Cambridge
Forum

RESEARCH ARTICLE

Rethinking the judicial duty to state reasons in the age of automation?

Victoria Hendrickx

KU Leuven, Centre for IT & IP – imec, Leuven, Belgium
Email: victoria.hendrickx@kuleuven.be

Abstract

As judges increasingly rely on algorithmic and AI systems in their decision-making process, concerns have been raised about their impact on the judicial duty to state reasons – a duty that is essential to the right to a fair trial and the rule of law. This paper examines whether existing legal frameworks adequately ensure this duty and its underlying normative goals in the age of automation and explores how best to safeguard this duty. Central to the analysis is the question of whether the duty to state reasons itself should be redefined and impose stricter reasoning requirements on judges using AI, requiring them to articulate both 'pragmatic' and technical explanations of the systems' role in their decisions. While this approach may enhance the duty's normative goals of legitimacy, transparency and accountability, the paper also considers the potential drawbacks of imposing such a more robust reasoning obligation in AI-assisted judicial processes.

1. AI's expanding yet unconstrained role in the judiciary

Although judges historically adopted a cautious approach to accepting new technologies, they have increasingly embraced the digitalisation shift and have started integrating algorithmic and AI systems into their courts. Initially, experiments in the legal domain centred around rule-based and knowledge-based systems, known as legal information retrieval systems, and other simple applications grounded in an 'if-then'-logic. These early systems were designed to assist judicial actors by automating simple, repetitive tasks, mainly for administrative purposes, such as case management, electronic filing, precedent and evidence analysis, case scheduling and assignment, or admissibility assessments. However, technological advancements have since propelled the development and use of increasingly sophisticated AI and machine learning applications, which learn from large amounts of data and experience (Cohen, 2021; Liebowitz, 1986; Susskind, 1986). Algorithmic and AI systems can nowadays provide substantive support, assisting judges in calculating average sentences for crimes, assessing the likelihood of someone reoffending, or predicting case outcomes (Corvalan, 2020; Fabri, 2024; Medvedeva, 2023; Reiling, 2020; Smuha & Hendrickx, 2023). Since the launch of ChatGPT by OpenAI in 2022, generative AI models have further broadened AI's role in the judiciary, as they can offer high-quality human-like text that can assist judges with drafting their judgements, summarising documents or providing legal advice across different legal domains (Labour Circuit of Cartagena 2023; Rechtbank Gelderland, 2024; Farah, 2023; Gutiérrez, 2024b; Smith, 2024; Taylor,

2023), going from tax and family law to criminal law (Corvalan, 2020; Dhungel, 2024; Koukoulioti, 2024; Saied-Tessier, 2024). This progression shows how algorithmic and AI systems are assuming an increasingly pivotal role in shaping judgements, and the judicial decision-making process in more general – thereby not only impacting individuals but society alike.

The continuing drive to integrate these technologies into the judicial decision-making process can largely be explained by the ambition to enhance the efficiency of justice systems. As courts are faced with an overwhelming influx of cases and persistent backlog, the turn to technologies seems to hold the promise of expediting case resolution, reducing costs, and making procedures more transparent and accountable (Carneiro Rocha, 2021; CEPEJ, 2024b; Schindler, 2024). However, over time, it has become clear that this efficiency discourse and techno-solutionism paradigm are not a silver bullet (Hedler, 2022; Paul, 2022; Skaug Saetra & Selinger, 2023). The reliance on algorithmic and AI systems in the judiciary comes with many risks and challenges, particularly concerning the right to a fair trial and the rule of law (Dessers & Valcke, 2020; Dhungel, 2024; Dymitruk, 2019; Smuha, 2024). Issues have surfaced regarding the negative impact on judicial independence (Gentile, 2022; Schmitz-Berndt, 2024) or the authenticity and admissibility of AI-generated evidence (Grossman, 2023; Seng, 2021). Despite the growing concerns, only recently attention has turned towards understanding how algorithmic and AI systems can affect the judicial duty to state reasons, which refers to the obligation of judges to provide reasons whenever they rule in a case and a crucial component of the right to a fair trial and rule of law (Barry, 2024; Dymitruk, 2019; Hendrickx, 2024, 2025).

Considering the impact of algorithmic and AI systems on the judicial duty to state reasons and its underlying normative goals (Hendrickx, 2024, 2025), this paper examines how best to safeguard this duty in the age of automation, in particular, by assessing whether existing legal frameworks are adequate to safeguard this duty and, if not, whether we should rethink the duty's core and scope. Given the duty's fundamental role in ensuring the right to a fair trial, upholding the rule of law and promoting different important functions, it is essential to determine whether and how this duty should be adapted or strengthened to remain effective when reliance on AI in the judiciary continues to increase. This paper is divided into six sections. The current section provided an overview of the growing presence of algorithmic and AI systems within the judiciary (Section 1).

The second section reviews how the reliance on algorithmic and AI systems can affect the judicial duty to state reasons, drawing on previous research on this exact topic. To this end, I provide a concise conceptualisation of the judicial duty to state reasons and identify its underlying normative goals being legitimacy, transparency and accountability of judicial decision-making. I then examine how different AI systems can impact these normative goals (Section 2).

In the third section, I evaluate whether existing legal frameworks, in particular the right to explanation under the General Data Protection Regulation (GDPR) and the AI Act, provide safeguards to the judicial duty to state reasons by requiring explanations from judges when they rely on AI systems. While certain provisions might safeguard the duty to some extent, it seems that the current legal frameworks are inadequate to uphold the normative goals of the duty in the age of automation (Section 3).

Building on this assessment, I explore whether a reconsideration of the judicial duty to state reasons is necessary in the sense of imposing stricter reasoning requirements on judges who rely on AI. Specifically, I explore whether the duty should be expanded to require judges to provide both 'pragmatic' and technical explanations. Pragmatic explanations would require judges to disclose whether and how they used an AI system, in which part of the process, the extent to which they integrated AI-generated outputs, the role the system played in forming the final judgement and so on. In addition, technical explanations would offer insights into the logic and inner workings of the AI system, which connects with the broader debate on explainable AI (XAI) (Section 4).

The fifth section critically examines the potential drawbacks of imposing such heightened reasoning obligation in AI-assisted judicial processes (Section 5). The last section concludes the paper by summarising the key findings and implications of the research (Section 6).

2. The impact of algorithmic and AI systems on the judicial duty to state reasons

2.1 Introduction

Before addressing the main research question of this paper – whether we should rethink the judicial duty to state reasons in order to best safeguard it in the age of automation – it is first important to show how the reliance on algorithmic and AI systems in the judiciary can impact this duty, since that assessment not only contextualises but also justifies the current research. This section therefore addresses the question how and to what extent the duty can be impacted whenever judges rely on these technologies. I will first provide an overview of the judicial duty to state reasons, followed by an analysis of how it can be impacted. The conceptualisation of the duty is restricted to the European level. This is because at European level, minimum standards apply across all Member States. While Member States may impose additional requirements, it shows that if the reliance on algorithmic and AI systems in the judiciary already affects the duty at European level, it will even more so have an impact at national level.

Rather than looking at how the duty itself can be impacted, I believe at this stage in time, it is more insightful to look at how the underlying normative goals of the duty can be impacted. I argue that these goals – legitimacy, transparency and accountability of judicial decision-making – are fundamental and universal objectives that are intrinsic to the duty and should remain safeguarded, regardless of how the duty itself is framed or operationalised across different jurisdictions.

2.2 The judicial duty to state reasons

The judicial duty to state reasons refers to the obligation of judges to provide reasons or motivations for the decisions they take whenever they rule in a case. The duty obliges judges to describe factual and legal circumstances of the case, the manner of interpretation of the arguments of parties during the trial and legal reasoning (Dymitruk, 2019). At the European level, the duty is grounded in Article 6 of the European Convention on Human Rights and Article 47 of the European Charter of Fundamental Rights. Although not explicitly mentioned in the text of these articles, both the European Court of Human Rights (ECtHR) and the European Court of Justice (ECJ) have recognised that it is an essential part of the right to a fair trial (ECtHR 21 January 1999; ECJ 15 October 1987). The duty is often also referred to as the right to a reasoned judgement, meaning that individuals have a right to obtain reasons for decision that affect them (ECtHR, 2024).

The judicial duty to state reasons constitutes an essential part of the rule of law and the right to a fair trial. The latter aims to ensure fair litigations through a range of procedural safeguards, including this duty that requires judges to provide justifications for their decisions. Judges' explanations enhance the fairness of judicial proceedings and ensure that decisions are perceived as legitimate and just (Dymitruk, 2019). The ECtHR and numerous scholars have consistently affirmed that the right to a reasoned judgment is also part of the rule of law in liberal democracies. On the one hand, reason-giving is an important component of the procedural conception of the rule of law, serving as a safeguard against arbitrary decisions, irrationality and unreasonability. On the other hand, it serves the substantial conception of the rule of law as providing rules and reasons leads to better public decisions and enhanced quality (ECtHR 6 March 2006; ECJ 6 September 2012; Cohen, 2009; Neil, 1998; Simmons, 2018).

Although no uniform definition of the judicial duty to state reasons exists, both the ECtHR and ECJ have provided insights into its scope and characteristics through their case law. First of all, it is important to distinguish between a formal and substantive duty (Hendrickx, 2024, 2025). A formal

duty requires only that a judgement is reasoned without assessing the correctness or accuracy of the reasoning. Such formal duty exists at European level (ECtHR 18 May 2010). A substantive duty calls for more detailed and robust reasoning for judges' decisions, as seen in jurisdictions like Brazil and Mexico.[1]

Whether it be a formal or substantive duty, as a minimum standard, the duty obliges judges to address the essential arguments raised by the parties that are determinant to the case's outcome (ECtHR 9 December 1994(a)). Judges thus not necessarily have to answer to all arguments raised by the parties, but only to the ones that can affect the resolution of the dispute. The extent of how elaborate the reasoning should be depends on the nature of the decision and circumstances of the case. Factors to take into account include '*the diversity of the submissions that a litigant may bring before the courts and the differences existing in the Contracting States with regard to statutory provisions, customary rules, legal opinion and the presentation and drafting of judgment*' (ECtHR 9 December 1994(b)). If a context is clear, some arguments may not be addressed or even indirectly rejected (ECtHR 9 December 1994(a); ECtHR 9 December 1994(b)). Another example is that judges may fulfil the duty by simply endorsing the reasoning of a lower court's decision in certain situations, whereas in other circumstances not. Sufficient procedural safeguards and parties' ability to understand the decision may counterbalance the lack of reasons (ECtHR 19 February 1998). For example, if a convicted person was present during all hearings and heard all the essential arguments, it can be assumed that the person can be reasonably aware of the reasons for the decision even without explicit reasoning in the judgement (ECtHR 30 September 2022; ECtHR 7 September 2023). A violation occurs when adequate reasons lack, for instance, if the judgment does not mention anything about evidence or statements that are crucial to acquit of convict someone – which must be assessed on a case-by-case analysis (ECtHR 8 April 2008). In 2008, the Consultive Council of European Judges (CCJE) adopted Opinion No 11 on the quality of judicial decisions (CCJE, 2008), in which they proposed several recommendations to enhance the quality of decisions, including through the judicial duty to state reasons. The CCJE stated that decisions should be clear, intelligible, drafted in a simple language, accessible to anyone, and that reasons should be consistent, clear, unambiguous, free from contradictions or insulting or unflattering remarks about the parties. While the recommendations set a high standard, they remain non-binding guidelines. In practice, neither the ECtHR nor the ECJ has embraced these standards, but instead adhere to a more minimal approach to the duty to state reasons.

Beyond these characteristics, case law and literature indicate that the judicial duty to state reasons fulfils several important functions in liberal democracies. By stating reasons, judges not only conclude and resolve the dispute, they also demonstrate that they have heard the parties' arguments and provide them with insights as to why a decision was made, enabling parties to make informed choices about appeals (ECJ 15 October 1987; ECtHR 12 February 2004). Reason-giving serves as a safeguard against arbitrary powers, since it curbs the improper exercise of judicial discretion and thus enhancing accountability vis-à-vis judges. Stating reasons fosters transparency within the judicial decision-making process and the justice systems as a whole (ECtHR 30 September 2022). Research has shown that when judges' reasoning is accessible, it increases acceptance of judicial decisions, contributes to legal certainty, and renders the whole procedure more fair (ECtHR 14 February 2007; Opdebeek & De Somer, 2016). The duty empowers individuals and constitutes a prerequisite to their ability to exercise their rights, such as the right to be heard, right to a reasoned judgement and right to a fair trial in general (ECtHR 27 September 2001).

In a subsequent step, I argue that the judicial duty to state reasons pursues important normative goals, referring to goals or requirements that must be met whenever judges rule on cases –

---

[1]Art. 489 §1 of the Brazilian code of civil procedure. https://www.lawyerinbrazil.com/wp-content/uploads/2019/06/BRAZILIAN_CODE_OF_CIVIL_PROCEDURE-1.pdf; Art. 402 of the Mexican National Code of Criminal Procedures, https://www.wipo.int/wipolex/en/legislation/details/17432.

including when they rely on algorithmic and AI systems to do so. These goals not only grasp the essence of what the duty aims to achieve but also what is should pursue. I develop this normative framework for different reasons. As mentioned, the conceptualisation of the duty remains relatively underexplored in legal scholarship. By examining its underlying normative goals, this research contributes to a more comprehensive theoretical foundation. From a practical perspective, an assessment of AI's impact on the duty must be preceded by a conceptual clarification of the duty's normative goals. In addition, focussing on the normative goals of the duty – rather than the duty itself – allows for a more universally applicable framework. The normative goals are fundamental objectives that should be safeguarded regardless of how the duty is framed or operationalised in different jurisdictions. The theoretical framework allows to examine how AI systems affect the normative goals of the duty without being constrained by how particular jurisdictions interpret the duty itself.

Drawing on case law and literature, I detect three normative goals pursued by the duty: legitimacy, transparency and accountability of judicial decision-making. While each of the goals is a fundamental objective of the judicial process in its own right, I discuss them in connection to the judicial duty to state reasons. Nevertheless, a comprehensive analysis of these goals is beyond the extent of this paper. For the purpose of this discussion, I will briefly introduce and conceptualise them to establish the necessary theoretical foundation for the subsequent impact assessment. Please note that the normative goals can show some overlap.

The first underlying normative goal of the duty to state reasons is legitimacy of judicial decision-making. Public justifications are considered as a prerequisite for legitimacy, since reasons reinforce public trust and allow the general public to perceive the judiciary as worthy of their institutional role as well as being appropriate, proper and just (Cohen, 2010; Forst, 2013; Hendrickx, 2024). While judges are often not elected, they nonetheless exert authority over individuals, society and other governmental branches. Providing reasons legitimises their role vis-à-vis other branches of government and shows that their decisions are grounded in law rather than personal preferences, thereby fostering normative legitimacy and public trust (Katz & Zamir, 2024; Merill, 1993). In addition, providing reasons allows judges' reasoning to be subject to scrutiny (Cohen, 2011), and thus operate as a standard for identifying legitimate exercises of judicial power. Legitimacy, in turn, ensures acceptance and tolerance of outcomes, and is crucial to make parties and society respect and execute judicial decisions. It can also foster social trust in courts and public confidence in the judicial system, and facilitate acceptance of (controversial) court decisions (Chronowski, 2021; Mentovich, 2023; Tyler, 2006; Ulenaers, 2020).

Second, reason-giving enhances transparency of judicial decision-making. Transparency can be understood in different ways depending on its context, and is often linked to concepts such as explainability, fairness, interpretability and human oversight. In this paper, transparency is specifically connected to judges' reasoning. Most straightforwardly and from an epistemological perspective, reasons illuminate not only the final outcome of the decision but also the process that led to it, thereby making the judicial process more understandable to both parties and the general public (Bentham, 1790; Postema, 2014). This is particularly important given that judicial decisions can have implications beyond individuals and affect the broader society. Being transparent about the reasons for a decision can act as a tool against opacity, facilitate explainability and procedural justice and fairness, and allow for judicial review, evaluation, audit and vetting (Beckman et al., 2024; Shapiro, 1992; Simmons, 2018). It also contributes to legal certainty (Hazelhorst, 2017) and facilitates the right to appeal for litigants (Dreyer, 2021).

Lastly, this transparency goal, in turn, supports the third normative goal of accountability vis-à-vis judges. Stating reasons puts a hold against the arbitrary exercise of judges' discretionary powers. By making judgments public, it encourages judges to act more fair, consistent and impartial. Public scrutiny serves as an incentive for judges to engage in more rigorous reasoning and deliberative thinking, and in turn less intuitive decision-making. This contributes to better quality reasoning, and

demonstrates respect for litigants (Cohen, 2011; Katz & Zamir, 2024). Accountability also strengthens public confidence in the judiciary and reinforces acceptance of both the process and outcome of the decision-making process (France, 2019). Ultimately, it maximises responsibility (Bentham, 1790; Cohen, 2010, 2015; Postema, 2014; Richardson, 2003; Staszewski, 2009).

2.3 Impact of algorithmic and AI systems on the judicial duty to state reasons and its underlying normative goals

Given the important functions the judicial duty to state reasons fulfils and its underlying normative goals, the continued reliance on algorithmic and AI systems within the judiciary has raised concerns about their impact on this duty and its goals (Albright, 2023; Araujo, 2020; Beckman et al., 2024; Chronowski, 2021; Barry, 2024; Hendrickx, 2024). This section provides an overview of key findings on how these technologies may undermine the duty to state reasons and its underlying normative goals.[2] Importantly, the extent of the impact largely depends on the concrete ways judges engage with the AI systems. When judges simply replicate the system's output – such as risk scores, average sentences or predictions – without critical analysis or independent judgement, the adverse effects on the duty and its goals will be more pronounced. In contrast, when judges thoughtfully reflect on the system's output and write the judgement themselves, the adverse effects will be less significant. Moreover, the nature of the duty itself influences how AI affects judicial reasoning: a substantive duty may be more profoundly impacted by the reliance on AI than a formal duty, as the former requires stronger reasoning that the use of AI could potentially dilute (Hendrickx, 2024).

Before outlining potential adverse impact, it is important to note the distinction between algorithmic and AI systems assisting judges from systems completely replacing judges. The former can be indicated as 'judicial decision-support systems' (JDSS) and refers to systems that support judges in their decision-making while formally leaving the final judgement in the hands of the judge. This contrasts with automated decision-making systems, which operate autonomously without human participation. As the latter are not yet fully developed or lack sufficient accuracy, this analysis focuses on JDSS. Moreover, if reliance on JDSS shows to have a negative impact on the judicial duty to state reasons and its normative goals, similar issues are even more likely to arise in case of automated decision-making systems.

Several factors can erode the underlying normative goals of the judicial duty to state reasons. Although this issue merits its own research in itself, the following section illustrates how each of the normative goals can be negatively impacted by drawing on case studies of algorithmic and AI systems currently being deployed within the judiciary. These examples are not exhaustive but rather try to highlight key challenges.

First of all, the underlying normative goal of legitimacy of the duty can be adversely impacted when judges rely on JDSS. Legitimacy of judicial decision-making concerns the question whether the public perceives the judiciary as worthy of their institutional role as well as being appropriate, proper and just. Providing clear reasons for judicial decisions helps courts maintain this legitimacy. However, reliance on algorithmic and AI systems can compromise the goal in multiple ways. For example, when judges rely on risk assessment tools or systems calculating average sentences, they are relying on outputs that are generated based on large datasets that often lack diversity and representative data. If judges based their decision on such 'tainted' data, it leads to judgements and reasoning that can undermine both the fairness and reliability of judicial decisions, and erode public trust in the judiciary's legitimacy (Hendrickx, 2024).

---

[2]While research also indicates that algorithmic and AI systems can have a positive impact on the duty and its underlying normative goals, this paper focuses on highlighting their potential adverse impact.

Legitimacy may also be compromised by the involvement of private companies in the design, development and deployment of these systems. Many algorithmic and AI systems used in the judiciary are developed by large technology companies,[3] whose products often reflect values and ideologies of their developers (Buyl, 2024). These embedded values and biases may subtly shape the system's output, which, in turn, influence the decisions made by judges who rely on these output. This can undermine the judiciary's duty to provide transparent and justifiable decisions, and affect the legitimacy of judicial decision-making as a normative goal underlying the judicial duty to state reasons.

Transparency, the second underlying goal of the judicial decision-making process, may also be challenged by the judiciary's use of AI. Many algorithmic and AI systems are characterised by their 'black box' nature, making it difficult or even impossible to fully understand their internal workings (Dymitruk, 2019). Empirical research has shown that algorithms are unable to accurately perform the complex legal reasoning required in judicial decision-making and cannot provide legally meaningful explanations for their output (Kolkman et al., 2024). This lack of transparency directly undermines the judicial duty to state reasons, as judges may not be able to understand AI's conclusions or suggestions, which they may integrate in their decisions. The opacity limits litigants and the general public to understand the reasoning behind a decision and thereby hinders litigants' right to appeal. Consequently, the lack of transparency risks impairing both the reasoning process and the duty to state reasons, as judges cannot adequately explain and justify their reliance on AI systems' output. The inability to understand and validate the systems' decision-making process can both erode confidence in the reliability of the judicial decision and infringe on the transparency goal.

The third underlying goal of the judicial duty to state reasons, accountability, can also be negatively impacted when judges rely on AI systems. For instance, if judges use generative AI systems to formulate their reasoning, or similarly, rely on the suggestions for a recidivism risk score or average sentence recommendation, they may (appear to) delegate parts of their decision-making to systems that lack democratic pedigree. This delegation can dilute judges' accountability for both the reasoning and the outcome of the decision. Furthermore, the opacity of these systems' inner workings can prevent judges from providing clear reasoning since they do not understand how the system generated its output, thus hindering individuals to scrutinise judicial reasoning and conduct effective judicial review (Hendrickx, 2024, 2025; Posner & Saran, 2025).

These examples illustrate how reliance on AI technologies can negatively impact the judicial duty to state reasons and its underlying normative goals. While other concerns also arise from this reliance, the purpose of this section was to demonstrate that AI reliance can indeed affect the judicial reasoning and the dynamic of the duty.

3. How the current legal frameworks (fail to) safeguard the judicial duty to state reasons and its underlying normative goals

3.1 Introduction

The preceding discussion has demonstrated that the existing procedural rules are insufficient to safeguard the normative goals of the duty. While they establish certain minimum 'criteria' for the duty, they offer only limited substantive guidance and remain open-ended. For instance, when judges rely on AI to formulate legal arguments or assist in drafting judgements, existing procedural rules do not specify whether such reliance must be disclosed, which parts of the decision are based on the AI tool, or how the AI system operates.

This sets the stage for the following analysis, namely whether alternative legal frameworks can impose additional obligations towards judges to articulate their reasoning when relying on AI systems. I examine two relevant legal instruments at the European level, namely the right to explanation

---

[3]For instance, generative AI systems are currently designed by big tech companies, like ChatGPT by OpenAI, Gemini by Google, Copilot by Microsoft or DeepSeek by High-Flyer.

under the GDPR and the AI Act. Although neither instruments are specifically designed to regulate the duty, they may nonetheless offer useful insights or partial safeguards in this context.

3.2 Right to explanation under the GDPR

The right to explanation under the GDPR may provide some relevant safeguards for the judicial duty to state in the context of JDSS when it obliges judges to provide explanations when they rely on AI systems.

While there is an ongoing debate regarding the existence, scope and precise nature of the right to explanation under the GDPR (in favour: Metikos & Ausloos, 2025; Metikos, 2024a; Almada, 2025; Malgieri & Comandé, 2017 – contra: Wachter, Mittelstadt & Floridi, 2017), it is generally assumed that such a right can be inferred. Specifically, the right to explanation can be deduced from Article 22(3), read together with Recital 71 and Articles 13, 14 and 15 GDPR. Without engaging in this broader debate, and assuming the right exists, let me briefly explain the right's main characteristics. The right to explanation requires data controllers to provide individuals with meaningful information about the logic involved in automated decisions, as well as the significance and envisaged consequences of such processing for the data subject that significantly affect them (Juliussen, 2025; Metikos & Ausloos, 2025). The right to explanation is rooted in transparency and accountability, and aims to ensure that individuals receive clear and understandable explanations about how their data is processed and can have meaningful control over their data, especially in automated decision-making. The right also allows individuals to assess the fairness and accuracy of data processing, detect biases, understand how decisions are made, get insights into the systems' logic, and thereby fostering trust and confidence. At first glance, the right could indeed impose additional reasoning obligation towards judges when they rely on AI systems in their decision-making process. When judges, for instance, take notes during court hearing, draft and publish judgements concerning private individuals while using AI systems in the process, they are engaging in personal data processing under the scope of the GDPR. This would imply that data subjects have a right to request information about the inner workings and logic of the AI system used in judicial decision-making. The required explanations can enhance the three normative goals. For instance, it enhances transparency of the decision-making process, by detecting and preventing the use of opaque and potentially biased algorithms. Providing more reasons prompts judges to critically assess the AI systems and their reliance on these systems, thereby enhancing judges' accountability. And in general, more reasoning can enhance legitimacy of judicial-decision making.

Nevertheless, the many limitations and unclarities of the right to explanation restrict the effectiveness of this right to impose additional reasoning obligations towards judges. Fundamentally, the GDPR is a data protection framework that aims to enhance transparency and accountability in the processing of personal data, for instance, to allow data subjects to assess whether the processing of their personal data was fair and to understand the logic involved in automated decision-making – without so much relevance to judicial reasoning. The most important limitation lies in the scope of the right to explanation. Under Article 22(1) GDPR, the right applies only to decisions that are solely based on automated processing and that produce legal or similarly significant effects on individuals (ECJ 7 December 2023; Metikos & Ausloos, 2025). While judicial decisions undoubtedly affect individuals, they do not meet the criterion of *solely* automated. It is rather the opposite in the context of JDSS: judges remain involved in the final decision-making process. The European Data Protection Board (EDPB) has clarified though that not any human intervention suffices: it should be meaningful rather than a mere formality. A human overseer must have both the authority and competence to change the decision (EDPB, 2017). Relevant factors include the amount of time available to oversee a task, the qualifications of the overseer, their liability, the support received to exercise the oversight, the agency of the overseer, the access to information, and the AI's system adaptability to human intervention (Wagner, 2019). Given these context-dependent considerations, it will thus depend on the

concrete circumstances and the extent a judge *de facto* relies on JDSS recommendations to determine whether judicial reliance on JDSS would fall outside the scope of this right.[4] As mentioned, in principle, since it concerns JDSS, judges' intervention is not limited to a mere formality. However, research suggests that AI-based recommendations might become binding in practice when people, including judges, rarely deviated from them, especially if the recommendations align with their preexisting views (Harbarth et al., 2024). In such cases, one could argue that judicial decisions are *solely* based on automated processing.

In case the right to explanation under the GDPR would be effectively triggered, judges would be required to provide explanations. Nevertheless, the scope of the right is rather vague: What is a meaningful or useful explanation? The GDPR itself does not specify how such meaningful explanations should take shape. However, in a recent case, the ECJ provided some guidance (ECJ 27 February 2025). The Court affirmed that data subjects have the right to receive sufficiently clear explanations of any automated decision that affects them, including details on how their data is processed. These explanations must be comprehensible to laypersons, which implies that a mere communication of an algorithm's technical elements is neither concise nor intelligible enough. They should allow individuals to verify and contest the outcome. Controllers must outline which personal data was used in the assessment, how the logic of the system weighs that data, and whether modifying certain variables could lead to different conclusions.

The new clarifications helps indeed interpreting of what meaningful explanations constitute. They indicate that merely disclosing the algorithm is insufficient and must be supplemented with additional information so that individuals can understand and contest the automated decisions. Nevertheless, several unclarities remain. For instance, while explanations must allow individuals to contest a decision, what kind of explanations would allow to achieve this? What form should these explanations take, given that individuals have different levels of expertise and interests? And what kind of technical details should be provided beyond the mere disclosure of the algorithm? Should explanations focus on the system's overall functionality, including its general logic, significance, and envisaged consequences of automated systems, such as model structures and decision criteria? Or should they be more decision-specific, such as the weighting of features, case-specific decision rules or information about references or profile groups? In addition, should explanations be provided ex ante or ex post? The latter distinction refers to the common temporary distinction made between ex ante explanations, which focus on system functionality before decisions are made, and ex post explanations, which can encompass both system functionality and case-specific explanations after a decision has been rendered.

Recital 71 of the GDPR suggests that ex post explanations of specific decisions should be provided, as it mentions suitable safeguards must be taken in case of automated decision-making and 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*' (own emphasis). However, as recitals are not legally binding, they do not establish any legitimate expectation or enforceable right. Over the years, scholars have tried to interpret this further. It is argued that basic information about the system's logic should be provided. The explanations can be descriptive, but should allow data subjects to verify the lawfulness of processing. Although technical explanations should be provided, its specific nature depends on the context. It should go beyond stating that the process applies AI but should not detail the different vectors or include mathematical formulas. It should not entail full transparency regarding the underlying system, but rather a description of the underlying system and it the functioning (Almada, 2025; Metikos & Ausloos, 2025; Wachter et al., 2017).

---

[4]Also note that Article 22(2) limits the scope of the right of explanations. Explanations are not required if one of three conditions is met, namely (a) if the decision is necessary for entering into, or performance of, a contract, (b) if the decision is authorised, or (c) if the decision is based on the data subject's explicit consent.

In conclusion, the right to explanation under the GDPR provides individuals the right to understand the logic, significance, and consequences of automated decisions that significantly affect them. It can indirectly support the normative goals of duty in the age of automation as it obliges judges to provide more reasons regarding the systems they relied upon. Nevertheless, its effectiveness in the judicial context remains limited. The GDPR is primarily a data protection framework. The key aim of the right is to ensure that data subjects understand how their personal data is used in automated decision-making. In addition, the right has a limited scope as it applies only to decisions that are *solely* based on automated processing, which rarely is the case in the context of JDSS. Even when the right is triggered, its focus is on providing insights into the logic of the AI system itself, and not about explaining how judges specifically used the system or how these recommendations influenced their reasoning. In addition, the right to explanation is only open to data subjects directly affected by the decision, excluding the general public. Given that judicial decision not only affect individuals but often have broader societal implications, this limits the right's potential to enhance the judicial duty to state reasons and its normative goals.

3.3 Right to explanation under the AI Act

Besides the GDPR, the recently adopted European AI Act may provide safeguards for the normative goals underlying the judicial duty to state reasons by imposing a right to explanation. Again, although the regulation is not specifically designed to govern this duty, it introduces provisions on transparency and the disclosure of the logic behind AI systems that could oblige judges to provide additional reasoning when they rely on AI systems.

The AI Act adopts a risk-based approach, meaning that the obligations imposed on AI systems vary depending on the level of risk they pose to individuals' health, safety and fundamental rights. Accordingly, the AI Act sets out different transparency obligations for different types of AI systems. Two provisions are particularly relevant to this context.

First, Article 86(1) AI Act introduces a right to explanation of individual decision-making in the context of high-risk AI systems. Specifically, affected individuals subject to a decision made by a *deployer*[5] on the basis of the output from a high-risk AI systems listed in Annex III that produces legal effects or similarly adversely affects their health, safety or fundamental rights, have a right to obtain clear and meaningful explanations. The explanations must cover both the role of the AI system in the decision-making process and the main elements of the decision. It closely resembles the right to explanation under the GDPR. However, Article 86(2) AI Act clarifies that this right applies only to the extent no other European or national law already grants a similar right. Hence, when data subjects already have a right to receive meaningful information under the GDPR, Article 86 AI Act does not apply. Rather than changing existing rights under the GDPR, the AI Act adds an additional right that individuals can invoke independently (Juliussen, 2025). Individuals involved in judicial proceedings where judges rely on JDSS can invoke this right if they can demonstrate that their health, safety or fundamental rights – such as the right to a fair trial – are at risk. According to Article 3(4) and Recital 13, judges can indeed be considered *deployers* when they use an AI system under their authority.[6] By providing affected individuals the right to clear and meaningful explanations on the role of the AI system in the decision-making process and main element of the decision, the normative goals of the judicial duty to state reasons can be enhanced similarly as explained under the previous sub-section.

Second, Article 50 AI Act introduces transparency requirements for different types of AI systems. Paragraph 4, subparagraph 2, is particularly relevant for this research, and specifically for the use case

---

[5]Indeed, affected individuals can only ask explanations from the deployer of the AI system and not, for instance, from the provider.

[6]Given financial and expertise constraints courts often face, it is indeed more feasible to consider courts or judicial agencies as deployers and not providers (the latter referring to the persons developing the AI systems).

of judges relying on generative AI systems to draft their judgements. It requires deployers – in this case, judges – to disclose that content has been artificially generated or manipulated when the text is published with the purpose of informing the public on matters of public interests. If judicial decisions are considered matters of public interest, for instance, because they inform how law is interpreted (Gils, 2024), then this provision could impose an obligation on judges to disclose their reliance on generative AI in drafting their judgements.

While these two transparency rights in the AI Act may require judges to provide additional reasoning when relying on AI systems in their decision-making, thereby reinforcing the underlying normative goals of the judicial duty to state reasons, their practical impact remains uncertain due to their limitations.

As for the right to explanation of individual decision-making enshrined in Article 86(1) AI Act, individuals can only request explanations from the deployer – in this case the judge – when the decision – in this case, the judgement – is based on the output from a high-risk AI system as mentioned in Annex III. This presupposes that the AI system in question qualifies as a high-risk AI systems under the AI Act. At first glance, this seems to be the case: Annex III, point 8(a) classifies AI systems used in the administration of justice as high-risk, specifically those intended to be used by the judiciary to assist them in researching and interpreting facts and the law and in applying the law to a concrete set of facts. Recital 61 clarifies that this should not extend to AI systems for purely ancillary administrative activities that do not affect the actual administration of justice in individual cases, such as anonymisation or pseudonymisation of judgements, documents or data, communication between judicial personnel, or other administrative tasks. This is to ensure that only the AI systems posing an adverse impact on democracy, rule of law and fundamental rights are subject to legal obligations (JuLIA, 2024). However, while judicial decisions based on outputs from high-risk AI systems can thus trigger the right to explanation, research has shown that the scope of point 8(a) is not only vague but also narrow. Are the criteria in Point 8(a) cumulative, requiring that the AI systems assist in both researching and interpretating facts and the law *and* in applying the law to the facts? What is meant by 'researching facts': how broad should this be understood? And what about AI systems that are marketed to both the judiciary and practitioners and its intended use is not limited to the judiciary? While certain AI system clearly fall under the scope, such as systems assisting in drafting judgements, for many others is remains unclear, such as precedent analysis tools, smart retrieval databases, or automated case allocation systems (Schwemer, Tomada & Pasini, 2021). These unclarities can significantly hinder the applicability of the right to explanation. The right is further restricted by the discretion of providers under Article 6(3) AI Act. Providers can unilaterally decide that their AI system does not qualify as high-risk AI systems if they argue that the system does not pose a significant risk the health, safety or fundamental rights. This discretion provides considerable leeway to circumvent the scope of high-risk AI systems (Metikos, 2024b; Metikos & Ausloos, 2025). If a system is consequently not considered as high-risk AI system, no specific obligations arise when judges rely on it. Lastly, as briefly mentioned, if data subjects already have a right to receive meaningful information about the logic of the AI system under the GDPR, Article 86 AI Act does not apply, further limiting its scope.

Similarly, Article 50(4), subparagraph 2, contains exceptions that restrict the requirement for deployers to disclose that a text was artificially generated. The exceptions apply either when the use is authorised by law to detect, prevent, investigate or prosecute criminal offences, or when the AI-generated content has undergone a process of human review or editorial control and where a natural or legal person holds editorial responsibility for the publication of the content. The latter exception might indeed be triggered in this context. When judges rely on generative AI to formulate their judgements, they should – as required by the definition of JDSS – have the final authority over judicial decisions. It can therefore be argued that the AI-generated content has undergone human review or editorial control. Nevertheless, the precise scope of such review is unclear: Does

a minimal spelling and grammar check suffice, or must the review encompass accuracy or coherence checks (Gils, 2024)? A case-by-case evaluation will likely be needed to determine whether this criterion is truly met. Besides such human review or editorial control, the exception requires a natural or legal person to hold editorial responsibility for the publication of the content. This refers to a person clearly designated as a point of contact in case of questions, which would likely fall to the courts issuing the judgements. Since the two conditions of the exception are likely met when judges rely on generative AI, the disclosure obligation in Article 50(4), subparagraph 2, will not be applicable.

In case the right to explanation under Article 86(1) AI Act would be effectively triggered, the next question is what kind of explanations must be provided to the affected individuals. Although the provision appears relatively straightforward at first glance, its practical application is vague. Article 86(1) specifies that deployers must provide clear and meaningful explanations regarding the role of the AI system in the decision-making procedure and the main elements of the decision taken. According to Recital 171, these explanations must be clear and meaningful and allow the affected persons to exercise their rights. Consequently, it seems that both pragmatic explanations – clarifying the AI system's role in the decision-making process – and more technical explanations concerning the system's main elements should be provided. In theory, this could enhance the normative goals of the judicial duty to state reasons. However, does this obligation require deployers to explain the specific role the AI system's output played in the decision, or does it extend to the logic of the system itself? Some scholars argue that deployers should not only provide explanations on the role of the AI output in the decision-making process but also include information about the system's data, algorithm type, and other technical aspects (Juliussen, 2025). It most likely does not require mathematical formula. Though it is unclear whether deployers should provide details about the system's main parameters. If so, would such technical explanations be meaningful to affected individuals, and who would be responsible to verify their correctness (De Mulder & Valcke, 2021)?

In case Article 50(4), subparagraph 2, is applicable, the question arises as to what the disclosure obligation entails. Article 50(5) specifies that that information should be delivered in a clear and distinguishable manner before the first exposure. According to Gils, this requires a case-by-case analysis to ensure that explanations are both understandable by the target audience and distinct from the wider context in which the AI content is used. However, the provision remains vague regarding the level of detail required, leaving uncertainty about the appropriate amount of information necessary to fulfil the disclosure obligation (Gils, 2024). More general, the scope of the disclosure obligation, which applies when content is artificially generated, is ambiguous. It is unclear whether the obligation applies when only a small part of the judgment is artificially generated, or whether the entire judgement should be generated by AI.

In conclusion, the right to explanation under the AI Act grants individuals a right to obtain explanations regarding the role of the AI system in the decision-making process and the main elements of the system. In doing so, it has the potential to enhance the normative goals of the duty. Nevertheless, its scope is limited: it applies only to high-risk AI systems, meaning that judges are not always required to provide additional explanations regarding the AI systems they use. Even when the right applies, ambiguity remains regarding the precise content of the required explanations. Similar as to the GDPR, it should be noted that the right to explanation under the AI Act is only available to individuals adversely affected by the decision, excluding the general public from requesting any information. Given that judicial decision not only affect individual but society alike, this restriction limits the right's usefulness in strengthening the judicial duty to state reasons and its normative goals. As for Article 50(4), subparagraph 2, its scope of application is limited as well. It applies only when judges rely on AI-generated text – a specific and narrow use case – and even when triggered, the precise requirements for disclosure remain vague. As a result, while transparency obligations in the AI Act may contribute to the normative goals underlying the duty to state reasons, their utility is inherently limited due to the scope and ambiguities of these provisions.

4. Rethinking the judicial duty to state reasons in the age of automation

4.1 Introduction

The current legal frameworks – both the procedural rules, and the GDPR and AI Act – do not adequately safeguard the judicial duty to state reasons and its underlying normative goals in the age of automation. However, given the duty's fundamental role in ensuring the right to a fair trial, the rule of law and its important functions, it is crucial to consider how it can remain effective. My proposal is to assess whether a more robust reason-giving obligation towards judges is required, and whether the requirements and conditions governing the duty to state reasons should be enhanced. In fact, I argue to rethink the core of duty, which currently adheres to formalistic and minimal standards. Instead, I propose to expand the duty and make it more substantive, by requiring judges to provide both pragmatic and technical reasons when relying on AI systems.

4.2 Judicial candour

Before elaborating on these additional reasoning requirements, it is important to clarify what this paper does not argue. The proposed heightened duty should not be mistaken for an endorsement of judicial candour. The proposed approach to require more substantive reasoning vis-à-vis judges closely relates to the discourse on judicial candour or sincerity. Judicial candour or sincerity, as outlined by scholars like Shapiro and Cohen, centres around the idea that judges must disclose their actual motivations, either through an internalist view (revealing their true motivations), or an externalist perspective (stating reasons believed to justify the outcome, even if those reasons were not the considerations that actually motivated them) (Cohen, 2010; Shapiro, 1987). At present, there is no universal consensus requiring judges to be bound by norms of sincerity. While not arguing for the disclosure of judges' sincere or candour reasoning, my argument does build upon the idea that judicial reason-giving should entail more than the minimum procedural rules that exists today, especially in case of the formal duty. However, instead of requiring actual motives for decisions, I propose that judges be required to provide additional explanations specifically addressing the use of AI systems, the system's role in the reasoning and decision-making as well as technical aspects of the systems themselves.

4.3 Pragmatic reasons

Returning to my argument, I propose that a more robust duty to state reasons should first and foremost include pragmatic reasons. Pragmatic reasons refer to judges explaining their interaction with and reliance on AI systems in the decision-making process. There are different ways in which such reasons could take shape. Emerging guidelines on the use of AI in courts may serve as inspiration. In recent months, an increasing number of guidelines on the use of AI systems in courts have been developed at both the national (e.g. Felsky & Eltis, 2024; Courts of New Zealand, 2023; UK Courts and Tribunals Judiciary, 2023) and international level (e.g. CEPEJ, 2024a; CCJE, 2023; UNESCO, 2024), while others still in development (e.g. Le ministère de la Justice France, 2025; Vanderstichele, 2024). A common feature among these guidelines is the emphasis on the 'transparent use' of AI systems in courts. Although the specific requirements vary, transparent use entails the idea that judges should inform properly and timely when and how AI systems are used and how these tools work. For instance, the CEPEJ Information Note recommends judges to be transparent and explicitly indicate whether generative AI has contributed to content or analysis. The UK Guidelines adopt a narrower approach by simply requiring judges to disclose the use of generative AI systems. The Canadian Guidelines, in contrast, emphasise that AI tools must be able to provide understandable explanations for their decision-making output. However, the UNESCO guidelines set the most concrete standards for transparency. They require judges to provide meaningful information about when AI tools are used, how their use may affect individuals involved in judicial proceedings, and whether materials

are produced based on these tools. In addition, the principle of opportunity to review decision and contestability adds to that that judges must provide information on how the AI system operates, how it is trained, the inputs, and the extent to which its outputs have informed the decision. Specifically for generative AI, judges must disclose its use and indicate which parts are produced by AI, for instance, through quotation marks or a citation system.

Building on these guidelines, particularly those from UNESCO, I propose that pragmatic reasons should be understood in an extensive manner, including the following requirements under the judicial duty to state reasons:

- Judges should disclose whether they use algorithmic or AI systems in their decision-making process.
- Judges should disclose what AI system version they rely on, and specify whether it is proprietary or open-source.
- Judges should disclose any known limitations, biases or potential errors associated with the AI system they use.
- Judges should disclose at what stage of the judicial decision-making process they use an AI system. Kolkman et al. identifies five stages of the process in which AI systems can be used: (i) the inventory phase, where judges read the case and determine the applicable legislation, (ii) the selection phase, where judges establish the facts and points of dispute, (iii) the assessment phase, where judges analyse and assess the dispute, (iv) the decision phase, where judges apply the legal rules to the case and decide, and (v) the editing phase, where judges motivate their decision and write the judgement (Kolkman et al., 2024). Judges should, for instance, clarify whether they use percent analysis software in the inventory phase, risk assessment tools in the assessment phase, or generative AI systems in the decision or editing phase.
- Judges should explain why they chose to rely on a particular AI tool, whether for efficiency, complexity, or other reasons.
- Judges should inform how the output produced by the AI system informed the decision-making process and their reasoning. They should specify whether it merely provided background information without being adopted was partially integrated in the reasoning or was fully integrated (i.e. copied verbatim).
- Judges should clarify whether engaging with the AI system changed their initial legal assessment.
- Judges should indicate whether they took measures to review the output, such as independent verification, consultation with human colleagues or cross-checking with jurisprudence.
- In case of generative AI, judges should indicate which parts of their reasoning or analysis are based on the output of AI, for instance, through quotation marks or citations.
- In case of generative AI, judges should disclose the prompts used to interact with the AI system.
- In general, judges should indicate whether they have received training on the use and limitations of AI systems in the judiciary to ensure responsible use.

By implementing these pragmatic reasons, the judicial duty to state reasons would evolve into a more substantive obligation. Requiring judges to provide pragmatic reasons regarding the role AI systems play in their decision-making process can strengthen the normative goals of this duty. In the first place, a more robust duty would encourage judges to assess more rigorously the quality and relevance of AI-generated output. It also helps mitigate automation bias and overreliance on AI, as judges must engage in more reflective reasoning (Klingbeil, 2024; Kolkman et al., 2024; Miller, 2023). This, in turn, strengthens accountability, as judges must critically evaluate the AI's contributions rather than relying on its output unexamined. In addition, openly describing the role of AI in judicial reasoning would allow parties and the general public to scrutinise the decision-making process (Barry,

2024). Pragmatic reasons also align with research suggesting that thorough explanations are essential for procedural fairness, a key factor of judicial legitimacy. It shows that individuals care about the decision-making process as much as the outcome itself (Allan Lind and Tyler, 1988). A detailed explanation on the role and function of AI in the judicial decision-making process is thus essential for maintaining trust and legitimacy in the judiciary (Edwards & Veale, 2017; Lim, 2009; Richard & Johnson, 1995; Westphal, 2023).

4.4 Technical reasons

Second, besides pragmatic reasons, the judicial duty to state reasons should include technical reasons that provide insights into the functioning and inner workings of AI systems used in the judicial decision-making. The emerging guidelines on the use of AI in courts, such as those from UNESCO, have indeed indicated that individuals should be able to review and contest AI-assisted decisions, which necessitates judicial transparency regarding how an AI system operates, how it was trained, and what inputs were used. This aligns with the broader discourse on XAI. XAI aims to render AI's internal workings more explainable and transparent (Abusitta, 2024; Wang & Ming, 2021), focusing among others on the design of explainable systems, explanations for legal reasoning, rationale discovery, and computational models of arguments (Collenette, 2023; Ross, 2017; Steging et al., 2021). The XAI discourse has developed alongside the rapid growth of machine learning applications, particularly in response to the black box character of many of these systems. XAI aims to realise a deeper understanding of how AI systems operate, as it is believed that it would enhance fairness and foster transparency, allowing both users and the public to scrutinise the systems (Barry, 2024; Speith, 2022). Information on AI systems and its inner workings is increasingly seen as a safeguard against arbitrary interventions. Consequently, explainability has evolved into a critical requirement in AI system design because of the need to justify decisions in ways that are accessible and comprehensible to both experts and laypersons.

Within the XAI discourse, explanations can be categorised at different levels. At the macro (global) level, explanations address how a system functions generally, offering global insights. This can also be referred to as ante-hoc explanations. It can include explanations on the design, training process, and model architecture. In contrast, at the micro (local) level, explanations provide insights into how the system arrives at specific outputs in response to specific inputs, often referred to as post-hoc explanations (Juliussen, 2025). They analyse why particular inputs led to specific outputs and may involve model-agnostic techniques, which apply broadly across different AI models, or model-specific techniques, which provide tailored insights into a particular system. Different approaches exist. On the one hand, structural approaches can integrate explainability into the system design itself, and make AI processes inherently interpretable. This could entail avoiding certain techniques in the design and development of an AI system if it makes external review impossible. On the other hand, artefactual approach can be used: by providing access to technical artefacts, insights can be given into the system. Think of source code, technical documentation or model summaries. While artefactual approach offer flexibility in the sense that different artefacts can be given depending on the stakeholders, its consistency can vary significantly (Almada, 2025).

To ensure more robust judicial reason-giving, courts should extend their duty to state reasons beyond merely acknowledging AI reliance and its role in decision-making and also include technical explanations of the AI system relied upon. XAI techniques can be used to support judges in formulating technically robust explanations and translate complex AI logic into more comprehensible formats. Different technical explanations can be provided in judicial reasoning. Judges could disclose general system features, such as its design and intent, metadata on training data, or the performance metrics underlying the system's reliability. Alternatively, they could provide global insights into the model's mechanisms and structure, case-specific justifications, detailing why particular inputs led to specific outputs. Or they could also opt for a hybrid approach, combining general model explanations

with case-specific explanations (Binns, 2018; Edwards & Veale, 2017). At a minimum, I argue that global explanations should always be required, supplemented when necessary by local case-specific reasoning. In line with Almada, I also argue for a pluralistic and socio-technical approach: not only purely technical explanations are required, but also explanations on key design choices, governance structure, and the broader organisational context in which the AI system is deployed (Almada, 2025).

The judicial duty to provide technical explanations when relying on AI systems can strengthen the normative goals of the judicial duty to state reasons. It promotes judicial understanding of the tools they employ and fosters a greater sense of responsibility and awareness of systems' limitations. This strengthens accountability, as judges must engage more critically with AI-generated outputs rather than relying on them unconditionally. Beyond judges themselves, technical explanations can empower litigants and the general public to better grasp and assess the role of AI in judicial decision-making. Research has indeed indicated that clear explanations can improve individuals' understanding of AI-generated outcomes, foster public trust and acceptance, and shape positive attitudes towards automation in legal settings (Binns, 2018; Kizilcec, 2016; Shulner-Tal, 2022).

5. Critical reflections

Whereas the previous section proposed rethinking and strengthening the judicial duty to state reasons by requiring additional reasons, several challenges and drawbacks arise in implementing this approach effectively – prompting the question whether this approach is feasible and suitable. I outline some key obstacles.

The first and perhaps biggest hurdle concerns whether it is realistic to expect judges to provide the level of explanations required under this enhanced duty to state reasons, especially with regard technical explanations. Even the inclusion of pragmatic reasons may already require some degree of technical literacy, let alone the expectation that judges explain the inner workings of AI systems. Judges are usually not trained in technologies. In this regard, Article 4 AI Act introduces a requirement for AI literacy. It requires that providers and deployers of AI systems, such as judges, shall take measures that persons dealing with the systems have a sufficient level of AI literacy. It intends to equip these people with awareness of the risks and opportunities of AI systems. However, practical challenges arise in implementing effective AI literacy initiatives. While some training and resources have started to emerge, ensuring that these efforts are effective remains a complex task. Successful AI literacy requires that training programmes are targeted and responsive to rapid technological advancement. It also remains unclear how to ensure that these literacy efforts are genuinely internalised by participants, which seems to require rigorous oversight and continuous evaluation, as well as tailored trainings that emphasise the importance of the duty to state reasons in AI-driven decisions. Moreover, Article 4 uses open-ended notions, for instance, what is meant by a 'sufficient level' of literacy? More importantly, even with AI literacy initiatives, judges will likely never possess the depth of technical knowledge required to explain complex AI models fully. This raises the concern whether requiring pragmatic and technical reasons from judges is truly realistic.

Even if judges were cable of providing pragmatic and even technical explanations, the question arises whether XAI methods are going to be effective. There are certain challenges with XAI, such as technical shortcomings regarding their robustness, the inability of their explanation to remain consistent and accurate across a range of inputs, and the fact that XAI explanations remain proxies, rather than precise representations of the actual decision-making process (Panigutti, 2023). Hence, given the limitations of XAI, an alternative approach could be to prioritise reviewability rather than explanations alone. Cobbe et al. argue indeed that opacity in algorithmic systems limits meaningful explanations, either due to illiterate opacity, where the system's technical complexity renders it inaccessible to most users, or intrinsic opacity, where the mathematical nature of machine learning is even difficult for experts to interpret. Therefore, they propose to focus on reviewability of the automated decision-making process, referring to a comprehensive record-keeping mechanism that tracks

both technical and organisational details necessary for meaningful review. They argue that most people are not concerned about machine learning's inner-workings, but rather with broader aspects of automated decision-making process, such as its purpose, roles and outcomes, from commission to design, deployment, use and consequences of the process (Cobbe, 2021). This ties in with the question of whether we should prioritise explaining algorithms or, instead, the entire decision-making process, including explanations how systems are designed and developed and who is involved (De Bruijn, 2022).

Another consideration is whether different levels of reasoning are required depending on the phase in which an AI system is used in judicial decision-making. At each stage – from design to development and actual decision-making – different reasoning may be relevant. Therefore, a straightforward answer to the question what characterises an effective AI-explanation lacks. In addition, meaningful explanations can also depend on the audience (Brughmans et al., 2024). For explanations to be useful, they must be comprehensible and appropriate to the audience, as they are social and contextual. However, individuals' responses to explanation styles can vary, with some preferring technical precision while others favour simplified explanations. This mean that different audiences – judges, lawyers, parties or general public – might need different explanations. This raises the question whether we need tailored, audience-sensitive explanations – particularly since AI explanations, whether global or case-specific, are inherently complex (Dodge, 2019). Hence, for explanations to serve judicial fairness, it seems they must strike a balance, being neither too simplified to be meaningless nor too technical to be inaccessible. However, if tailored, the reasoning process becomes even more complex. Related, one may also question whether the enhanced duty to state reasons should be required for every instance of judicial reliance on AI systems, regardless of the type of technology used, or whether a distinction should be made based on the specific AI system involved. The latter seems likely, but warrants more research.

Requiring technical explanations also touches upon the broader debate over transparency of the source code and whether big tech companies should be obligated to open their software to judicial or public review (Edwards & Veale, 2017). Some have argued that the release of the source code to the public is necessary so that the public knows how the systems work (Berry, 2006).

A specific concern regarding pragmatic reasons is whether judges are under an obligation to provide counterfactual explanations if they have to disclose the extent to which an AI system influenced their decision and whether their judgement would have been different had they not used AI (Yacoby, 2022). Besides being difficult to provide, counterfactual explanations may not be effective in fostering critical thinking. In the context of AI-assisted decision-making, counterfactual decision do not significantly alter people's choices, so they may not actually help judges critically assess the AI's impact on their reasoning. If not useful, then imposing them as part of the judicial duty to state reasons may add an unnecessary complexity without achieving the intended benefits.

A more general drawback concerns the 'transparency paradox' and the phenomenon of information overload. While more reasoning and information is often assumed to enhance transparency, there is evidence that overwhelming details can decrease clarity and accessibility, especially for laypersons (Greenstein, 2022). The judiciary thus faces a challenge in ensuring that additional (and often technical) explanations do not obscure the decision-making process. For example, if judges provide extensive technical details, these could inadvertently lead to more opacity where irrelevant or overly complex information distracts from the core reasoning. Studies confirm that, under information overload, people may disregard crucial details, confuse relevant with irrelevant information, or even stop processing entirely (Edwards & Veale, 2017; Kizilcec, 2016; Stohl, 2016). It can indeed be asked whether longer, e.g. 80-page versus 20-page, judgements are necessarily better. Does the length really make a difference in the degree of acceptance and perceived legitimacy? It seems that as long as there is an answer to the crucial questions, more elaborate reasoning might not be imperative.

While it has been argued that explanations can help mitigate the risk of overreliance on AI or reduce automation bias, it has also been shown that this effect may not be the case. For simple tasks,

explanations have little impact on overreliance, as individuals may rely on their own judgement. Conversely, with complex tasks – such as court decisions, explanations can be difficult to interpret, leading users to follow the AI's suggestions. For instance, one study found that when participants were asked to assess an AI-generated summary for reading comprehension, they opted to trust the AI rather than engage with the explanations. This tendency may also emerge in the judicial context, where the mental efforts required to interpret AI explanations may lead judges to accept AI outputs without deeper scrutiny (Miller, 2023).

Another notable challenge is that requiring more robust reasoning may impose constraints on judicial flexibility. Frederick Schauer argues that when the judiciary is tasked with documenting every rationale thoroughly, it may restrict the judge's flexibility in decision-making, especially in cases where the opaque black box nature of AI precludes transparent explanations (Schauer, 1995). Judicial flexibility is especially crucial as the interpretation of law through case law evolves over time to reflect changes in society. A rigid requirement for exhaustive reasoning could impede this and lead to a situation where law is disconnected from social realities and the judiciary as being perceived outdated or out of touch (Gomez, 2015).

Moreover, requiring more reasons may also imply an undesirable accountability shift from private companies or other entities responsible for these systems to judges. Rather than these entities being held accountable for their systems and decisions, judges would be expected to justify the use of a particular system, their interaction with it, and potentially even its functioning.

A more robust reasoning requirement also risks transforming inadequate reasoning into grounds for appeal, a trend observed in Brazil and Mexico, where the substantive duty to state reasons has led to frequent appeals based on disagreement with reasoning alone (Ho, 2000). Historically, concerns about overly rigid reason-giving requirements date back to the 19th century, where US Lord Mansfield famously advised against giving reasons for every judgment, because '*for your judgement will probably be right, but the reasons will certainly be wrong*' (Campbell, 1973). He warns for the risk that overly strict reasoning requirements may affect the soundness and perceived correctness of judicial decisions.

A practical but one of the most important downsides is the added workload for judges (Gutiérrez, 2024a). Insisting on rigorous and thorough examination of any content generated by these systems defeats the exact purpose of time-efficiency arguments in favour of AI systems.

Finally, even if we were to agree that a more robust duty to state reasons is desirable, practical questions arise on how to best integrate such duty into the judicial practice, as well as constitutional considerations that may arise when adapting existing frameworks.

6. Way forward: to rethink the duty or not?

This research demonstrates that the judicial duty to state reasons acts both as a principle under pressure and a potential solution to the challenge posed by AI in the judiciary.

The question mark behind the paper's title 'Rethinking the judicial duty to state reasons in the age of automation?' is indeed intentional. Whereas rethinking the judicial duty to state reasons in the sense of requiring additional pragmatic and technical reasons from judges may seem to bring benefits, this approach comes with several challenges, prompting questions about whether requiring additional explanations is truly the way forward. At its core, the debate resolves around what we can reasonably expect from judicial reasoning in an era where decisions are increasingly informed by AI systems.

While it is clear that the judicial duty to state reasons must be safeguarded and even strengthened in the age of automation, this research has raised – but not definitely answered – questions about the most effective approach to do so. There remains a need to further explore the type of reasoning that can best support the judicial duty to state reasons' normative goals, and whether more robust explanations are in fact the best way forward.

References

Regulation (EU) No 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). http://data.europa.eu/eli/reg/2024/1689/oj

Regulation (EU) No 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). http://data.europa.eu/eli/reg/2016/679/oj

ECJ 15 October 1987, C-222/86, Heylens.

ECJ 6 September 2012, C-619/10, Trade Agency.

ECJ 7 December 2023, C-634/21, SCHUFA.

ECJ 27 February 2025, C-203/22, Dun & Bradstreet Austria.

ECtHR 9 December 1994(a), No. 18390/91, Ruiz Torija/Spain.

ECtHR 9 December 1994(b), No. 18064/91, Hiro Balani/Spain.

ECtHR 19 February 1998, No. 134/1996/753/952, Higgins and Others/France.

ECtHR 21 January 1999, No. 30544/96, Garcia Ruiz/Spain.

ECtHR 27 September 2001, No. 49684/99, Hirvisaari/Finland.

ECtHR 12 February 2004, No. 47287/99, Perez/France

ECtHR 6 March 2006, No. 19960/04, Popov/Moldova.

ECtHR 14 February 2007, No. 6923/03, Melnic/Moldova

ECtHR 8 April 2008, No. 7170/02, Grădinar/Moldova.

ECtHR 18 May 2010, No. 36552/02, Vetrenko/Moldova.

ECtHR 30 September 2022, No. 15269/13, Rusishvili/Georgia.

ECtHR 7 September 2023, No. 43627/16, Okropiridize/Georgia.

Abusitta, A. E. A. (2024). Survey on Explainable AI: Techniques, challenges and open issues. *Expert Systems with Applications*, *255*, 1–18.

Albright, A. (2023). The hidden effects of algorithmic recommendations. https://apalbright.github.io/pdfs/Algo_Recs_July_2023.pdf

Allan Lind, E., & Tyler, T. R. (1988). *The Social Psychology of Procedural Justice*. Springer.

Almada, M. (2025). Technical AI transparency: A legal view of the black box. *SRRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5096913

Araujo, T. E. A. (2020). In AI we trust? perceptions about automated decision-making by artificial intelligence. *AI and Society*, *35*, 611–623.

Barry, B. M. (2024). AI for assisting judicial decision-making: Implications for the future of open justice. *Australian Law Journal*, *98*, 656–669.

Beckman et al. (2024). AI and democratic legitimacy. The problem of publicity in public authority. *AI and Society*, *39*, 975–984.

Bentham, J. (1790). *Draught of a New Plan for the Organisation of the Judicial Establishment in France*. https://archive.org/details/bim_eighteenth-century_draught-of-a-new-plan-fo_bentham-jeremy_1790/page/n2/mode/1up

Berry, D. M. E. A. (2006). Free and open-source software: Opening and democratising e-government's black box. *Information Polity*, *11*(1), 21–34.

Binns, R. E. A. (2018). 'It's reducing a human being to a percentage': Perceptions of justice in algorithmic decisions. *arXiv*: 1801.10408v1

Brughmans, D. E. A. (2024). Disagreement amongst counterfactual explanations: How transparency can be misleading. *TOP*, *32*, 429–462.

Buyl, M. E. A. (2024). Large language models reflect the ideology of their creators. *arXiv*: 2410.18417v1

Campbell, L. (1973). The lives of the chief justices. *Lord Mansfield*, *3*, 481.

Carneiro Rocha, C. M. E. A. (2021). Descriptive e-justice framework to support studies at administration of justice. *Conference: Administration of Justice Meeting.*

CCJE (2008). *Opinion No. 11 of the Consultative Council of European Judges on the Quality of Judicial Decisions.* https://rm.coe.int/16807482bf

CCJE (2023). *Opinion No. 26. Moving forward: The use of assistive technology in the judiciary.* https://rm.coe.int/ccje-opinion-no-26-2023-final/1680adade7

CEPEJ (2024a). *Information Note: Use of Generative Artificial Intelligence (AI) by judicial professionals in a work-related context.* https://rm.coe.int/cepej-gt-cyberjust-2023-5final-en-note-on-generative-ai-1680ae8e01

CEPEJ (2024b). *Special file - Report "European judicial systems - CEPEJ Evaluation report - 2024 Evaluation cycle (2022 data).* https://www.coe.int/en/web/cepej/special-file

Chronowski, N. E. A. (2021). Artificial intelligence, justice, and certain aspects of right to a fair trial. *Acta Univ. Sapientiae, Legal Studies*, *10*(2), 169–189.

Cobbe, J. E. A. (2021). Reviewable automated decision-making: A framework for accountable algorithmic systems. *ACM Conference on Fairness, Accountability, and Transparency (FAccT '21).*

Cohen, M. (2009). The rule of law as the rule of reasons. *SSRN.* https://papers.ssrn.com/abstract=1518006

Cohen, M. (2010). Sincerity and reason-giving: When may legal decision makers lie? *Depaul Law Review*, *59*(4), 1091–1150.

Cohen, M. (2011). The Social Epistemology of Public Institutions. In M. D. Mar (Ed.), *New Waves in Philosophy of Law* (185–208). Springer.

Cohen, M. (2015). When judges have reasons not to give reasons: A comparative law approach. *Washington and Lee Law Review*, *72*(2), 483–571.

Cohen, S. (2021). The evolution of machine learning: Past, present, and future. In C. S (Ed.), *Artificial Intelligence and Deep Learning in Pathology*, (1–12). Elsevier. doi:10.1016/B978-0-323-67538-3.00001-4

Collenette, J. E. A. (2023). Explainable AI tools for legal reasoning about cases: A study on the European Court of Human Rights. *Artificial Intelligence*, 317. doi:10.1016/j.artint.2023.103861

Corvalan, J. G. (2020, May 1). *Prometea experience. Using AI to optimize public institutions.* CERIDAP. https://ceridap.eu/prometea-experience-using-ai-to-optimize-public-institutions/?lng=en

Courts of New Zealand. (2023). *Guidelines For Use of Generative Artificial Intelligence in Courts and Tribunals.* (20231207–GenAI–Guidelines–Judicial.pdf).

De Bruijn, H. E. A. (2022). The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making. *Government Information Quarterly*, *39*(2), 1–8.

De Mulder, W., & Valcke, P. (2021). The need for a numeric measure of explainability. *2021 IEEE International Conference on Big Data (Big Data)*, 2712–2720.

Dessers, V. K., & Valcke, P. (2020). Judicial analytics on trial: An assessment of legal analytics in judicial systems in light of the right to a fair trial. *Maastricht Journal of European and Comparative Law*, *27*(6), 759–773.

Dhungel, A. K. (2024). "This verdict was created with the help of generative AI..?" On the use of large language models by judges. *Digital Government: Research and Practice*, *6*(1), 1–8.

Dodge, J. E. A. (2019). Explaining models: An empirical study of how explanations impact fairness judgment. *arXiv*: 1901.07694

Dreyer, V. (2021). The Epistemology and Science of Justified Reason *Philosophia*, *50*, 503–532.

Dymitruk, M. (2019). The right to a fair trial in automated civil proceedings. *Masaryk University Journal of Law and Technology*, *13*(1), 27–44.

ECtHR (2024, February 29). *Guide on Article 6 of the European Convention on Human Rights. Right to a fair trial (civil limb).* https://ks.echr.coe.int/documents/d/echr-ks/guide_art_6_eng-pdf

EDPB (2017). *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679.* WP251rev.01). https://ec.europa.eu/newsroom/article29/redirection/document/49826

Edwards, L., & Veale, M. (2017). Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for. *Duke Law & Technology Review*, *16*, 18–84.

Fabri, M. (2024). From court automation to e-justice and beyond in Europe. *International Journal for Court Administration*, *15*(3), 1–23.

Farah, H. (2023, February 15). Court of appeal judge praises 'jolly useful' ChatGPT after asking it for legal summary. The Guardian. https://www.theguardian.com/technology/2023/sep/15/court-of-appeal-judge-praises-jolly-useful-chatgpt-after-asking-it-for-legal-summary

Felsky, P., & Eltis, K. (2024). *Guidelines for the Use of Artificial Intelligence in Canadian Courts.* Canadian Judicial Council. https://ajrn.org/2024/10/25/canadian-judicial-council-outlines-guidelines-for-use-of-ai-in-courts/

Forst, R. (2013). *The Right to Justification.* Colombia University Press.

France, G. (2019, September 30). *Transparency of court proceedings.* Transparency International Anti-Corruption Helpdesk Answer. https://knowledgehub.transparency.org/assets/uploads/helpdesk/Transparency-of-court-proceedings_2019_PR.pdf

Gentile, G. (2022). AI in the courtroom and judicial independence: An EU perspective. *SSRN.* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4198145#

Gils, T. (2024). A Detailed Analysis of Article 50 of the EU's Artificial Intelligence Act. In C. N. Pehlivan, N. Forgo & P. Valcke (Eds.), *The EU Artificial Intelligence (AI) Act: A Commentary*. Kluwer.

Gomez, M. I. G. (2015). The contribution of judges to society. *Archives for Philosophy of Law and Social Philosophy*, *101*(3), 332–353.

Greenstein, S. (2022). Preserving the rule of law in the era of artificial intelligence (AI). *Artificial Intelligence and Law*, *30*, 291–323.

Grossman, M. R. (2023). The GPTJudge: Justice in a generative AI world. *Duke Law & Technology Review*, *23*(1). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4460184#

Gutiérrez, J. D. (2024a). AI technologies in the judiciary: Critical appraisal of Large Language Models in judicial decision-making. In R. Paul, E. Carmel & J. Cobbe (Eds.), *Handbook on Public Policy and AI*. Edward Elgar Publishing.

Gutiérrez, J. D. (2024b). *UNESCO Global Judges' Initiative: Survey on the use of AI systems by judicial operators, CI/DIT/2024/JI/01*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000389786

Harbarth, L., Grosswein, E., Bodemer, D., & Schnaubert, L. (2024). (Over)trusting AI recommendations: How system and person variables affect dimensions of complacency. *International Journal of Human–Computer Interaction*, *41*, 391–410.

Hazelhorst, M. (2017). The Right to a Fair Trial in Civil Cases. In M. Hazelhorst (Ed.), *Free Movement of Civil Judgments in the European Union and the Right to a Fair Trial*. Springer.

Hedler, L. (2022). Algorithms, efficiency and the two faces of courts: A case study of the Brazilian Superior Court of Justice (STJ). *Soziale Systeme*, *26*(1–2), 370–395.

Hendrickx, V. (2024). The judicial duty to state reasons in the age of automation. The impact of generative AI systems on the legitimacy of judicial decision-making. *Erasmus Law Review*, *3*, 1–13.

Hendrickx, V. (2025). AI and the Judicial Duty to State Reasons. In M. Zalnieriute & A. Limante (Eds.), *The Cambridge Handbook on AI and Technologies in Courts*. Cambridge University Press forthcoming.

Ho, H. L. (2000). The judicial duty to give reasons. *Legal Studies*, *20*(1), 42–65.

JuLIA (2024). Artificial intelligence, judicial decision-making and fundamental rights, 1–217. https://www.julia-project.eu/sites/default/files/2025-02/JuLIA_handbook_Justice_final.pdf

Juliussen, B. A. (2025). The right to an explanation under the GDPR and the AI act. *Lecture Notes in Computer Science*, *15523*, 184–197.

Katz, O., & Zamir, E. (2024). Law, justice and reason-giving. *Journal of Empirical Legal Studies*, *22*, 243–266. forthcoming.

Kizilcec, R. F. (2016). How much information? Effects of transparency on trust in an algorithmic interface. *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. doi: 10.1145/2858036.2858402

Klingbeil, A. E. A. (2024). Trust and reliance on AI—An experimental study on the extent and costs of overreliance on AI. *Computers in Human Behavior*, *160*, 1–10.

Kolkman et al (2024). Justitia ex machina: The impact of an AI system on legal decision-making and discretionary authority. *Big Data and Society*, *11*(2), 1–14.

Koukoulioti, V. (2024, January 17). *Robot judges and the use of predictive justice in tax cases*. TaxJournal. https://www.taxjournal.com/articles/robot-judges-the-use-of-predictive-justice-in-tax-cases.

Labour circuit of Cartagena 30 January 2023. *Case 032*. https://pupilacdn.nyc3.cdn.digitaloceanspaces.com/diariojudicial.public/documentos/000/106/904/000106904.pdf

Le ministère de la Justice France (2025, February 5). *Développer l'intelligence artificielle (IA) pour la Justice: entre innovation et sécurisation*. https://www.justice.gouv.fr/actualites/actualite/developper-lintelligence-artificielle-ia-justice-entre-innovation-securisation

Liebowitz, J. (1986). Expert systems in law: A survey and case study. *Telematics and Informatics*, *3*(4), 263–271.

Lim, B. Y. (2009). Why and why not explanations improve the intelligibility of context-aware intelligent systems. *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

Malgieri, G., & Comandé, G. (2017). Why a right to legibility of automated decision-making exists in the general data protection regulation. *International Data Privacy Law*, *7*(4), 243–265.

Medvedeva, M. E. A. (2023). Rethinking the field of automatic prediction of court decisions. *AI and Law*, *31*(1), 195–212.

Mentovich, A. E. A. (2023). Legitimacy and online proceedings: Procedural justice, access to justice, and the role of income. *Law & Society Review*, *57*(2), 189–213.

Merill, T. W. (1993). Judicial opinions as binding law and as explanations for judgments. *Cardozo Law Review*, *15*, 43–79.

Metikos, L. (2024a). Explaining and contesting judicial profiling systems. beyond a procedural right to an explanation. *Technology and Regulation*.

Metikos, L. (2024b). The AI act: Weak, weaker, weakest. *SSRN*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4873496

Metikos, L., & Ausloos, J. (2025 forthcoming). *The Right to an Explanation in Practice: Insights from Case Law for the GDPR and the AI Act*. Law, Innovation and Technology.

Miller, K. (2023). *AI Overreliance Is a Problem. Are Explanations a Solution?* HAI Stanford University. May 13.

Neil, P. (1998). The Duty to Give Reasons: The Openness of Decision-Making. In C. Forsyth & I. Hare (Eds.), *The Golden Metwand and the Crooked Cord: Essays in Honour of Sir William Wade QC* (161–184). Oxford University Press.

Opdebeek, I., & De Somer, S. (2016). The duty to give reasons in the European legal area: a mechanism for transparent and accountable administrative decision-making? A comparison of Belgian, Dutch, French and EU Administrative Law. *Rocznik Administracji Publicznej, 2,* 97–148.

Panigutti, C. E. A. (2023). The role of explainable AI in the context of the AI Act. *FAccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency,* 1139–1150. doi: 10.1145/3593013.3594069

Paul, R. (2022). Can critical policy studies outsmart AI? Research agenda on artificial intelligence technologies and public policy. *Criminal Policy Studies, 16*(4), 497–509.

Posner, E. A., & Saran, S. (2025). Judge AI: Assessing large language models in judicial decision-making. *University of Chicago Coase-Sandor Institute for Law & Economics Research Paper,* 25–03.

Postema, G. (2014). The Soul of Justice: Bentham on Publicity, Law, and the Rule of Law. In Q. Zhai, and M. Quinn (Eds.), *Bentham's Theory of Law and Public Opinion.* Cambridge University Press, 40–62.

Rechtbank Gelderland 7 June 2024, ECLI:NL:RBGEL:2024:3636. https://linkeddata.overheid.nl/front/portal/document-viewer?ext-id=ECLI:NL:RBGEL:2024:3636

Reiling, A. D. (2020). Courts and artificial intelligence. *International Journal for Court Administration, 11*(2), 1–8.

Richard, Y. L., & Johnson, P. E. (1995). The impact of explanation facilities on user acceptance of expert systems advice. *MIS Quarterly, 19*(2), 157–172.

Richardson, H. S. (2003). *Democratic Autonomy: Public Reasoning about the Ends of Policy.* Oxford University Press.

Ross, A. E. A. (2017). Right for the right reasons: Training differentiable models by constraining their explanations. *arXiv:* 1703.03717v2

Saied-Tessier, A. (2024). AI in the family justice system. *Nuffield Family Justice Observatory.* https://www.nuffieldfjo.org.uk/wp-content/uploads/2024/05/NFJO_AI_Briefing_Final.pdf

Schauer, F. (1995). Giving Reasons. *Stanford Law Review, 47*(4). https://www.jstor.org/stable/1229080

Schindler, E. (2024, January 8). *Judicial systems are turning to AI to help manage vast quantities of data and expedite case resolution.* IBM. https://www.ibm.com/blog/judicial-systems-are-turning-to-ai-to-help-manage-its-vast-quantities-of-data-and-expedite-case-resolution/

Schmitz-Berndt, S. (2024). Interfering with judicial independence? The legal constraints in the realm of AI-powered judicial decision-making. *Aire, 1*(2), 199–211.

Schwemer, S. F., Tomada, L., & Pasini, T. (2021). Legal AI systems in the EU's proposed Artificial Intelligence Act. *Proceedings of the Second International Workshop on AI and Intelligent Assistance for Legal Professionals in the Digital Workplace (LegalAIIA 2021).*

Seng, D. K. B. (2021). Artificial intelligence and evidence. Singapore academy of law Special issue on law and technology. *SSRN.* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3924762

Shapiro, D. L. (1987). In defense of judicial candor. *Harvard Law Review, 100*(4), 731–750.

Shapiro, M. (1992). The Giving Reasons Requirement. *University of Chicago Legal Forum, 1992*(1), 179–220.

Shulner-Tal, A. E. A. (2022). Fairness, explainability and in-between: Understanding the impact of different explanation methods on non-expert users' perceptions of fairness toward an algorithmic system. *Ethics and Information Technology, 24*(2), 1–13.

Simmons, R. (2018). Big data, machine judges, and the legitimacy of the criminal justice system. *UC Davis Law Review, 52,* 1067–1118.

Skaug Saetra, H., & Selinger, E. (2023). The siren song of technological remedies for social problems: Defining, demarcating, and evaluating techno-fixes and techno-solutionism. *SSRN,* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4576687

Smith, A. E. A., (2024, May 30). *AI in the courtroom: Judges enlist ChatGPT help, critics cite risks.* CS Monitor. https://www.csmonitor.com/USA/Justice/2023/0530/AI-in-the-courtroom-Judges-enlist-ChatGPT-help-critics-cite-risks

Smuha, N. A. (2024). *Algorithmic Rule By Law. How Algorithmic Regulation in the Public Sector Erodes the Rule of Law.* Cambridge University Press.

Smuha, N. A., & Hendrickx, V. (2023, October 2). AI and the Administration of Justice: Taking 'precedent analysis' as a use case to assess the adequacy of the AI Act. *KU Leuven AI Summer School.* https://www.law.kuleuven.be/ai-summer-school/blogpost/Blogposts/AI-administration-justice

Speith, T. (2022). A review of taxonomies of explainable artificial intelligence (XAI) methods. *FAccT.* doi:10.1145/3531146.3534639

Staszewski, G. (2009). Reason-giving and accountability. *Minnesota Law Review, 545,* 1253–1326.

Steging, C. E. A. (2021). Discovering the rationale of decisions: experiments on aligning learning and reasoning. *arXiv:* 2105.06758

Stohl, C. E. A. (2016). Managing opacity: information visibility and the paradox of transparency in the digital age. *International Journal of Communication, 10,* 123–137.

Susskind, R. E. (1986). Expert systems in law: A jurisprudential approach to artificial intelligence and legal reasoning. *The Modern Law Review, 49*(2), 168–194.

Taylor, L. (2023, February 3). *Colombian judge says he used ChatGPT in ruling.* The Guardian. https://www.theguardian.com/technology/2023/feb/03/colombia-judge-chatgpt-ruling.

Tyler, T. R. (2006). *Why People Obey the Law*. Princeton University Press.

UK Courts and Tribunals Judiciary (2023, December 12). *Artificial Intelligence (AI) Guidance for Judicial Office Holders*. https://www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf

Ulenaers, J. (2020). The impact of artificial intelligence on the right to a fair trial: Towards a robot judge? *Asian Journal of Law and Economics*, *11*(2), 20200008.

UNESCO. (2024). Draft UNESCO guidelines for the use of AI systems in courts and tribunals. *UNESCO*. https://unesdoc.unesco.org/ark:/48223/pf0000390781

Vanderstichele, G. (2024, October 14). *Voorstel van richtlijnen voor de zetel voor het gebruik van generatieve AI (grote taalmodellen)*. Substack. https://gvanderstichele.substack.com/p/voorstel-van-richtlijnen-voor-de

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International Data Privacy Law*, *7*(2), 76–99.

Wagner, B. (2019). Liable, but not in control? Ensuring meaningful human agency in automated decision-making systems. *Policy & Internet*, *11*(1), 104–122.

Wang, X., & Ming, Y. (2021). Are explanations helpful? A comparative study of the effects of explanations in AI-assisted decision-making. *IUI '21: Proceedings of the 26th International Conference on Intelligent User Interfaces*, 318–328. doi: 10.1145/3397481.3450650

Westphal, M. E. A. (2023). Decision control and explanations in human-AI collaboration: Improving user perceptions and compliance. *Computers in Human Behavior*, *144*, 107714.

Yacoby, Y. E. A. (2022). "If it didn't happen, why would I change my decision?": How judges respond to counterfactual explanations for the public safety assessment. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing. arXiv*: 2205.05424

Victoria Hendrickx is a PhD researcher at KU Leuven, CiTiP, where she examines the impact of emerging digital technologies on the judiciary. Under the research title 'Algorithmic Justice: Safeguarding the Judicial Duty to State Reasons in the Age of Automation', she researches the impact of algorithmic and AI systems on the judicial duty to state reasons and its underlying normative goals.

Victoria is an academic assistant of the KU Leuven Summer School on the Law, Ethics and Policy of AI. She is also an assistant editor of the corresponding Law, Ethics & Policy of AI Blog. Victoria is a member of both The KU Leuven Institute for Artificial Intelligence (Leuven.AI) and The KU Leuven Digital Society Institute (DigiSoc). Her research interests include AI in the judiciary, the intersection of technology and law, interdisciplinary studies, AI governance and AI ethics.