem ocorrer, mas sim de quando e como, o Regulamento Geral sobre a Proteção de Dados, no seu artigo 22, pretendeu estabelecer um mecanismo que colocasse os direitos, liberdades e garantias dos titulares dos dados em primeiro plano. A questão que permanece, e que pretendemos responder, prende-se com a utilização destas decisões automatizadas impossibilitar tanto essa desejada proteção, ao ponto

[*] Consultora na área da Proteção de Dados. Licenciada em Direito pela Escola de Direito da Universidade do Minho, frequenta o Master Degree in Law – Specialization in Law and Technology na Faculdade de Direito da Universidade Nova de Lisboa, em fase de dissertação.

[**] Licenciada em Direito pela Faculdade de Direito da Universidade de Lisboa, frequenta o Master Degree in Law- Specialization in International and European Law na Faculdade de Direito da Universidade Nova de Lisboa, em fase de dissertação.

de dever ser suprimida, pese embora os supostos benefícios que podem advir para as entidades que as empregam.

Palavras-chave: *Algoritmos; Decisões automatizadas; Definição de perfis; Enviesamento; Regulamento Geral sobre a Proteção de Dados.*

Introduction

In a digital and data driven era, the use of algorithms and data analytics has become a common business practice, especially towards consumers, that has also spread to public entities and state services. Currently, companies base their probabilities of attracting more consumers and achieving more efficiency on technology, and algorithms play one of the main roles for their potential success. Profiling individuals, either in public or private institutions, has proven to be the desirable key for progress.

Indeed, profiling methods regarding the use of algorithms can, and generally do, give a basis for automated decision-making (also mentioned in this article as "ADM"), which consists in the ability to, using technological means, make decisions with none (solely automated) human involvement. These methods use, among others, personal data, which in nature can become a highly privacy-invasive process. Moreover, artificially intelligent agents, often based on machine learning systems, can be quite opaque on their procedures, sometimes carrying inherent biases that can put data subjects in a very unsafe position.

As technology has been evolving quite rapidly and strongly over the last two decades, due, inter alia, to the exponential increase in computing power, the concerns with automated decision-making processes, already addressed on article 15[1] of the Data Protection Directive 95/46/EC (DPD), published in 1995, were reiterated, and updated, in the General

[1] "Member States shall grant the right to every person not to be subject to a decision which produces legal effects concerning him or significantly affects him and which is based solely on automated processing of data intended to evaluate certain personal aspects relating to him, such as his performance at work, creditworthiness, reliability, conduct, etc."

Data Protection Regulation[2] (hereinafter "GDPR" or "Regulation"), published in 2016, with its article 22, a successor of DPD's article 15, aiming to define a set of rules to protect data subjects from the risks posed by ADM and uphold human dignity through the process[3].

This article thus seeks to dive into provision 22 of the GDPR, conceptualizing the undefined concepts, deconstructing the demandable requirements, always with a conscious data subject protection against algorithmic bias, aiming to be a beacon to companies and organizations that are lost in the vast sea of data protection.

1. Principles Applicable to Automated Processing

Based on the conception of privacy and data protection as a fundamental right, enshrined in articles 7 and 8 of the Charter of Fundamental Rights of the European Union (CFREU)[4], the GDPR intends, also respecting article 16 of the Treaty of the Functioning of the European Union (TFEU), to lay down the data protection rules for the processing of personal data within the scope of Union law[5]. As the Regulation's material scope comprises, among other, the processing of personal data wholly by automated means[6], its data protection rules apply, as well, to the automated decisions under consideration here.

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

[3] Mendoza Isak, Bygrave Lee A., "The Right Not to Be Subject to Automated Decisions Based on Profiling", *University of Oslo Faculty of Law Research Paper*, No. 2017-20, 2017.

[4] Article 7 of the CFREU establishes the right to respect for everyone's private and family life, home and communications, as article 8 recognizes an explicit right to the protection of everyone's personal data, who must be processed fairly for specified purposes.

[5] Article 16 of the TFEU not only recognizes the right to the protection of everyone's personal data, but also obliges the EU to "lay down the rules relating to the protection of individuals with regard to the processing of personal data by Union institutions, bodies, offices and agencies, and by the Member States when carrying out activities which fall within the scope of Union law, and the rules relating to the free movement of such data".

[6] As defined in GDPR's article 2.

As such, it is on article 5 that the GDPR lays down the principles that generally apply to the processing of personal data. Concerning automated decision-making, certain principles govern the use and creation of algorithms. One of those is the transparency, lawfulness and fairness principle, present in article 5 (1) (a), which requires controllers to take the appropriate measures to keep data subjects informed about how their data is being used. Thus, when using ADM systems, data controllers shall inform the data subject, about the logic behind the algorithms. The Article 29 Data Protection Working Party's (hereinafter "WP29") Guidelines on Automated Individual Decision-Making and Profiling for the purposes of Regulation 2016/679 gives an example[7] of the appropriate information that shall be given when automated processing exists. The information given needs to address which data, that it is being processed under those means, was collected and the consequences of it, aligning, as well, with the information and access requirements present in articles 13 to 15 of GDPR. As such, data subjects have the right to be informed by data controllers about the existence of ADM and to be given information about the logic involved and the envisaged consequences of such automated processing [as provided in articles 13 (2) (f) and 14 (2) (g)], as well as to be given details about their personal data that is being used for ADM [as determined in article 15 (1) (h)].

This information and access rights can be perceived as a right to an *ex ante* generic explanation about the system functionality and its consequences to the data subject, though a right to an *ex post* explanation is not included in the provision[8]. However, as we will see further on, although this information may be given, every so often data subjects may not engross the information.

In addition, it is also important to mention the principle of data protection by design and by default, which includes those two complementary concepts that can jointly fortify each other, and ultimately, the protection of personal data. According to GDPR's Article 25, data controllers must consider the protection of personal data, both at the

[7] WP29, p.10. Example of the insurance company that offers insurance according to the profiling of the individuals, based on their driving behaviour.

[8] A brief analysis about the existence of a right to explanation on the GDPR is provided further on, in point 5.3.

design stage of the processing activities and during the processing itself, by implementing the appropriate technical and organizational measures and default settings to meet the demands of the Regulation's principles[9]. Regarding ADM, such measures should ensure the accuracy and quality of the data, to minimize the possibility of false, non-representative or biased outputs and, also, the respect for the fairness principle, under which personal data cannot be processed in a manner that is unjustifiably detrimental and discriminatory, by allowing, for example, the necessary human intervention to uncover machine bias and review the fairness of the algorithms used. Therefore, controllers must carefully consider the use of ADM, when designing its processing activities, applying the necessary safeguarding measures at that stage and ensure that, by default, data subjects' personal data is protected.

Other principles that significantly apply to ADM are the principle of purpose limitation, in article 5 (1) (b), which encompasses that the data collected for a specific purpose shall never be processed for a different one; the principle of data minimisation, in article 5 (1) (c), that respects to the minimum necessary extent to which the data shall be processed and applies either to quantity and quality, meaning that one cannot process an excessive amount of data, and equally, cannot go beyond the limit that was established as necessary to process it; the principle of accuracy, in article 5 (1) (d), that requires the data processed to be accurate and kept to date, relating to the right of data rectification; and finally the principle of storage limitation, in article 5 (1) (e), according to which personal data shall be stored for an amount of time that is considered essential for the task of processing and, in line with Recital 39[10], controllers shall establish a time limit during which these data will be stored.

[9] EDPD Guidelines 4/2019 on Article 25, Data Protection by Design and by Default; available at: <https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf>. Last visited on 17.04.2021.

[10] Recital 39 GDPR, "(...) In order to ensure that the personal data are not kept longer than necessary, time limits should be established by the controller for erasure or for a periodic review. Every reasonable step should be taken to ensure that personal data which are inaccurate are rectified or deleted. (...)".

2. Deepening the analysis of Article 22

Article 22 consists of four paragraphs: in short, paragraph 1 states that individuals shall not be subject to the automated processing of personal data, as a general rule; paragraph 2 states the exception, specifying three situations in which ADM processing is allowed; the 3rd paragraph alludes to the safeguards that must be applied when ADM processing can occur, to ensure protection of data subjects' rights and interests; and finally, paragraph 4 refers to special categories of data present in article 9. To fully comprehend article 22 and how it needs to be addressed, each paragraph will be considered.

2.1. The construction of the data subject's "right"

Prior to a deeper analysis of the requirements of paragraph 1, it is important to address the proper construction of this "right" of the data subject. Indeed, as Maja Brkan[11] further developed, the right hereby in scrutiny can be understood either as an active right, dependent on the data subjects' effective exercise, or as a passive one, that the data controller in charge of an automated decision must observe without their active claim.

Construing this as an active right would make its exercise solely dependent on the data subject's choice and will. This could lead, on a worst-case scenario, to data controllers lawfully taking automated decisions, having the characteristics described in paragraph 1, without the necessary safeguards to protect data subjects' rights, freedoms, and legitimate interests, as described in paragraph 3 (further analysed below). The detrimental effects of that omission present a clear burden on the data subject's back, who is probably not sufficiently knowledgeable to understand the impact of such omission of conduct. Another issue raised here are the effective legal consequences that derive from its exercise – does this right translate into a right to object automated decisions? Or as a right to request human intervention in the decision?

[11] Brkan Maja, "Do Algorithms Rule the World? Algorithmic Decision-Making in the Framework of the GDPR and Beyond", *Conference Terminator or the Jetsons? The Economics and Policy Implications of Artificial Intelligence*, 2017, p.1-29.

Consequently, interpreting article 22 (1) as giving the data subject the burden of actively exercising this right could, therefore, go against the aim of this provision, which is to protect data subjects from a general possibility of an automated decision being applied to them. Systematically, article 22 implies that all decisions which fulfil the requisites of paragraph 2 must be accompanied with paragraph 3's safeguards, otherwise will not be authorised by the GDPR.

Taking all this into account, scholars such as Mendoza and Bygrave[12], claim (and correctly, in our opinion) that, to achieve the ultimate goal of this provision, it is more appropriate to construct this "right" of the data subjects as a general prohibition to data controllers of fully automated decision-making. Actually, such interpretation of article 22 (1) aligns with the wording of 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 and provide appropriate safeguards for the data subjects' rights and freedoms.

In sum, construing the data subjects' "right" as a general prohibition of certain types of automated decisions is, in our opinion, the better way to ensure the protection of their rights, freedoms and legitimate interests.

2.2. Paragraph 1

According to the first paragraph of article 22, "The 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.", the data subject shall not be subordinated to ADM systems[13]. This prohibition has specific criteria – the ADM needs to be based *solely* on automated processing, and it must produce *legal effects* or *similarly significant effects*. To grasp better the meaning of this paragraph we need to analyse each criterion.

[12] Mendoza Isak, Bygrave Lee A., "The Right Not to Be Subject to Automated Decisions Based on Profiling", *University of Oslo Faculty of Law Research Paper*, No. 2017-20, 2017.

[13] Office of the Data Protection Ombudsman, *Automated Decision Making and Profiling*, Finland. Available at: <https://tietosuoja.fi/en/automated-decision-making-and-profiling>. Last visited on 23.09.2020.

Firstly, an automated decision-making system, based solely on automated means will be any processing that is operated without human intervention, and that leads to a decision upon personal data. This lack of human intervention means that, even though a human input in the system may have existed or a human may have interpreted the decision[14] in the end, the decision itself, did not have any human interference. The focal point in this definition is that it not only needs to be solely based on algorithms, but it also needs to be a full decision. After all, if we have systems that only prepare the basis for human intervention or systems that help in the interpretation of the decision humanly made, those will not be under article 22, because the decision, ultimately, is carried out by a human. Human involvement must be meaningful to align with the definition as a decision solely based on automated means. Otherwise, if the processing does not fulfil the criteria, it will not be relevant for article 22, since it will not jeopardize the data subject in a significant manner, thus being permitted. However, that does not mean that it will not fall under the scope of GDPR – if it includes personal data processing it is concerned by the Regulation.

When we read article 22, we see that one example given of ADM is profiling. Profiling is probably the most common reference to ADM that we find; nevertheless, the provision opens the scope for any other type of ADM. Hence, what can we define as profiling?[15] The GDPR describes it in article 4 (4) as the processing of personal data by automated means

[14] Intervention in the decision process is different from the interpretation of a decision. As described, for automated decision making to not be considered as solely automated, human intervention must comprise the making of a decision, rather than just its interpretation. When making a decision, the human has a meaningful intervention on the process, with decisive involvement and power to change the course of the automated process. On the contrary, the interpretation of an already-made automated decision would reduce the human involvement to an *ex-post* action, without the ability to change, in fact, the decision, which, in our opinions, cannot be considered as an intervention per se and, therefore, does not fall under the scope of article 22.

[15] The European Commission dedicates on its website some webpages to enlighten citizens on their data protection rights. The webpage dedicated to automated individual decision-making, including profiling is available at: <https://ec.europa.eu/info/law/law-topic/data-protection/reform/rights-citizens/my-rights/can-i-be-subject-automated-individual--decision-making-including-profiling_en>. Last visited on 23.09.2020.

to assess, infer or predict individual aspects of a data subject such as health or working performance, personal interests, or economic situation, among other examples. Profiling creates the always desired but never achieved possibility to predict the behaviours of individuals, which can be a very useful tool for companies. These data can be collected from social media, online forms, video surveillance, among other sources. Valeria Ferraris[16] explains in her work that we can have individual or group profiling, the latter consisting of gathering and assessing data of a community or of a group of people that share the same attributes. Profiling uses algorithms to complete the correlations between the personal data that was collected and the intended result of it, like, finding a pattern in the economic behaviour of a certain group of people, regarding the opening of a specific store. Some companies use profiling to perform their recruitment, others for marketing purposes (one of the main uses of profiling); police departments can use profiling to predict certain behaviours and act upon it; and doctors can make use of it to know the right treatments to apply to a patient[17].

Following the analysis of paragraph 1, it is necessary to define "legal effects" and "similarly significant effects", for the provision to be understood completely. Regarding the first one, "producing legal effect", will be every decision or action that affects someone's rights or legal status, or even their rights under a contract. Examples of such situations could comprehend impacts on the right to vote, the right to receive a monthly pension for disability, or the ability of someone to enter in a country. Moreover, legal effects also play a role in contracts, for example, if a contract is terminated due to ADM.

The criterion of "significant similar effects" opens a broad scope for the application of article 22 (1) which might lead to confusing and uncertain situations for data controllers, when they are deciding on the use of ADM, aiming to maintain GDPR compliance. Similar effects to the legal ones will be those consequences that although do not create

[16] Ferraris Valeria, Bosco Francesca, Cafiero Gioacchino, D'Angelo Elena, Y Suloyeva and Koops Bert-Jaap, "Working Paper: Defining Profiling", 2013. Available at: <https://www.academia.edu/4834070/Defining_Profiling>. Last visited on 23.09.2020.

[17] Bietti Elettra, "Data is power: Towards additional guidance on profiling and automated decision-making in the GDPR", *Institute for Research and Publication Journal*, 2017.

an impact on someone's legal rights, will, either way, have a significant weight on their lives[18]. The effect will be relevant enough if the data subject, due to ADM, finds him/herself in a situation where protection is needed because it influenced one's choices, behaviours, or circumstances. Recital 71 of GDPR gives examples of what can be thought as similar effects: the refusal of an online credit application that, even though it does not imply any right, thwarts the expectations of an individual; or recruitment that happens without any human interference and one may feel, when looking for a job, that being analysed and further on chosen or not by an e-recruiter system will have a significant impact on his/her life. WP29 refers as well to examples of situations that can significantly affect someone, such as, on an education level, in the case of someone not entering in their desired university based on an automated decision. Some of the most extreme cases of similarly significant effects can be those that lead to discriminatory outcomes. As written on the provision, a similar effect needs to be significant and here is where the threshold becomes difficult to meet. Consequently, not every automated decision will have a significant impact – for example, the recommendations of music on an app based on what you want, like or hear the most, occur due to profiling and are not prone to have a significant impact. On the contrary, there is a case for targeted advertisement that in principle cannot be considered as pertinent to be forbidden under 22 (1)[19]. According to what WP29 refers to, to understand if targeted advertising can be acknowledged as having significant similar effects, we need to see how much intrusiveness is present due to profiling or which are the particular vulnerabilities of the data subjects, in a case-by-case basis. A practical example will be if a person that is in financial debt and is known to have a gambling problem keeps being targeted with an advertisement for gambling. The vulnerabilities may lead to discriminatory outcomes and those are the cases that need to be avoided.

[18] European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European Data Protection Law, 2018 Edition*, April 2018.

[19] The European Data Protection Supervisor referred that targeted advertisement is an activity without significant importance to the audience that is targeted by it.

2.3. Paragraph 2

Tackling the second part of article 22, we are presented with exceptions to paragraph 1, based on the lawful processing basis that the GDPR entails in its article 6. The first one, in subparagraph (a), is the processing necessity to enter in or to perform a contract between the data subject and data controller. WP29[20] states that the data controller must be able to show this necessity, demonstrating that the usual way to conduct that contract would have been prejudicial or impractical. If any other way, that creates fewer risks on the fundamental rights of the data subject, is possible to be exercised, that shall be used.

The ICO[21] states that this necessity does not have to be considered essential, but shall be a reasonable way to achieve the parties' contractual goals. The essentiality referred by the ICO, and as stated by the EBDP on its Guidelines on the processing of personal data under article 6 (1) (b)[22], relates to the objective necessity of the processing for a "purpose that is integral to the delivery of that contractual service to the data subject".

As such, the processing must be more than useful for the performance of the contract, but it does not have to be the only way. Thus, the necessity will be determined considering the personal data and processing operations concerned and their impact on the performance or non-performance of the contractual service.

In subparagraph b) we have the permission to use ADM, if allowed either by EU Law or Member States law, to which the data controller is subject. Recital 71 gives some hints about these laws, such as the monitoring and prevention of fraud and tax evasion, as well as systems that are designed to safeguard the security of a specific service provider[23]. For

[20] WP 29, p. 23.

[21] Statement made on the ICO's website, in its dedicated guideline to organisations: <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection –regulation-gdpr/automated-decision-making-and-profiling/when-can-we-carry-out-this-type-of-processing/>. Last visited on 23.9.20.

[22] EBDP Guidelines 2/2019 on the processing of personal data under article 6 (1) (b) GDPR in the context of the provision of online services to data subjects. Available at: <https://edpb.europa.eu/sites/edpb/files/ files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_en.pdf>. Last visited on 17.4.21.

[23] WP29, p. 21.

this to be allowed, either the EU or the Member States law shall consider the protection of the freedoms and legal interests of data subjects and create safeguards when applying it. The ICO's point of view is that, when approaching companies and institutions that may wish to perform with ADM systems, even though this exception is predicted in the GDPR, the data controller needs to show that it is reasonable to do so.

The last exception regards to the consent of the data subject. The definition of consent is present in article 7 and it needs to be explicit, which means that the consent cannot be inferred from the silence of the individual. For this specific consent to be explicit, the data subject needs to be informed that the decision will be based entirely on automated systems. Dreyer and Schulz note that in the case of this specific consent, we face an intricate question[24]. Indeed, regarding ADM, data subjects will not only consent to the processing of their personal data, but they will consent as well to the automatic performance of the decision, hence this consent needs to be an extended and complex declaration. This consent should include all information required by articles 13 or 14 (depending if personal data was or was not collected from the data subject), and, specifically, the information about the existence of automated decision--making. This should include, at least, "meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject", as demanded by article 13 (2) (f) or article 14 (2) (g).

3. Risks and Benefits

Although ADM is forbidden under the criteria expressed, paragraph 2 provides an exception in three situations. We believe that this occurs because, even though ADM can create some risks to data subjects, it also brings numerous benefits, especially to businesses. The prohibition from

[24] Dreyer Stephan and Schulz Wolfgang, "The General Data Protection Regulation and Automated Decision-making: Will it deliver?", *Bertelsmann Stiftung*, 2019. Available at: <https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/GDPR.pdf>. Last visited on 23.09.2020.

22 (1) arises mainly due to the necessity of protecting the legal interest and rights of data subjects.

One of the main risks associated with ADM is the possibility of discrimination. Algorithms are designed for people, so, like humans themselves, they can carry inherent (or not inherent) biases that can provoke discriminatory outcomes regarding data subjects' characteristics, such as, their ethnic origin, sexual orientation, economic situation, gender, among others. Discrimination may arise out of the design but also from improper datasets that, for example, contain inaccurate data or in which data sampling is flawed, due to having over or underrepresent groups in the training data[25]. Recital 71 specifically addresses the question of discrimination, stating that ADM can only be admissible if it prevents discriminatory outcomes rather than provoking them.

Besides this major risk, we also have the question that, quite a lot of times, ADM is incomprehensible for individuals. Even supposing that individuals may have some information about it, as for its technological features, which involve numerous scientific methods that most of us are not familiar with, in the end, it is a black box matter. In short, the Black Box phenomenon, usually associated with AI built by machine-learning algorithms[26], considered by ICO as "one of the technical mechanisms that underpins and facilitates AI"[27], concerns to the human inability to fully understand the process of decision-making of these systems, as they are capable to arrive to determined solutions or decisions based on specific patterns of massive amounts of data, that humans, even the ones

[25] As an example, Buolamwini and Gebru found that facial detection technologies had higher error rates for minorities, particularly for darker females, probably due to under-representative face data sets: Buolamwini Joy and Gebru Timnit, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification", *Proceedings of Machine Learning Research: Conference on Fairness, Accountability, and Transparency*, 81, 2018, p.1–15.

[26] Machine learning algorithms are differentiated from other ones, due to their ability to learn from data, test the probabilities and make a decision, without human pre-written instructions. For a more in-depth analysis on machine learning algorithms and the black-box problematic, see Bathaee Yavar, "The artificial intelligence black box and the failure of intent and causation", *Harvard Journal of Law & Technology*, Vol. 31, 2, 2018.

[27] ICO, "Big data, artificial intelligence, machine learning and data protection", available at: <https://ico. org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf>. Last visited on 28.04.2021.

who created the system, cannot perceive. This human inability, created, among other, by algorithmic opacity and unpredictability[28], can lead to non-transparent decisions, which consequently are more difficult to audit and review, and, ultimately, present a threat to data subjects' rights. Those same data subjects that, in a more vulnerable position than the creators of such systems, who also may not understand them, are clueless consenting with ADM.

Nevertheless, ADM advantages, especially to businesses, involve an increase in efficiency and in innovation as it allows for further innovation and less bureaucracy. These advantages are visible as well in public institutions because it also can help the judicial sector, the educational sector, healthcare, social security, and police investigations for it hands in reducing the time to collect the pieces of evidence.

4. DPIA – Data Protection Impact Assessment

To mitigate the above-mentioned risks, the GDPR provides a major tool for data controllers that allows them to ensure their processing is compliant with the regulation and to guarantee that no data breaches will occur or will not expectedly occur. The DPIA, defined in article 35 GDPR, comprises an assessment of the potential and real impact of determined processing operations on the protection of data subjects' personal data. It is considered a mandatory assessment when processing operations, particularly when new technologies are used, and considering their nature, scope, context and purposes, are likely to result in a "high risk to the rights and freedoms of natural persons"[29]. As for ADM, article 35 (3) (a) demands a DPIA when the personal aspects relating to natural persons are subject to a systematic and extensive evaluation based on automated processing, including profiling, that serve as a base for decisions that "produce legal effects concerning the natural person or similarly significantly affect the natural person".

[28] Burrell Jenna, "How the machine "thinks": Understanding opacity in machine learning algorithms", Big Data & Society, 2016.

[29] GDPR's article 35 (1).

The concept "systematic", though not defined in the GDPR, is interpreted by the WP29 Guidelines on DPOs[30] as meaning a systematic processing based on a system, and/or with a methodical or organised method, and/or taking place as part of a general plan for collecting data and/or carried out as part of a strategy.[31] Whereas the concept "extensive" also not defined in the GDPR, is interpreted by ICO as a processing which involves a large-scale area, a wide range of data or that affects a large number of data subjects.[32]

Nevertheless, considering the risks and impacts already mentioned[33], even if no extensive and systematic evaluation based on automated processing is conducted, we believe that it is highly likely, due to its (more or less) opaque nature, that any automated decisions which fall within the scope of article 22, will be required a DPIA, given the potential high risk to the data subjects' rights and freedoms.

WP29 points out that this provision does not refer to "solely" automated means, but rather to systems "based on automated means", which indicates that this assessment shall be conducted not only when using ADM systems as of article 22 (1), but as well when using those that are not solely automated. Subsequently, if a company already knows that its system fits under article 22 requirements, a DPIA must be conducted to assess the risks. If it falls under article 22 (1) and there are no exceptions, it cannot be admitted. This practice allows companies to move towards good policies and procedures and to consider significantly the possible dangers that may arise with their processing activities to data subjects. Specifically, regarding ADM, we understand from this provision that the data controller does not need to refrain from using it at all, but it may need to take some precautions when using some specific algorithm.

[30] WP29 Guidelines on Data Protection Officers (DPO), available at: <https:// ec.europa.eu/ information_society/newsroom/image/document/2016-51/wp243en40855.pdf?wb48617274=CD63BD9A> Last visited on 17.04.2021.

[31] The WP29 interprets the word "systematic" as meaning one or more of the definitions provided. These definitions are alternative but might also be cumulative.

[32] Statement on the ICO's website regarding the concepts of systematic and extensive on GDPR: <https://ico.org.uk/for-organisations/guide-to-data-protection/guide--to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/when-do-we-need-to-do-a-dpia/#when8>. Last visited on 18.04.2021.

[33] Cf. Paragraph *3. Risks and Benefits.*

5. Introducing paragraph 3

Article 22, as previously mentioned, provides that there is a general prohibition of solely automated decision-making, including profiling, which produces legal or similarly significant effects on the data subject. Though we have exceptions to this rule, suitable measures that safeguard the data subject's rights, freedoms and legitimate interests should be in place. Paragraph 3 comes along regarding those suitable measures, stating the following:

"In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision." Although not mentioned, if the basis for processing is article 22 (2) (c), it is desirable that the Member State law that authorises such processing, provides for appropriate safeguarding measures as well.

5.1. The alignment with the information and access rights

Transparency, as we have stated, is one of the core principles underpinning the GDPR and acts as the background rationale for a significant number of its provisions. Specifically, regarding the ones directed to ADM, transparency poses as a foundation for the data controller's duties, as they must make sure they explain clearly and intelligibly to data subjects these processes, its consequences and provide them with tools to act against them, if they intend too[34]. Though we are focusing on a deeper analysis of this last duty, it is important to address that the safeguards provided by paragraph 3 come as a complement and reinforcement of the information and access rights stated in GDPR's articles 13, 14 and 15 as they act almost in symbiosis. Understanding the Information and

[34] Regarding this topic, Recital 60 of the GDPR states, "The principles of fair and transparent processing require that the data subject be informed of the existence of the processing operation and its purposes."

Access Rights' logic can help us get a better grasp of the rationale behind paragraph 3.

Indeed, the GDPR, through its articles 13 (2) (f) and 14 (2) (g), requires controllers, when using ADM processes, to explicitly inform data subjects that they are employing these types of activities and meaningfully explain what the logic involved is and the specific consequences of such processing, in a way that is intelligible to them[35]. This does not mean that the technical process behind that automated decision must be explained to the data subject, either because that (probably) belongs to companies' trade secrets[36] or because (and mainly) the data subject will not understand nor need that kind of information, to comprehend the effects of the decision. Article 15 (h) (1), on the other hand, provides a special type of right to the data subject when giving them the possibility to obtain confirmation on whether or not personal data concerning them is being processed. If that confirms, the data subject has also the right to access that personal data and all the information already mentioned for article 13 and 14.

Providing this information to data subjects will ultimately enable controllers to ensure they are meeting the required safeguards referred to in article 22 (3) and its connected Recital 71, regarding a decision based on automated processes, as they are already equipped with meaningful information to pursue their contesting intentions.

In fact, complementing the information duties pertinent to the information and access rights with specific mechanisms that provide data subjects with the possibility to effectively exercise their rights will,

[35] Dreyer Stephan and Schulz Wolfgang, "The GDPR and algorithmic decision-making – Safeguarding individual rights, but forgetting society", Völkerrechtsblog, 2019. Available at: <https://voelkerrechtsblog.org/articles/the-gdpr-and-algorithmic-decision-making/>. Last visited on 25.09.20.

[36] The WP29 Guidelines, on page 17, states, "Recital 63 provides some protection for controllers concerned about revealing trade secrets or intellectual property, which may be particularly relevant in relation to profiling. It says that the right of access 'should not adversely affect the rights or freedoms of others'. (...) Controllers should not use this as an excuse to deny access or refuse to provide any information to the data subject. These rights should be considered in context and balanced against individuals' rights to have information." Indeed, as a protective mechanism for their interests, companies can make use, e.g., of non-disclosure agreements.

ultimately, reinforce article 22 (3)'s purpose of rendering automated decisions contestable.

5.2. Deepening the analysis of paragraph 3

Moving on to a more in-depth analysis of paragraph 3, it is important to understand that the 'suitable measures' required by it are only a minimum standard to be met by data controllers, but any other that can complement them in safeguarding the data subjects' rights could, of course, be considered as a good practice. Another relevant point is the fact that, although these measures seem defined as independent, they can actually be interdependent, especially the right to express his or her point of view, that is seen as a subsequent step to both the right to obtain human intervention and the right to contest the decision, as we will see.

5.2.1. Right to obtain human intervention

Human intervention is a key element in the whole paragraph. The WP29 Guidelines clearly state that any review of any automated decision must be carried out by someone who has the appropriate authority and capability to change the decision, assertively opening the possibility of human intervention when an automated decision is taken.

In a way, human intervention is meant to provide input and solve problems raised by these decisions that cannot be solved or addressed by current machine capabilities, but it can also be sustained in the necessity of preserving human dignity. Regarding the prior, since we have experience-based knowledge and intuitions, that are challenging for algorithms to represent, a human reviewer can serve as a machine-error antidote and identify mistakes committed by machines. This kind of quality control is quite crucial due to the possible large-scale harms that these decisions can cause to data subjects. Even though technology is developing at full speed and systems that can accurately codify human intuitions are already seen as a possibility, for the time being it is not something we can count on. Up until we have machines that are able to internalize the effects of their decisions and judgements on humans, human oversight and intervention must remain a possibility for data subjects.

Notwithstanding, when it comes to decisions only based on data analysis, human intervention can be quite limited in altering the result, unless we only take into consideration the statistical correlation. Consequently, unless the human reviewer has a minimal knowledge of data analysis, in order to distinguish relevant from irrelevant correlations to the automated decision, as well as to reduce false positives[37], this human intervention may only be a formal requisite in the future. Towards a better accomplishment of this duty, a multidisciplinary team with data analysts will be essential.

5.2.2. Right to express his or her point of view

As a complement to the prior safeguard, and to allow the data subject intervention in an automated decision concerning him or her, the right to express his or her point of view is also vital.

Nevertheless, neither the provision nor the WP29 Guidelines mention clearly when, in the automated decision-making timeline, are data subjects able to communicate their point of view. It is the understanding of some scholars[38] that in a machine learning context, the data subject should be consulted prior to the final definition of the automated decision, guaranteeing in that way a more efficient process. This interpretation is supported by GDPR's article 25 (1) which establishes that the protection of the data subjects' rights – including this one – should be pursued "both at the time of the determination of the means for processing and at the time of the processing itself", therefore requiring, in this case, data controllers to provide for suitable measures at every step of the

[37] Regarding the occurrence of false positives, Antoni Roig in his contribution on the European Journal of Law and Technology [Vol 8, No 3, 2017] titled "Safeguards for the right not to be subject to a decision based solely on automated processing (Article 22 GDPR)", states that "The possibility of having false positives due to meaningless statistical correlations is a major risk scenario to be tackled by human expert data analyst intervention. Obviously, even without false positives the tool can also discriminate and have negative effects on citizens."

[38] Leenes Ronald, Van Brakel Rosamunde, Gutwirth Serge and De Hert Paul, *Data Protection and Privacy: The Age Of Intelligent Machines*, Hart Publishing Ltd, Oxford, 2017.

ADM process. However, this may be quite challenging in situations where decisions are taken in response to data in real time. More clarity regarding this subject matter, either from the European Data Protection Board (EDPB) or the European Court of Justice (ECJ), in a future decision, could be helpful in this case.

5.2.3. Right to contest the decision

The wording of the GDPR using the term "contest" connotes more to a right of recourse rather than to a mere opposing, requiring data controllers at least an obligation to hear the merits of the appeal and to provide data subjects the legitimising grounds behind the decision. Indeed, the expression "right to contest the decision" makes a clear statement on the obligation to render automated decisions contestable or cease them at all. As such, more than disclosure or meaningful information, it is required a mechanism to ensure that the outcome decision will be sufficiently interpretable and the logic behind the system tractable enough, at least, to be argued against a human arbiter. For that, the controller needs to provide a simple way for the data subject to exercise these rights, or he/she will not be able to contest without fully understanding how the decision has been made and on what grounds. And that is what we find tricky in this provision.

First of all, though article 22 does not specify if the decision hereby in scrutiny has to be the final or if it could be an intermediate one in the whole automated processing spectrum, it has been discussed that a broader interpretation of the provision, in alignment with Recital 71[39], allows for contestation of an 'interim' decision or measure.

Secondly, the provision remains unclear to who the data subjects must appeal when they want to contest the decision. On this matter, the GDPR does not specify that the contestant has to appeal to a human or if that can be made to a machine. It appears however, from the approach taken in article 22, that the data subjects must at least be granted the possibility

[39] GDPR's Recital 71 clearly states that decision may include a 'measure' and if so, in a broad interpretation, we could include intermediate or 'interim' decisions: "The data subject should have the right not to be subject to a decision, which may include a measure".

of requiring human intervention in the decision-making process and, if requested by them, a human should be tasked with reviewing the decision. Having said that, it stands unclear who this human should be and how he/she will be able to review a decision or its process that may have been based on third party algorithms or on opaque machine learning systems. Nor is it clear if this human reviewer could be the same person who firstly provided this decision to the data subject, still potentially consciously or subconsciously biased towards him or her.

One might ask, taking into consideration the ambiguity involved in human-subjected appeals, if it is fairer for data subjects to be able to appeal to a machine instead. Though machines can inherently carry bias with them, as explained above in point 3, technology development is opening the possibility for it to be designed in order to disregard certain sensitive characteristics (e.g., race, age, religion, etc.), hence leading to machine learning algorithms achieving higher levels of objectivity and neutrality that humans would effectively do. This does not mean that indirect discrimination is impossible to occur, due to the correlations that can result between inputs (e.g., one may infer a person's race by their address, if they live in a specific race-limited neighbourhood) or as a result of shadowing certain groups of people on account of under representative datasets, but fairer results could be achieved using machines trained as mentioned. Nonetheless, WP29 considers human intervention as a key element in the revision of automated decisions and recommends that any review must be carried out by "someone who has the appropriate authority and the capability to change the decision", appearing to be inclined for human (and not machine) revision of automated decisions.

Lastly, neither the GDPR nor the WP29 or other EU-provided resources make a stand regarding the legal effects of this right to contestation on the decision itself. The question that remains unsolved is what happens after a decision goes through an appeal? Taking a look into 'traditional appeals', when customers disagree with a company's decision, they can either contest that decision directly to the company's responsible department or to a government provided service. They also have the possibility to take the decision to an Arbitral Centre that the company has adhered to. When they disagree with the appeal's decision, the possibility of recourse to a supervisory authority of that specific market is always available. It is our understanding that we can take inspiration from these procedural methods, adapting them to ADM systems. Thus, it is important that data controllers provide for a specific department responsible for analysing these decisions (whether human or machine-controlled ones) and that Governmental Services, Arbitral Centres and Supervisory Authorities are technically prepared to analyse appeals that refer to ADM.

Ultimately, article 79 of GDPR grants data subjects the right to an effective judicial remedy against a controller or processor, if they consider that their rights under the GDPR have been infringed due to non-compliance with the Regulation. Data subjects also have the right to receive compensation if the infringement of the GDPR has caused material or non-material damage on them, from such controller or processor, who shall be held liable for that damage, in accordance and subject to the conditions of GDPR's article 82.

5.3. What about a right to an explanation?

Having stated all that, contesting a decision without at least a simple but meaningful explanation on the grounds behind it, would probably remain difficult for the data subjects to enforce their rights. Considering the pattern in the 'traditional world', where decisions are solely made by humans, if the concerned party disagrees with a decision attributed to him or her and intends to appeal against it, the minimum requisite for such contestation is an explanation on the reasoning of the decision. Moreover, the lawmakers behind the GDPR thought the same, at least before they released the final version of the Regulation where they included the right to an explanation as a suitable measure[40]. As we saw, the final version of the provision does not include that, however, the wording of the paragraph on "at least" indicates that other measures can be included.

It is our understanding that, the Recitals here, serve not only as a guidance for interpretation but also for a broader perception of the

[40] Prior to the release of GDPR's final version, the right to an explanation was included in the number 5 of article 20 as a suitable measure to safeguard data subjects' rights, freedoms and legitimate interests. Vid. European Parliament legislative resolution of 12 March 2014. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52014AP0212&from=PT>. Last visited on 17.04.2021.

minimum suitable safeguards, including in fact a right to an explanation of the decision reached after such assessment[41].

The discussion around the existence and enforceability of this right to an explanation has been quite inflammatory leading to lots of pages being written about it[42]. Scholars, specifically Wachter et al.[43], have argued that, while the GDPR grants a right to an *ex ante* generic explanation about the system functionality, which is almost equivalent to the traditional right to be informed (and, therefore, does not add to the information rights already in place), a right to an *ex post* specific explanation about the decision's rationale is not clearly expressed in the GDPR, besides the mention on Recital 71, which, as an interpretative mechanism, is not legally binding[44]. Nevertheless, at this point, the majority of the literature on this topic, to which we subscribe, seems to agree that this reasoning is erroneous and that, in fact, we can perceive this right to an explanation as a suitable and enforceable safeguard[45]. The WP29 upholds this interpretation as

[41] Recital 71 states for the matter "In any case, such processing should be subject to suitable safeguards, which should include specific information to the data subject and 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."

[42] This matter, as mentioned, has been thoroughly discussed in literature and plenty of pages could be written about it. However, for the purpose of this paper, and in order to limit the analysis to the essential of the topic, we refer, for a more in-depth investigation, to Kaminski, Margot E., "The Right to Explanation, Explained", *Berkeley Technology Law Journal*, Vol. 34, 189, 2019, p. 190-217.

[43] Wachter Sandra, Mittelstadt Brent and Floridi Luciano, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation", *International Data Privacy Law*, 2, 2017, p. 76–99.

[44] Idem. Though the position of Wachter et al. is more complex and nuanced than what was briefly explained, for the purpose of limiting to the essentiality of the topic, and as mentioned in a previous footnote, we refer, for further development, to the authors analysis.

[45] Goodman Bryce and Flaxman Seth, "European Union regulations on algorithmic decision-making and a 'right to explanation'", *AI Magazine*, Vol. 38, 3, 2017, p. 50-57 (They recognize the existence of the right to explanation in the GDPR, stating that "The law [referring to the GDPR] will also effectively create a "right to explanation," whereby a user can ask for an explanation of an algorithmic decision that was made about them."). Brkan Maja, "Do Algorithms Rule the World? Algorithmic Decision-Making in the Framework of the GDPR and Beyond", *International Journal of Law and Information Technology*, Vol. 27, 2, 2019, p. 91-121 ("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

well, when referring to the need for this transparency mechanism since an individual can only challenge a particular decision or express his or her view if he/she actually understands "how it has been made and on what basis.". Meaning that this right to an explanation is essential to enable data subjects to invoke the other rights explicitly enumerated in article 22 (3).

Accordingly, the kinds of information that should be provided by data controllers in this case are exemplified in those same guidelines. Indeed, it is mentioned that individuals should be provided with the categories of data that have been used in the process and why those categories are considered relevant. Furthermore, information on "factors taken into account for the decision-making process, and (...) their respective 'weight' on an aggregate level (...)" are expected to be provided as well as a simple explanation on how the profiles are built, why they are relevant to the decision-making process and how it is used for it[46].

On an ending note, companies should also make an effort to provide a simple and intelligible explanation to individuals that, mainly, are not well educated on this topic, using, for instance, visual schemes and user-friendly information.

6. Brief look over Paragraph 4

In the last paragraph of article 22, there is the reference to specific categories of data, such as health data or ethnic data which in general are

case law which regularly uses recitals as an interpretative aid.") and Selbst Andrew D. and Powles Julia, "Meaningful Information and the Right to Explanation", *International Data Privacy Law*, Vol. 7, 4, 2017, p. 233-242 ("Recital 71 is not meaningless, and has a clear role in assisting interpretation and co-determining positive law."). Malgieri Gianclaudio and Comandé Giovanni, "Why a right to legibility of Automated Decision-Making exists in the General Data Protection Regulation, *International Data Privacy Law*, Vol. 7, 4, 2017, p. 243-265 ("The right to obtain an explanation of the decision reached after the assessment should always be exercisable."). Kaminski Margot E., "The right to explanation, explained", *Berkeley Technology Law Journal*, Vol. 34, 1, 2019, p. 189-218 ("an individual has a right to explanation of an individual decision because that explanation is necessary for her to invoke the other rights – e.g., to contest a decision, to express her view – that are explicitly enumerated in the text of the GDPR.").

[46] WP29 p. 27-31.

forbidden to be processed, as quoted in article 9 (1). However, article 9 (2) sets some exceptions to such prohibition, and two of them, items (a) and (g), are also referred in 22 (4). If a controller uses ADM and it falls under the exceptions of article 22 (2) but the data is, for example, genetic data, it can only have a processing and a solely automated decision upon it, if the processing of that specific data is allowed. The permission will occur when the data subject has given his/her explicit consent [item (a)] or because it is a matter of public interest [item (g)]. This dual protection aims to safeguard the rights of data subjects, since this data is more sensitive and creates higher perils to individuals and, for that reason, more effective safeguards are paramount.

7. Children and Profiling

When the GDPR aims to protect data subjects from the automated decision-making systems, it does not specify in article 22 which subjects it refers to, therefore we could admit that it equally includes children. Under Human Rights Law, children are under extreme protection and their rights always need to be safeguarded. Respectively, Recital 71 refers that children shall not be subject to ADM when any of the exceptions predicted in the second paragraph occur. Even though it is projected in the recital, it is not binding, which leaves companies and organizations in a glassy floor, where doubt has the main role. To give some light to data controllers WP29 advises that, if the ADM fits in one exception of article 22(2) and the processing's end is the welfare of children, such as health care or education, it can be admissible. Anyhow, the safeguards of article 22 (3) and the best interest of the child shall be figured perpetually.

8. Good Practices for Data Controllers on the use of ADM

More than a mere explanation, it is inherent in the formulation of the rights conceded by paragraph 3, that this provision requires the implementation of the necessary mechanisms to ensure its ultimate goal – rendering automated decisions contestable.

As an effort to ease the concerns of data controllers, companies and organizations on having to completely disregard automated decisions, WP29 provided suggestions on good practices to apply, to ensure the compliance with the GDPR, such as, the implementation of regular quality assurance checks on their systems; the conduction of internal or external audits to the algorithms used or developed by machine learning systems, depending on the level of risk of the decisions on the individuals' sphere; the application of anonymisation or pseudonymisation techniques to ensure a higher level of protection; the strict compliance with the data minimisation principle, by establishing clear and strictly necessary retention periods of the personal data processed[47].

In line with the concerns brought by companies, we could say that the GDPR could have defined the "suitable safeguards" in the provision, to restring or impose determined measures, but did not. Their choices provide a possibility for companies to argue that the definition of those safeguards was meant to be flexible and adaptable to the market. It is our point of view that with the implementation of these and other good practices they can still conduct automated decision-making processes when allowed by the legislation.

Conclusion

Automated decision-making systems that have zero human intervention are considered great threats to data subjects' rights and freedoms. Article 22 (1) must be understood as a general prohibition of certain types of automated decisions and a passive right that controllers have to observe when taking them, without an active claim from the data subject. The criteria of this prohibition are quite specific, however, the threshold becomes thinner when the significant similar legal effects need to be assessed, as it is a broad undefined concept that can be assessed under different perspectives to determine if an automated decision falls within the scope of the article. The data controller shall conduct this assessment whenever an automated decision is at place, according to article 35 (3) (a).

[47] Other suggestions are provided in Annex 1 of WP29, p. 31-32.

Data Protection Impact Assessment is an excellent mechanism for ensuring compliance with the GDPR and it should be carried out every time a processing takes place.

The article analysed also requires, in paragraph 3, the implementation of suitable safeguards, when the automated decision falls under the exceptions of paragraph 2. Those include the right to obtain human intervention, to express his or her point of view and to contest the decision. Though the legal consequences of these rights remain quite unclear and require further development by a competent entity, like the EDPB or the ECJ, in the context of a new decision, these are seen as a minimum standard to be met by data controllers, that can and should provide more measures to ensure data subjects' protection.

On that note, though not explicit in the provision itself, a right to an explanation is considered to play an essential role in the exercise of the other suitable safeguards. Indeed, to counteract the maleficence algorithmic decisions, it is our understanding that all decisions based on automated decision making must be possible to be explained and understood not only by the subjects of those decisions but also by those who work side by side with this technology.

In sum, this article argued that data controllers, companies and organizations can still conduct these automated decision-making processes, when allowed by the legislation, and if they implement some good practices as a meaning to achieve the ultimate goal of this provision – protecting the data subjects' rights, freedoms and legitimate interests.