Volume **7** (1) 2025 https://doi.org/10.5204/lthj.3812

Bias in Adjudication and the Promise of AI: Challenges to Procedural Fairness

Giovana Peluso Lopes
Tilburg Institute for Law, Technology, and Society (TILT), Tilburg University, The Netherlands

Abstract

Empirical research demonstrates that judges are prone to cognitive and social biases, both of which can reduce the accuracy of judgements and introduce extra-legal influences on judicial decisions. While these findings raise the important question of how to mitigate the effects of judicial bias, they have also been used to argue in favour of incorporating artificial intelligence (AI) into adjudication, either as decision aids or, in a more extreme way, to fully automate judicial tasks. The argument goes as follows: if human judgement is susceptible to biases, and if the human psyche is also inscrutable, would it not be better to replace it with AI? After all, AI promises greater accuracy and consistency and can replace biased human decisions with objective automated ones. However, the use of AI by courts requires careful deliberation, as it potentially introduces new challenges, particularly concerning procedural fairness. This article seeks to explore how the use of AI in the administration of justice can challenge some of the foundational elements of the right to a fair trial, as enshrined in Article 6 of the European Convention on Human Rights (ECHR). This analysis is conducted through the theoretical framework of procedural justice, arguing that the use of AI for judicial decision-making can negatively impact perceptions of procedural fairness in ways that traditional human adjudication does not. It therefore seeks to debunk the narrative that, at least where bias is concerned, human and artificial decision-making are equally problematic.

***Keywords**:* Artificial intelligence; procedural fairness; AI in courts; judicial bias.

1. Introduction: From Biased Judges to Artificial Intelligence

Artificial intelligence (AI) systems are rapidly becoming an important part of judicial decision-making processes worldwide. An AI system can be defined as 'a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments'.[1] Within this broader definition, these systems are commonly classified between deterministic (rule-based) systems and probabilistic systems. The former operate using predefined rules and logic, ensuring consistent and predictable outputs for the same inputs. They excel at well-defined tasks but struggle with adaptability, requiring manual updates when conditions change. In contrast, probabilistic AI, such as large language models (LLMs), relies on statistical patterns and probability to generate responses, making them more flexible and capable of handling complex, ambiguous tasks. However, their outputs can vary, and they depend heavily on large datasets for training.[2]

Both deterministic and probabilistic AI systems have their place in the digitalisation of justice. While the use of AI by courts initially revolved around administrative tasks, including triaging, case allocation and electronic court filing, these technologies

[1] Council of Europe, "Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law." This is one amongst many definitions of AI. For a discussion, see Clarke, "The Re-conception of AI"; Russell, "Artificial Intelligence."
[2] Lehr, "Playing With the Data."

are increasingly being deployed for more substantive purposes, such as predicting litigation outcomes, automating online dispute resolution and, importantly, supporting judicial decision-making.[3] For instance, even if risk-assessment tools have long been adopted within some courtrooms to assist with several decision points throughout the criminal justice system – from pretrial release to post-conviction sentencing, probation and parole – some of them now incorporate AI in their analyses.[4] More recently, generative AI systems[5] have started to be used by some judges to facilitate different parts of the decision-making process, including summarising case law, conducting legal research and even drafting judgments.[6] A recent survey from the United Nations Educational, Scientific and Cultural Organization (UNESCO) has revealed that 44 per cent of judicial operators are utilising generative AI tools, such as ChatGPT and Google Bard, to assist their work.[7] To deal with this increasing trend, some judicial institutions have started to issue guidelines and opinions on the appropriate use of AI by the judiciary.[8]

The use of AI in courtrooms is promising for several reasons. AI can automate routine judicial tasks, reducing the effort cost required to search through vast amounts of legal documents or to seek out relevant legal provisions,[9] thus allowing judges to focus on more complex decision-making and critical case deliberations. If, for instance, 'an AI system can do preliminary work, such as providing judges with a selection of previous and similar cases, and – based on these results – suggest a decision in the case at hand, the human judge can focus on the legal reasoning and justification of the decision'.[10] Additionally, applications such as advanced case-law search engines, online dispute resolution or document categorisation and screening can possibly lower the cost of dispute resolution and help courts to address their backlog of cases, many of which are low-volume, low-value and low-complexity matters. This might help improve access to justice by effectively establishing and enforcing the right to see one's dispute resolved without prohibitive cost or inordinate delay.[11] Another well-established benefit of (probabilistic) AI systems lies in their capacity to produce predictions that are challenging for humans to make, particularly because of their limited ability to recognise patterns in complex situations. Hence, AI can 'potentially increase judicial accuracy by providing new information that cannot be detected by the naked eye or by improving the analysis process'.[12] For instance, by detecting patterns in previous cases, predictive analytics tools can contribute to the development of consistent and coherent case-law.[13]

Using AI to partially or fully automate certain judicial tasks is also appealing due to its potential to improve the objectivity and consistency of judges' decision-making. In this context, the adoption of AI for the administration of justice is often justified by putting forward a narrative in which the fallibility of human judges, who are often biased and limited in their cognitive capacities, is contrasted with AI's purported ability to enhance the accuracy, objectivity and consistency of judicial decisions. Indeed, a large body of research demonstrates how judges – like jurors and laypeople – are prone to various implicit biases, which can subtly shape their perceptions, judgements and decision-making processes throughout various stages of legal proceedings.[14] These include cognitive biases, which involve broadly erroneous forms of reasoning such as hindsight and confirmation bias, and social biases, which involve reasoning based on stereotypes, including racial and gender bias.[15] For instance, during pretrial proceedings, implicit biases have been found to influence judicial decisions regarding arraignment and setting bail amounts. During the trial, they have been shown to impact the assessment of witness credibility and the interpretation of evidence in a way that confirms pre-existing beliefs or assumptions. During sentencing, they have been demonstrated to affect guilt assessments, decisions concerning sentence severity and damage awards. While these findings raise the important question of how to effectively mitigate the effects of bias in judicial decision-making, they have also been used to argue in favour of the automation of judicial systems, either by implementing AI as decision-aids or, in its most extreme

[3] See Laptev, "Application of Artificial Intelligence in Justice"; Reiling, "Courts and Artificial Intelligence."
[4] One example of a risk assessment tool that incorporates a machine learning approach is Correctional Offender Management Profiling for Alternative Sanctions, or COMPAS, used by some US courts to assess the likelihood of recidivism: see Van Dijck, "Predicting Recidivism Risk."
[5] Generative AI is a branch of AI that leverages machine learning models, particularly deep learning techniques such as neural networks, to generate high-quality text, images, and other content based on the data on which they were trained: Martineau, "What is Generative AI?"
[6] Taylor, "Colombian Judge"; Smith, "AI in the Courtroom"; Farah, "Court of Appeal Judge."
[7] UNESCO, "UNESCO Survey Uncovers Critical Gaps."
[8] For example, UK Courts and Tribunals Judiciary, "Artificial Intelligence (AI)"; European Commission for the Efficiency of Justice, "Use of Generative Artificial Intelligence (AI)"; UNESCO "Draft UNESCO Guidelines"; Courts of New Zealand, "Guidelines for Use of Generative Artificial Intelligence"; Canadian Judicial Council, "Guidelines for the Use of Artificial Intelligence"; Consultative Council of European Judges, "Opinion No. 26."
[9] Barysė, "Algorithms in the Court."
[10] Ulenaers, "The Impact of Artificial Intelligence," 11.
[11] Steponenaite, "Judicial Analytics on Trial."
[12] Barysė, "Algorithms in the Court," 118.
[13] Ulenaers, "The Impact of Artificial Intelligence," 6–7.
[14] For an overview, see Wistrich and Rachlinski, "Implicit Bias."
[15] Zenker, "De-Biasing Legal Fact Finders."

version, as replacements for judges.[16] According to this line of reasoning, if human judgement is *already* susceptible to biases, and the reasons leading judges to reach a certain decision are already inscrutable, it would be better to replace it with AI.[17] After all, 'one of the benefits of automated decision-making is that it can reduce the arbitrariness of human decisions', since 'whereas many evaluative decisions made by humans are based on unconscious group biases and intuitive reactions, algorithms follow the parameters set out for them'.[18]

Yet the use of AI by courts requires careful deliberation since it potentially introduces several new risks. In practice, these systems are filled with errors and biases, leading to decisions that are based on incorrect information and that aggravate inequalities. These are not only hard to detect, since they may arise from unknown associations done by (often opaque) systems, but there is also a risk-magnifying potential associated with AI that is not present with human decision-making.[19] In this article, I explore how these and other risks can challenge the right to a fair trial, as established, on a European level, in Article 6 of the European Convention on Human Rights (ECHR). There are many requirements for the fulfilment of the right to a fair trial, such as a reasonable length of procedures, public hearings, legal certainty and the presumption of innocence in criminal matters. Some of these have been explored elsewhere in the context of automating judicial decision-making.[20] Here, my analysis is limited to the requirements of judicial independence and impartiality, adversarial proceedings and equality of arms, as well as the duty to motivate judgments. These are explored through the theoretical framework of procedural justice, in relation to the main factors taken into consideration by litigants to evaluate the fairness of a trial. I argue that the use of AI for judicial decision-making can negatively impact perceptions of procedural fairness in ways that traditional human adjudication does not. After briefly discussing the advantages justifying the adoption of AI in courts, as well as the problematic narrative of biased judges being pitted against objective and consistent AI systems, I proceed with presenting results from social psychology research on procedural justice highlighting the importance of fair procedures for the acceptance and legitimacy of legal institutions, as well as the criteria that people use to evaluate the legal procedures they experience. Subsequently, I explore how these expectations are connected to the foundational elements of the right to a fair trial and how the use of AI for the administration of justice potentially challenges them. Methodologically, this research is essentially theoretical and bibliographic, drawing on direct and indirect sources for a comprehensive analysis of the theme. By integrating insights from social psychology with jurisprudence, this article offers an interdisciplinary analysis of the often-overlooked implications of AI for procedural justice and some of the foundational elements of the right to a fair trial.

2. Procedural Justice and the Right to a Fair Trial

When AI systems are deployed to resolve a dispute, as an instrument to assist in judicial decision-making, or to give guidance to the public, 'it is essential to ensure that they do not undermine the guarantees of the … right to a fair trial'.[21] The right to a fair trial is at the core of human rights protection, being one of the most fundamental principles of any democratic society under the rule of law[22]. The right is enshrined in the International Covenant on Civil and Political Rights (ICCPR)[23] and regional human rights instruments such as the American Convention on Human Rights (ACHR)[24] and the European Convention on Human Rights (ECHR),[25] along with numerous national and regional constitutions and bills of rights.[26] The right to a fair trial as prescribed in international human rights conventions does not favour any particular judicial system over another, but regardless of the system adopted (e.g. inquisitorial or adversarial, collegial or single-judge), it requires that certain foundational elements are met.[27]

As explained in the introduction, the focus of this article will be the European Convention on Human Rights, which establishes in Article 6 the right to a fair trial. The first paragraph states that 'in the determination of his civil rights and obligations or of any criminal charge against him, everyone is entitled to a fair and public hearing within a reasonable time by an independent

[16] See, for example, Doleac, "Let Computers Be the Judge."
[17] See Chatziathanasiou, "Beware the Lure of Narratives."
[18] Chesterman, "Through a Glass, Darkly."
[19] Dietterich, "Robust Artificial Intelligence."
[20] See, for example, Ulenaers, "The Impact of Artificial Intelligence"; Steponenaite, "Judicial Analytics on Trial"; Terzidou, "The Use of Artificial Intelligence."
[21] European Commission for the Efficiency of Justice (CEPEJ), "European Ethical Charter."
[22] *Pretto and Others v Italy*. App. No. 7984/7 (ECtHR, 8 December 1983).
[23] United Nations, "International Covenant on Civil and Political Rights," art. 14.
[24] Organization of American States, "American Convention on Human Rights," art. 8.
[25] Council of Europe, "Convention for the Protection of Human Rights and Fundamental Freedoms," art. 6.
[26] For instance, European Union, "Charter of Fundamental Rights of the European Union," art. 47; United States, "Constitution of the United States," Sixth Amendment; Germany, "Basic Law for the Federal Republic of Germany," art. 103.
[27] Molbæk-Steensig, "Artificial Intelligence and Fair Trial Rights."

and impartial tribunal established by law'.[28] Key elements of the right to a fair trial can already be identified in the wording of Article 6(1), including judicial independence and impartiality, a reasonable duration of proceedings and the openness of court hearings. Additionally, while not explicitly stated, some other elements can nevertheless be derived from the article, such as the right to access the courts, equality of arms and the right to receive a reasoned judgment. These will be further elaborated on throughout the next sections of this article.

While Article 6(1) of the Convention applies to administrative, civil and criminal processes, the remaining two paragraphs apply only to criminal proceedings. They establish that defendants must be presumed innocent until proven guilty (Article 6(2)). They must also be promptly informed, in a language they understand, of the charges against them, as well as have sufficient time and resources to prepare a defence, including access to free legal aid if they cannot afford it (Article 6(3)). Additionally, they have the right to present favourable witnesses and receive interpretation services if needed.

The right to a fair trial is procedural in nature, and its basic guarantees exist to secure procedural justice. Procedural justice, also referred to as 'procedural fairness' or 'procedural due process', refers to the idea of fair processes and how people's perception of fairness is significantly influenced not only by the end result of their experiences, but also by their quality. Within the field of legal philosophy, there is an extensive debate on whether procedural guarantees serve only to ensure that legal rules governing matters of substance are accurately applied to the appropriate cases, or whether procedural propriety has any intrinsic value.[29] Here, I will focus on findings from social psychology research that point towards the latter – that is, demonstrate that people's acceptance of the decisions of legal authorities, as well as their approval of legal rules and, more generally, their endorsement of the legal system are most impacted by how fair they perceive procedures to be, rather than whether their outcomes were favourable.

Pioneers in the field, Tom R. Tyler and E. Allan Lind, explain how, when evaluating an experience with a legal authority, there are three aspects that people might take into consideration: the first is outcome favourability, meaning whether they benefit or lose from the experience; the second is outcome fairness, meaning whether what they receive seems fair to them, thus having to do with substantive justice; and the final aspect is procedural fairness, meaning whether the decisions were made in ways that they regard as fair.[30] The literature on the influence of these three aspects indicates that judgements about legal authorities, legal decisions and legal rules are most impacted by the procedural justice of experiences.[31] Naturally, this does not imply that opinions based only on self-interest are meaningless. When assessing levels of pleasure or discontent, the outcome's favourability plays a significant role because, as expected, people are typically less satisfied when they lose than when they win. However, research indicates that as long as they consider the processes involved fair, they do not carry over those unfavourable impressions to their views about the law, the judge or the legal system. Furthermore, regardless of the favourability of the outcome, research also indicates that people more readily accept and obey decisions resulting from a procedure they believe to be fair.[32]

These counter-intuitive findings were initially made prominent through the research of psychologist John Thibaut and law professor Laurens Walker.[33] They demonstrated in laboratory studies that procedural differences (between the adversarial and inquisitorial models of process) led to different evaluations of fairness, irrespective of the outcome of the trial. Participants considered outcomes arising from an adversarial procedure to be fairer than those emanating from an inquisitorial procedure, mainly due to differences in control over the presentation of evidence and arguments, which led authorities to pay more attention to disputants' 'voice'. Subsequent studies by these researchers later confirmed that the degree to which a procedure ensured the disputant's voice in hearings and trials did, in fact, have a significant impact on the perceived fairness of the procedure, with high levels of disputant voice translating into a greater sense of perceived fairness. Subsequent field studies focusing on real-world trials and hearings supported their findings, showing that when people believe court procedures to be fair, they are more satisfied, more likely to voluntarily accept and obey legal decisions, and more likely to evaluate legal authorities as legitimate.[34]

Procedural justice researchers have devoted considerable attention not only to demonstrating the value of procedures for evaluations of fairness, but also to identifying the criteria that people use to evaluate the legal procedures they experience. According to Professor Tom R. Tyler, procedural fairness is evaluated on the basis of four basic expectations.[35] These constitute

[28] Council of Europe, "European Convention on Human Rights," Article 6, paragraph 1.
[29] See Meyerson, "Procedural Justice and the Law."
[30] Tyler, "Procedural Justice."
[31] Tyler, "A Relational Model of Authority in Groups."
[32] Tyler, "Procedural Justice."
[33] Thibaut, "Procedural Justice."
[34] For an overview, see Tyler, "Procedural Justice."
[35] Tyler, "Why People Obey the Law."

a group-value theory of procedural justice, which emphasises the symbolic and psychological implications of procedures for feelings of inclusion in society and for the belief that the institution using the procedure holds the person in high regard.[36] The first expectation, which has already been mentioned, is process control or 'voice' – the ability to present evidence and arguments, to participate in the case by expressing their viewpoint, even when that is unlikely to influence its outcome. The second is 'trust', which consists in feeling that authorities are trustworthy and benevolently disposed towards the person making the justice judgement. In this sense, 'people are affected by the degree to which they feel that the authorities with whom they deal are motivated to try to be fair to them and to others in the group'.[37] This perception is essential to informing how people feel about authorities, because it not only reflects the character of the individual authority figure with whom they are dealing, but is also the basis for predicting their future behaviour. The third expectation is 'standing', which encompasses feelings that one is viewed by authorities as a full-fledged member of society, thus being treated politely and with dignity, and having their rights and opinions respected. Finally, there is 'neutrality', which is the belief that one is accorded even-handed, non-discriminatory treatment:

> Neutrality reflects the degree to which people feel that authorities are creating a 'level playing field' by engaging in even-handed treatment of all. Neutrality involves honesty, unbiased treatment, consistency, and factual decision-making. Prejudice, the idea of discrimination based on group membership, is perhaps the strongest evidence of a lack of neutrality, since people are not given an equal opportunity to have access to social resources.[38]

Additionally, according to the group-value theory of procedural justice, because satisfactory interpersonal treatment by group authorities symbolically communicates that we possess value or status in the eyes of our community, people generally associate these interpersonal aspects of procedures with a fair treatment. This, in turn, supports a sense of self-worth or self-respect. On the other hand, when interpersonal aspects of procedures are not up to par, people feel as though they have been deprived of procedural justice since it undermines their self-esteem and puts doubt on their status in the group.[39] In sum, the picture that emerges from the empirical findings of social psychology research is that, in their dealings with authorities, four factors are particularly important to people: that authorities 'treat citizens in a fair and respectful way, make neutral and unbiased decisions, display trustworthy motives, and allow the citizen a voice in their interactions'.[40]

In the next section, I will explore how these basic expectations[41] are connected to some of the foundational elements of the right to a fair trial, and how the use of AI for the administration of justice potentially challenges them. Before proceeding, one important caveat must be considered: that whether a particular AI system should be used in the context of judging is dependent on several factors such as the purpose and context of its deployment, the approach or methodology (e.g. deterministic or probabilistic), the system's performance, and so forth. Other factors to be taken into consideration revolve around judicial values, such as open justice, accountability, impartiality and equality before the law, procedural fairness, access to justice and efficiency.[42] This analysis is highly context-dependent, making it impossible to provide a singular answer to the general issue of whether AI should be used by courts. Nevertheless, some of these questions will be explored in the following sections.

3. Neutrality: Judicial Independence and Impartiality

Article 6 ECHR enshrines the right to a fair trial *by an independent and impartial* tribunal. Judicial independence and impartiality are essential elements of the right to a fair trial, being two different but closely related and complementary components. This close relation between the guarantees of an 'independent' and an 'impartial' tribunal often leads the ECtHR to consider the two requirements together in the cases it analyses.[43] An independent judicial system is one of the pillars of the rule of law, with independent judges being considered a fundamental instrument to establish and implement a system of impartial and fair rules. The requirements of judicial independence and impartiality are connected to litigants' expectations of neutrality in judicial proceedings. Neutrality, we have seen, concerns individuals' perceptions of whether authorities act impartially and treat them fairly, without bias stemming from personal motives or group characteristics such as gender, race, ethnicity or religion. For example, police authorities demonstrate neutrality by treating all individuals and groups in the community equally, avoiding actions based on biases or preconceived notions when making decisions. Over-policing or under-

[36] Tyler, "Procedural Justice."
[37] Tyler, "Procedural Justice," 76.
[38] Tyler, "Procedural Justice," 75–76.
[39] Blader, "A Four-Component Model."
[40] Jackson, "Norms, Normativity."
[41] Tyler, "Why People Obey the Law."
[42] Moses, "Artificial Intelligence."
[43] *Kleyn and Others v the Netherlands* App nos. 39343/98, 39651/98, 43147/98, and 46664/99 (ECtHR, 6 May 2003), §192; *Findlay v the United Kingdom* App no 22107/93 (ECtHR, 25 February 1997).

policing specific communities can be seen as a failure to provide neutral treatment across society.[44] Likewise, judicial authorities might be seen as failing to provide neutral treatment between litigants due to conflicts of interest, having a personal connection to one of the parties or suffering undue influences from other branches of power or third parties. Next, I will discuss how the use of AI systems in the courtroom might rightfully lead the parties and the public to perceive the justice system as lacking neutrality, particularly when these systems are biased, when judges tend to over rely on their outputs and when they represent the interests of private for-profit companies.

Article 6 of the ECHR requires judicial independence not only from the other branches of power – that is, the executive and the legislature – but also from the parties.[45] To perform their functions effectively, and therefore be considered legitimate by the parties to the case, the judge is expected to adjudicate independently, according to the rules of the legal system, without expectation of benefits or fear of reprisals.[46] This necessary independence may refer to the relations between the judiciary and the other branches of government (external independence) or focus on guarantees aimed at protecting individual judges from undue pressures from higher-ranking judges (internal independence).[47] When it comes to the criteria for assessing judicial independence, the ECtHR has considered the manner of appointment of the members of a court and the duration of their term of office, the existence of guarantees against outside pressures, and whether the court of tribunal presents an appearance of independence.[48] In sum, independence is mainly associated with the existence of institutional guarantees and safeguards that allows judges to free themselves from external and internal pressures when deciding, including the neutrality of the appointment procedure, the stability of the position, autonomy from other branches of government, a reasonable sphere of immunity and the inviolability of their salary, among others.[49]

Although judicial external independence usually refers to a lack of subordination to any organ of the state, it can also be interpreted as requiring no undue influence from non-state actors on judicial autonomy. In this context, it is possible that judicial independence is threatened by the use of AI in courts, since various third parties, including developers, owners or shareholders, may be interested in certain case outcomes and thereby act to ensure that the system is in line with their interests.[50] Outsourcing can compromise judicial independence, as companies driven by profit aim to provide efficient algorithmic systems tailored to meet their clients' needs, potentially neglecting values such as fairness. In the justice sector, 'technology companies are doing more than "just" building the digital infrastructure of the courts of the future; they are also importing their own values in the process'.[51] For example, the developer of a precedent analysis tool might fail to include into its dataset an important line of case law with outcomes that are not in their favour, while the judge relying on the system is unable to identify this omission.[52] Such a situation in which a judge would be completely unaware about how the system operates and the interests behind it could arguably result in a violation of the independence requirement.[53]

Additionally, algorithm developers take on a governmental role by making critical decisions about the technical parameters of AI tools used in judicial administration, particularly when these systems influence decision-making processes that should be the exclusive domain of the judiciary. As legal technology becomes more advanced, judges using that technology grow increasingly reliant on the expertise of the technology developers.[54] In a comprehensive analysis of algorithmic decision-making tools for criminal justice authorities, Yeung and Harkens explore how seemingly 'technical' choices made by developers have serious legal and constitutional implications.[55] They explain how ML-based algorithmic tools are conventionally built and implemented without reference to their larger context of application, with technical developers making abstraction decisions that intentionally ignore context-specific features of the real-world domain in which the tool will operate. As a consequence, public law principles and the legal duties to which they give rise are often regarded as irrelevant 'noise'. This results in the circumvention of vital institutional safeguards against the arbitrary or otherwise unjust exercise of power by public authorities.

---

[44] Tyler, "Profiling and Police Legitimacy."
[45] *Ninn-Hansen v Denmark* App no 28972/95 (ECtHR, 18 May 1999); *Beaumartin v France* App no 15287/89 (ECtHR, 24 November 1994), §38; *Sramek v Austria* App no 8790/79 (ECtHR, 22 October 1984), §42.
[46] Guarnieri and Piana, "Judicial Independence."
[47] *Agrokompleks v Ukraine* App no 23465/03 (ECtHR, 6 October 2011), §137; *Parlov-Tkalčić v Croatia* App no 24810/06 (ECtHR, 22 December 2009), §86.
[48] ECtHR, "*Langborger v Sweden*," § 32; ECtHR, "*Kleyn and Others v the Netherlands*," § 190.
[49] Papayannis, "Independence, Impartiality and Neutrality in Legal Adjudication."
[50] Steponenaite and Valcke, "Judicial Analytics on Trial."
[51] Helberger, "The Rise of Technology Courts."
[52] Smuha and Hendrickx, "AI and the Administration of Justice."
[53] Steponenaite and Valcke, "Judicial Analytics on Trial."
[54] Helberger, "The Rise of Technology Courts."
[55] Yeung and Harkens, "How Do 'Technical' Design Choices Part I"; Yeung and Harkens, "How Do 'Technical' Design Choices Part II."

Judicial internal independence might be threatened if judges are indirectly pressured into using certain AI tools in their everyday decision-making processes or if they are subtly coerced into relying on specific AI systems for routine judicial tasks.[56] The use of the risk assessment system COMPAS in the United States is a paradigmatic example, 'given the pressure within the judicial system to use these assessments'.[57] Indeed, some US states already require the use of risk-assessment tools in sentencing proceedings, which is also encouraged by advocates in academia and the judiciary. This 'widespread endorsement' of risk-assessment systems not only threatens judicial independence, but also 'communicates to judges that the tools are, in fact, reliable'.[58]

While judicial independence is concerned mainly with the autonomy of the judiciary, impartiality is usually associated with objective decision-making or the absence of prejudice or bias towards one of the parties.[59] It is commonly stated that independence is a necessary condition for impartiality, but not a sufficient one.[60] At a first glance, impartiality requires that the judge is an independent third party who is not involved in the process and who occupies a transcendent position with respect to the disputing parties. Additionally, it requires that judges refrain from taking sides in their decisions, maintaining objectivity towards the interests of the parties. This implies an attitude of neutrality, meaning the judge does not have a relationship with any of the parties and does not perform acts that may indicate prior positioning regarding the matter to be decided. When it comes to the criteria for assessing judicial impartiality, the ECtHR employs a subjective test, where regard must be had to the personal conviction and behaviour of a particular judge (i.e. whether the judge held any personal prejudice or bias in a given case), and also an objective test, where it ascertains whether the tribunal itself and, among other aspects, its composition, offered sufficient guarantees to exclude any legitimate doubt in respect of its impartiality.[61] The ECtHR has recognised the hardship of establishing a breach of Article 6 on account of subjective partiality, given the difficulty of procuring evidence with which to rebut the presumption of impartiality,[62] and has therefore emphasised how the requirement of objective impartiality provides a further important guarantee.[63]

The importance of impartiality in the administration of justice has led to the development of rules and procedures aimed at safeguarding legal institutions and practices from bias, including potential perceptions of it. Indeed, these institutional requirements are a formal component of impartiality – that is, a set of restrictions regarding the actions of the judge, which tend to reduce the effects of bias, thus setting out how they should proceed with respect to the parties involved: when they ought to be heard, how the evidence is to be presented and how the judge is to justify the final decision after considering the arguments and evidence produced.

An obvious challenge that the adoption of AI systems to perform judicial tasks poses to impartiality relates to the fact that these systems are often biased, particularly if the datasets on which they are trained are not representative of the population, and can lead to discriminatory outcomes based on, for example, race or ethnicity.[64] In this context, Yeung and Harkens explain that datasets inevitably reflect underlying biases in the historic social practices to which the data pertains so that, if used to generate prediction models, the resulting outputs will reflect and reinforce these biases. Historically marginalised groups are thus subjected to a higher risk of unjust discrimination relative to individuals from majority groups because the resulting outputs systematically discriminate unjustly (rather than being arbitrary, they are systematically patterned for reasons we can point to) and may violate the right to be free from unjust discrimination in the determination of opportunities and burdens.[65]

Although this is a well-known problem, the narrative being discussed in this article employs as a counter-argument the fact that – as highlighted in the introduction – judges are *also* biased. It argues that, even if there is a risk of bias in employing AI for judicial decision-making, this cannot possibly be worse than the cognitive and social biases that already affect judges.[66] This argument is problematic for several reasons. First, it fails to acknowledge that human decision-making has a self-regulative feature that AI systems lack, meaning that when individuals justify their decisions by providing reasons, they often engage in

[56] European Commission for the Efficiency of Justice (CEPEJ), "European Ethical Charter."
[57] Harvard Law Review, "*State v Loomis*," 1535.
[58] Harvard Law Review, "*State v Loomis*," 1536.
[59] *Wettstein v Switzerland* App no 33958/96 (ECtHR, 21 December 2000), §43; *Micallef v Malta* App no 17056/06 (ECtHR, 15 October 2009), § 93.
[60] Ibánez, "Imparcialidad Judicial e Independencia Judicial."
[61] *Micallef v Malta* App no 17056/06 (ECtHR, 15 October 2009).
[62] In applying the subjective test, the court has consistently held that 'the personal impartiality of a judge must be presumed until there is proof to the contrary' (*Le Compte, Van Leuven and De Meyere v Belgium* App nos 6878/75, 7238/75 (ECtHR, 23 June 1981), § 58).
[63] *Micallef v Malta* App no 17056/06 (ECtHR, 15 October 2009), §§ 95 and 101.
[64] Malek, "Criminal Courts' Artificial Intelligence."
[65] Yeung, "How Do 'Technical' Design Choices Part II," 13.
[66] Chatziathanasiou, "Beware the Lure of Narratives."

a self-regulatory process, adjusting their thoughts and actions to align with their stated justifications.[67] Second, it overlooks the enhanced power of these technologies, the instant reproducibility, transfer and storage of digital data and their capacity to operate automatically at scale, yet in a highly opaque manner.[68] This risk-magnifying potential associated with AI[69] is exemplified by the Horizon IT scandal, considered one of the United Kingdom's worst miscarriages of justice, where the Post Office wrongly prosecuted over 900 sub-postmasters between 1999 and 2015 based on faulty data from the Horizon accounting system, leading to wrongful convictions, financial ruin and severe personal suffering.[70] Finally, even when oversight mechanisms are in place to ensure the proper use of AI to enhance the abilities of human decision-makers, this argument ignores the susceptibility of the latter to automation bias – a topic that will be addressed in the following section.

3.1. Automation Bias

To understand how the use of certain decision-support systems in judicial settings can negatively impact judicial neutrality, it is helpful to revisit the key points of the 2016 case *State v Loomis*.[71] Eric Loomis was sentenced to six years in prison after being assessed as high risk by the COMPAS software, following a guilty plea in a drive-by shooting case. Loomis argued, among other things, that COMPAS violated his due process rights, claiming it relied on group data, not his individual circumstances, thus undermining his right to an individualised sentence. However, the Wisconsin Supreme Court rejected his arguments. The court stated that the risk score was only one of several factors considered in sentencing, not the sole or primary factor. The court indicated that a fair trial argument might have been valid if the risk score had been the decisive or only factor considered by the judge. The clear issue is that it is impossible to ascertain how much a judge actually relied on an automated risk score when determining a defendant's sentence – or, more generally, how much any recommendation made by an AI system has weighed on a judge's final decision. In reality, it is unrealistic to expect a judge, after reading the risk scores attributed to the defendant, 'to exercise discretion without any predetermined views of, or even bias against, the defendant'.[72]

Other examples of algorithmic risk assessment tools used in criminal justice settings include the London Gangs Matrix, Durham Constabulary's Harm Assessment Risk Tool (HART) and the Dutch System Risk Indication (SyRI). The London Gangs Matrix, developed by the Metropolitan Police Service, classifies individuals based on perceived gang involvement, using a secretive algorithmic harm score that has been criticised for racial bias and unlawful data practices. HART, previously used by Durham Constabulary, employed machine learning to predict an individual's likelihood of reoffending, categorising arrestees as high, moderate or low risk, but was found to mislabel individuals and reinforce biases. SyRI, a Dutch government tool designed to detect social welfare fraud, combined extensive administrative data to flag individuals for investigation but was ultimately ruled unlawful for violating privacy rights and disproportionately targeting marginalised communities.[73]

The difficulty faced by decision-makers in ignoring some of these automated systems' recommendation can be attested by a large body of cognitive psychology research showing how difficult it is for humans to override their initial heuristic response in favour of alternative solutions, particularly when it confirms one's preconceptions, views and expectations. To do so, the agent must be able to use situational cues to detect the need to override the heuristic response and sustain the inhibition of the heuristic response while analysing alternative solutions. These alternative solutions must, of course, have been learned and previously stored in memory. In sum, it requires having the appropriate knowledge of solutions and strategic rules to substitute for a heuristic response, as well as having the cognitive disposition necessary to override heuristic processing.[74]

This problem is further aggravated by the existence of automation bias, consisting in the human 'tendency to use automated cues as a heuristic replacement for vigilant information seeking and processing'.[75] It describes the phenomenon whereby 'judges accept the guidance or recommendations of an automated decision-making system, and cease searching for confirmatory evidence, perhaps even transferring responsibility for decision-making onto the machine'.[76] Automation bias can be the source of omission errors, which take place 'when decision makers fail to notice problems because an automated aid fails to detect them', or of commission errors, which occur 'when people inappropriately follow an automated decision aid directive or

[67] Peters, "Explainable AI Lacks Regulative Reasons."
[68] Yeung, "How Do 'Technical' Design Choices Part II."
[69] Dietterich, "Robust Artificial Intelligence and Robust Human Organizations."
[70] Sweney, "What is the Post Office Horizon IT Scandal All About?"
[71] Wisconsen Supreme Court, "*State v Loomis*."
[72] Liu, "Beyond *State v Loomis*."
[73] For a detailed analysis of the three risk assessment tools, see Yeung, "How Do 'Technical' Design Choices Part II."
[74] Stanovich, "On the Relative Independence."
[75] Mosier, "Human Decision Makers."
[76] Gravett, "Judicial Decision-Making."

announcement'.[77] Concerning the former, it would be the case where an AI system fails to take account of a relevant factor when generating a risk score or fails to incorporate an important precedent in its analysis of the relevant case law – or even 'hallucinates' a false case[78] – and the judge is unable to identify these flaws. With regard to the latter, it would be the case where the system provides accurate information and judges solely or mostly rely on it, overlooking other relevant sources when they should take those into account.

The presence and extent of automation bias are influenced by several factors, referred to as effect mediators.[79] One is trust: when people are informed that an automated system is designed to perform a particular task, by default they will trust it to do so correctly.[80] Depending on their interactions with the system, this initial position may change. A different factor is cognitive load: the higher it is, whether due to the existence of multiple tasks simultaneously competing for the individual's attention[81] or due to the complexity of a single task,[82] the more likely it is that automation bias will arise. This over-reliance on the automated system in high cognitive load scenarios happens even if the trust that the user has regarding it is low, thereby suggesting that the influence of cognitive load is independent of the influence of trust.[83] This means that, even if judges do not consider the automated system to be fully reliable, they may still rely excessively on it due to high cognitive load – which judges arguably experience on a daily basis, often regarding themselves as being overburdened and feeling that they are often pressed to decide issues faster than they may have wished.[84]

Finally, another relevant effect mediator is the influence of pre-existing stereotypes, with automation bias relating to it in the following way: 'where algorithmic advice departs from pre-existing stereotypes, decision-makers may be more likely to override it; but where algorithmic advice conforms to pre-existing stereotypes, they may be more likely to conform to it'.[85] For instance, in the context of bail decisions, the degree to which judges relied on algorithmic risk assessments varied according to racial bias,[86] with the algorithm's suggestion being more likely to be disregarded in favour of stricter bond conditions for Black defendants compared with their white counterparts. This is particularly significant because it debunks the narrative presented earlier according to which, if human judgment is already susceptible to biases, using algorithmic and AI systems – even if they are somewhat flawed – would *still* lead to more objective and consistent results, thus helping address the problem of bias in judicial-decision making. If, however, pre-existing attitudes and stereotypes affect automation bias, the use of artificial intelligence for decision support might aggravate the problem of bias in judicial settings, rather than ameliorate it.

3.2. Lack of Accountability

A related issue is that automated decision-making systems may potentially 'provide cover for human agents'.[87] In this context, judges may opt for a cautious strategy – for instance, imposing stricter penalties on defendants identified as 'high risk' by the software to avoid the personal and societal risks associated with a repeat offender committing another crime. This is something that already takes place with actuarial risk tools, most of which do not yet incorporate AI technologies. Drawing on data collected from interviews with over a hundred legal practitioners in Canada, including probation/parole officers, prosecutors, defence lawyers and judges, researcher Kelly Hannah-Moffat found that:

> Practitioners routinely admitted that risk assessment tools were not 'reliable' predictors, but that they preferred them because using them minimized their own risk of being blamed for subsequent consequences, and that of the prosecutor's office or correctional authorities. The use of these tools suggests the curtailment of discretion and subjective judgments, which are often viewed negatively and associated with errors.[88]

We have seen that automation bias results from people's tendency to ascribe greater trust in the analytical capabilities of an automated system than their own.[89] And because algorithms are deployed with the express purpose of reducing human bias and error, they are further seen as authoritative sources, with more knowledge than the humans who interpret them. As a result,

[77] Mosier, "Human Decision Makers."
[78] Bohannon, "Lawyer Used ChatGPT."
[79] Goddard, "Automation Bias."
[80] Dzindolet, "The Role of Trust."
[81] Parasuraman, "Complacency and Bias."
[82] Lyell, "Automation Bias."
[83] Kazim, "Automation Bias."
[84] Zenker, "Debiasing and Rule of Law."
[85] Kazim, "Automation Bias," 13.
[86] Cowgill, "The Impact of Algorithms."
[87] Chesterman, "Through a Glass, Darkly," 279.
[88] Hannah-Moffat, "The Uncertainties of Risk Assessment," 244.
[89] Chesterman, "Through a Glass, Darkly."

human users such as judges often follow the algorithm's decisions, even though doing so could be detrimental to others.[90] This is unsurprising considering how people tend to weigh expert empirical assessments more heavily than non-empirical evidence – for instance, in the judicial context, judges often give forensic-based evidence heavier weight than other factors, notwithstanding their lack of understanding about some of the methods used.[91]

Returning to algorithmic decision-making, as Theo Araújo and colleagues explain, attitudes and perceptions about it are not only influenced by the technological solution they offer, or their actual performance, but also the way in which these processes are framed or communicated to the user or subject of the decision.[92] The authors highlight how these systems are often carefully and strategically articulated as impartial and objective socio-technical actors in the discourse surrounding their implementation and usage in different scenarios, leading to an assumption of neutrality and objectivity. Yet they are 'created for purposes that are often far from neutral: to create value and capital; to nudge behaviour and structure preferences in a certain way; and to identify, sort and classify people'.[93] Unlike state actors, who contract them and bear direct responsibility for the impacts of the system, however, private companies are not perceived by the public as directly accountable for errors of AI systems that may lead to – for example, the production of discriminatory outcomes.[94]

This is an example of what Sharon and Gellert describe as 'sphere transgressions', as technology companies take advantage of their superior technological capabilities to influence the sphere of justice – an illegitimate influence 'insofar as tech corporations do not have the domain expertise proportional to their new level of influence in these different societal spheres, and insofar as they are not accountable in the way that public sector actors are'.[95] Transposing this analysis to the justice sector and the growing importance of technology companies over it, Helberger further elaborates that, when administering justice, judges and courts function within a framework of checks and balances, procedural safeguards and obligations to uphold the rule of law and human rights. In contrast, these safeguards do not apply to technology companies, which, as private entities, are primarily accountable to their shareholders and CEOs.[96]

4. Process Control: Adversarial Proceedings and Equality of Arms

Another factor directly influencing people's perceptions of the fairness of proceedings is process control or voice, which refers to the ability to present evidence and arguments, and to participate in the case by expressing their viewpoint, even when that is unlikely to influence the outcome. The fact that the algorithmic and AI systems used by courts are often commercial and proprietary products makes it difficult for the individuals affected by their outputs to understand and contest them, therefore presenting a challenge to the adversarial principle and to the principle of equality of arms.

The right to a fair trial comprises the fundamental right to adversarial proceedings, the requirements of which are in principle the same in both civil and criminal cases.[97] The right to adversarial proceedings generally refers to the opportunity for the parties in a criminal or civil trial to be aware of and comment on all evidence presented or statements filed, even those from an independent member of the national legal service, with the aim of influencing the court's decision.[98] The need to save time and speed up proceedings cannot justify overlooking such a fundamental principle as the right to adversarial proceedings.[99] Moreover, parties must be able to exercise this right in satisfactory conditions – that is, by having the possibility of familiarising themselves with the evidence before the court, as well as the possibility of commenting on its existence, contents and authenticity in an appropriate form and within an appropriate time.[100]

[90] Gravett, "Judicial Decision-Making."
[91] Spellman, "Reasoning about Forensic Science Evidence."
[92] Araújo, "In AI We Trust?"
[93] Kitchin, "Thinking Critically."
[94] Terzidou, "The Use of Artificial Intelligence."
[95] Sharon, "Regulating Big Tech Expansionism?" 2653.
[96] Helberger, "The Rise of Technology Courts."
[97] *Werner v Austria* App no 138/1996/757/956 (ECtHR, 24 November 1997), § 66.
[98] *Ruiz-Mateos v Spain* App no 12952/87 (ECtHR, 23 June 1993), § 63; *McMichael v the United Kingdom* App no 16424/90 (ECtHR, 24 February 1995), § 80; *Vermeulen v Belgium* App no 19075/91 (ECtHR, 20 February 1996), §33; *Lobo Machado v Portugal* App no 15764/89 (ECtHR, 20 February 1996), §31; *Kress v France* App no 39594/98 (ECtHR, 7 June 2001), §7 4.
[99] *Nideröst-Huber v Switzerland* App no 18990/91 (ECtHR, 18 February 1997), § 30.
[100] *Krčmář and Others v the Czech Republic* App no 35376/97 (ECtHR, 3 March 2000), § 42; *Immeubles Groupe Kosser v France* App no 38748/97 (ECtHR, 21 March 2002), §26.

Importantly, the actual effect on the court's decision is of little consequence for the assessment of this right's fulfilment.[101] This is in line with the basic expectation of process control, which we have seen refers to the idea that individuals feel more satisfied and perceive a process to be fair when they are given an opportunity to express their views, present evidence or have a say in the proceedings. The feeling of being able to participate, or 'having a voice', is crucial because it conveys respect and allows individuals to feel that their perspectives have been considered in the decision-making process. Even when the ability to present evidence and arguments, or to express a viewpoint, is unlikely to influence the outcome, the mere opportunity to participate can enhance individuals' sense of fairness and legitimacy in the process. This participation reinforces their sense of being treated with dignity and empowers them to accept the outcome, regardless of its favourability. Indeed, according to the ECtHR, litigants' confidence in the workings of justice is based on, *inter alia*, the knowledge that they have had the opportunity to express their views.[102]

Returning to the case of *State v Loomis*, the defendant contended that he could not evaluate the scientific validity of his score because COMPAS is a proprietary tool and trade secret, preventing the disclosure of how the risk scores are calculated or how the factors are weighted. The Wisconsin Supreme Court, however, rejected Mr Loomis's argument, claiming that, since the 'risk assessment is based upon his answers to questions and publicly available data about his criminal history',[103] the defendant had the opportunity to verify the accuracy of the questions and answers listed on the report. However, even though the data are publicly available, the defendant cannot verify how the system has processed and weighted the data to reach its conclusion. Again, this issue is particularly significant when a private company develops the system and claims intellectual property rights, thus refusing to reveal the workings of the system. This calls into question the use of risk-assessment tools and AI systems in general, which are created by for-profit and private firms, in the exercise of public power. It also highlights the importance of the parties having access to information regarding algorithms and the way they produce their decisions in order to have meaningful control over the process, or to exercise their voice.

The Court of Justice of the European Union (CJEU) recently issued a ruling[104] that clarifies essential aspects of transparency in automated decision-making under the General Data Protection Regulation (GDPR). The case in question involves Yettel Bulgaria, a mobile telecommunications provider, and a consumer who disputed both the calculation of their telephone bill and the transparency of Yettel's automated billing system. Yettel utilises automated systems and algorithms that collect data from users' mobile devices to determine billing charges based on time and megabyte consumption. However, this process is highly opaque, making it nearly impossible – even with expert analysis – to determine how decisions regarding charge calculations are made. The lack of transparency led the consumer to stop paying their phone bills, arguing that they were unable to verify how the charges had been determined.

In its ruling, the CJEU held that companies employing ADM are not required to disclose their algorithms to data subjects. However, they may need to explain how changes in input data would affect the billing outcome. Transparency obligations must comply with Article 12(1) of the GDPR, ensuring that the information provided is concise, clear, accessible and easy to understand. Regarding the role of supervisory authorities, the judgment emphasizes the need to balance data-protection rights with other fundamental rights, such as the freedom to conduct a business. While trade secrets and intellectual property – such as copyright-protected software – must be respected, certain information nevertheless needs to be disclosed to supervisory authorities or courts if withholding it would infringe upon the rights of others, such as the right to appeal a decision. The ruling reaffirms that, while companies are not obligated to reveal their algorithms, they *must* provide explanations about the logic of their systems, particularly how variations in input data might affect outcomes, unless such explanations would expose trade secrets.[105]

The principle of equality of arms, in turn, requires a 'fair balance' between the parties, meaning that each side must be able to present their case in court without facing significant disadvantage compared with the other.[106] As with the adversarial principle, this requirement is inherent in the broader concept of a fair trial, being applicable to both civil and criminal cases.[107] An example of failure to observe the equality of arms principle relates to one of the parties enjoying significant advantages regarding access

[101] *Nideröst-Huber v Switzerland* App no 18990/91 (ECtHR, 18 February 1997), § 27; *Ziegler v Switzerland* App no 33499/96 (ECtHR, 21 February 2002), § 38.
[102] *Pellegrini v Italy* App no 30882/96 (ECtHR, 20 July 2001), § 45.
[103] Wisconsen Supreme Court, "*State v Loomis*," § 55.
[104] CJEU, "*Dun & Bradstreet Austria* (C-203/22)."
[105] CJEU, "*Dun & Bradstreet Austria* (C-203/22)."
[106] *Regner v Czech Republic* App. No. 35289/11 (ECtHR, 19 September 2017); *Dombo Beheer B.V. v the Netherlands* App no 14448/88 (ECtHR, 27 October 1993), § 33.
[107] *Feldbrugge v the Netherlands* App no 8562/79 (ECtHR, 29 May 1986), § 44.

to relevant information, occupying a dominant position in the proceedings and wielding considerable influence with regard to the court's assessment.[108]

Artificial intelligence challenges these principles because, first, based on the features of the system being used, a party can encounter a black box problem. To comment on the system's outcome in a meaningful way, a party must have the knowledge of its functioning, and in order to be able to comment on the relevant materials, a party must first understand the information. It is not possible to assume, however, that a party would be able to understand the systems or their functioning, much less provide insightful commentary on them, considering how these are often incomprehensible to a layperson, including judges. Second, crucial information may be withheld to protect conflicting interests – such as the need to safeguard trade secrets, intellectual property or the possibility of data manipulation – even in the unlikely event that a party *is* able to understand the system and comment on it appropriately.[109] In this context, the CJEU decision in *Yettel Bulgaria* is of particular importance since it underscores that, while organisations are not required to disclose their algorithms, they must ensure meaningful transparency in how ADM functions.

These challenges are also exemplified in a 2023 ruling of the ECtHR dealing with the encrypted messaging app ByLock.[110] The ECtHR determined that Turkish courts had violated Article 6 of the ECHR by convicting the applicant of being a member in a terrorist organisation responsible for the 2016 *coup d'état* attempt. This conviction was primarily based on the applicant's use of ByLock, which authorities alleged was exclusively used by members of the group. This was despite the app being globally available and the authorities' unwillingness to reveal any potentially incriminating content from it. Instead, merely downloading or using ByLock was enough for individuals to be accused of membership in a terrorist organisation. Concerning the violations to the right to a fair trial,[111] the ECtHR identified that there were inadequate safeguards to ensure that the applicant had a meaningful opportunity to contest the evidence against him and present his defence effectively and on equal terms with the prosecution. The intelligence services withheld the raw ByLock data from the applicant, preventing him from challenging the conclusions drawn from its use. Additionally, domestic courts refused to allow an independent examination to verify the data's content and integrity. Concerns were also raised about the reliability of the ByLock data, including inconsistencies between different user lists compiled by intelligence services and discrepancies in the number of identified and prosecuted users. Other shortcomings included the fact that ByLock was available for public download from app stores and websites until early 2016, with no restrictions for nearly two years. Consequently, the app was accessible to anyone, not just members of a terrorist group, yet the domestic courts failed to acknowledge this.[112]

5. Trust: Reasoned Verdicts

In deciding cases at the domestic level, the courts in Council of Europe member states are usually obliged under domestic law to provide detailed reasoning to their judgments. And the ECHR, as interpreted by the ECtHR in its case law, also gives rise to substantial obligations concerning the content of that reasoning. According to the latter, the guarantees enshrined in Article 6(1) of the Convention include the obligation for courts to give sufficient reasons for their decisions.[113] Although courts have some margin of appreciation when choosing arguments and admitting evidence, they are obliged to justify their activities by providing reasons for their decisions.[114] The provided reasons should allow the parties to effectively exercise any available right of appeal.[115] Not every argument advanced by the parties needs to receive a detailed answer by the court;[116] however, where a party's submission is decisive for the outcome of the proceedings, a specific and express reply is required.[117] The courts are therefore required to examine, at the very least, the litigants' main arguments[118] and pleas concerning the rights and freedoms guaranteed by the Convention and its Protocols, with particular rigour and care.[119]

---

[108] *Yvon v France* App no 44962/98 (ECtHR, 24 July 2003), §37.
[109] Steponenaite, "Judicial Analytics on Trial."
[110] *Yüksel Yalçınkaya v Türkiye* App no 15669/20 (ECtHR, 26 September 2023).
[111] In addition to the violations of Article 6 ECHR, the ECtHR also identified violations to Article 7 (no punishment without law) and Article 11 (freedom of assembly and association) of the Convention.
[112] *Yüksel Yalçınkaya v Türkiye* App no 15669/20 (ECtHR, 26 September 2023).
[113] *H. v Belgium* App no 8950/80 (ECtHR, 30 November 1987), § 53.
[114] *Suominen v Finland* App no 37801/97 (ECtHR, 1 July 2003), § 36.
[115] *Hirvisaari v Finland* App no 49684/99 (ECtHR, 27 September 2001), §30 *in fine*.
[116] ECtHR, "*Van de Hurk v the Netherlands*," § 61; *García Ruiz v Spain* App no 30544/96 (ECtHR, 21 January 1999), § 26; *Perez v France* App no 47287/99 (ECtHR, 12 February 2004), § 81.
[117] *Ruiz Torija v Spain* App no 18390/91 (ECtHR, 09 December 1994), § 30; *Hiro Balani v Spain* App no 18064/91 (ECtHR, 9 December 1994), §28.
[118] *Buzescu v Romania* App no 61302/00 (ECtHR, 24 May 2005), § 67; *Donadzé v Georgia* App no 74644/01 (ECtHR, 7 June 2006), § 35.
[119] *Wagner and J.M.W.L. v Luxembourg* App no 76240/01 (ECtHR, 28 June 2007), § 96.

A number of benefits theoretically achieved by giving reasoned judgments have been identified, which are summarised according to the three central values promoted by reason-giving in the courtroom: participation, accountability and accuracy. Fundamentally, imposing an obligation to give reasons for a judgment aims to secure litigants' involvement in the judicial process while ensuring that the court is held accountable and accurate decisions are reached. To demonstrate that they are receptive to the evidence and arguments put out by the parties, judges must provide an explanation for their rulings. A judge conveys the degree to which the parties' arguments have been comprehended, accepted or served as the foundation for the verdict by explaining their decisions.[120] As previously seen, the perception that the decision-maker has given due consideration to the respondent's views and arguments (i.e. voice) is crucial to acceptance of both the decision and the authority of the institution that imposes the decision.[121]

Giving reasoned verdicts also works as an accountability-enhancing mechanism. It limits judicial discretion by ensuring that written decisions, or at least some record of the proceedings, can be read and reviewed by higher courts, as well as the public in general, and by encouraging judges to treat similarly situated cases alike and differently situated cases differently. This is closely related to the expectation of neutrality. Lastly, giving reasons for judgments enforces a form of self-discipline that is thought to improve the quality of the decisions. By requiring judges to substantiate their decisions based on facts and legal arguments, the accuracy of judicial decision-making is improved. Giving reasons guarantees that 'judicial decisions are not made arbitrarily or based on speculation, suspicion, or irrelevant information' and that 'the deciding court has considered all relevant factors, researched the applicable law and given the case the thought it deserved'.[122] When judges present litigants and the general public with the reasons why they decided in a certain way, they are more inclined to accept and comply with the decision, and thus overall trust in the judicial system rises. This connection between giving reasons, litigants' trust and the legitimacy of courts has been emphasised by the ECtHR, according to which a reasoned decision shows the parties that their case has truly been heard,[123] thereby fostering public confidence in judicial outcomes. As Tyler and Lind explain:

> People are also affected by the degree to which they feel that the authorities with whom they deal are motivated to try to be fair to them and to others in the group … The perception of a motivation to be fair is crucial to people's feelings about authorities, since it both reflects the character of the individual authority and is the basis for predicting his or her future behaviour.[124]

One of the main ways in which judicial authorities convey to the public that their case has been heard – that is, their trustworthiness – is by providing reasons for their decisions. AI systems can pose significant risks to the judicial duty to provide reasons,[125] particularly due to the quality and scope of their training data. As explained in Section 1, the accuracy of AI-generated outputs depends on the diversity, representativeness and timeliness of the data used to train the models. If these systems are not trained on sufficiently comprehensive legal sources tailored to specific jurisdictions, their outputs may be incomplete or unreliable, ultimately undermining the reasoning required in judicial decisions. Even seemingly straightforward applications, such as summarising case law, can produce misleading results if the AI has been trained on a limited dataset. Additionally, since models such as ChatGPT are predominantly trained on English-language sources, they may not accurately reflect the legal principles of non-English-speaking jurisdictions. Copyright restrictions further limit the effectiveness of AI-generated legal reasoning. While court rulings are typically public and can be used for training, legal literature – an essential resource for judges – is often copyrighted and inaccessible to AI systems. This restriction may reduce the accuracy and depth of AI-generated outputs, making it harder for judges to provide well-reasoned justifications for their decisions.[126] Additionally, biases embedded in under-representative training data can perpetuate inaccuracies, further compromising judicial reasoning.

The 'black box' nature of AI[127] adds another layer of concern, as the lack of transparency about how these systems generate responses makes it difficult for judges to explain their reasoning when relying on AI tools:

> Even if the sources used for training are known and accessible, the internal mechanisms often remain obscure. This opacity and lack of transparency pose challenges for the judicial duty to state reasons, as judges must be able to clearly explain the

[120] Cohen, "When Judges Have Reasons."
[121] Cohen, "The Social Psychology."
[122] Cohen, "When Judges Have Reasons," 511–512.
[123] *H. v Belgium* App no 8950/80 (ECtHR, 30 November 1987), § 53.
[124] Tyler, "Procedural Justice," 76.
[125] For a comprehensive analysis of the impact of generative AI systems on the judicial duty to state reasons, see Hendrickx, "The Judicial Duty."
[126] Hendrickx, "The Judicial Duty."
[127] Pasquale, "The Black Box Society."

> basis for their decisions. Without understanding how the AI reached its conclusions, judges cannot adequately justify their reliance on such tools, which undermines both the reasoning process and the broader perception of judicial legitimacy.[128]

Compared with human decision-makers, AI systems can be seen as inscrutable, thereby impacting people's willingness to accept the system's decisions or recommendations.[129] In a recent study about public perceptions of algorithmic judges, researchers found that, despite acknowledging various advantages associated with algorithms (e.g. speed and cost), court users tend to trust human judges more, and show greater intentions to go to court when a human judge adjudicates, as opposed to an algorithmic one. Additionally, the extent to which people trust algorithmic and human judges was also found to vary according to the nature of the case, with trust for algorithmic judges being particularly low in emotionally complex cases, compared with technically complex or uncomplicated cases.[130] This might be related to people perceiving AI systems as lacking empathy or a full understanding of the complexity of human relations – a topic that will be explored in the next section.

It is often argued that the judicial duty to give reasons can lead judges to be less reliant on the recommendations put forward by an algorithmic or AI system. For instance, returning to the case of *State v Loomis*, the Wisconsin Supreme Court established that a judge ought to explain which factors, in addition to the COMPAS risk assessment, independently support the sentence imposed,[131] in an attempt to ensure that the results of automated risk assessments are weighed appropriately. What this line of reasoning fails to consider, however, is that the justification put forward by the judge might simply be an *ex post* rationalisation of a decision already made on other grounds – namely, on the output provided by the system. Social psychology research on judgement and decision-making suggests that our reasoning is heavily influenced by accountability pressures: if people believe that they might eventually be called upon to explain themselves, they will reason much more carefully – not necessarily trying to find out the truth, but trying to figure out what is justifiable and defendable:[132]

> [A] central function of thought is making sure that one acts in ways that can be persuasively justified or excused to others. Indeed, the process of considering the justifiability of one's choices may be so prevalent that decision makers not only search for convincing reasons to make a choice when they must explain that choice to others, they search for reasons to convince themselves that they have made the 'right' choice.[133]

This also applies to judges, whose written decisions may be scrutinised by appellate courts, the parties and the public. In this sense, there is a risk that judges might simply look for arguments that support a conclusion originally reached on the grounds of, for example, a risk-assessment report, and favour those conclusions for which arguments can be found.[134] If the public believes judges are using AI systems they consider untrustworthy, the legitimacy of the courts could be undermined.

6. Standing: The Role of Courts

Article 6 ECHR enshrines that everyone is entitled to a fair and public hearing by an independent and impartial *tribunal* established by law. Throughout this article, I have discussed several of the essential guarantees pertaining to the right to a fair trial, without discussing what is meant by and what is the role of a tribunal or court. According to the ECtHR, a court or tribunal is characterised in the substantive sense of the term by its judicial function – that is, determining matters within its competence based on rules of law and after proceedings conducted in a prescribed manner.[135] An authority that is not formally recognised as a court within a state may still fall under the concept of a 'tribunal' in the substantive sense of the term for the purposes of Article 6(1).[136] Additionally, the fact that it performs various functions (e.g. administrative, advisory, adjudicative) does not preclude an institution from being a deemed a 'tribunal'.[137] An inherent characteristic of a court or tribunal is the authority to issue a binding decision that cannot be changed by a non-judicial body to the detriment of an individual party.[138] In order to have the legitimacy required in a democratic society, a tribunal must always be 'established by law'.[139] This encompasses not only the legal basis for a tribunal's existence, but also compliance by the tribunal with the particular rules that govern it.[140] By

[128] Hendrickx, "The Judicial Duty," 9.
[129] Yeomans, "Making Sense of Recommendations."
[130] Yalcin, "Perceptions of Justice."
[131] Wisconsin Supreme Court, "*State v Loomis*."
[132] Haidt, "Moral Psychology and the Law."
[133] Lerner, "Bridging Individual, Interpersonal, and Institutional Approaches," 874.
[134] Mercier, "Why Do Humans Reason?"
[135] *Cyprus v Turkey* App no 25781/94 (ECtHR, 10 May 2001), § 233.
[136] *Sramek v Austria* App no 8790/79 (ECtHR, 22 October 1984), § 36; *Belilos v Switzerland*, no. 10328/83, 29 April 1988, Series A no. 132
[137] *H. v Belgium* App no 8950/80 (ECtHR, 30 November 1987), § 50.
[138] ECtHR, "*Van de Hurk v the Netherlands*," § 45.
[139] ECtHR, "*Lavents v Latvia*," § 81.
[140] *Sokurenko and Strygun v Ukraine* App nos 29458/04 and 29465/04 (ECtHR, 20 July 2006), § 24.

definition, the lawfulness of a court or tribunal also includes its composition, including the procedures governing the appointment of judges.[141] Finally, a 'tribunal' must meet additional requirements, including independence – particularly from the executive – as well as impartiality, the duration of its members' terms of office and procedural safeguards, several of which are explicitly stated in Article 6(1).[142]

The role of courts, however, is being affected by the introduction of digital technologies. According to Helberger, 'technology is not only changing the materiality of courts – moving from physical buildings to digital portals – but also affecting their symbolic function as public institutions'.[143] Regarding the former, the author explains that courthouses have an important symbolic function as the place where one goes to see justice being done, which allow for different forms of communication and access than, for example, virtual courtroom meetings. With the arrival of the internet and digital technologies, the materiality of courts as a physical location was one of the first characteristics that were let go by society as a defining feature of the sphere of courts – a process further accelerated by the global pandemic. Helberger notices a general trend to 'digitise courts and court proceedings to offer various court services online, streamline and make processes more efficient, enable remote hearings, and improve online accessibility 24 hours a day' and argues that these initiatives 'effectively erode our idea of courts as a location – a place one has to go to get justice'.[144] In this context, others have also presented concerns regarding this shift, mainly relating to procedural justice and reduced possibilities for interaction and effective communication.[145]

The ECtHR judgment in *Xavier Lucas v France*[146] serves as an appropriate example of the risks associated with the digitalisation of justice. The case centred on the mandatory use of the 'e-Barreau' electronic platform for filing appeals. Lucas's lawyer attempted to challenge an arbitration award on paper, but the French courts rejected it, insisting that it must be submitted electronically – even though the platform lacked the necessary legal categories for accurate submission. The ECtHR found that requiring electronic filing without addressing practical obstacles placed an excessive and disproportionate burden on the applicant. The ruling criticised France's rigid, formalistic approach, stating that it undermined access to justice rather than ensuring legal certainty. Consequently, the court held that France had failed to balance procedural formalities with the fundamental right to a fair trial.

In addition to changing the materiality of courts, Helberger further argues that digital technologies are affecting a second requirement that, until recently, identified a court: expertise and judicial authority:

> Courts have been described as a place where we see 'justice being done.' They are places where we can see justice being done because in courts, we expect to find expertise and experts that can be trusted to adhere to judicial values and procedures. The personification of this expertise is the figure of the judge. Judges are experts who adhere to and possess the necessary substantive and procedural knowledge to apply the law and operationalize judicial values such as impartiality, knowledge of and adherence to principles of due process, respect for transparency, and the ability to reason in line with the judicial method. Typically, admission to the position of a judge is conditional upon the successful completion of a study of the law. But expertise in itself is not sufficient. Public perception and confidence in the public that judges and courts can comply with this task are also important elements of their authority.[147]

With digitisation, the traditional perception of judges as the ultimate symbol of judicial authority is fading. While judges once held a central role in courtrooms, overseeing legal proceedings, digital societies are increasingly adapting to new forms of authoritative decision-making. Judicial authority is now being substituted by alternative decision-makers with distinct areas of expertise, including citizens, corporations and even AI.[148]

From the perspective of procedural fairness, the shift of authority being increasingly expressed algorithmically is problematic not only in relation to accessing courts but particularly when it comes to expectations of standing. This, we have seen, encompasses feelings that one is viewed by authorities as a full-fledged member of society, thus being treated politely and with dignity, and having their rights and opinions respected. It is connected to the extent to which authorities are polite and respectful in their dealings with litigants. Meyerson and Mackenzie explain that the reason why people associate these interpersonal aspects of procedures with fair treatment is because:

---

[141] *Buscarini and others v San Marino* App no 24645/94 (ECtHR, 18 February 1999); *Guðmundur Andri Ástráðsson v Iceland*.
[142] *Le Compte, Van Leuven and De Meyere v Belgium* App nos 6878/75, 7238/75 (ECtHR, 23 June 1981), § 55; *Cyprus v Turkey* App no 25781/94 (ECtHR, 10 May 2001), § 233.
[143] Helberger, "The Rise of Technology Courts," 1.
[144] Helberger, "The Rise of Technology Courts," 2–3.
[145] Donoghue, "The Rise of Digital Justice"; Moore, "Digital Government, Public Participation and Service Transformation."
[146] ECtHR, "*Xavier Lucas v France*."
[147] Helberger, "The Rise of Technology Courts," 4.
[148] Helberger, "The Rise of Technology Courts."

> satisfactory interpersonal treatment by group authorities symbolically communicates the information that we possess value or status in the eyes of our community, which in turn supports a sense of self-respect or self-worth. Conversely, people feel that they have been denied procedural justice when the interpersonal aspects of procedures are unsatisfactory, because poor interpersonal treatment casts doubt on their standing in their group and damages their self-respect.[149]

The shift towards automation may dehumanise the court experience, reducing litigants to mere numbers or probabilities,[150] Additionally, when it comes to the judicial role, human traits such as empathy are often described as crucial, especially for how litigants perceive the process. For instance, presenting a case to a judge who can empathise with the parties involved and deliver a judgment informed by an understanding of the decision's impact is vital for a fair and meaningful process (especially in emotionally complex cases). But empathy is not the only thing lost when essential aspects of judicial decision-making are automated.

In this context, Moses identifies three characteristics inherent to the role of judges that can hardly be replaced by AI: the ability to exercise judgement, being attuned to the morality of the community in which decisions are made and being subject to law in a meaningful sense.[151] Regarding the first characteristic, she explains that the exercise of judgement in reaching a decision is more critical to the function of judging than the mere production of text or predictions. She argues that what judges do, even in higher courts, extends beyond merely producing text with valid doctrinal arguments. According to her, the most important aspect is that judges are exercising judgement, which differs from both prediction – where the expected outcome of litigation is determined using probability – and simulation, such as what ChatGPT does when asked to generate a judgment. She emphasizes that the manner in which a decision is made is as critical as its content.[152] The other characteristics relate to the purpose of the rule of law in tempering power. When judges view themselves as subject to the same law they are interpreting and applying, this belief serves as a safeguard against arbitrariness. In theory, if a judge committed the same offence as the person before them, they would be vulnerable to being sentenced by someone in a similar position. Similarly, if the judge or an entity with which they were associated was involved in a civil dispute akin to the one before them, the same legal principles and interpretations would apply. This awareness likely makes the judge less prone to acting arbitrarily, unlike a despotic ruler who is not bound by the same rules. An automated system, however, lacks this awareness because it does not experience the law in the same way. To Moses, 'if the entity making, interpreting or enforcing rules experiences those rules fundamentally differently, then the rule of law as a means of tempering power breaks down'.[153] While this is not the same as empathy, it is nevertheless linked to considerations of standing, and the overall identification (or lack thereof) of the general public with the entity making decisions affecting them.

7. Concluding Remarks

In this article, I have presented how the adoption of AI systems to support judicial decision-making is often justified by putting forward a narrative in which the fallibility of human judges, who are often biased and limited in their cognitive capacities, is contrasted with AI's purported ability to enhance accuracy, objectivity and consistency in judicial decision-making. Indeed, there are clear benefits to the use of AI for the administration of justice, but the risks are also substantial and therefore require a careful analysis of which forms of automation are useful, appropriate and consistent with the rule of law and with the foundational elements of the right to a fair trial. In this context, the main goal of this article was to explore how the use of AI by courts can challenge these elements, and how they can negatively impact people's perceptions of procedural justice in ways that normal human adjudication – even when done by 'fallible judges' – does not. The way litigants and society perceive the fairness of procedures merit attention because it is intimately associated with the legitimacy of judicial institutions.

Although it is beyond the scope of this article to offer detailed solutions to the challenges that have been explored, some strategies are worth further investigation. One promising approach is improving AI literacy[154] among judges. As mentioned in the introduction, the UNESCO Draft Guidelines for the Use of AI Systems in Courts and Tribunals were developed following the unnerving discovery that, even though 44 per cent of judicial operators already use AI tools for work-related activities, only 9 per cent of them report that their organisations have issued guidelines or provided AI-related training.[155] A lack of guidance and proper understanding about the AI's functionalities and capabilities can result in situations where judges heedlessly accept

[149] Meyerson, "Procedural Justice and the Law," 7.
[150] Martinho, "Surveying Judges."
[151] Moses, "Artificial Intelligence."
[152] Moses, "Artificial Intelligence," 127.
[153] Moses, "Artificial Intelligence," 128.
[154] See Tsz Kit Ng, "Conceptualizing AI Literacy."
[155] UNESCO, "UNESCO Survey Uncovers Critical Gaps."

the system's output. This need for AI literacy has been recognised in recent attempts to regulate the technology, with both the Council of Europe Framework Convention on Artificial Intelligence[156] and the European Union AI Act enshrining obligations to promote digital literacy and skills, particularly among professionals deploying AI systems.

Enhanced education on emerging technologies for legal professionals, along with increased collaboration with researchers, engineers and developers of AI, can help judges gain a clearer grasp of the strengths and limitations of AI systems. Collaboration with AI developers is also crucial for enhancing the design and functionality of automated systems in the judiciary. Ideally, the development of AI tools for use in judicial settings should shift from outsourcing to private companies toward state-built or non-profit models.[157] However, given the considerably high cost of designing an AI application and the fact that judicial institutions – unlike companies – often lack the resources and expertise to efficiently do so, the outsourcing of external private parties is often a more viable option than the in-house development of AI systems[158]. Still, by working closely with developers, judicial institutions can ensure that AI tools are tailored to the specific needs of the courts, aligning with legal principles and ethical standards. This collaboration can help to address potential biases and errors in AI systems (for example, by improving data quality), making them more reliable and transparent. Finally, in relation to the problem of judges' bias in adjudication, a promising possibility consists in the use of AI methods to help detect and counteract judges' implicit biases, such as through predictive judicial analytics in the form of machine learning.[159] AI can help identify the situations where judicial bias is likely to take place, based on the analysis of covariates that, despite being legally irrelevant, have been shown to influence judicial decisions. By identifying the instances in which bias commonly arises, it is possible not only to alert judges but also to target debiasing interventions, such as educating judges on the subject and offering feedback on their work.[160] Given the high stakes involved in judicial decision-making, and the importance of having litigants' basic expectations concerning fair processes met, these are avenues of investigation that are worth pursuing.

[156] Council of Europe, "Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law," article 20; European Union, "Regulation (EU) 2024/1689 … Artificial Intelligence Act."
[157] As suggested by McGill, "Judging by Numbers."
[158] Terzidou, "The Use of Artificial Intelligence."
[159] See, for example, Chen, "Judicial Analytics."
[160] See Lopes, "Artificial Intelligence."

Bibliography

Araujo, Theo, Natali Helberger, Sanne Kruikemeier and Claes H. De Vreese. "In AI We Trust? Perceptions about Automated Decision-Making by Artificial Intelligence." *AI & Society* 35, no 3 (2020): 611–623. https://doi.org/10.1007/s00146-019-00931-w.

Barysė, Dovilė and Roee Sarel. "Algorithms in the Court: Does It Matter Which Part of the Judicial Decision-Making is Automated?" *Artificial Intelligence and Law*, January 8, 2023. https://doi.org/10.1007/s10506-022-09343-6.

Blader, Steven L. and Tom R. Tyler. "A Four-Component Model of Procedural Justice: Defining the Meaning of a 'Fair' Process." *Personality and Social Psychology Bulletin* 29, no 6 (2003): 747–758. https://doi.org/10.1177/0146167203029006007.

Bohannon, Molly. "Lawyer Used ChatGPT in Court – and Cited Fake Cases. A Judge is Considering Sanctions." *Forbes*, June 8, 2023. https://www.forbes.com/sites/mollybohannon/2023/06/08/lawyer-used-chatgpt-in-court-and-cited-fake-cases-a-judge-is-considering-sanctions.

Canadian Judicial Council. "Guidelines for the Use of Artificial Intelligence in Canadian Courts" (2024). https://cjc-ccm.ca/sites/default/files/documents/2024/AI%20Guidelines%20-%20FINAL%20-%202024-09%20-%20EN.pdf.

Chatziathanasiou, Konstantin. "Beware the Lure of Narratives: 'Hungry Judges' Should Not Motivate the Use of 'Artificial Intelligence' in Law." *German Law Journal* 23, no 4 (May 2022): 452–464. https://doi.org/10.1017/glj.2022.32.

Chen, Daniel L. "Judicial Analytics and the Great Transformation of American Law." *Artificial Intelligence and Law* 27, no 1 (2019): 15–42. https://doi.org/10.1007/s10506-018-9237-x.

Chesterman, Simon. "Through a Glass, Darkly: Artificial Intelligence and the Problem of Opacity." *The American Journal of Comparative Law* 69, no 2 (2021): 271–294. https://doi.org/10.1093/ajcl/avab012.

Clarke, Roger. "The re-conception of AI: Beyond Artificial, and beyond Intelligence". IEEE Trans. on Technology and Society 4(1) (2023): 24.

Cohen, Mathilde, "When Judges Have Reasons Not to Give Reasons: A Comparative Law Approach." *Washington & Lee Law Review* 72 (2015): 483.

Cohen, Ronald L., E. Allan Lind and Tom R. Tyler, "The Social Psychology of Procedural Justice," *Contemporary Sociology* 18, no 5 (1989): 758. https://doi.org/10.2307/2073346.

Consultative Council of European Judges. "Opinion No. 26 (2023) Moving Forward: The Use of Assistive Technology in the Judiciary". Strasbourg: Council of Europe, 2023. http://www.europeanrights.eu/public/commenti/CCJE_Opinion_No._26_(2023)_final.pdf.

Council of Europe. "Convention for the Protection of Human Rights and Fundamental Freedoms" (European Convention on Human Rights, as amended) (ECHR). Strasbourg: Council of Europe, 1950.

Council of Europe. "Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law." Strasbourg: Council of Europe, 2024.

Court of Justice of the European Union (CJEU). "*Dun & Bradstreet Austria*, Case C-203/22 of 27/02/2025" (2025).

Courts of New Zealand. "Guidelines for Use of Generative Artificial Intelligence in Courts and Tribunals." (2023). https://www.courtsofnz.govt.nz/assets/6-Going-to-Court/practice-directions/practice-guidelines/all-benches/20231207-GenAI-Guidelines-Judicial.pdf.

Cowgill, Bo. "The Impact of Algorithms on Judicial Discretion: Evidence from Regression Discontinuities". (2024). http://www.columbia.edu/~bc2656/papers/RecidAlgo.pdf.

Dietterich, Thomas G. "Robust Artificial Intelligence and Robust Human Organizations." *Frontiers of Computer Science* 13, no 1 (2019): 1–3. https://doi.org/10.1007/s11704-018-8900-4.

Doleac, Jennifer. "Let Computers Be the Judge: The Case for Incorporating Machine Learning into the U.S. Criminal Justice Process." *Medium*, April 20, 2017. https://medium.com/@jenniferdoleac/let-computers-be-the-judge-b9730f94f8c8.

Donoghue, Jane. "The Rise of Digital Justice: Courtroom Technology, Public Participation and Access to Justice." *The Modern Law Review* 80, no 6 (2017): 995–1025

Dzindolet, Mary T., Scott A. Peterson, Regina A. Pomranky, Linda G. Pierce and Hall P. Beck. "The Role of Trust in Automation Reliance." *International Journal of Human-Computer Studies* 58, no 6 (2003): 697–718. https://doi.org/10.1016/S1071-5819(03)00038-7.

European Commission for the Efficiency of Justice (CEPEJ). "European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment." (2018). https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.

European Commission for the Efficiency of Justice (CEPEJ). "Use of Generative Artificial Intelligence (AI) by Judicial Professionals in a Work-related Context." (2024). https://rm.coe.int/cepej-gt-cyberjust-2023-5final-en-note-on-generative-ai/1680ae8e01.

European Union. "Regulation (EU) 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)" (2024).
Farah, Hibaq. "Court of Appeal Judge Praises 'Jolly Useful' ChatGPT After Asking It for Legal Summary." *The Guardian*, September 15, 2023. https://www.theguardian.com/technology/2023/sep/15/court-of-appeal-judge-praises-jolly-useful-chatgpt-after-asking-it-for-legal-summary.
Goddard, Kate, Abdul Roudsari and Jeremy C. Wyatt. "Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators." *Journal of the American Medical Informatics Association* 19, no 1 (2012): 121–127. https://doi.org/10.1136/amiajnl-2011-000089.
Gravett, Willem H. "Judicial Decision-making in the Age of Artificial Intelligence." In *Multidisciplinary Perspectives on Artificial Intelligence and the Law*, edited by Henrique Sousa Antunes, Pedro Miguel Freitas, Arlindo L. Oliveira, Clara Martins Pereira, Elsa Vaz De Sequeira and Luís Barreto Xavier, 281–297). Springer.
Guarnieri, Carlo and Daniela Piana. "Judicial Independence and the Rule of Law: Exploring the European Experience." In *The Culture of Judicial Independence*, edited by Shimon Shetreet and Christopher Forsyth, 111–124. Leiden: Brill, 2012. https://doi.org/10.1163/9789004215856_007.
Haidt, Jonathan. "Moral Psychology and the Law: How Intuitions Drive Reasoning, Judgment, and the Search for Evidence." *University of Alabama Law Review* 64, no 4 (2013): 867–880.
Hannah-Moffat, Kelly. "The Uncertainties of Risk Assessment." *Federal Sentencing Reporter* 27, no 4 (2015): 244–247. https://doi.org/10.1525/fsr.2015.27.4.244.
Harvard Law Review. "*State v Loomis*: Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessment in Sentencing." *Harvard Law Review* 130, no 5 (2017): 1530–1537. https://harvardlawreview.org/2017/03/state-v-loomis.
Helberger, Natali. "The Rise of Technology Courts, or: How Technology Companies Re-Invent Adjudication for a Digital World." *Computer Law & Security Review* 56 (2025): 106118. https://doi.org/10.1016/j.clsr.2025.106118.
Hendrickx, Victoria. "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* 17, no 3 (2024). https://doi.org/10.5553/ELR.000275.
Ibánez, Perfecto Andrés. "Imparcialidad Judicial e Independencia Judicial." in *La Imparcialidad Judicial*, 41–71. Madrid: Consejo General del Poder Judicial, 2009.
Jackson, Jonathan. "Norms, Normativity, and the Legitimacy of Legal Authorities: International Perspectives." *Annual Review of Law and Social Science* 14 (2018): 145–165.
Kahneman, Daniel. *Thinking, Fast and Slow*. Harmondsworth: Penguin, 2012.
Kazim, Tatiana and Joe Tomlinson. "Automation Bias and the Principles of Judicial Review." *Judicial Review* 28, no 1 (2023): 9–16. https://doi.org/10.1080/10854681.2023.2189405.
Kitchin, Rob. "Thinking Critically About and Researching Algorithms." *Information, Communication & Society* 20, no 1 (2017): 14–29. https://doi.org/10.1080/1369118X.2016.1154087.
Laptev, Vasiliy and Daria Feyzrakhmanova. "Application of Artificial Intelligence in Justice: Current Trends and Future Prospects." *Human-Centric Intelligent Systems* 4 (2024): 394–405. https://doi.org/10.1007/s44230-024-00074-2.
Lehr, David and Paul Ohm. "Playing with the Data: What Legal Scholars Should Learn About Machine Learning." *UC Davis Law Review* 51, no 653 (2017): 653–717. https://lawreview.law.ucdavis.edu/sites/g/files/dgvnsk15026/files/media/documents/51-2_Lehr_Ohm.pdf.
Liu, H.-W., C.-F. Lin and Y.-J. Chen. "Beyond State v Loomis: Artificial Intelligence, Government Algorithmization, and Accountability." *International Journal of Law and Information Technology* 27, no 2 (2019): 122–141. https://doi.org/10.1093/ijlit/eaz001.
Lopes, Giovana. "Artificial Intelligence and Judicial Decision-making: Evaluating the Role of AI in Debiasing." *TATuP – Journal for Technology Assessment in Theory and Practice*, 33, no 1 (2024): 28–33. https://doi.org/10.14512/tatup.33.1.28.
Lyell, David and Enrico Coiera. "Automation Bias and Verification Complexity: A Systematic Review." *Journal of the American Medical Informatics Association* 24, no 2 (2017): 423–431. https://doi.org/10.1093/jamia/ocw105.
Martineau, Kim. "What is Generative AI?" IBM Blog. https://research.ibm.com/blog/what-is-generative-AI.
Martinho, Andreia. "Surveying Judges About Artificial Intelligence: Profession, Judicial Adjudication, and Legal Principles." *AI & Society* (2024). https://doi.org/10.1007/s00146-024-01869-4.
Malek, Md. Abdul. "Criminal Courts' Artificial Intelligence: The Way It Reinforces Bias and Discrimination." *AI and Ethics* 2 (2022): 1–13. https://doi.org/10.1007/s43681-022-00137-9.
McGill, Jena and Amy Salyzyn. "Judging by Numbers: Judicial Analytics, the Justice System and Its Stakeholders." *Dalhousie Law Journal* 44, no 1 (2021): 249–284.
Mercier, Hugo and Dan Sperber. "Why Do Humans Reason? Arguments for an Argumentative Theory." *Behavioral and Brain Sciences* 34, no 2 (2011): 57–74. https://doi.org/10.1017/S0140525X10000968.

Meyerson, Denise and Catriona Mackenzie. "Procedural Justice and the Law." *Philosophy Compass* 13, no 12 (2018): e12548. https://doi.org/10.1111/phc3.12548.

Meyerson, Denise. "The Moral Justification for the Right to Make Full Answer and Defence." *Oxford Journal of Legal Studies* 35, no 2 (2015): 237–265. https://doi.org/10.1093/ojls/gqu024.

Molbæk-Steensig, Helga and Alexandre Quemy. "Artificial Intelligence and Fair Trial Rights." In *Artificial Intelligence and Human Rights*, edited by Molbæk-Steensig and Alexandre Quemy, 265–280. Oxford: Oxford University Press, 2023. https://doi.org/10.1093/law/9780192882486.003.0018.

Moore, Sarah. "Digital Government, Public Participation and Service Transformation: The Impact of Virtual Courts." *Policy & Politics* 47, no 3 (2019): 495–509. https://doi.org/10.1332/030557319X15586039367509.

Moses, Lyria B. "Artificial Intelligence: Affordances and Limits in the Context of Judging." *Journal & Proceedings of the Royal Society of New South Wales* 157 (2024):123–129.

Mosier, Kathleen L. and Linda J. Skitka. "Human Decision Makers and Automated Decision Aids: Made for Each Other?" In *Automation and Human Performance: Theory and Applications*, edited by Raja Parasuraman and Mustapha Mouloua, 201–220. Boca Raton, FL: CRC Press, 1996.

Organization of American States (OAS). "American Convention on Human Rights (ACHR)." (1969).

Papayannis, Diego M. "Independence, Impartiality and Neutrality in Legal Adjudication." *Revus*, 28 (2016): 3352. https://doi.org/10.4000/revus.3546.

Parasuraman, Raja and Dietrich H. Manzey. "Complacency and Bias in Human Use of Automation: An Attentional Integration." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 52, no 3 (2010): 381–410. https://doi.org/10.1177/0018720810376055.

Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press, 2016.

Peters, Uwe. "Explainable AI Lacks Regulative Reasons: Why AI and Human Decision-making are Not Equally Opaque." *AI Ethics* 3 (2023): 963–974. https://doi.org/10.1007/s43681-022-00217-w.

Reiling, A.D. (Dory). "Courts and Artificial Intelligence." *International Journal for Court Administration* 11, no 2 (2020): 8. https://doi.org/10.36745/ijca.343.

Russell, Stuart and Peter Norvig. *Artificial Intelligence: A Modern Approach*, 3rd ed. Boston: Pearson Education, 2016.

Sharon, Tamar and Raphaël Gellert. "Regulating Big Tech Expansionism? Sphere Transgressions and the Limits of Europe's Digital Regulatory Strategy." *Information, Communication & Society* 27, no 15 (2024): 2651–2668. https://doi.org/10.1080/1369118X.2023.2246526.

Smith, Adam, Anastasia Moloney and Avi Asher-Schapiro. "AI in the Courtroom: Judges Enlist ChatGPT Help, Critics Cite Risks." *CS Monitor*, May 30, 2023. https://www.csmonitor.com/USA/Justice/2023/0530/AI-in-the-courtroom-Judges-enlist-ChatGPT-help-critics-cite-risks.

Smuha, Nathalie and Victoria Hendrickx. "AI and the Administration of Justice: Taking 'Precedent Analysis' as a Use Case to Assess the Adequacy of the AI Act." *The Law, Ethics & Policy of AI Blog*. (2023). https://www.law.kuleuven.be/ai-summer-school/blogpost/Blogposts/AI-administration-justice.

Spellman, Barbara and Adele Quigley-McBride. "Reasoning About Forensic Science Evidence." In *Legal Reasoning and Cognitive Science: Topics and Perspectives*, edited by Marco Brigaglia and Corrado Roversi, 439–458. Milan: Diritto & Questioni Pubbliche | Recognise, 2023.

Stanovich, Keith E. and Richard F. West. "On the Relative Independence of Thinking Biases and Cognitive Ability." *Journal of Personality and Social Psychology* 94, no 4 (2008): 672–695. https://doi.org/10.1037/0022-3514.94.4.672.

Steponenaite, Vilte and Peggy Valcke. "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, no 6 (2020): 759–773. https://doi.org/10.1177/1023263X20981472.

Sweney, Mark. "What is the Post Office Horizon IT Scandal All About?" *The Guardian*, January 7, 2024. https://www.theguardian.com/business/2024/jan/07/what-is-the-post-office-horizon-it-scandal-all-about.

Taylor, Luke. "Colombian Judge Says He Used ChatGPT in Ruling." *The Guardian*, February 3, 2023. https://www.theguardian.com/technology/2023/feb/03/colombia-judge-chatgpt-ruling.

Terzidou, Kalliopi. "The Use of Artificial Intelligence in the Judiciary and Its Compliance with the Right to a Fair Trial." *Journal of Judicial Administration* 31 (2022): 154–168.

Thibaut, John W. and Laurens Walker. *Procedural Justice: A Psychological Analysis*. New York: Wiley, 1975.

Tsz Kit Ng, Davy, Jac Ka Lok Leung, Samuel Kai Wah Chu and Maggie Shen Qiao. "Conceptualizing AI Literacy: An Exploratory Review." *Computers and Education: Artificial Intelligence* 2 (2021): 1–11. https://doi.org/10.1016/j.caeai.2021.10004.

Tyler, Tom R. *Why People Obey the Law*. New Haven, CT: Yale University Press, 1990.

Tyler, Tom R. and E. Allan Lind. "Procedural Justice." In *Handbook of Justice Research in Law*, edited by Joseph Sanders and V. Lee Hamilton, 65–92. Boston: Kluwer Academic, 2001. https://doi.org/10.1007/0-306-47379-8_3.

Tyler, Tom R. and E. Allan Lind. "A Relational Model of Authority in Groups." In *Advances in Experimental Social Psychology*, 25 (1992): 115–191. https://doi.org/10.1016/S0065-2601(08)60283-X.

Tyler, Tom R. and Cheryl Wakslak. "Profiling and Police Legitimacy: Procedural Justice, Attributions of Motive and Acceptance of Police Authority." *Criminology* 42, no 2 (2004): 253–282. https://colab.ws/articles/10.1111%2Fj.1745-9125.2004.tb00520.x#.

Ulenaers, Jasper. "The Impact of Artificial Intelligence on the Right to a Fair Trial: Towards a Robot Judge?" *Asian Journal of Law and Economics* 11, no 2 (2020): 10. https://doi.org/10.1515/ajle-2020-0008.

UNESCO. "Draft UNESCO Guidelines for the Use of AI Systems in Courts and Tribunals." (2024). https://unesdoc.unesco.org/ark:/48223/pf0000390781.

UNESCO. "UNESCO Survey Uncovers Critical Gaps in AI Training Among Judicial Operators." 14 June 2024. https://www.unesco.org/en/articles/unesco-survey-uncovers-critical-gaps-ai-training-among-judicial-operators-0.

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

United Nations. "International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171 (ICCPR)."

Van Dijck, Gijs. "Predicting Recidivism Risk Meets AI Act." *European Journal on Criminal Policy and Research* 28, no 3 (2022): 407–423. https://doi.org/10.1007/s10610-022-09516-8.

Wistrich, Andrew J. and Jeffrey J. Rachlinski. "Implicit Bias in Judicial Decision Making: How It Affects Judgment and What Judges Can Do About It." *SSRN Electronic Journal*, 2017. https://doi.org/10.2139/ssrn.2934295.

Yalcin, Gizem, Erlis Themeli Evert Stamhuis, Stefan Philipsen and Stefano Puntoni. "Perceptions of Justice by Algorithms." *Artificial Intelligence and Law* 31, no 2 (2023): 269–292. https://doi.org/10.1007/s10506-022-09312-z.

Yeomans, Michael, Anuj Shah, Sendhil Mullainathan and Jon Kleinberg. "Making Sense of Recommendations." *Journal of Behavioral Decision Making* 32, no 4 (2019): 403–414. https://doi.org/10.1002/bdm.2118.

Yeung, Karen and Adam Harkens. "How Do 'Technical' Design Choices Made When Building Algorithmic Decision-Making Tools for Criminal Justice Authorities Create Constitutional Dangers? (Part I)." *Public Law* (April, 2023): 265–286.

Yeung, Karen and Adam Harkens. "How Do 'Technical' Design Choices Made When Building Algorithmic Decision-Making Tools for Criminal Justice Authorities Create Constitutional Dangers? (Part II)." *Public Law* (April, 2023): 448–472.

Zenker, Frank. "De-Biasing Legal Factfinders." In *Philosophical Foundations of Evidence Law*, 395–410. Oxford: Oxford University Press, 2021.

Zenker, Frank and Christian Dahlman. "Debiasing and Rule of Law." In M. Klatt (ed.), *Legal Argumentation and the Rule of Law*, 217–229. The Hague: Eleven International, 2016.