UNSW Law & Justice Research Series

Judicial Impartiality: AI in Courts

Felicity Bell and Michael Legg

[2025] *UNSWLRS* 4
2025 *Cambridge Handbook of AI in Courts*
(forthcoming)

UNSW Law & Justice
UNSW Sydney NSW 2052 Australia

E: LAW-Research@unsw.edu.au
W: http://www.law.unsw.edu.au/research/faculty-publications
AustLII: http://www.austlii.edu.au/au/journals/UNSWLRS/
SSRN: http://www.ssrn.com/link/UNSW-LEG.html

Judicial Impartiality: AI in Courts

Felicity Bell and Michael Legg

Forthcoming in the *Cambridge Handbook of AI in Courts*

Impartiality, broadly meaning the absence of bias and according equal treatment before the law, is a foundational element of judicial decision-making around the world. In this chapter we consider how the goal of judicial impartiality may be either enhanced and supported, or undermined by, the use of artificial intelligence. Key developments in legal AI include innovations directed toward courts and decision-makers. These may be process-driven – for example, triaging or decision supporting systems; in the case of pre-trial processes, judges may need to manage technology-facilitated document discovery. AI systems may also be involved in the production of evidence submitted to the court. Finally, courts and judges themselves may be the subjects of AI tools, such as those which identify patterns in decision-making. As this chapter explores, these different uses all have implications for the way that judicial impartiality is enacted and tested.

1. Introduction

In many jurisdictions, impartiality is a core component of the exercise of judicial power.[1] Indeed former Law Lord Patrick Devlin called 'impartiality and the appearance of it…the supreme judicial virtue'.[2] Judicial impartiality is also regarded as an element of the right to a fair trial.[3] However, this crucial and revered judicial virtue must be considered in light of the development of Artificial Intelligence (AI).

[1] Robert French, 'Essential and Defining Characteristics of Courts in an Age of Institutional Change' (Speech, Supreme and Federal Court Judges Conference, 21 January 2013) 2 <https://www.hcourt.gov.au/assets/publications/speeches/current-justices/frenchcj/frenchcj21jan13.pdf> accessed 25 November 2024. See also James Allsop, 'Courts as (Living) Institutions and Workplaces' (Speech, Joint Federal & Supreme Court Conference, 23 January 2019).
[2] Patrick Devlin, 'Judges and Lawmakers' (1976) 39(1) MLR 1, 4.
[3] Universal Declaration of Human Rights (adopted 10 December 1948 UNGA Res 217 A(III)) art 10; Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as amended) art 6.

AI, as a term or field of computer science, is employed where processes are used to carry out tasks which, if performed by a human, would be seen as evidence of intelligence – i.e. the processes mimic, imitate or simulate intelligence. AI is also an umbrella term. There are different branches of AI, such as expert systems, machine learning and generative AI.[4]

In brief, expert systems, the oldest and most simple form of AI, are pre-programed systems which can guide users through a sequence or series of steps, similar to a decision tree. The system involves obtaining and deconstructing human expert knowledge into a computable form that can then be accessed more cheaply and widely.[5] Machine learning refers to 'an automated process of discovering correlations (sometimes alternatively referred to as relationships or patterns) between variables in a dataset, often to make predictions or estimates of some outcome'.[6]

Generative AI is an AI system capable of generating 'content', such as text, computer code, images or music, in response to prompts. The most commonly discussed type of generative AI in the legal sphere are those generating text based on large language models (LLMs) which employ machine learning/neural networks to predict an output – the next word in a sentence – based on an input, such as a sequence of words. The model recognises patterns in the dataset that the LLM is trained on and uses those patterns to produce a response. The responses are refined through feedback from humans and adjustments to the weights in the LLM to minimize error and remove objectionable responses.[7]

[4] For further description of the various forms of artificial intelligence see Michael Legg and Felicity Bell, *Artificial Intelligence and the Legal Profession* (Hart 2020) ch 2.
[5] Legg and Bell (n 4) 28.
[6] David Lehr and Paul Ohm, 'Playing with the Data: What Legal Scholars Should Learn About Machine Learning' (2017) 51(2) UC Davis L Rev 653, 671; citing K P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 1.
[7] Stephen Wolfram, 'What is ChatGPT doing … and Why does it work?' (*Stephen Wolfram Writings*, 14 February 2023) <https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/> accessed 25 November 2024; Genevieve Bell and others, 'Rapid Response Information Report: Generative AI: language models and multimodal foundation models' (Australian Council of Learned Academies, 24 March

This chapter is structured as follows: the first part discusses the judicial value of impartiality, and what this requires of judges. As explained below, impartiality is also connected to judicial independence and to accountability. The following part considers how the value of impartiality interacts with AI, in four subsections: where AI appears as evidence in courts; where judges themselves use AI; where AI is used to scrutinise judges and/or their decision-making; and finally, whether judges can be replaced by AI. The final part concludes.

2. Impartiality

Impartiality is central to at least two conceptions of the rule of law. First, the rule of law requires that government is limited by or subject to law, so that contraventions are determined by the courts which necessitates the judiciary being independent from the rest of government.[8] Second is 'the rule of law not of man' perspective, where the law is applied objectively and is not subject to the foibles or biases of the individual(s) in power, or indeed the judge themselves,[9] requiring judges to interpret and apply the law in an unbiased manner.[10]

Impartiality is often discussed in connection with other core judicial values: independence (separation of powers), and the maintenance of public confidence in the judiciary, as the 'third arm' of government.[11] It has an institutional connection to independence, achieved through processes such as appointment, security of tenure and remuneration, and enabling

2023). Examples include Chat-GPT, Google Bard now Google Gemini, Microsoft Bing and 365 Copilot, and Anthropic Claude.

[8] Brian Tamanaha, *On the Rule of Law: History, Politics, Theory* (Cambridge University Press 2004) 117.

[9] Henry Ballantine, *Blackstone's Commentaries* (Blackstone Institute 1915) 366, 471; Tom Bingham, *The Rule of Law* (Penguin Books 2011) 51 ('The job of judges is to apply the law, not to indulge their personal preferences.').

[10] Tamanaha (n 8) 122–5; Raphael Franck, 'Judicial Impartiality in Politically Charged Cases' (2018) 29(2) Constitutional Political Economy 193, 194.

[11] Jessica Conser, 'Achievement of Judicial Effectiveness through Limits on Judicial Independence: A Comparative Approach' (2005) 31(1) NC J Int'l L 255, 260–61; Richard Devlin and Adam Dodek, 'Regulating Judges: Challenges, Controversies and Choices' in Richard Devlin and Adam Dodek (eds), *Regulating Judges: Beyond Independence and Accountability* (Edward Elgar 2016) 9.

impartiality.[12] Writing of India, Sengupta has said that 'independence and accountability are two independent variables which are both relevant to impartial adjudication';[13] Cartier has said that judicial independence is a 'necessary condition of impartiality'.[14]

Individual impartiality, meaning freedom from improper influence and being open-minded, is the duty of a judge.[15] This is supported by the rule against bias, which requires that a judge not hear a case if a fair-minded lay observer might reasonably apprehend that the judge might not bring an impartial mind to the resolution of the question that the judge is required to decide.[16] The US Supreme Court found a clear example of bias where a judge seeks to preside over a case in when he/she has a 'direct, personal, substantial, pecuniary interest'.[17] Another is where a judge has been so personally 'enmeshed in matters' concerning one party that the judge is biased against that party.[18]. Impartiality and the appearance of impartiality are necessary for the maintenance of public confidence in the judicial system.[19] Yet, as the Australian Law Reform Commission commented in a report into judicial impartiality, 'Judges are human, and the public knows it: Judges, and the public they serve, have recognised that

[12] *MacKeigan v Hickman* [1989] 2 SCR 796, [80]-[81] (Supreme Court of Canada); Murray Gleeson, 'Who Do Judges Think They Are?' (1998) 22 Criminal Law Journal 10, 11; cf. Caterina Mugelli, 'Judicial Independence in China: A Comparative Perspective' (2013) 54(1) Acta Juridica Hungarica 40, 56–57 ('there is no specific legal provision providing judges with a guaranteed income' in China).
[13] Arghya Sengupta, 'Judicial Independence and the Appointment of Judges to the Higher Judiciary in India: A Conceptual Enquiry' (2011-2012) 5 Indian Journal of Constitutional Law 99, 107.
[14] Emmanuel Cartier, 'Judicial Independence: A French Paradoxical Pattern in the European Field' in Shimon Shetreet, Hiram Chodosh and Eric Helland (eds), *Challenged Justice: In Pursuit of Judicial Independence* (Brill 2021) 181; quoting Charles Eisenmann, *La justice constitutionnelle et la Haute-Cour constitutionnelle d'Autriche* (rev edn, Economica 1986) 176. See also Nathaniel Yong-Ern Khng, 'Judicial Independence and the Singapore Judiciary' [2012] Lawasia Journal 53, 60 (making the same point about Singapore).
[15] *Valente v R* [1985] 2 S.C.R. 673 (Supreme Court of Canada) (Impartiality refers to a state of mind or attitude of the tribunal in relation to the issues and the parties in a particular case.); Conser (n 11); David Malcolm, 'Independence of the judiciary in the Asia-Pacific Region' (2004) 78(7) Australian Law Journal 458, 461; Bingham (n 9) 93.
[16] Abimbola Olowofoyeku, 'Regulating Supreme Court Recusals' [2006] Singapore Journal of Legal Studies 60, 85; attributing the approach to *Liteky v US* 510 US 540 (1994); *Taylor v Lawrence* [2002] EWCA Civ 90 (CA); *Lawal v Northern Spirit Ltd*. [2003] UKHL 35; *Laws v Australian Broadcasting Tribunal* (1990) 170 CLR 70 (HCA); *Johnson v Johnson* [2000] HCA 48; *Jaipal v State* [2004] ZASCA 45 (S. Afr. SC); *Valente v The Queen* [1985] 2 SCR 673 (Canada SC).
[17] *Tumey v Ohio*, 273 US 510, 523 (1927).
[18] *Johnson v Mississippi*, 403 US 212, 215 (1971).
[19] Conser (n 11); *North Australian Aboriginal Legal Aid Service v Bradley* (2004) 218 CLR 146, 162–3 (McHugh, Gummow, Kirby, Hayne, Callinan and Heydon JJ).

human decision-making can never be completely neutral.'[20] 'Perfect' impartiality is not required, as some degree of bias is part of everyone's psychological make-up. Rather, the Commission explained, the purpose of the law is to 'define the point at which the risk or appearance of an improper influence on decision-making is unacceptable to maintaining public confidence in the administration of justice'.[21] And as Higgins and Levy have pointed out: 'Despite their differences, legal bias and psychological bias have at least one crucial objective in common: ensuring decisions are not influenced by irrelevant factors'.[22] Giving reasons for decisions is one way of guarding against bias in decision-making.

3. Impact of AI

AI systems may operate as a support system, providing information, assistance, and guidance to individuals within the justice system, including judges. AI may also develop to a point where it can replace tasks traditionally performed by humans, of relevance here, the exercise of judicial power to determine guilt and resolve disputes.[23] Typically these uses of AI are examined from the perspective of dangers to judicial impartiality. This chapter includes that perspective. However, AI may also offer a check on human bias in the judicial process, as well as enable scrutiny of patterns of judicial decision-making, bolstering judicial impartiality.

[20] Australian Law Reform Commission, 'Without Fear or Favour: Judicial Impartiality and the Law On Bias – Summary Report' (ALRC Report 138, December 2021) 15.
[21] Ibid 15.
[22] Andrew Higgins and Inbar Levy, 'Judicial Policy, Public Perception, and the Science of Decision Making: A New Framework for the Law of Apprehended Bias' (2019) 38(3) CJQ 376, 382.
[23] See, eg, Vera Lúcia Raposo, 'The Digital "To Kill A Mockingbird": Artificial Intelligence Biases in Courts' (2024) 54 California Western International Law Journal 459, 461; Christoph K Winter, 'The Challenges of Artificial Judicial Decision-Making for Liberal Democracy' in Piotr Bystranowski, Bartosz Janik and Maciej Próchnicki (eds), *Judicial Decision-Making: Integrating Empirical and Theoretical Perspectives* (Springer Nature 2022) 179, 180; Benjamin Minhao Chen, Alexander Stremitzer and Kevin Tobia, 'Having Your Day in Robot Court' (2022) 36 Harv J L & Tech 127, 128.

3.1. Judges Assessing AI

In 2016, online news agency ProPublica published a story which propelled machine learning in the US criminal justice system to global headlines.[24] The story concerned AI programs used for risk assessments, which in turn, were used to inform judicial sentencing decisions.[25] The US is not alone – the use of such tools has been described as 'an international trend'.[26] However, the EU AI Act has prohibited AI systems that '[assess] the risk of an individual committing criminal offenses solely based on profiling or personality traits' though with the significant caveat 'except when used to augment human assessments based on objective, verifiable facts directly linked to criminal activity'.[27]

The issues that ProPublica identified with the use of the 'algorithmic risk assessment' were that the program was biased against particular racialized groups and its predictive validity was middling.[28] The creation of AI models involves a multistep process beginning with the selection of appropriate data. Machine learning algorithms, trained on data that often reflects societal biases and inequalities, risk replicating and reinforcing these issues. The ProPublica story also generated discussion about the nature of 'bias' and 'fairness' and how these terms should be used, in machine learning and in wider society.[29] Kleinberg and others have

[24] Julia Angwin and others, 'Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks', (*ProPublica*, 23 May 2016) <www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 25 November 2024.
[25] Angwin and others (n 24), discussing Correctional Offender Management Profiling for Alternative Sanctions (COMPAS). See also the response of Anthony W Flores, Kristin Bechtel and Christopher T Lowenkamp, 'False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks"' (2016) 80(2) Fed Probation 38.
[26] Kelly Hannah-Moffat, 'Actuarial Sentencing: An "Unsettled" Proposition' (2013) 30(2) JQ 270, 271 (citations omitted).
[27] 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] OJ L 12.7.2024, Article 5.
[28] See generally Seena Fazel, 'The Scientific Validity of Current Approaches to Violence and Criminal Risk Assessment' in Jan W de Keijser, Julian V Roberts and Jesper Ryberg (eds), *Predictive Sentencing: Normative and Empirical Perspectives* (Bloomsbury Publishing 2019).
[29] E.g., Laurel Eckhouse and others, 'Layers of bias: A unified approach for understanding problems with risk assessment' (2019) 46(2) Criminal Justice and Behavior 185.

explained that there are different ways to conceive of, and measure, 'fairness', and as there are 'inherent trade-offs' involved in each, the different methods cannot be reconciled with one another.[30] However, there is also a broader context, as Solow-Niederman et al explained, going beyond data to the many human decisions which are made in the course of developing the tool, as well as the very fact of using the data of multiple other individuals to make predictions about one individual:

> The use of an algorithm might be more standardized than the older clinical assessment based on, say, the look in the person's eye. Yet it does not make the predictive enterprise objective; rather, it turns on a spate of human choices about what data to use, what statistical model to adopt, how to "tune" the model, and how to apply the findings.[31]

A further issue raised is that where AI programs are not transparent (either because they are commercially protected, or by lacking interpretability), neither judges, nor defence lawyers, or even prosecutors, can possibly understand exactly how a prediction or recommendation was arrived at.[32] This is predominantly an issue of due process, noting the critique of Solow-Niederman that transparency would not solve the issue of individuals lacking the means to challenge the operation or results of AI systems.[33] As Solow-Niederman points out, invoking Tom Tyler's work on procedural justice, it is hard for people to have faith in the operation of the law in such circumstances.[34]

[30] Jon Kleinberg, Sendhil Mullainathan and Manish Raghavan, 'Inherent Trade-Offs in the Fair Determination of Risk Scores' (2016) Proceedings of Innovations in Theoretical Computer Science, arXiv:1609.05807.
[31] Alicia Solow-Niederman, YooJung Choi and Guy Van den Broeck, 'The Institutional Life of Algorithmic Risk Assessment' (2020) 34(3) Berkeley Tech L J 705, 713.
[32] Han-Wei Liu, Ching-Fu Lin and Yu-Jie Chen, 'Beyond State v Loomis: Artificial Intelligence, Government Algorithmization and Accountability' (2019) 27(2) IJLIT 122.
[33] Alicia Solow-Niederman, 'Algorithmic Grey Holes' (2023) 5 Journal of Law and Innovation 116.
[34] Ibid 123–26.

In relation to decision-making, therefore, it has been argued that judges relying on biased (or potentially biased) risk assessments, and the inability (of anyone) to question and test this evidence, are at risk of their decision also being infected by bias or having the appearance of being biased, compromising their impartiality. Indeed, if the operation of an AI system is opaque and cannot be explained or tested it may not be permissible to adopt it as part of the judicial process. Yet jurisdictions struggling with heavy caseloads and inadequate government resourcing are likely to do so.

3.2. Judges Assisted by AI

As well as producing evidence that judges may need to consider, AI has the capacity to augment, structure, support or enhance human decision-making. 'Decision-support' systems might be deployed to ensure more consistency in judicial decision-making, or to make sure that judges are aware of relevant considerations when making determinations.

AI prediction using case law has a long history, and there are numerous examples where machine learning systems can achieve good accuracy.[35] Yet many factors, such as the partial nature of the datasets used to train such systems, also mean that such systems are not infallible. However, examining issues of procedural fairness, Volokh has said that: 'Our question should not be whether AI judges are perfectly fair, only whether they are at least as fair as human judges.'[36] The same question can be raised in relation to impartiality or absence of bias. A common response in relation to complaints about bias and inaccuracy in AI systems is to raise issues of human bias. It is argued that AI systems may actually be better placed to take account of masses of data, thereby generating predictions which are more accurate (and

[35] See Legg and Bell (n 4) ch 4 for a summary.
[36] Eugene Volokh, 'Chief Justice Robots' (2019) 68 Duke L J 1135, 1169.

therefore less biased, and more impartial) than those of people.[37] In the case of sentencing, it is suggested that the use of AI promotes consistency, excludes irrelevant factors and can reduce appeals.[38]

Psychological studies suggest that human decision-making, including that of judges, is susceptible to unconscious bias and dependent on heuristics (cognitive short cuts).[39] This is more likely to occur when we are exercising quick, automatic judgment (fast thinking), as opposed to considered, effortful thought (slow thinking).[40] The late Daniel Kahneman, and colleagues, explained that 'human predictions tend to be "noisy": given the same input, different people (or even the same person at different times) will make vastly different predictions'.[41] Using statistical formulae can remove this 'noise', or irrelevant factors, from the decision-making.[42] Hence, it is argued that properly designed and tested automated systems can enable human biases to be controlled for, or removed from, the decision-making process.[43] The pursuit of consistency has prompted China to declare the integration of AI into

[37] Mirko Bagaric and Dan Hunter, 'Enhancing the Integrity of the Sentencing Process through Artificial Intelligence Systems' in Jesper Ryberg and Julian Roberts (eds), *Principled Sentencing and Artificial Intelligence* (OUP 2022).
[38] David Pannick, 'Why No Offender Wants to Face a Judge Who Is Tired, Hungry or Disappointed' *The Times* (UK, 19 January 2017) <https://www.thetimes.co.uk/edition/law/why-no-offender-wants-to-face-a-judge-who-is-tired-hungry-or-disappointed-6bdxbm2w0> accessed 25 November 2024; cited by Chen, Stremitzer and Tobia (n 23), 133.
[39] Chris Guthrie, Jeffrey J Rachlinski and Andrew J Wistrich, 'Blinking on the Bench: How Judges Decide Cases' (2007) 93 Cornell L Rev 1, 13–19 (judges are as susceptible as others to relying on intuition rather than deliberative judgment); Shai Danziger, Jonathan Levav and Liora Avanim-Pesso, 'Extraneous Factors in Judicial Decisions' (2010) 108(17) Proceedings of the National Academy of Sciences 6889; Jon Kleinberg and others, 'Discrimination in the Age of Algorithms' (2018) 10 J Legal Analysis 113.
[40] Kahneman's 'system 1' and 'system 2' mental systems, respectively: see Daniel Kahneman, *Thinking Fast and Slow* (Penguin 2011).
[41] Daniel Kahneman and others, 'Reducing Noise in Decision Making' (2016) 94(12) HBR 18, 18; J Nathan Matias 'Bias and Noise: Daniel Kahneman on Errors in Decision-Making' (*Medium,* 18 October 2017), <www.medium.com/@natematias/bias-and-noise-daniel-kahneman-onerrors-indecision-making-6bc844ff5194> accessed 25 November 2024.
[42] Jongbin Jung and others, 'Simple Rules for Complex Decisions', (Stanford University Working Paper, 2017) <https://arxiv.org/pdf/1702.04690.pdf>.
[43] Jay Thornton, 'Cost, Accuracy, and Subjective Fairness in Legal Information Technology: A Response to Technological Due Process Critics' (2016) 91(6) NYU L Rev 1821, 1840, 1849; Nigel Stobbs, Dan Hunter and Mirko Bagaric, 'Can Sentencing Be Enhanced by the Use of Artificial Intelligence' (2017) 41 Crim LJ 261.

judicial processes a national priority, introducing, for example, precedent recommendation systems that assist human judges by formulating judgments based on past decisions.[44]

Proponents also anticipate efficiency gains.[45] Aletras and colleagues used machine learning to predict the outcomes of cases before the European Court of Human Rights.[46] These authors suggested that such a program could be used to triage cases going before judges, prioritising those with the greatest chance of success:

> The system may be used to rapidly identify cases and extract patterns that correlate with certain outcomes. It can also be used to develop prior indicators for diagnosing potential violations of specific Articles in lodged applications and eventually prioritise the decision process on cases where violation seems very likely. This may improve the significant delay imposed by the Court and encourage more applications by individuals who may have been discouraged by the expected time delays.[47]

It is reported that in Spain, similar processes are being used in selective types of case to generate decisions which judges may accept or reject.[48]

Humans are, however, also prone to biased thinking due to framing and anchoring effects – where a person's judgment is influenced by a reference point or 'anchor'. If judges believe

---

[44] Jinting Deng, 'Should the Common Law System Welcome Artificial Intelligence?: A Case Study of China's Same-Type Case Reference System' (2019) 3 Geo L Tech Rev 223, 224–6; see also Supreme People's Court, 'White Paper on Chinese Courts and Internet Judiciary' (Supreme Court of China White Paper, 5 December 2019); Ray Worthy Campbell, 'Artificial Intelligence in the Courtroom: The Delivery of Justice in the Age of Machine Learning' (2020) 18 Colo Tech LJ 323, 343; Jiahui Shi, 'Artificial Intelligence, Algorithms and Sentencing in Chinese Criminal Justice: Problems and Solutions' (2022) 33(2) Crim LF 121.

[45] Frank Fagan, 'Big Data Legal Scholarship: Toward A Research Program and Practitioner's Guide' (2016) 20(1) Virginia J Law and Tech 1, 2; Arthur Rizer and Caleb Watney, 'Artificial Intelligence Can Make Our Jail System More Efficient, Equitable, and Just' (2018) 23(1) Texas Review of Law and Politics 181.

[46] Nikolaos Aletras and others, 'Predicting judicial decisions of the European Court of Human Rights: A Natural Language Processing perspective' (2016) PeerJ Computer Science DOI:10.7717/peerj-cs.93.

[47] Ibid; See also the response of Frank Pasquale and Glyn Cashwell, 'Prediction, Persuasion, and the Jurisprudence of Behaviourism' (2018) 68(Supp 1) UTLJ 63.

[48] Tania Sourdin, 'Replacing, Supporting or Enhancing Judges? Judge AI Considerations for the Future' (2024) 98 Aust LJ 696; citing Juan-Luis Gómez Colomer, *El Juez Robot La independencia judicial en peligro* (Tirant lo Blanch Madrid 2023); José Martin Pastor, 'Retos de la justicia Digital' in Fernando Conde, Julio Banacloche and F Gascón Inchausti (eds), *Logros y retos de la justicia civil en España* (Tirant lo Blanch Madrid 2023).

that an AI program has a 90 per cent accuracy rate when it comes to classifying applications as either "meritorious" or "unmeritorious", are they likely to overturn a machine-generated decision and reinstate an unmeritorious case?

The rise of GenAI means that judges also have new options available to them when it comes to drafting judgments, with implications for impartiality. Research suggests that lawyers are less likely to trust material they believe has been authored by AI.[49] In 2023, an English Court of Appeal judge stated that he had used ChatGPT to draft a section of a judgment, praising it as a useful tool.[50] Conversely, it is suggested that judges using GenAI to draft judgments risks creating perceptions of bias and undermining public trust in their impartiality.[51]

The giving of reasons also acts as a protection for impartiality and addressing bias and heuristics. The extent of a judge's duty to state reasons for a decision is related to the function to be served by the giving of reasons, the importance of the point involved and the likely effect of the decision on the rights of parties to the proceedings. Reasons may be stated succinctly and need not be elaborate, but they need to be a sufficient explanation of a judge's reasoning process for a reader to understand why or in what way a conclusion was reached.[52] The provision of reasons can assist in ensuring that biases are negated. This is because:

[49] Jakub Harasta, Tereza Novotna and Jaromir Savelka, 'It Cannot be Right if it was Written by AI: On Lawyers' Preferences for Documents Perceived as Authored by an LLM vs a Human' (2024) Artificial Intelligence and Law, arXiv:2407.06798.
[50] Jane Dalton, 'Judge admits using 'jolly useful' ChatGPT to write court ruling', *The Independent* (UK, 15 September 2023) <https://www.independent.co.uk/news/uk/home-news/chatgpt-ai-judge-chatbot-ruling-b2412378.html> accessed 25 November 2024.
[51] Chief Justice Andrew Bell, 'Truth Decay and Its Implications for the Judiciary: An Australian Perspective' (Speech, Durham University, 23–26 April 2024); see also A S Bell, 'Guidelines For New South Wales Judges in Respect of Use of Generative AI' (Supreme Court of NSW, 21 November 2024).
[52] Michael Legg, *Case Management and Complex Civil Litigation* (2nd ed, Federation Press 2023) 342–44.

> Taking the time to write down our reasons, as many of you already know, assists us to gather and organise our thoughts coherently, logically and defensibly. It also provides the necessary pause, enabling us to engage in higher level, deliberative thinking.[53]

Deliberative thinking is the considered, 'slow thinking' identified above. That written reasons act as a debiasing tool has also been supported by experimental research.[54]

The giving of reasons also allows for another check on bias – appeal.

3.3. Judges Scrutinised by AI

Judges or courts may use AI to scrutinise the applications and cases before them. Yet, increasingly, judges and their decisions are likely to also become the scrutinised. There is a long history of 'predicting' the outcome of legal cases using statistical and machine learning analysis,[55] based on analysing patterns of judicial decision-making and characteristics of judicial officers themselves.[56] As well as predicting outcomes, analyses of this type might be used to illustrate or bring to light the existence of patterns of anomalous decision making among the judiciary.[57] Ghezelbash et al explain that, from a case management perspective, rigorous use of statistical analysis might also demonstrate where judges are overworked or under-resourced, noting that there is also potential correlation with bias: 'Numerous other studies demonstrate that time pressures and stress in the context of decision-making leads

[53] Chief Justice Gibbs Salika, 'Ethics - The Construct of Impartiality and Unconscious Bias' (Pacific Judicial Integrity Program Judicial Fraud & Anti-Corruption Workshop, Port Moresby, Papua New Guinea, 28 November 2022) 10. See also Frank Kitto, 'Why write judgments?' (1992) 66 Aust LJ 787; Chad M Oldfather, 'Writing, Cognition, and the Nature of the Judicial Function' (2008) 96 Geo L J 1283; Mathilde Cohen, 'When Judges Have Reasons Not to Give Reasons: A Comparative Law Approach' (2015) 72 Wash and Lee L Rev 483.
[54] Zhuang Liu, 'Does Reason Writing Reduce Decision Bias? Experimental Evidence from Judges in China' (2018) 47 J Legal Stud 83.
[55] See Legg and Bell (n 4) ch 6 for a summary.
[56] See, e.g., Daniel Martin Katz, Michael J Bommarito II and Josh Blackman, 'A General Approach for Predicting the Behavior of the Supreme Court of the United States' (2017) 12(4) PLoS ONE e0174698; Mihai Surdeanu and others, 'Risk Analysis for Intellectual Property Litigation' (Conference Paper, International Conference on Artificial Intelligence and Law, 6–10 June 2011).
[57] Daniel Ghezelbash, Keyvan Dorostkar and Shannon Walsh, 'A Data Driven Approach to Evaluating and Improving Judicial Decision-Making: Statistical Analysis of the Judicial Review of Refugee Cases in Australia' (2022) 45(3) UNSWLJ 1085.

people to consider alternatives less systematically and completely, and correlate with less accurate decisions.'[58]

It also seems increasingly likely that, given the sophistication now available, computational methods will found successful claims of bias or lack of impartiality among judges.[59] Resistance to this path is found in, for instance, the French prohibition on 'judge analytics',[60] and the Australian Federal Courts' rejection of statistics as capable of founding claims of bias.[61] Clopton and Huq also detail the partial nature of publicly-available US federal judicial data: 'Statistical reports issued by the Administrative Office and other entities are far from comprehensive. For example, the courts do not provide aggregate information on sentencing that reveals the identity of the sentencing judge or that is searchable by judge.'[62] They argue that the failure to make this information available seems to be based on protecting judges from public criticism, but that this is an insufficient reason to withhold information.[63]

Even without overturning individual judicial decisions, it is possible that such analyses might 'undermine broader concepts of judicial impartiality'.[64] A more positive use, however, it put forward by Chen, who

> ... cites research that shows that decisionmakers tend to reduce bias when alerted. Machines, equipped with a capacity to observe many variables at once, can detect and

[58] Ibid 1122; citing Jerry Kang, 'What Judges Can Do about Implicit Bias' (2021) 57(2) Court Review 78, 84.
[59] Ghezelbash et al (n 57) 1117; citing Matthew Groves, 'Bias by the Numbers' (2020) 100 AIAL Forum 60.
[60] *Justice Reform Act,* art 33 (*loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice (1)*) <www.legifrance.gouv.fr/eli/loi/2019/3/23/2019-222/jo/article_33>.
[61] See *ALA15 v Minister for Immigration and Border Protection* [2016] FCAFC 30, [38]–[44] (Allsop CJ, Kenny and Griffiths JJ); Law Council of Australia, 'Judicial Impartiality: Consultation Paper Australian Law Reform Commission ' (Submission No 37 to ALRC Review of Judicial Impartiality, 8 July 2021).
[62] Zachary D Clopton and Aziz Z Huq, 'The Necessary and Proper Stewardship of Judicial Data' (2024) 76 Stan L Rev 893, 922.
[63] Ibid, 946–47.
[64] Monika Zalnieriute and Felicity Bell, 'Technology and the Judicial Role' in Gabrielle Appleby and Andrew Lynch (eds), *The Judge, the Judiciary and the Court: Individual, Collegial and Institutional Judicial Dynamics in Australia* (Cambridge University Press 2020) 116, 136.

> uncover various biases unbeknownst to the judge who perpetuates them. Once notified, that judge can engage in introspection and self-correction.[65]

In this view, the fact that judges may be scrutinised in new ways should lead to judges taking even more care to act in an unbiased fashion.

3.4. Judges Replaced by AI?

In this final section, we consider the direction of travel. Is automated decision-making on the way to replacing judges? And would this be a good thing for impartiality, or not?

Volokh used a thought experiment to discuss whether AI could replace judges (assuming the AI judge could produce reliable judgments) by focussing on the quality of the proposed AI judge's product, not on the process that yields that product, and assessing quality through a metric of persuasiveness.[66] Volokh argues that society should focus on the results not the process. Further, the AI judge would need to provide reasons supporting the decision but not its actual reasoning, because its actual reasons are really the explanation of the operation of its algorithm that produced the result and AI does not reason/function in the same way that a human does, and a human may find the actual reasons incomprehensible.[67] The reasons need to persuade expert evaluators, i.e., human judges. Volokh adds that if the 'evaluators are persuaded that the AI judge's decisions are just, equal, and impartial, it shouldn't matter whether this stems from programming or from an oath'.[68]

[65] Summary of Frank Fagan, 'Natural Language Processing for Lawyers and Judges' (2021) 119 Mich L Rev 1399, 1407; citing Daniel Chen, 'Machine Learning and the Rule of Law' in Michael A Livermore and Daniel N Rockmore (eds), *Law As Data: Computation, Text, and the Future of Legal Analysis* (Santa Fe Institute Press 2019) ch 16.
[66] Volokh (n 36) 1192.
[67] Ibid 1164.
[68] Ibid 1162–63.

The ‘experiment’ apparently did not last long as in 2019, in Hangzhou, China, an AI powered ‘judge’ named Xiaozhi resolved a lending dispute. In just thirty minutes, Xiaozhi questioned the parties, analysed their evidence, and issued a ruling.[69]

From an impartiality perspective, persuasive reasons miss the point. Impartiality guards against improper influences on decision-making. The focus on persuasiveness to overcome the AI black box does not address bias. AI might be tainted by the biases of its programmers, by the data it was trained on, so that it recycles previous injustices, or it may develop bias.[70] Moreover, persuasive ‘reasons’ may act or seek to hide bias in aid of persuasiveness.

In short, we may not be able to tell if the AI judge is impartial or biased. A human judge is told that they must be unbiased, and if they are biased, or it may appear so to the reasonable observer, then they should recuse themselves. It seems unlikely an AI judge would have a conscience upon which the obligation to do the right thing could operate. Equally, a party before a human judge who believes that a judge is partial may seek to have that judge recused based on evidence and argument as to the existence or perception of bias. A party would have no way of challenging an AI judge in relation to bias.

4. Conclusion

In this chapter we have discussed the fundamental importance of judicial impartiality and its connection to other key judicial values such as independence. We adopted a definition of impartiality which focused on an absence of bias in decision-making. We then considered four ways that judges and decision-makers might (and in some instances, do) interact with AI systems where that interaction may impact on impartiality.

[69] David Horton, ‘Forced Robot Arbitration’ (2024) 109 Cornell L Rev 679, 680; citing Nu Wang, ‘“Black Box Justice”: Robot Judges and AI-based Judgment Processes in China's Court System’ (2020) 2020 IEEE Int'l Symposium on Tech & Soc 58, 62.
[70] Volokh (n 36) 1167–69; Horton (n 69) 683.

As we explained, fairness and absence of bias are themselves contested concepts. We can think of fairness as occurring in layers, where the base layer references the idea of making decisions about individuals based wholly on aggregations of data about others.[71] In many judicial systems, this is not acceptable, as individualised case features must be considered. Further, where new decisions are based wholly on past data, they can naturally only be iterations of what has previously occurred.

The goal of promoting consistency in decision-making must give way to the need to consider each case on its own merits. Hence, the EU AI Act deems as 'high risk' 'AI systems intended to be used by a judicial authority or on their behalf to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts, or to be used in a similar way in alternative dispute resolution'.[72] However, a system which is 'intended to detect decision-making patterns or deviations from prior decision-making patterns and is not meant to replace or influence the previously completed human assessment, without proper human review' is exempted (i.e., is not deemed high risk).[73] This would allow the use of AI to assist or support decision-makers, but with various safeguards in the case of high risk systems.

Judges are human, and they may have their own biases. Yet they also inhabit a system designed to put a check on unacceptable expression of those biases. Judicial impartiality requires a decision-maker to approach each decision without prejudgment, to weigh up incommensurable and value-laden concepts, and to explain their reasoning – not just provide a plausible justification.

[71] Eckhouse and others (n 29) 185–209.
[72] Artificial Intelligence Act (n 27), art 6(2) and annex III: High-Risk AI Systems Referred to in Article 6(2).
[73] Ibid art 6(3)(c).