

Transparency vs Explanation of Machine Learning Algorithms: Perspectives from Recent Legal Proceedings

Uchenna Nnawuchi¹ Dr Carlisle George² Dr Florian KammueLLer³

¹ PhD Student, ALERT Research Group, Department of Computer Science, Faculty of Science and Technology, Middlesex University, London UK
UN055@live.mdx.ac.uk

² Associate Professor, ALERT Research Group, Department of Computer Science, Faculty of Science and Technology, Middlesex University, London UK
c.george@mdx.ac.uk

³ Associate Professor, ALERT Research Group, Department of Computer Science, Faculty of Science and Technology, Middlesex University, London UK
f.kammueLLer@mdx.ac.uk

Abstract The extensive use of Machine Learning (ML) algorithms has significantly altered decision-making processes, societal structures, and power dynamics. These algorithms, known for their efficiency, are increasingly making critical decisions. However, concerns about their opacity and biases persist. Some legal scholars and data scientists argue that the GDPR includes a right to explanation to address ML opacity, while others advocate for transparency as a better solution. This study explores the debate between transparency and explanation within the GDPR's regulatory framework, focusing on Articles 15, 22 and recent EU court rulings. By examining judicial developments and scholarly perspectives, the study highlights the GDPR's efforts to promote transparency and accountability in automated decision-making. Despite the GDPR's provisions, the lack of an explicit right to explanation has led to ongoing legal and ethical discussions. This research aims to provide insights into how transparency, explanation, and individual rights intersect in the realm of algorithmic decision-making under the GDPR.

Keywords: Machine Learning, Right to explanation, Transparency, Automated decision-making, GDPR.

1 Introduction

In recent years, the pervasive integration and proliferation of Machine Learning (ML) algorithms into various spheres of human society have transformed decision-making processes, societal structures, and power dynamics. These algorithms, renowned for their promise of greater efficiency and accuracy, play a crucial role ranging from personalised recommendations on online platforms to complex decision-making processes in finance, healthcare, and criminal justice sectors [1]. However, alongside the promises

of enhanced efficiency, ML algorithms also raise profound challenges about opacity, fairness [2], transparency, accountability, and individual rights [3].

Central to the discourse surrounding algorithmic decision-making is whether individuals possess an explicit legal right to explanation when these challenges associated with ML algorithms significantly impact their rights and interests. While some legal scholars and data scientists have claimed that the right to explanation is a valid concept domiciled within the General Data Protection Regulation (GDPR) [4] and a compelling strategy for addressing and rectifying ML opacity [5,6,7], others contend that what is obtainable in the GDPR is transparency rights and not an explicit right to explanation right [8].

This study attempts to delve into the nuanced arguments surrounding the existence of a legal right to explanation within the GDPR framework. By exploring several perspectives from legal scholars on both sides of the divide and recent court decisions, it seeks to elucidate the evolving landscape of algorithmic decision-making and its implications for individual rights. This study examines the provisions of the GDPR, particularly Art 15 and 22 [4] which pertain to transparency and automated decision-making.

The first section of this study briefly evaluates scholarly opinions regarding the explicit recognition of a right to explanation within the GDPR. Through a thorough analysis of relevant GDPR articles and recitals, this section examines arguments both supporting and challenging the existence of such a right in the GDPR and finds that the GDPR does not provide an explicit legal right to explanation.

Subsequently, the paper analyses recent court decisions within the EU that address issues related to automated decision-making and the provision of meaningful explanations to individuals in the GDPR and the Passenger Name Record (PNR) Directive [9]. Through a close examination of these rulings, it seeks to elucidate the court's stance on the right to explanation and its implications for data subjects, or decisional subjects in the context of algorithm decision-making.

In addition, this study explores the rationale for advocating for an explicit legal right to explanation in the context of ML algorithms and automated decision-making systems. As algorithmic decision-making continues to proliferate, it becomes imperative to address the challenges posed by opacity and lack of accountability. Hence, this study proposes the development of a comprehensive and operative framework, specifically, a right to explanation and explainable AI, which is crucial for ensuring the ethical and responsible use of ML algorithms while upholding fundamental principles of fairness, non-discrimination, and human dignity in the digital age.

Ultimately, this study aims to contribute to the ongoing discourse surrounding the intersection of technology, law, and ethics, offering insights into the complex challenges and opportunities presented by algorithmic decision-making in the digital age.

1.1 The GDPR and a Right to explanation

The assertion of a right to explanation purportedly enshrined in the GDPR has ignited a persistent controversy regarding its existence within the regulatory framework [10]. Debates surrounding this issue pivot on whether it establishes a novel right to explanation, engendering a spectrum of contrasting viewpoints and opinions. The expansive scope of the GDPR has precipitated an escalating discourse on Art 15, 22, and Recitals 71 [4] concerning the nature, requisites, functionalities, and legal status of this purported right [11]. An analysis of the arguments from various scholars on both sides of the divide is discussed below to illuminate the intricate complexities inherent in the multifaceted discourse among legal scholars before delving into the analysis of the positions adopted by the courts undertaken to navigate through the intricacies of this conundrum, thereby shedding light on the nuanced landscape surrounding the purported right to explanation within the GDPR.

1.2 Assessing Arguments on the Existence of a Legal Right to Explanation

Mazur[11] references GDPR Art 13-15 and 22, initially claiming these provisions ensure a right to explanation for privacy and data protection. However, she later contradicts herself, stating these articles provide grounds for debate on the right's existence. Instead, she advocates for a collective right to access information, positing that such information must be held by the authority, be of public interest, and serve to inform the public. This approach is restrictive and does not conclusively establish the right to explanation in the GDPR, highlighting Mazur's uncertainty about its existence.

Goodman and Flaxman [5] argue that a right to explanation can be inferred from GDPR Art13-15 and 22, suggesting that the GDPR reaffirms the Data Protection Directive 1995's stance on explanation which does not recognise this right. They assert that an adequate explanation would make a trained model understandable to humans, relating input features to predictions. However, their argument lacks detailed reference to GDPR language, weakening their conclusion that a right to explanation exists in the GDPR.

The pair of Kim and Routledge [6] made frantic efforts not to explicitly state whether a right to explanation exists in the GDPR. They focus on developing a moral foundation for such a right, based on trust and informed consent, likening it to a promisee/promisor or trustee/trustor relationship. They argue that informed consent requires ex-post explanation to foster trust among data subjects. However, their approach is limited to algorithmic decisions involving personal data and does not thoroughly explore the GDPR's legal framework.

Brkan [7], in his paper, argued boisterously that a right to explanation exists in the GDPR, particularly grounded in Art 22(1), which prohibits decisions based solely on automated processing with significant effects. Although Brkan acknowledges that the right to explanation is not explicitly mentioned in the Articles of the GDPR but only in

Recital 71. He interprets this omission as a political compromise, leaving the final interpretation to the Court of Justice. Additionally, he suggests that the right to explanation is implied through the right to contest decisions under Art 22(3), as contesting without understanding would be ineffective. *Au contraire*, a right to contest can exist and function independently without implying the existence of a right to explanation. Conversely, the right to explanation would be ineffective without an accompanying right to contest. If a right to explanation were provided without the right to contest, the latter could be implied to ensure practical utility; otherwise, the right to explanation would be meaningless.

However, some other authors have argued that a right to explanation does not exist in the GDPR either expressly or impliedly.

Wachter et al. [8] argued that the GDPR does not implement a legal right to explanation but rather a right to be informed. They contend that the alleged right to explanation is only alluded to in Recital 71 and that its omission from the main text indicates a deliberate decision by the framers. They describe Art 13-15 as providing transparency rights, or what they termed the ‘Magna Carta’ of data subjects’ rights to obtain information about the data held about them. They recommended including a right to explanation explicitly in Art 22(3) and clarifying the GDPR’s language. While the recommendations of Wachter et al., are commendable, their work did not define the boundaries, scope, and nature of what is expected of this right.

Edward and Vaele [10], aligning with Wachter et al., assert that the GDPR does not create a legal right to explanation. They argue that Art 13-15 and 22 do not explicitly provide for such a right, and Recital 71 is non-binding. Edwards emphasises the GDPR’s limitation on personal data, pointing out that significant algorithmic decisions might involve non-personal data (such as traffic or road condition-related data), which the GDPR does not address. They underscore the lack of precise language and clear definitions within the GDPR, making it difficult to construe any existence of a legal right to explanation [10].

The debate on the right to explanation in the GDPR reflects a complex interplay of interpretations and legal reasoning. While scholars like Mazur, Goodman and Flaxman, Kim and Routledge, and Brkan offer various insights supporting its existence, their arguments vary in clarity and conviction. Conversely, Wachter et al. and Edwards provide strong counterarguments, emphasising the GDPR’s lack of explicit provisions and the non-binding nature of Recital 71. However, these scholars are *ad idem* notwithstanding their diverse opinions that there is a need for such a thing as a legal right to explanation – even Courts are *ad idem* on the need to have a legal right to explanation to minimise algorithmic opacity issues [12].

2. Legal Analysis of the recent Decisions of Courts on the right to explanation and automated decision-making

While there are several cases of automated decisions, only very few have considered Art 15 and 22 of the GDPR. Thus, this study now investigates some cases on automated decision-making.

The Amsterdam District Court recently addressed issues related to Article 15 of the GDPR in the cases of Ola Netherlands BV [13] and Uber BV [14]. These cases, although decisions of a court of first instance and subject to appeal, are significant. Additionally, the European Court of Justice (ECJ) addressed similar concerns of automated decision-making of AI systems in Breyer v Federal Republic of Germany (2019) [15], SCHUFA Holding AG [16], and Ligue des droits humains v. Conseil des ministres [17].

2.1 Analysis of the Court's Decision in Ola Netherlands BV

In the court's recent decision in Ola Netherlands BV, the applicants requested access to the existence of automated decision-making and profiling based on Article 15 paragraph 1 opening lines and under (h) GDPR. The applicants contested that its agreement with Ola contained a document titled *'How we process your data'* wherein Ola informed on which decisions she uses automated decision-making and profiling for and which personal data she uses for this processing [13]¹ However, the applicants claimed that the information contained in the document was incomplete because Ola did not provide information on other automated decision-making such as fraud probability score, earning profile, guardian software, assigning trips, and imposing fines and discounts etc. The court stated in its judgement that the terms "useful information about the underlying logic, as well as the significant and envisaged consequences" are not explained in the GDPR, but referred to the Guidelines of Working Party 29 (WP29, now European Data Protection Board) and stated [13]:²

'the Guidelines state that the point is that the data subject can understand the underlying idea of the decision or on the basis of which criteria that decision was taken. The concepts of 'interest' and 'expected consequences' indicate that information must be provided about the intended or future processing and about how the automated decision-making could affect the data subject. With due observance of the explanation in the Guidelines, the court understands by 'useful information about the underlying logic' that the most important assessment criteria and their role are communicated to the data subject in the automated decision, so that he can understand on the basis of which criteria that decision is has been taken and is able to verify the correctness and lawfulness of the data processing'. (Underlined for emphasis)

¹ paragraph 4.43 of the decision.

² paragraph 4.41 of the decision

Having reached the above interpretation, the court held that regarding the information on ‘imposing fines and discounts’ Ola must provide applicants with transparent and verifiable information about the data used and the assumptions underlying the automated decision-making process. This includes communicating the main assessment criteria and their role in the decision-making so that applicants can understand and verify the correctness and lawfulness of the data processing [13].³ Some scholars[18,19] have interpreted this decision as supporting the existence of a right to explanation within the GDPR. However, this study argues that the GDPR does not explicitly establish a legal right to explanation. Instead, it provides an ex-ante right to explanation referred to as the Magna Carter of information rights [8]. The court’s decision does not mandate the existence of an explicit legal right to explanation in the GDPR, notwithstanding the language, but one, founded on the need to grant data subject’s request to access personal data (processed), *id est* transparency rights, under Art 15 (1h) and to rectify or correct the personal data.

The case involved incomplete information contained in Ola’s document on ‘*How we Process your Data*’ concerning the data used in automated decision-making. The applicants activated Article 15(1)(h). The court in its wisdom held that notwithstanding the explanation contained in the WP29 Guideline on automated decision-making, the main criteria forming the basis of the decision should be provided to verify the correctness and lawfulness of the processing. With due observance of the decision of the court, and the opinions of a few scholars in Gellert et al, Bruggen and Nieuwenhuizen, [18] the import of the decision is not a recognition of an inferred right to explanation, much less so an explicit right to explanation. Such a conclusion from the court’s decision would be highly misconceiving; the incomplete document, the basis of requesting access to personal data used in the automated decision-making was made before the algorithmic decision took place. Therefore, “the useful information about the underlying logic, as well as the significant and envisaged consequence” as stated by the court of automated decision-making in imposing fines and sanctions when processing the applicant’s data is merely futuristic [8]. This is information about the likely consequences of personal data processed or how the personal data used will be processed which antecede the decision-making and not an explanation (Why) behind a specific decision, albeit the access request was made post the algorithmic decision.

Thus, the application of Art 15(1h) is based on how Ola processed their data. Hence, the court’s decision that data subjects be informed about the main assessment criteria for an automated decision does not translate to Ola recognising an explanation right not explicitly included in the GDPR. Instead, it aligns with Ola’s “*How We Process Your Data*” document, which details the decisions and personal data used for automated decision-making and profiling.

Deconstructing this decision, the court directed “...Ola provides the main assessment criteria...” ‘...their role in the automated decision, so applicants can understand the criteria on the basis on which the decisions were taken...’ Informing the data subject

³ paragraph 4.52 of the decision.

on the main criteria and its role does not amount to an explanation of how and why it was pivotal to the decision-making. Undeniably, the court's statement underscores the importance of transparency, however, it refrains from unequivocally guaranteeing individuals a right to a functional and explicit explanation of such a decision. Instead, this decision emphasises furnishing "useful information about the underlying logic" to enable data subjects to apprehend the decision-making criteria and most importantly assess the legitimacy of the data processing. In essence, individuals are informed of significant decision-making factors and are not necessarily entitled to a detailed breakdown of every aspect of how these factors affect the decision-making process. The focus of this court's decision is to accentuate the importance of data subjects having access to information to comprehend the reasoning behind a decision to verify its accuracy and compliance with relevant laws and legal standards and not to challenge or contest the decision that a right to explanation facilitates. Additionally, having access to the information - *id est* information rights - is not a comprehension of the inert workings of an algorithm in the decision-making process. By the foregoing, the court's stance demonstrates a nuanced understanding of transparency as contained in the GDPR, wherein the provision of meaningful information facilitates understanding and verification; more so the latter and not the former.

Thus, this decision does not support the existence of an explicit right to explanation in the GDPR but emphasises the need for transparency in automated decision-making processes, thereby creating room for interpretative elasticity within the tenets of GDPR.

Albeit arguable that transparency rights and explanation rights are similar or that explanation is a process that succeeds transparency, it is suggested that both rights are relatable, though mutually exclusive. Thus, an explanation in this instant case would inform the data subject on the main assessment criteria, and its role in reaching a particular outcome (justification of how and why the criteria influenced the outcome); however, this was not the decision of the court. An explanation must be functional [20], as it permits a positive action by the data subject e.g. contesting an algorithmic decision, not merely to verify the legitimacy of the data processed. Information on the factors considered in decision-making alone is by no means functional and explanatory. Such information merely grants data subjects access to the assessment criteria for verification purposes, which is a passive act; to reach a contrary conclusion would be an imagination stretched too far. However, such information supports transparency and accountability rights.

Additionally, this decision evidence transparency in informing the data subject of the process (How) of decision-making. The how here is 'procedural how' and not substantive; it describes the personnel, language, type, and role of data used for decision-making,⁴ as opposed to 'substantive how' which involves describing the role of the data used and how the data contributed to the decision. Thus, procedural how equals transparency, whereas substantive 'how' is part of the explanation process alongside the 'why'.

⁴ not necessarily a detailed explanation of how the role of the data influenced the decision.

Finally, on this point, the court held that the main assessment criteria and their role are communicated to the data subjects so that they "...are able to check the correctness and lawfulness of the data processing..." The intention of the court in its decision is to promote transparency rights in an algorithmic setting or at best ex-ante explanation. The court's reasoning behind granting the data subjects access rights is to enable the data subjects to verify the correctness and lawfulness of the processing. Admittedly, it is possible that explanatory rights, as well as transparency rights, provide the data subject with the ability to verify the correctness of personal data processed and whether the data was lawfully processed, however, explanation, unlike transparency provides a lot more. *Id est*, it enables the data subject to understand why and how (substantive) a decision was made and to challenge the decision. Transparency of an algorithmic decision does not *ipso facto* mean or include an explanation, but an explanation of an algorithmic decision includes transparency because it provides information on the processing of personal data, automated decision and profiling, and how and why a specific outcome was reached.

A holistic read of the provisions of Article 15(1h) and its application in the above decision would lead to the conclusion that the court's decision merely reinforces data subjects' transparency and information rights, as has now become elementary knowledge, a case is an authority for what it decides.

2.2 Analysis of the Court's decision in the Uber BV

In the Uber BV case, the 2nd and 4th applicants complained of the deactivation of their account without human intervention (in essence solely automated {Art 22 (1)}) or their consent, and was not in performance of the agreement {Art 22 (2)}. They also claimed that Uber violated the principle of transparency under Articles 14 (2f) and 15 (1h) for not informing them of the application of automated decision/or profiling. Also, and at least in cases where automated decisions are made, useful information about the underlying logic, as well as the importance and the expected consequences of that processing. The court held [14]⁵ that 'Uber should provide access to their data under Art 15 considering they were the basis for the decision to deactivate their accounts, in such a way that they can be able to verify the correctness and lawfulness of the processing of their data'.

Like Ola Netherland case [13], the court neither confirms that the applicants have an explicit legal right to explanation nor did the applicants request that the decision of the algorithm be explained. This is not antithetical to what has already been canvassed in the analysis of Ola Netherlands case [13] but merely supports it. The court stated that where personal data forms the basis of the decision to deactivate the account of the applicants, they should be availed of the information on their data to verify if the same

⁵ Paragraph 4.29 of the court's decision.

is accurate. Thus, in cases where non-personal data forms part of the basis of the decision, the applicants would have no right under Art 15. Additionally, the court's position is informative on the essence of granting the applicant's request which is to enable them to verify the correctness, accuracy, and lawfulness of the data being processed. Hence, the focus is not to challenge the decision before a court of competent jurisdiction which is the main purpose of the right to explanation but to be informed and to ensure transparency.

Thus, from the above decision, this case, like *Ola Netherlands* case [13], appears to focus its reasoning on transparency and not explanation rights. The court in its reasoning perceives a situation where the information on the personal data which is the basis of the decision is accompanied by an explanation of how the personal data influenced the algorithmic decision and why the decision was reached as opposed to merely informing the data subjects of what personal data was relied on to ascertain its accuracy. Ironically the latter is the court's holding.

It is clear, even from the court's decision that an explicit right to explanation availed to a data subject behind a specific decision is not the same as granting a data subject access to personal data in cases where they are the basis of the decision. Additionally, the court in these two cases above did not specifically allude to or reference where an explicit legal right to explanation is provided for in the GDPR.

2.3 Analysis of the ECJ's decision in *Breyer*

In *Breyer's* case [15], the issue was whether data subjects should be informed of the reason for the collection and storage of their dynamic IP addresses by government websites. *Breyer* argued that these addresses should be considered personal data under the Data Protection Directive 1995 [21]. Although the case was under the 1995 Directive, similar to GDPR Articles 12(b) and 15, the ECJ ruled that individuals have a right to explanation when automated decision-making is used in administrative processes, even if not explicitly provided in the Directive.

2.4 Analysis of *Schufa Holding*

In this case, the CJEU had to decide whether Art 22(1) GDPR means that calculating a probability value for a data subject's ability to repay a loan, using personal data, constitutes a decision based solely on automated processing. This includes profiling, which has legal effects or similarly significant impacts, especially when this value is used by a third-party controller to decide on a contractual relationship with the data subject?

The CJEU stated [16]⁶ that where the probability value determined by a credit information agency and subsequently communicated to a bank plays a decisive role in the credit granting process, the establishment of such a value must be regarded as a decision

⁶ Paragraph 50 of the CJEU's decision.

that produces 'legal effects' or similarly significantly affects the data subject within the meaning of Art 22(1). Additionally, on the scope of the right to information concerning automated decision-making, the court stated that the controller is required to provide information under its obligations in Art 13(2)(f), 14(2)(g), and 15(1)(h). Specifically, the controller must furnish meaningful information about the logic involved, the significance, and the envisaged consequences of such processing. The court emphasised that these enhanced requirements concerning the lawfulness of automated decision-making, the information obligations of the controller, and the related rights of data subjects are explained by the purpose of Art 22, which aims to protect individuals against risks to their rights and freedoms posed by the automated processing of personal data, including profiling.

The Advocaat General (AG) Pikamäe, in this case, discussing the scope of the right to information under Art 15(1h), specifically '*meaningful information about the logic involved*', stated that it includes ..."the calculation method used by a credit information agency to establish a score... [22]"⁷ In buttressing his point, he opined that ..."the obligation to provide 'meaningful information about the 'logic involved' must be understood to include sufficiently detailed explanations of the method used to calculate the score and the reasons for a certain result... in general, the controller should provide the data subject with general information, notably on factors taken into account for the decision-making process and on their respective weight on an aggregate level, which is also useful for him or her to challenge any 'decision'...[22]"⁸

Notably, the CJEU was silent on whether the phraseology '*meaningful information about the logic involved*', translates to a right to explanation notwithstanding it had read the opinion of the AG. The CJEU overall described Art 15 as an information obligation of the controller in line with lawful processing and protecting rights of the data subjects. This means the protection of the rights conspicuously and explicitly provided in the GDPR concerning lawful processing and automated decision-making. An attempt to draw in Recital 71, specifically on explanation, is futile: Recitals do not add or create a legal obligation but merely explain the explicit provisions in a legal text. Thus, the non-inclusion of this right represents the intention of the lawmakers to not have a right to explanation. However, some authors [23] have claimed that the AG's opinion of Art 15 (1h) demonstrates the widely applicable right to explanation in the GDPR. *Au contraire*, the decision of the CJEU is the extant interpretation of the law on Art15 (1h); the CJEU neither alluded to the AG's opinion on that interpretation nor did it venture to *suo motu* infuse Art 15(1h) with a right to explanation. Thus, the opinion of the AG is just that, an opinion and nothing more. Undoubtedly such opinions are influential, but the CJEU is not bound by an AG's opinion and need not even address them [24], as reflected in this case. The opinion of an AG can shed light on the direction of the CJEU's decision, but it cannot add or take its place, especially when none of the opinions are adopted or included in the judgement. The CJEU's silence on reading a right to explanation into Art 15(1h) is informative of its position. Thus, a right to explanation cannot be read into the provisions of the GDPR.

⁷ Paragraph 54 AG Pikamäe's opinion.

⁸ paragraph 58 AG Pikamäe's opinion.

2.5 Analysis of the Ligue des droits humains Case

The applicants sought to annul the Law of December 25, 2016, before the Belgian Constitutional Court. This law requires passenger transport carriers to transfer data to a Belgian database for security purposes. The Ligue des droits humains argued it breached privacy and data protection rights under the Belgian Constitution, GDPR, and the EU Charter of Fundamental Rights (CFR). The Belgian government defended the law as crucial for combating terrorism. The court referred the matter to the ECJ for guidance on GDPR applicability, fundamental rights compatibility, and the legality of intra-EU data processing. The ECJ ruled that the Belgian law must comply with the GDPR and CFR. Specifically, the ECJ held that the law's data collection, transfer, processing, and retention must be strictly necessary and proportionate, and the five-year retention period was deemed excessive. Additionally, the court held that individuals affected in the processing of personal data must be given information to understand how the pre-determined assessment criteria and programs work to enable them to decide whether to exercise their right to judicial redress under Article 13(1) of the PNR Directive [17].⁹ *Id est*, member states must ensure passengers have the same data protection rights, including access, rectification, erasure, restriction, compensation, and judicial redress [25],¹⁰ as per Union and national law, and Framework Decision 2008/977/JHA [25].

The decision of the ECJ, in this case, is based on the analogy in *Minister van Buitenlandse Zaken case* [26]. The court was confronted with whether a person should be informed of the reason for the refusal of a visa. The court held that the member state refusing the visa is required to indicate, in that decision, the identity of the Member State which raised that objection, the specific ground for refusal based on that objection, accompanied, where appropriate, by the essence of the reasons for that objection, and the authority which the visa applicant may contact to ascertain the remedies available in that other Member State [26].¹¹

It appears that the *Minister van Buitenlandse Zaken case* recognises that visa applicants are entitled to a right to explanation when their visa application is refused. However, it is important to note that the court in this case considered Art 47 CFR. Equally, the ECJ's point on informing individuals to understand how the assessment criteria works was based on the analogy adopted by the court in *Minister van Buitenlandse Zaken case which considered* Art 47 CFR.

Art 47 CFR provides for a right to an effective remedy and a fair trial. Fair trial which includes fair and public hearings requires a person affected by a decision to be availed all the necessary elements and materials to contest a decision or a wrong which includes notice of a decision, the grounds of that decision, the right to know the evidence on which a decision is based, a hearing before an unbiased tribunal, and a right to present arguments against a decision and a statement of reasons [27], in essence a right to explanation. Although Art 47 CFR may not provide nor establish an explicit right to an

⁹ Paragraph 210 of the ECJ ruling

¹⁰ Art 17, 18, 19, 20 Framework Decision 2008/977/JH

¹¹ Paragraph 56 the ECJ ruling.

explanation of automated decisions, it has been inferred to provide elements of a right to explanation [27]. It is based on Art 47 and its interpretation that the cases of *Ligue des droits humains* and *van Buitenlandse Zaken* are distinguishable from the other cases of Ola Netherlands, Uber BV, Schufa holding etc. These other cases did not consider Art 47 CFR, only the GDPR. On the sole authority of the GDPR, there is no explicit right to explanation contained in its dry letters. Additionally, the legal principle "*expressio unius est exclusio alterius*" (the express mention of one thing excludes all others) suggests that the deliberate exclusion of the right to explanation from the original text indicates the framers' intention not to include it in the law.

Recently, the EU Council and Parliament agreed on the AI Act [28], regulating AI systems in the EU market, with Art 86 establishing a right to explanation for high-risk AI systems [29].¹² However, Art 86 is criticised for being non-functional and lacking practical efficacy due to the absence of guidelines for its application and enforcement, mirroring the clarity issues of the GDPR. Additionally, the Council of Europe is also proposing a Convention on AI [30], however, Paragraph VII(34) of its first draft, which provides a rationale for a right to explanation, does not include this right. To address these shortcomings and AI opacity, in another paper we proposed a functional and operative right to explanation as a protocol to the CoE Convention on AI [30]. Notwithstanding, it is imperative to justify why a right to explanation has become indispensable in the ML context.

3. Justification for a Legal Right to Explanation

ML algorithms — colloquially engaging public consciousness simply as ‘algorithms’, or “AI”— have been criticised extensively due to concerns within their systems reflecting both historical and present real-world prejudices [10]. The use of complex models like Deep Neural Networks (DNNs) exacerbates these issues due to their opacity [31], making it difficult to understand their behaviour and the rationale behind their decisions or predictions, which often have disparate effects on data subjects [31].

This opacity in AI decision-making presents significant operational and reputational risks for organisations, while also frustrating data subjects or customers who struggle to trust systems they cannot understand or challenge. This inscrutability affects human rights by undermining dignity, liberty, fair treatment, and equality, and by entrenching discrimination, thus reinforcing power imbalances between data subjects and controllers or developers [31]. ML algorithms have been criticised for reinforcing human biases, inequalities, and other vulnerabilities [32], raising concerns about safety, security, and fairness. For instance, a 2016 ProPublica article highlighted that the COMPAS algorithm used by judges to predict recidivism was racially biased [33], impacting parole decisions and, consequently, the human rights of individuals. Similarly, algorithms determining kidney transplant recipients involve life-and-death decisions based on prioritised lists, which can be morally contentious [34].

¹² Art 1, EU AI Act

The opacity of ML algorithms makes it nearly impossible to understand how decisions are made, harming privacy protection, human rights, and democratic [35], social, and ethical values. These algorithms can process both personal and non-personal data, manipulate individuals, and automate decisions without public scrutiny [35]. While the GDPR focuses on privacy, it does not require firms to disclose how they use algorithms in ways that protect privacy or ensure their use of data is not prejudicial. This lack of disclosure is often due to trade secret protections for AI systems [35].

The capability of ML algorithms as decision-support and decision-making systems highlights the urgent need for explanation and explainability to safeguard data subjects' interests and rights [35]. Understanding the intricacy encircling the internal workings on how ML outputs are generated is essential, making the development of a right to explanation (R2X) and a right to explainable AI (R2XAI) paramount. These rights build trust, allow stakeholders to examine underlying models and data (look “under the hood”), clarify decision processes, and provide coherent explanations [35], thus protecting human rights such as dignity, non-discrimination, liberty, fair hearing, and the right to contest. Explanatory rights ensure data subjects can challenge ML decisions which is vital for accountability and transparency; hence, necessitating a reinforcement of fundamental rights.

4. Conclusion

The debate over the existence of a legal right to explanation within the GDPR has generated significant discussion among legal experts. While some argue that the GDPR provides transparency rights rather than an explicit right to explanation, recent court rulings highlight the importance of transparency in automated decision-making. However, an explicit autonomous right to explanation, although recently established in the AI Act, is lacking in clarity, and practical efficacy.

Court cases have primarily focused on data protection, privacy rights, and fair trial principles, rather than explicitly referencing GDPR Articles 15 or 22 for a right to explanation. The GDPR addresses personal data processing rights but lacks a clear right to explanation in the machine-learning context. Despite this, the opacity of algorithmic decision-making underscores the need for further consideration. Algorithms can reinforce biases and raise ethical concerns, undermining privacy protection. Thus, this study advocates for the development of an explicit right to explanation within data protection and human rights laws to empower individuals, foster trust, and ensure accountability in automated systems. In essence, while the GDPR does not specifically acknowledge a right to explanation, evolving AI technologies demand proactive measures to safeguard fundamental rights and ensure ethical use of decision-making systems. Developing a Right to Explanation (R2X) and right to Explainable AI (R2XAI) is essential to address these challenges and uphold fairness, non-discrimination, and human dignity in the digital age.

References

1. Dormehl, L.: Thinking machines: the quest for Artificial Intelligence—and where it's taking us next (New York: Penguin–TarcherPerigee, 2017)
2. Chin-Rothmann, C.: Fairness in Algorithmic Decision-Making (Brookings, 9 March 2022) <<https://www.brookings.edu/research/fairness-in-algorithmic-decision-making/>> accessed 23 April 2024
3. Rodrigues, R.: Legal and Human Rights Issues of AI: Gaps, Challenges and Vulnerabilities (Journal of Responsible Technology, 16 October 2020) <<https://www.sciencedirect.com/science/article/pii/S2666659620300056>> last accessed 23 April 2024
4. The European General Data Protection Regulation 2016
5. Goodman, B., Flaxman, S.: Eu regulations on algorithmic decision-making and a right to explanation https://www.researchgate.net/publication/304548505_EU_regulations_on_algorithmic_decision-making_and_a_right_to_explanation last accessed 2 October 2021
6. Wan Kim, R., Routledge, B.: Why a Right to Algorithm decision-making should exist; A trust-based approach, Cambridge University Press <https://www.cambridge.org/core/journals/business-ethics-quarterly/article/why-a-right-to-an-explanation-of-algorithmic-decisionmaking-should-exist-a-trustbased-approach/C620A6A5FCB781384D20E08BE4CD09BC> last accessed 18 October 2021
7. Maja Brkan, Do algorithm rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond, https://techpolicyinstitute.org/wp-content/uploads/2018/02/Brkan_do-algorithms-rule.pdf#page=1&zoom=auto,-34,842 last accessed 20 march 2022
8. Wachter, S., Mittelstadt, B., Floridi, L.: Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. 7 International Data Privacy Law 76 (2017)
9. Directive (EU) 2016/681 of the European Parliament and of the Council of 27 April 2016 on the use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime, <https://eur-lex.europa.eu/eli/dir/2016/681/oj>, last accessed 3rd June 2024
10. Edwards, L., Veale, M.: Slave to the Algorithm? Why a Right to Explanation Is Probably Not the Remedy You Are Looking For. SSRN Electronic Journal [2017]
11. Mazur, J.: Right to access information as a collective based approach to GDPR's right to explanation in European Law. SSRN Electronic Journal [2021]
12. NJCM ET AL V. THE Dutch State (2020) the hague distric court ecli : nl: RBDHA:2020;1897; see also the case of Hous. Fed'n of Teachers, Local 2415 v. Hous. Indep. Sch. Dist., 251 F. Supp. 3d 1168 (S.D. Tex. 2017)
13. Ola Netherlands BV, Case number/application number: C / 13/689705 / HA RK 20-258.
14. Uber BV, Case number/application number: C / 13/692003 / HA RK 20-302.
15. Breyer v. Federal Republic of Germany, CJEU - C-582/14 -
16. SCHUFA Holding AG, Case C-634/21
17. Ligue des droits humains v. Conseil des ministres, Case C-817/19
18. Gellert, R., van Bekkum, M., Frederik Zuiderveen Borgesius.: 'The Ola & Uber Judgments: For the First Time a Court Recognises a GDPR Right to an Explanation for Algorithmic Decision-Making' (The Ola & Uber judgments: for the first time a court recognises a GDPR right to an explanation foralgorithmic decision-making, 1 January 1970) <<http://eulawanalysis.blogspot.com/2021/04/the-ola-uber-judgments-for-first-time.html>> last accessed 30 August 2023

19. van der Bruggen, L., Nieuwenhuizen, A.: Amsterdam Court of Appeal Rules in Favour of Uber and Ola Cabs' Drivers (Fieldfisher, 19 April 2023) <<https://www.fieldfisher.com/en/insights/amsterdam-court-of-appeal-rules-in-favour-of-uber-and-ola-cabs-drivers>> last accessed 30 August 2023
20. Hempel, C.: Aspects of Scientific Explanation, Free Press, New York, 1965, quote from p. 423.
21. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31995L0046> last accessed 30 August 2023
22. AG Pikamäe's opinion, <https://curia.europa.eu/juris/document/document.jsf?text=&docid=271343&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=446998> last accessed 20th August 2024
23. Kelder, K.: On the Relative Importance of the AI Act Right to Explanation — the Digital Constitutionalist' (digi-con.org 24 April 2024) <<https://digi-con.org/on-the-relative-importance-of-the-ai-act-right-to-explanation/>> last accessed 2 July 2024
24. Ritter, C.: A New Look at the Role and Impact of Advocates-General – Collectively and Individually". Columbia Journal of European Law. 12 (3). (2006-03-29).
25. Council Framework Decision 2008/977/JHA of 27 November 2008 on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters 'EUR-Lex - 32008F0977 - EN - EUR-Lex' (Europa.eu 2018) <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32008F0977>>.
26. Minister van Buitenlandse Zaken case C-225/19
27. Friendly, HJ.: Some Kind of Hearing, 123 U.Pa.L.Rev. 1267 (1975)
28. Cuypers, A.: The Right to Explanation in the AI Act: A Right to Interpretable Models?' (CiTiP blog 12 March 2024) <<https://www.law.kuleuven.be/citip/blog/the-right-to-explanation-in-the-ai-act-a-right-to-interpretable-models/>> accessed 1 July 2024
29. European Union Artificial Intelligence Act 2021
30. Nnawuchi, U.: 'Proposing a Protocol to the CoE Convention on Artificial Intelligence, Establishing and Implementing a Legal Right to Explanation of Machine Learning' [2024] Social Science Research Network <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4747094> accessed 13 July 2024.
31. Rauber, A., Trasarti, R., Giannotti, F.: Transparency in Algorithm Decision Making', Transparency in Algorithmic Decision Making - Introduction to the Special Theme (er-cim.eu) last accessed 8 October 2021
32. Risks from artificial intelligence', <https://www.cser.ac.uk/research/risks-from-artificial-intelligence/> last accessed October 2021
33. Angwin, J.: Machine Bias (ProPublica, 23 May 2016) <<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>> accessed 30 April 2024
34. Robinson, D.: Voices on the Code, ' Russell Sage Foundation, (New York, 2022)
35. Lu, S.: Why the GDPR may fail to protect individuals from privacy risks produced by artificial intelligence applications and a new transparency based approach to AI governance may help, <https://sites.duke.edu/thefinregblog/2022/03/01/why-the-gdpr-may-fail-to-protect-individuals-from-privacy-risks-produced-by-artificial-intelligence-applications-and-a-new-transparency-based-approach-to-ai-governance-may-help/> last accessed October 2021