Two years in: Does the GDPR already need updates?
A question brought by algorithmic decision-making

Beatriz Santiago Trindade[*]

Resumo: Hoje, mais do que em outro momento, vive-se uma (r)evolução tecnológica constante. Numa altura em que os dados pessoais são perspetivados como bens transacionáveis, valiosíssimos, também já referidos como "o novo petróleo", cumpre-nos verificar se, num contexto determinado, os textos legislativos neste âmbito providenciam uma proteção adequada. Estaremos, porventura, a ser excessivamente cautelosos a apontar para a necessidade de alterar o Regulamento Geral de Proteção de Dados? Ou será que a solução cabe apenas às entidades que desenvolvem as novas tecnologias de Inteligência Artificial?

Palavras-chave: *Proteção de Dados; Inteligência Artificial; Machine Learning; Dados Pessoais; Algoritmos*

Abstract: Nowadays, more than ever, we live a constant technological (r)evolution. And, today, in a time when personal data are prospected as very valuable "exchange" goods, hence already called "the new oil", it is up to us the need to certify ourselves whether present legislative texts are able to provide an adequate protection to it. Or are we already, perhaps in a very cautious way, facing the necessity to update the General Data Protection Regulation? Maybe the solution lies with the entities that develop the new Artificial Intelligence technologies...?

Keywords: *Data Protection, Artificial Intelligence; Machine Learning; Personal Data; Algorithms*

[*] Licenciada em Direito pela Faculdade de Direito da Universidade de Coimbra. Mestre em Direito Internacional e Europeu pela Faculdade de Direito da Universidade Nova de Lisboa. DPO certificada pelo Centro Europeu de Privacidade e Cibersegurança da Faculdade de Direito de Maastricht.

Introduction

Nowadays, there is absolutely no doubt that technology is all around us. It is in our house, our cars, our streets, through the use of internet, computers and other technologies, such as Internet of Things (IoT). Tech devices are practically omnipresent, namely in fields like Health, Finance or Education. The machines are mostly fed through data that us, Humans, insert in them, learning through the complex algorithms how to analyse data sets and make predictions based on them.

The technology we use has, in its inner workings, lots of mechanisms, namely *Machine Learning* and *Artificial Intelligence*[1], which works through the work of hand-coding the solution to each problem (for example, that can be helping someone going from A to B or translating text between two or more different languages). To better explain this, one shall first acknowledge that the machines collect and process big amounts of data, and amongst those data, there are personal data[2].

The machines achieve the solution which they are confronted by analysing the data we feed them and then finding patterns between those data. Through these actions, one tends to think there isn't any (human) bias. Nevertheless, it is urgent to remember that, by the sole use of data, the machine is not necessarily neutral.

Even with the best intentions, it is very difficult, if not practically impossible, to separate the developer from his/her own bias, which are inherent to every human being. Inevitably, our own human bias is transferred to the machine (well, mostly algorithms and software), since the idea that humans are biased by nature is supported by some

[1] Through the use of the terms of *Machine Learning*, as well as *Artificial Intelligence,* one pretends to refer to systems based on decision-making algorithms. In this sense, the machines are mainly used to help reach a decision or formulate some kind of recommendation for action. In short, we are going to analyze how decisions reached by a machine can affect its users, while also taking into account that data protection is, in its own nature, a fundamental right and has to be, therefore, balanced against many other fundamental rights that are established by the Charter of Fundamental Rights of the European Union (CFREU).

[2] Council of Europe, *"Handbook on European data protection law – 2018 edition"*, Luxembourg: Publications Office of the European Union, 2018, p. 347. Also available online <https://fra.europa.eu/sites/default/files/fra_uploads/fra-coe-edps-2018-handbook-data-protection_en.pdf>. Consulted on 15 October 2019.

authors[3]. Biases are often seen a form of optimizing the brain functions and have their root in individual experiences, as well as the social and cultural environment in which an individual is inserted[4].

This can be seen as a very complex issue, since technology must serve everyone. So the machine may "suffer" from different kinds of biases[5], such as interaction bias[6], latent bias[7] and selection bias[8].

The bigger question we now face is: as we develop more and more technology to make our lives better or easier, how do we keep our biases out of the algorithms we create, whilst protecting the data from numerous data subjects for the construction of the machine's sample?

Considering the aforementioned human bias, are the machines also necessarily (and/or inevitably) biased? And how does our legislation, namely the EU General Data Protection Regulation (GDPR)[9], keep our rights safe[10]?

[3] MOSKOWITZ, Gordon, *Are we all Inherently biased?,* Lehigh University. Article available in: https://www1.lehigh.edu/research/consequence/are-we-all-inherently-biased. Consulted on 12 February 2020.

This Author is quite clear when defending this idea: "(...) While I would say "no" to the question of whether stereotyping is inevitable, I would answer in the affirmative to the question of whether people are inherently biased. I see all human thought and action on the environment as always in the service of the goals of the person within that environment. These goals may be invisible to the naked eye, implicit (unconscious). But we are always pursuing a goal with every action, with every thought. Thus every action and thought is biased by these goals. (...)"

[4] XIANG, Mark (2019), *Human Bias in Machine Learning – What it means in our modern big data world*. Towards Data Science. Retrieved from https://towardsdatascience.com/bias-what-it-means-in-the-big-data-world-6e64893e92a1. Consulted on 12 February 2020.

[5] To understand this, the visualization of this short clip is advised: https://www.youtube.com/watch?v=59bMh59JQDo. These are not the only biases present in Machine Learning mechanisms, though.

[6] The algorithm becomes biased by the way the user interacts with it.

[7] These types of biases are mainly related with elements such as gender, income, race, or other characteristics.

[8] Selection bias happens when an algorithm favours one population or segment of population, at the expense of other subjects.

[9] 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)

[10] Cfr. DREYER, Stephan, SHULZ, Wolfgang, *The GDPR and algorithmic decision-making-Safeguarding individual rights, but forgetting society*. Völkerrechtsblog – International Law &

Despite the unquestionable benefits[11] brought by technologies designed with the intent of making our lives better and easier, one cannot forget that in order for those technologies to function properly or as intended, a feed of our data is needed[12]. Errors or bias, not only among the collected and shared data, but also in the – or as a result of the – automated decision-making process, can lead to inaccurate classifications, assessments based on imprecise projections and a negative impact on individuals[13].

It is not excessive to recall we are dealing with fundamental rights, and, therefore, worthy of the highest protection possible.

The present paper aims to confront the reader with the possible obstacles new technology poses to the community, whilst trying not to disregard the real benefits these innovations give us. It does not pretend to give a close answer to the general problem of automated decision-making, but at least bring the question to present minds as they assist to newer and more complex developments on this particular field.

Does the solution lie in the beginning of the process, when we insert data in the machines and work the algorithms out, or, on the contrary, it resides in the aftermath, the legislation that regulates the use that tech makes out of our data?

International Legal Thought, 2019. Retrieved from https://voelkerrechtsblog.org/the-gdpr--and-algorithmic-decision-making/. Consulted on 15 October 2019.

[11] For example, in the beginning of the new Corona virus (nCov-2019) outbreak, a Canadian health monitoring platform alerted to the possibility that an epidemy was about to begin. BlueDot uses an AI-algorithm that analyses and combines foreign language news reports and animal and plant networks, which ultimately led to the conclusion that the outbreak was about to take place. https://www.wired.com/story/ai-epidemiologist-wuhan-public-health-warnings/amp?fbclid=IwAR2LqZc2UB3DtDg--ccYPHuBzmB-voCOyFPp02nWuaELkJHstFupuExZhgYY. Consulted on 08 February 2020).

[12] Council of Europe, *"Handbook on European data protection law – 2018 edition"*, Luxembourg: Publications Office of the European Union, 2018, p. 347

[13] Article 29 Data Protection Working Party (W29), *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679.* Adopted on 3 October 2017 and last revised and adopted on 6 February 2018, page 27.

1. The relation between Artificial Intelligence (in general) and the GDPR

EU GDPR adopts a risk based approach towards data protection, and it is very easy to understand why: we are dealing with fundamental rights and, as one can infer from reading the Charter of Fundamental Rights of the European Union (2012/C 326/02), these rights must be balanced against each other.

GDPR focuses on the personal/individual dimension of the (data) subject, while Artificial Intelligence focuses more on the collective/group dimension. So, in a way, GDPR presents itself as inadequate for AI regulation. But, on the other hand, the GDPR is applicable when dealing with cases of the development of Artificial Intelligence where personal data and its use is destined to reach a decision about individual subjects[14].

This legislative instrument bases itself on data processing, which can be, in general lines, any operation or set of operations performed on personal data, whether by automated means or not, such as collection, recording, storage, organization, structuring, among others[15]. All of the actions just mentioned may be performed by an Artificial Intelligence device, and that is the approach adopted in the GDPR, through four strong challenges to the development of these innovations: data minimisation, purpose limitation, fairness and discrimination, transparency and right to information[16]. Although these principles are presented as difficult barriers to cross, one must conceive that, in reality, it is possible to use and develop AI technologies while safeguarding fundamental data protection rights. Nowadays, human behaviour is being highly scrutinized at the expense of decisions based on algorithms, which alerts individuals to the urgency of protecting them.

[14] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 15.

[15] Article 4(2) GDPR.

[16] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 4.

1.1. Data minimisation principle

To sum this principle in very few words, it resonates in the expression *the need to know is distinct from being nice to have*. In this way, data must be limited to what is necessary to the purpose, as it is clearly stated in article 5(1)(c) of the GDPR. This echoes the necessity for the data to be correct, updated and, most importantly, must not be retained for a longer period of time necessary for the purpose that justified the data collection in the first place.

Recital 156 of the GDPR is also of great importance in this matter. It states that the processing of personal data for purposes of public interest, scientific, historical or statistical research must be subject to appropriate safeguards regarding the rights and freedoms of the data subject – which means, not only, but also, that those safeguards ensure that technical and organisational measures must be put in place in order to guarantee the respect for the principle of data minimisation.

This can easily present as a challenge for individual automated decisions, since it is widely understood that the more training data is inputted in the machine, the better will the result be[17]. It is logical that the more data we feed the machines, the more accurate they will be, so their behaviour will be able to mimic our patterns with a shorter margin of error.

But how can we know where to draw a line? When and how do we know that enough is enough?

Regarding the data minimisation principle, we shall begin with a small sample of training data and then study the machine's learning curve to assess whether we need to input more data, and which data[18].

This principle does not limit itself to the regulation of the amount of data (to be) processed. Because we are dealing with the interference in a data subject's fundamental rights[19], the minimisation principle

[17] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 11.

[18] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 12.

[19] One can note that, according to Articles 7 and 8 of the Charter of Fundamental Rights of the European Union (CFREU), data protection and privacy are distinct, as processing of

also stipulates the proportionality of data processing[20]. About this question, the available solution reached to this moment relates to the extension to which is possible to identify the data subject. One can recognize that security measures such as encryption, anonymisation[21] and pseudonymisation[22] are very useful to preserve the data subject's identification, besides restricting the amount and nature of information used in the automated decision-making process[23]. These measures are not exhaustive, as other means exist, such as purposefully restricting the categories of data collected from a data bank, even if the data subject is still identifiable.

1.2. Purpose limitation principle

The definition of the purpose for processing personal data may also present as very complex due to the technical challenges inherent to system's development. At the moment data are collected, the underlying purpose must be already determined, also needing to be specific, explicit and legitimate[24].

personal data can be carried on without interfering with the subject's privacy. One may also interfere in a data subject's privacy without perform data processing.

[20] Proportionality is one of the key principles of EU law, as it assures that there is not an unnecessary disregard and violent compression of fundamental rights, like personal data rights – https://edps.europa.eu/data-protection/our-work/subjects/necessity-proportionality_en. Consulted on 17 October 2019.

[21] The Irish Data Protection Commission defines anonymisation in its *"Guidance Note: Guidance on Anonymisation and Pseudonymisation"*, June 2019, as the processing of data "(...) with the aim of irreversibly preventing the identification of the individual to whom it relates" (p. 2). This document is available on: https://www.dataprotection.ie/sites/default/files/uploads/2019-06/190614%20Anonymisation%20and%20Pseudonymisation.pdf.

[22] *Idem*, defining pseudonymisation as the replacement of "(...) any identifying characteristics of data with a pseudonym, or, in other words, a value which does not allow the data subject to be directly identified" (p. 3).

[23] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 18.

[24] Article 5(1)(b) of the GDPR.

Without adding any explanation, it may seem that the act of processing may later reveal itself forbidden if the purpose is changed in any way. One shall say, then, that the processing should stop immediately if it is incompatible with the original purpose.

This would present as a huge challenge to the development of new technologies in general, as many times it is very difficult – or even impossible – to outline a purpose for the collection and for the afterwards processing of data. The algorithms, when working in a black box AI and Machine Learning systems, can be very uncertain in what they can learn from data, thus the purpose can change as the machine is developed[25].

The key here will be to assess the compatibility between the initial purpose and the "new" purpose on a case-by-case basis. To do this assessment, one may analyse: the relationship between purposes for which the personal data have been collected and the purposes for further processing; the context in which the personal data have been collected and the reasonable expectations of the data subjects as to their further use; the nature of personal data and the impact of the further processing on the data subjects; the measure adopted by the controller[26] to ensure fair processing and to prevent any undue impact on the data subjects[27].

This means that it is possible to use the data beyond the original purpose in which the collection was based, but this usage cannot be unrestricted and out of control. This is of crucial importance especially

[25] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 18.

[26] In this sense, the controller must always be the one designated for this function by the owner of the system where the data in question is processed. Being the one who determines the how, when and what of processing, it must, however, be conceived that this person can use third parties to carry out the tasks required by the data processing.

However, in cases where there is no coincidence between that person and the person who conceives the algorithm, e.g. when the provision of this instrument is made by a third party without the controller mastering the internal operation of the algorithm, the powers and responsibilities arising from the controller's position as guarantor must cover the third party.

[27] Recital 50 of the GDPR.

regarding data subject's rights to access, rectification, erasure, among others[28].

So, in case that the new purpose presents itself in counterbalance with the purpose that allowed the collection of the data, the controller must assess its compatibility[29]. As said by the Article 29 Data Protection Working Party, before the GDPR came into force, "Further processing for a different purpose does not necessarily mean that it is incompatible: compatibility needs to be assessed on a case-by-case basis"[30]. If the purposes come to be incompatible between them[31], the controller may seek new consent to keep processing the data[32], or analyse which legal basis[33] appears to be more suitable. However, if the new purpose, which was not originally projected, reveals compatible with the original the controller may not need a new lawful basis[34].

[28] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 16.

[29] Article 6(4) of the GDPR established legal basis for further processing. If the controller has collected data on the basis of a contract, legal obligation, protection of vital interests of the data subject or performance of a task carried out in the public interest or in the exercise of official authority, the data can then be used for the new purpose if it reveals compatible with the original (https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/principles-gdpr/purpose-data-processing/can-we--use-data-another-purpose_en. Consulted on 12 February 2020).

[30] Article 29 Data Protection Working Party, *Opinion 03/2013 on purpose limitation*, adopted on 2 April 2013, page 3.

[31] UK's DPA highlights the cases regarding the changing of lawful basis very clearly: "You must determine your lawful basis before starting to process personal data. It's important to get this right first time. If you find at a later date that your chosen basis was actually inappropriate, it will be difficult to simply swap to a different one. Even if a different basis could have applied from the start, retrospectively switching lawful basis is likely to be inherently unfair to the individual and lead to breaches of accountability and transparency requirements." (Information Commissioner's Office, *Lawful Basis for Processing.* Retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection--regulation-gdpr/lawful-basis-for-processing/. Consulted on 12 February 2020).

[32] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 17.

[33] Article 6 of the GDPR, articulated with Article 9 if the processing refers to special categories of data.

[34] Information Commissioner's Office, *Lawful basis for processing*, retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-

Differently, failure to comply with the compatibility requirement will lead to an unlawful data processing, which therefore is not permitted[35].

The purpose limitation is especially important for the data subject in case the data subject desires to exercise control over their own personal information. The main objective of this principle relies on the avoidance of ambiguity of data processing, thus the specification of purposes must be understood by the concerned data subjects (as well as others, such as Data Protection Authorities), disregarding different and linguistic backgrounds, as well as any intellectual or special needs. Through this measure, it will be possible to reduce "(...) the risk that the data subjects' expectations will differ from the expectations of the controller. (...)".[36]

The rights of the data subject are the consequence of data protection being a fundamental right, hence the need of the controllers need to be transparent about how the deal with these rights in a concise, easily accessible manner, with clear and plain language.

This is the reason why the GDPR additionally contemplates the right to information[37], right of access[38], right to rectification[39], right to object[40],

data-protection-regulation-gdpr/lawful-basis-for-processing/. Consulted on 12 February 2020.

[35] Article 29 Data Protection Working Party (W29), *Opinion 03/2013 on purpose limitation*, adopted on 2 April 2013, pages 36 and 40.

[36] Article 29 Data Protection Working Party, *Opinion 03/2013 on purpose limitation*, adopted on 2 April 2013, page 17.

[37] Established in Articles 13 and 14, it is stated the controller needs to explain to the data subject what data is processed, for which purpose, by whom and with which parties that data are shared.

[38] Article 15 of the GDPR states that all data subjects have the right to access their own data, following authentication of their entity.

[39] GDPR, in its Article 16, establishes that any inaccurate or incomplete data can be rectified by the data subject.

[40] This right is only applicable in case of processing for public or legitimate interest, or direct marketing. Nevertheless, Article 21 states that it can be overridden by the data controller with compelling arguments.

right to erasure[41-42], right to restriction of processing[43] and right do data portability[44].

1.3. Fairness and non-discrimination

As previously mentioned, there is a tendency to forget that, in the basis of the construction and design of a device, there is a human-being, an engineer or developer, or even a whole team. Thus, we can conceive that

[41] Article 17 of the GDPR lays out the scenarios in which the data subject has the right to have his-her information erased from a database: when the data is no longer necessary for its underlying purpose, consent for processing is withdrawn, the right for object is exercised and cannot be overridden, the data have been processed unlawfully, a legal obligation to delete the data applies or data have been collected by *information society services* based on consent by a minor.

[42] A clear example of the importance of this principle is seen in Court of Justice of the European Union, *Google Spain v AEPD and Mario Costeja González* – C-131/12, 13 May 2014, even before the GDPR came into force. (Available on http://curia.europa.eu/juris/document/document.jsf?docid=152065&doclang=en). This decision by the Court of Justice of the European Union establishes the responsibility of an Internet search engine operator (in this case, Google) for the data processing of personal information that appears on web pages published by third parties. In 1998, the name of Mario Costeja González appeared in relation to a *La Vanguardia* article that focused the forced sale of properties arising from social security debts. Later, in 2009, Mario González contacted the newspaper, complaining that, whenever his name entered in the Google search engine, it led to these announcements. *La Vanguardia* replied saying that erasing was not appropriate due to a publication of the Spanish Ministry of Labour and Social Affairs. Thus, González contacted Google Spain asking for the announcements to be removed, while simultaneously lodging a complaint with the Spanish DPA, AEPD. In the end, the Court of Justice of the European Union ruled that whenever the data processing is "inadequate, irrelevant or excessive", it may be incompatible with the Directive 95/46/EC (which is the predecessor of the GDPR).

[43] This right, contemplated in Article 18 of the GDPR, allows the data subject to ask to restrict the processing of his/her data, if the data or lawfulness of processing is contested. Thus, if the process is restricted, the data cannot be further processed without consent of the data subject (or for legal defence).

If data are corrected, erased or the processing is restricted, the data controller has the responsibility to inform all recipients of the changes.

[44] Article 20 establishes that data subjects are entitled to take their data from one data controller to the other if processing is based on consent or contract or the data are processed by automated means. This right empowers data subject while ensuring free flow of data.

there can be discrimination intentionally or unintentionally embedded in the algorithms, especially if the training data generates biased results, and this use of personal data is in clear contradiction with the fairness principle[45].

Biases can originate from a variety of elements related to the development of Artificial Intelligence tech: methods (measurement, survey, pre-processing stages), datasets (social bias due to historical bias and/or misrepresentation of some categories), data sources (selection bias), data scientists (confirmation bias)[46].

So, the controller, owner of the technological device or algorithm or license holder of the algorithm owned by an external service provider, must be responsible – and able – for the implementation of measures that avoid reaching biased results. However, this may reveal itself insufficient to comply with the fairness principle. Thus, assessment and investigation by the controller is needed, in order to ensure this principle is respected[47].

1.4. Transparency and right to information

Transparency[48] is the key right of the data subject: it is in this way that it is assured to him/her that the relevant information is received, that they are given an explanation, at the time of the data collection, of what actions will take place and on what lawful basis the processing relies on. In conclusion, controllers must – again, to be able or in position to – inform the data subject about processing details (these cannot be limited to general lines, must also include rules, risks, safeguards and data subjects' rights)[49]. The information must be provided to data

[45] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 16.

[46] Information presented in Professor Alessandro Mantelero's talk "Personal Data Protection and AI – Challenges and Remedies" (available to watch in https://www.youtube.com/watch?v=Jp3LhIG6M1A).

[47] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 16.

[48] Article 5(1)(a) of the GDPR.

[49] Datatilsynet – The Norwegian Data Protection Authority, *"Artificial intelligence and privacy – Report, January 2018"*, p. 19.

subjects in a plain, clear and accessible form. Article 71 precisely states that the processing shall 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". Therefore, the controller must be able to explain to data subjects the rationale behind the algorithmic decision-making to the data subjects[50].

It is in this sense that the GDPR stipulates the right to information in its articles 13 and 14. From these rules it is possible to conclude that the obligation for the controller to explain, to the data subject, what data is processed, when not collected directly from the data subject, for which purpose(s), by whom and with which parties the data is shared is clearly established.

So, we can clearly understand how this presents as a challenge to algorithm decision-making advances. The more advanced the technology is, the more difficult it tends to be perceived, mostly because of the complexity of processes behind it[51]. It is also problematic in the sense that we may be dealing with Intellectual Property Rights and commercial secrets, like the Recital 61 of the GDPR establishes[52].

Although these challenges are very present and seem difficult to overcome, it is important to underline that these principles clearly regulate automated decisions, as clearly stated in article 22 of the GDPR. This article establishes that "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". This article, however, in its paragraph 2, presents three exceptions to that prohibition: when the decision is necessary for the performance of a contract between the data subject and data controller; when the decision is authorized by EU or national law to which the data controller is subject, since this legislation establishes

[50] W29, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679,* adopted on 3 October 2017, page 14.

[51] *Idem.*

[52] *Idem.*

measures to safeguard the data subject's rights; when the decision is based on the data subject's explicit consent.

However, there are already authors[53] who argue that the right of explanation established in the GDPR does not represent an answer to the problems mentioned above. Nevertheless, they defend that these explanations have positive aspects, such as helping data subjects/users to trust and make a better use of the systems and allowing them to project a draft of how it works[54]. In this sense, Edwards and Veale advocate that attention should be drifted from the data subjects to the intention of building better systems *ab initio*, as well as give powers to the competent agencies to analyse and eventually correct the algorithms bias, accuracy and integrity[55].

2. GDPR's regulation on automated decision-making

The GDPR may shape the development of Artificial Intelligence and Machine Learning in two distinct manners.

On one hand, this legislative instrument is truly focused on the enhancement of data security, as it states strict obligations to controllers and processors[56], knowing that Artificial Intelligence devices require extreme large data sets of varied nature to analyse, and personal data are, most of the time, among these[57].

All data subjects have the right[58] not to have their data processed uniquely by automated means when those involve decisions with specific effects on data subjects. This is the ultimate goal of article 22(1) of the

[53] EDWARDS, Lilian; VEALE, Michael, *Slave to the Algorithm? Why a 'Right to an Explanation' is probably not the remedy you are looking for*, Duke Law & Technology Review, May 2017. Available here: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972855.

[54] *Idem*, page 22.

[55] *Idem*, page 23.

[56] OLEKSIUK, Anna, *How to Train an AI with GDPR Limitations – Learn how AI companies can comply with the new European data protection regulation*, Intellias – Intelligent Software Engineering, 2019. Retrieved from https://www.intellias.com/how-to-train-an-ai-with-gdpr-limitations/. Consulted on 26 October 2019.

[57] *Idem*.

[58] This right is not new: it was already established in article 15 of the Directive 95/46/EC.

GDPR. By "solely automated" the legislator meant "(...) a decision-making process that is totally automated and excludes any human influence on the outcome. A process might still be considered solely automated if a human inputs the data to be processed, and then the decision-making is carried out by an automated system"[59].

The interpretation of article 22 must be in line with the fundamental principles of the GDPR, according to which the data subject has control over the use of their personal data[60]. This means that the prohibition established in this article can only be applied in limited circumstances; more precisely, only if the decision based solely on automated processing or profiling has a legal effect on the data subject[61]. Not to disregard, then, the logical requirement of safeguarding measures, namely the right to be informed (articles 13 and 14) and the right to challenge the decision [article 22(3)][62]. Nevertheless, is very difficult to draw a line regarding what decisions should be considered to "significantly affect him or her", although W29 gave some examples, such as decisions that can affect an individual's financial circumstances, employment opportunities or access to health services and education[63].

A process will not be considered solely automated if someone weighs up and interprets the result of an automated decision before applying it to the individual[64]. Essentially, if someone steps in at some point of

[59] Information Commissioner's Office, *What does the GDPR say about automated decision-making and profiling?*. Retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-does-the-gdpr-say-about-automated--decision-making-and-profiling/. Consulted on 13 February 2020.

[60] Article 29 Data Protection Working Party (W29), *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, adopted on 3 October 2017 and last revised and adopted on 6 February 2018, page 20.

[61] *Idem*.

[62] *Idem*.

[63] *Idem*, page 22.

[64] Information Commissioner's Office, *What does the GDPR say about automated decision-making and profiling?*. Retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-does-the-gdpr-say-about-automated--decision-making-and-profiling/. Consulted on 21 October 2019.

data processing, reviewing the decision, the process no longer fits in this definition.

The legislative instrument that is today the centre of data protection regulation, in its article 22, clearly states that it applies to automated individual decision-making and profiling with legal or similarly significant effects on the data subjects, and this type of processing is restricted[65]. Therefore, for this process of decision-making to be lawful, it must be "necessary for the entry into or performance of a contract; authorized by Union or Member state law applicable to the controller; or based on the individual's explicit consent"[66].

So, what one has to conclude after knowing this is that the controller must assess whether the processing activity fits itself in the scope of the mentioned article. If that is the case, the controller must provide all the processing information to data subjects, introducing simpler ways for them to require human intervention to contest any decision that affects them. The controller must perform all checks to assess the system's regular functioning[67].

Data protection, although being, as referred before, a fundamental right that deserves protection under the Charter of Fundamental Rights of the European Union[68], must not be seen as an unremovable obstacle to the use of innovative and data-driven technologies. The principles mentioned above (data minimisation, purpose limitation, fairness and non-discrimination, transparency and right to information) must serve as guidance for controllers regarding the processed personal data[69].

[65] Information Commissioner's Office, Rights related to automated decision making including profiling. Retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to--automated-decision-making-including-profiling/. Consulted on 21 October 2019.

[66] Article 22(2) of the GDPR.

[67] Information Commissioner's Office, Rights related to automated decision making including profiling. Retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to--automated-decision-making-including-profiling/. Consulted on 21 October 2019.

[68] Thus, it must be balanced against other fundamental rights contemplated in CFREU, such as the freedom of conducting a business (article 16).

[69] *Idem.*

3. What about Data Protection Impact Assessments?

The GDPR clearly states, namely when using new technologies, that controllers must carry out an assessment of impact of the processing regarding protection of personal data[70] before they begin to process data. These are particularly important when the data processing operations represent a menace or "(...) high risks to the rights and freedoms of natural persons. (...)"[71]. We are talking about a process that allows companies and organizations to identify and minimize risks[72], protecting themselves against possible future fines.

Then, it is easy to understand the need of Data Protection Impact Assessments (DPIAs)[73] while facing the development of decision-making tech. These are particularly useful, hence the fact that they tell if there is a high risk for the individuals' rights that cannot be mitigated, thus imposing[74] controllers the consultation with the competent Data Protection Authority.

DPIAs have proved to be extremely useful for controllers regarding the assessment of risks related to data processing, helping them to assure that their activities fall within the scope of Article 22(1), and in case of identifying an exception, to analyse which safeguarding measures must be applied[75].

[70] Article 35(1) of the GDPR; Recital 90 of the GDPR.

[71] Article 35(1) of the GDPR.

[72] Information Commissioner's Office, *Data protection impact assessments*. Retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/. Consulted on 27 October 2019.

[73] Data Protection Impact Assessments are processes designed to help regarding the identification and minimisation of risks related to data protection. In some cases, they appear as mandatory, specially facing scenarios involving high risk processing tasks that may harm data subject's rights. DPIAs are seen as crucial to mitigate these risks, even though it appears to be very difficult to eliminate these obstacles to data processing.

[74] In this particular case, prior consultation with the competent DPA is mandatory, not optional.

[75] Article 29 Data Protection Working Party (W29), *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/67,* adopted on 3 October 2017 and last revised and adopted on 6 February 2018, page 20.

DPIAs are the mirror of two key aspects (which are intertwined) brought by the GDPR: accountability and privacy by design. They are part of a risk-based approach mentioned above and must not be taken light-heartedly. They are presented as guidelines for controllers to go into detail about their processing activities, allowing them to exercise control and to demonstrate accountability for their systems[76].

The Information Commissioner's Office (ICO)[77] points out that "(...) DPIAs will force organisations to demonstrate the necessity and proportionality of any AI-related personal data processing; account for any detriment to data subjects that could follow from any bias or inaccuracy in a system; explain the rationale behind any trade-offs; and describe the relationships and the terms of any contracts with other processors or third party providers. DPIAs can also support organisations in thinking about the broader risks of harm to individuals or ethical implications for society at large. (...)[78]".

Although controllers may see DPIAs as a burden, they may reveal useful as mentioned above, as a preventive risk management measure[79]. These processes, through the audit of algorithms and regular reviews of the automated decision-making, allow controllers to assess and determine if there is the risk of any bias or errors and, in case this happens, to develop measures to minimize the potential harm[80].

[76] *Idem.*

[77] Information Commissioner's Office is UK's Data Protection Authority, which is an independent body with the goal of endorsing the information rights in the public interest. More information can be found in the following: link: https://ico.org.uk/about-the-ico/what--we-do/. Consulted on 08 February 2020.

[78] Information Commissioner's Office, https://ico.org.uk/about-the-ico/news-and-events/ai-blog-ai-auditing-framework-call-for-input-final-considerations-and-next-steps/. Consulted on 21 May 2020.

[79] *Idem.*

[80] Article 29 Data Protection Working Party (W29), *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, adopted on 3 October 2017 and last revised and adopted on 6 February 2018, page 28. The W29 also draws attention to the need for these processes to be used continuously, not only at the design stage.

A DPIA can also be used to inform the data subject about the underlying logic of an automated decision-making process, allowing him or her to oppose the decision and/or to express their point of view[81].

4. How do we proceed?

In light of everything that has been said, the main question that one can come across is how to make new inventions, new algorithmic decision-making devices, compliant and GDPR-friendly? Or, in a clearer way, how can Artificial Intelligence and GDPR co-exist harmoniously?

The W29 has already called for the need of establishing appropriate safeguards, in light of what is already established in Article 22(2)(a), 22(2)(c) and 22(3), in order to assure data subjects' rights, freedoms and legitimate interests[82]. The working party is very clear when defends that human intervention is crucial in this matter, since "(...) any review must be carried out by someone who has the appropriate authority and capability to change the decision. (...)"[83]. Recital 71 is also very important as it establishes that a data subject has the right not to be subject to a decision based solely on automated processing.

The W29 also highlights that there is a need for transparency regarding the processing of data, thus allowing the data subject to challenge the decision[84].

Artificial Intelligence devices are composed by algorithms that combine data (personal or not) inside their system, and then, after analysing and identifying patterns, give an answer to a proposed problem. As these systems are able to be trained in order to perform all kinds of

[81] *Idem,* page 30.

[82] *Idem,* page 27.

[83] *Idem.*

[84] *Idem.*

tasks, sometimes they operate as a black box[85], which raises obstacles in understanding how decisions are reached[86].

To be able to gather all the data the devices need, the controller, taking into account the rules established in GDPR, has to be able to assure that he/she has a lawful basis for the processing of that data. But since it is very difficult to fully understand how the data are being processed inside the machine, especially in the cases in which the systems operate as black box, which means that acquiring consent from the data subject, in its true form, may bring difficulties to perform his/her duties in this regard[87]. Controllers may say that, in a system based in black-box algorithms, "(...) only the algorithm itself can explain its decision-making"[88], though this is not valid for AI systems that function differently.

Controllers, for that matter, shall ensure that the technologies that are being developed by them are fully compliant, respecting the established in article 22 GDPR. Otherwise, they are at great risk of having to pay large fines[89], the amount depending on various criteria, namely the

[85] Black box is a concept in Machine Learning that defines the situations which not even the developers are able to explain how the system reached a certain conclusion. Not all Artificial Intelligence systems work this way. On other note, there are systems in which is possible to acknowledge its inner components and functions, thus allowing to get the full picture of how the system works – https://en.wikipedia.org/wiki/Black_box. Consulted on 11 February 2020.

[86] REESE, Hope, *Transparent machine learning: How to create 'clear-box' AI*. TechRepublic, 2016. Retrieved from https://www.techrepublic.com/article/transparent-machine--learning-how-to-create-clear-box-ai/?fbclid=IwAR1a1_O7wGMh3SAQP7Fq5dv76KAH7S8YhNf3vVM6Rz7AiX4D7_pLdVql3H0. Consulted on 11 February 2020.

[87] CAKEBREAD, Caroline, Can AI and GDPR Co-Exist? AI says, 'Give me more data!'; GDPR says, 'Slow down buddy!'. EMarketer, 2019. Retrieved from https://www.emarketer.com/content/what-gdpr-means-for-ai. Consulted on 23 October 2019.

[88] *Idem*. It may seem a viable solution for systems that if the decision-making system, which works as a black box, cannot be explained or totally understood, then the principles (data minimisation, purpose limitation, among others) must be enforced in the parts that we can analyse, which are the inputs (the data given) and the outputs (the answers extracted).

[89] OLEKSIUK, Anna, How to Train an AI with GDPR Limitations – Learn how AI companies can comply with the new European data protection regulation. Intellias – Intelligent Software Engineering, 2019. Retrieved from https://www.intellias.com/how-to-train-an--ai-with-gdpr-limitations/. Consulted on 23 October 2019.

nature of infringement, which type(s) of data was processed, mitigation measures, among others.

So how can we combine decision-making algorithms (and their need for big datasets) and the protection of the fundamental rights of data subjects?

From the point of view adopted in the present work, the solution may go through one of two paths:

1. the reform of GDPR, allowing it to better keep up with tech advances; or
2. development of GDPR-friendly decision-making algorithms.

The first measure may seem easier and/or more practical, because it is easy to see that European innovations will suffer while competing with nations that are not subjected to rules such as the ones established by the GDPR, like the United States and China[90]. The EU strategy presents itself as very distinct from the American and the Chinese approaches. While the US strategy relies in the development of private sector initiatives and self-regulation[91], the Chinese strategy is mainly designed by a strong coordination between the government and private and public investment in AI technologies[92].

In this sense, GDPR greatly limits the development of decision-making technologies, not only by the imposition of respect for the principles of data minimisation, purpose limitation, fairness, non-discrimination and transparency, but also because it requires giving explanations to data subjects that sometimes, neither the controllers/processor nor developers are able to give. This is easy to understand, as the complexity of solutions increases, the more difficult these can be to explain[93]. In this sense, it is

[90] CASTRO, Daniel, CHIVOT, Eline, *Want Europe to have the best AI? Reform the GDPR*. International Association of Privacy Professionals (IAPP), 2019. Retrievved from https://iapp.org/news/a/want-europe-to-have-the-best-ai-reform-the-gdpr/. Checked on 23 October 2019.

[91] For example, Microsoft is a company that has its own AI advisory board, while Google also drafted its own AI principles, which are available here: https://blog.google/technology/ai/ai-principles/. Consulted on 09 February 2020.

[92] European Parliament, *EU guidelines on ethics in artificial intelligence: Context and implementation,* European Union, 2019, page 3.

[93] *Idem*.

not convenient to limit the creation and development of new technologies for the sake of being able to explain to others how the machine and the algorithm work.

As referred before about the risk of being fined, the GDPR can present itself as an instrument that discourages the development of decision-making machines. The "fear" of being fined by DPAs for any violation may lead to the absence of innovations, making the EU less competitive worldwide, as mentioned before[94].

Regarding this, the German DPA alerts AI technologies for the need to "(...) observe fundamental rights in line with democratic and rule-of-law principles", as well to the fact that controllers must adopt organizational and technical approaches to make this possible[95].

The Hambach Declaration sets out the need of regulation regarding AI developments, however not being clear where or how to set said limitations, but knowing that there is the need to respect Ethical Principles[96].

The IAPP (International Association of Privacy Professionals), in the person of Daniel Castro and Eline Chivot, is very clear when it argues that the European policy makers have grounds to perform reforms in the GDPR. In their view, "(...) the EU should reform the GDPR for the algorithmic economy by expanding authorized uses of AI in the public interest, allowing the repurposing of data posing minimal risk, removing penalties for automated decision-making, permitting basic explanations of automated decisions, and making fines proportional to harm"[97].

Never to disregard that according to Article 97 of the GDPR[98], by 25 May 2020, and every 4 years after that date, the European Commission will issue a report on the evaluation and review of this legislative instrument to the European Parliament and to the Council[99].

[94] *Idem.*

[95] *Hambach Declaration on Artificial Inteligence – Seven Data Protection Requirements, Resolution adopted at the 97th Conference of the Independent German Federal and State Data Protection Supervisory Authorities*, Hambach Castle, 3 April 2019, pages 27 and 28.

[96] *Idem*, page 30.

[97] *Idem.*

[98] This capacity is generically established by Article 17 (2) of the Treaty of the European Union.

[99] To be able to draft these reports, the Commission has the faculty to ask informations to Member States and supervisory authorities. While exercising the competency to deliver these

On the other hand, if one thinks the solution lies within the design of GDPR-friendly algorithms, there are some suggestions already put in place[100]. Among them, there is one that stands out, which is the use of Generative Adversarial Networks (commonly referred as GANs)[101]. In the present work's perspective, these stand out because it related to the way Artificial Intelligence works.

Although it has not been referred before, "behind" Artificial Intelligence, machine and deep learning are neural networks, which are a system inspired in the operations that the human neurons perform. They are aimed to solve many problems, including signal processing and pattern recognition problems, adapting themselves along by the input of more and new information[102].

In this sense, knowing that this is a system that is embedded within the technology that is being developed, Generative Adversarial Networks can be envisioned as a viable solution because they aim to use a reduced amount of data, by using the training dataset more efficiently[103]. It is composed by two neural networks: the generator and the discriminator. Trying to explain in an accessible way, the discriminator is trained with a smaller dataset, but not so small as to introduce bias. Then, the generative network creates synthetic data, based on a randomized input and it learns as it is able to "fool" the discriminator or not. The discriminator through mechanisms of learning adjusts internal parameters proportionally to the error of the output. Thus, the more successful is the generative network in

reports, the Commission shall demonstrate causation between the progress of technologies and the need to review the GDPR.

[100] OLEKSIUK, Anna, How to Train an AI with GDPR Limitations – Learn how AI companies can comply with the new European data protection regulation. Intellias – Intelligent Software Engineering, 2019. Retrieved from https://www.intellias.com/how-to-train-an--ai-with-gdpr-limitations/. Consulted on 26 October 2019.

[101] A generative adversarial network consists of a class of machine learning systems that, when given a training set, learns how to generate new data while using the same statistics as before.

[102] ROUSE, Margaret, *DEFINITION – artificial neural network (ANN),* SearchEnterpriseAI, 2019. Retrieved from https://searchenterpriseai.techtarget.com/definition/neural-network. Consulted on 26 October 2019.

[103] *Idem.*

"fooling" the discriminator network, the greater will be the adjustments and the more the discriminator will learn without the need for real data.

However, GANs, alone, may not be the best solution to the problem that controllers are facing with the development of Artificial Intelligence and Deep and Machine Learning.

Other possible solutions are the Transfer Learning[104] and Explainable AI[105] methods, but, as the one described, face the need of big datasets to be trained.

To build accurate and complete models, one may have to combine several technical methods like these or even develop new ones. None of the current techniques presents itself as a definitive answer for the obstacles and questions raised with data protection in the European framework.

Conclusion

In the present work, there was an attempt to lay out some issues raised by the GDPR regarding the harmony between the development of decision-making algorithms and the protection of individual's personal data.

In light of what was analysed, one can come to the conclusion that there is not one right answer. At least, not yet.

Privacy and data protection rights cannot be ignored in any circumstance by the use of AI systems. Hence, it seems reasonable to argue that the best approach must be the one that combines both the need to reform some of the rules laid out by the GDPR and the development of systems that have embedded in them decision-making algorithms that obey the privacy by design and privacy by default principles.

[104] MACMAHAN, Brendan, RAMAGE, Daniel, *Federated Learning: Collaborative Machine Learning without Centralized Learning Data*. Google AI Blog, 2017. Retrieved from https://ai.googleblog.com/2017/04/federated-learning-collaborative.html. Consulted on 26 October 2019.

[105] OLEKSIUK, Anna, How to Train an AI with GDPR Limitations – Learn how AI companies can comply with the new European data protection regulation. Intellias – Intelligent Software Engineering, 2019. Retrieved from https://www.intellias.com/how-to-train-an-ai-with-gdpr-limitations/. Consulted on 26 October 2019.

Knowing that, possibly, it is asking too much, one could not choose a different way, without risking putting a burden either on the European legislator or on the developers alone.

I believe this is a teamwork, converging efforts, as, in the end, the main interest is to give proper protection to individuals who come across the scenario in which their data is collected and processed in ways which are too complex for the average person to understand.

For the time being, while there is not (yet!) any innovative leap in the AI and deep and machine learning fields, controllers must, at least, adopt some preventive measures, such as promoting digital literacy and making the algorithmic systems more easily understandable[106], as well as keep performing Data Protection Impact Assessments[107] regarding these systems, allowing to be aware of way the processing of data is being done, thus being possible to help prevent the black-box effect. It is not enough, not even by far, but it is the minimum that must be done to assess if the processing respects individuals' rights, or, if it is not the case, to give a warning sign for the need of performing corrections.

In the respect for fairness and accuracy, knowingly the data protection issues brought by decision-making algorithms, for the author there is not a singular right answer, but, instead, a series of plural efforts, both from the technological and legal fields, assuring data protection rights are respected.

[106] This was already recommended before the GDPR came into force, in 2017, by the CNIL (French DPA). Information available here: https://www.cnil.fr/en/how-can-humans-keep--upper-hand-report-ethical-matters-raised-algorithms-and-artificial-intelligence. Consulted in 11 February 2020.

[107] As mentioned above, according to Recital 90 and Article 35 of the GDPR, DPIAs are useful tools to use in any processing activities, although they are mandatory when the processing is likely to result in a high risk to data subjects' fundamental rights. An algorithmic decision-making process is very likely to result in a high risk regarding a subject's rights and freedoms.