NUS Law Working Paper 2026/005

Same Words, Different Worlds: The Illusion of Shared Judicial AI Principles

Ernest Lim

Ilya Akdemir

ernestlim@nus.edu.sg
ilya.akdemir@nus.edu.sg

Professor, Faculty of Law, National University of Singapore

Lecturer, Faculty of Law, National University of Singapore

[Uploaded April 2026]

University of Toronto Law Journal (2026) (forthcoming)

Same Words, Different Worlds[1]: The Illusion of Shared Judicial AI Principles

Ernest Lim* and Ilya Akdemir**

Abstract: Global AI governance frameworks increasingly deploy shared terminology such as "transparency", "bias" and "fairness", creating an illusion of alignment across jurisdictions. Through comparative analysis of AI deployment in common law courts (the U.S., the U.K., Australia) and Chinese courts, we expose how seemingly universal concepts mask fundamentally divergent interpretations and implementation.

This article makes three distinct contributions to legal scholarship. First, it provides a critical and comprehensive doctrinal and jurisprudential analysis of why and how judicial AI principles operate differently across legal systems. We demonstrate that common law courts define "reliability" as case-specific accuracy verified through adversarial contestation and ex post sanctions, while Chinese courts prioritize systemic reliability via centralized oversight. We show that "bias" threatens individual rights in adversarial systems but endangers social harmony in China. We explain that "transparency" enables party contestation in common law jurisdictions but renders processes legible to supervisory authorities in China.

*Chan Sek Keong Professor of Private Law, Co-Director, Center for Technology, Robotics, AI & the Law (TRAIL), Faculty of Law, National University of Singapore; Faculty Associate, Berkman Klein Center for Internet & Society, Harvard University

**Lecturer in Law, Faculty of Law, National University of Singapore

We would like to thank the participants at the Harvard Law School AI and Law conference for their insights and provocations, and in particular to the organizers Shafi Goldwasser, Martha Minow and Cass Sunstein for selecting this paper for presentation. For detailed, constructive comments on the paper, we are grateful to the two UTLJ's anonymous referees, Simon Chesterman, Justice Anselmo Reyes, Min Yan, Jeannie Paterson, Iris Eisenberger, Michaela Windisch-Graetz, Lin Lin, Julian Webb, Pengfei Shen, as well as participants at the Melbourne Law School Centre for Artificial Intelligence and Digital Ethics research workshop, the MLS-NUS judicial roundtable, the NUS Faculty Research Workshop, the Vienna-NUS Law and Technology Inaugural Workshop, and the Center for Transnational Legal Studies Global Conference. For helpful discussions that laid the foundation for this paper, we would like to express our appreciation to Sandy Steel, Catherine Valcke, Jia Qing Yap, and Alon Harel. Finally, we thank Cheok Teng Ann and Jonathan Gwee for research assistance funded by TRAIL.

[1] The title "Same words, different worlds" was inspired by De Chirico, Leonardo. SAME WORDS, DIFFERENT WORLDS: DO ROMAN CATHOLICS AND EVANGELICALS BELIEVE THE SAME GOSPEL? (InterVarsity Press, 2021).

Second, we move beyond comparative law by identifying three extra-legal forces driving these divergences: competing efficiency logics prioritizing different judicial functions; AI as "inherently political technology" presupposing particular institutional arrangements; and archival-epistemic structures determining authority over legal knowledge. We address the paradox of why common law judges deploy AI in sentencing and bail decisions but resist AI deployment for core adjudicatory functions. We also analyze the puzzle as to why they perceive AI as threatening their authority in contrast to Chinese judges. This analysis has broad applicability beyond AI to other transformative technologies in law.

Finally, our analysis has important practical implications for global AI governance. We demonstrate why international frameworks cannot succeed through harmonized terminology or technical standards alone. Due to discrepant frames underpinning common law and Chinese perspectives on AI, we suggest that successful global AI instruments require a frame-reflective reciprocal translational approach.

Table of Contents

I. Introduction

As AI governance becomes increasingly globalized, there is an implicit assumption that regulatory terms, ethical principles, and policy frameworks carry stable meanings across jurisdictions. Instruments such as the EU AI Act, the Colorado AI Act, the South Korea AI Basic Act, the UNESCO Guidelines for the Use of AI systems in Courts and Tribunals, the OECD AI Principles, and the G7 Hiroshima Process International Code of Conduct, give the impression that a shared vocabulary is emerging, whereby the circulation of common terms such as for instance "transparency", "bias" and "fairness" ensures a shared conceptual foundation. Yet such an assumption can be deeply misleading.

Identical terms can have sharply different meanings depending on semantic, ideological, political, institutional and other contexts. For example, in a celebrated study on political speech, Monroe et al. have empirically demonstrated that Republicans' and Democrats' Congressional speeches on the topic of 'abortion' populate distinct semantic fields and thus exhibit different lexical and semantic associations[2] – Republicans emphasize the "baby/child" dimension, whereas Democrats emphasize "women's choice" dimension.[3] In other words, the same word does not guarantee the same field of meaning - context, associated concepts, politics, ideology and innumerable other factors - all influence understanding and interpretation.

From this semantic field perspective, the construction of harmonized global AI regulations or frameworks on the basis of ostensibly shared concepts and common AI problems are not only insufficient but can be misleading. What matters are not just concepts, but the aspects that mobilize their use within each jurisculture[4] - the legal, institutional, sociolegal, historical, political, ideological and other conditions that

---

[2] Eva Feder Kittay, *Semantic Fields and the Individuation of Content*, in FRAMES, FIELDS, AND CONTRASTS 229 (Routledge 1992), defining semantic fields as "clusterings of lexicalized concepts." Dirk Geeraerts, THEORIES OF LEXICAL SEMANTICS 52 (OUP Oxford 2009), stating that "a lexical field, then, is a set of semantically related lexical items whose meanings are mutually interdependent, and which together provide conceptual structure for a certain domain of reality."

[3] Burt L. Monroe, Michael P. Colaresi & Kevin M. Quinn, *Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict*, 16 POL. ANAL. 372 (2008). Benoit Berthelier, *Division and the Digital Language Divide: A Critical Perspective on Natural Language Processing Resources for the South and North Korean Languages*, 47 KOREAN STUD. 243 (2023).

[4] Pierre Legrand, *European Legal Systems Are Not Converging*, 45 INT'L & COMP. L.Q. 52 (1996). Here, we understand jurisculture as legal culture – and according to Legrand, "rules are but the outward manifestation of an implicit structure of attitude and reference, they are a reflection of a given legal culture." Pierre Legrand *Foreign Law, the Comparatist, and Culture: How It Is*, in LAW, CULTURE AND IDENTITY IN CENTRAL AND EASTERN EUROPE 23, 30 (Cosmin Cercel, Alexandra Mercescu & Mirosław Michał Sadowski eds., Routledge 2023), stating that "everywhere, one finds sets of learned elements that are shared more extensively by people who interact with one another – and who have been interacting with one another over the very long term – and that differ from other sets of learned elements to be found in other people with whom there has not been such a significant level of interaction." It is important to note that according to Legrand, "a legal culture is not monolithic."

determine meaning across jurisdictions.[5] Thus, a deeper analysis of AI principles and concepts and how, but more importantly, why their meaning varies across jurisdictions is therefore essential. Without an understanding of the divergent semantic fields from which these concepts draw their meaning, and the ways in which they are subsequently framed, interpreted and deployed, any claim to a global regulatory instrument on AI will result, at minimum, in misunderstandings, intranslatability and ineffectiveness.

These stakes are not merely academic. When governance frameworks built on assumed shared meaning are enacted, concrete problems may follow. Jurisdictions may adopt international AI standards believing they have agreed on safeguards for "fairness" or "transparency," only to discover that implementation diverges so fundamentally that the standards provide no meaningful protection. Individuals subject to AI-assisted adjudication may find that the "transparency" provided for in international instruments affords them no ability to understand or challenge the algorithmic reasoning applied to their case. More broadly, superficial terminological consensus can mask what amounts to conceptual imposition: when one jurisdiction's understanding of "bias" or "reliability" is implicitly treated as universal, governance frameworks risk serving particular interests rather than enabling genuine cross-jurisdictional cooperation. Exploring how and why these divergences arise is accordingly important not only for scholarly advancement, but also for preventing governance failure with respect to the administration of justice as well as the rights, liberty, and dignity of those subject to AI-assisted judicial decisions.

For our analysis, we examine why and how AI concepts like "reliability", "bias", "fairness" and "transparency" are framed, understood and acted upon similarly or differently in the common law courts (in the U.S., the U.K., Australia) and Chinese courts. We examine the judicial guidelines, case law, policy documents, academic articles and other relevant materials from these jurisdictions concerning the use of AI in legal contexts, which allows us to explore the jurisdiction-specific fields of meaning within which the aforementioned AI-related concepts and principles are embedded, understood and operationalized. We compare common law with Chinese courts because China arguably represents the world's most advanced AI adoption in adjudication, "leapfrogging efforts elsewhere"[6] with systems like Faxin 2.0, Shanghai 206, Smart Court infrastructure and most recently, the use of large language models (LLMs) for assistance in writing legal decisions in Shenzhen province.[7] If the understanding and interpretation of AI governance principles

[5] For a similar problem in the behavioural sciences, see: Joseph Henrich, Steven J. Heine & Ara Norenzayan, *Most People Are Not WEIRD*, 466 Nature 29 (2010). The problem of false universality is not unique to law - the so-called "WEIRD" problem in behavioural sciences demonstrates how psychological findings are in fact derived almost entirely from Western, Educated, Industrialized, Rich, and Democratic populations (mostly American undergraduates) and are thus not representative of other populations of the world.

[6] Rachel E. Stern, Benjamin L. Liebman, Margaret E. Roberts & Alice Z. Wang, *Automating Fairness? Artificial Intelligence in the Chinese Courts*, 59 COLUM. J. TRANSNAT'L L. 515, 519 (2021); Straton Papagianneas & Nino Junius, *Fairness and Justice Through Automation in China's Smart Courts*, 51 COMP. L. & SEC. REV. 105897 (2023).

[7] John Zhuang Liu & Xueyao Li, *How Do Judges Use Large Language Models? Evidence from Shenzhen*, 16 J. LEGAL ANAL. 235 (2024).

diverge fundamentally between common law and the most advanced AI-adopting legal system, this may provide evidence for the difficulty of developing objective technical and legal standards. The comparison is also important as China is arguably "leading the world on AI governance"[8], and therefore, understanding how Chinese courts operationalize "transparency" or "fairness" is essential for assessing whether global standards can bridge or merely obscure irreconcilable differences.

Thus, this article makes three distinct contributions to legal scholarship. First, it provides a critical and comprehensive doctrinal and jurisprudential analysis of why and how judicial AI principles operate differently across legal systems. We show that, due to the underlying legal systems and philosophical commitments, perceptions of "reliability" reflect a divergence between individualized and aggregate frames in common law and Chinese courts, respectively. As for "bias," divergences arise from individualized versus systemic frames. Issues of "transparency" are understood as questions of external versus internal visibility of judicial conduct. And questions over "procedural fairness" hinge on the balance between due process and governance in common law and Chinese contexts respectively, whereby, for example, AI output similarity and convergence[9] can actually be seen as a positive or a negative phenomenon depending on the frame.

Our second contribution is that we move beyond comparative legal analysis by identifying three extra-legal forces driving these divergences: competing efficiency logics that prioritize different judicial functions; AI as an "inherently political technology" presupposing particular institutional arrangements; and archival-epistemic structures determining who holds authority over legal knowledge. We identify and address the paradox of why common law judges embrace AI in sentencing and bail decisions but resist AI when it comes to core adjudicatory functions. We also address the puzzle as to why common law judges perceive AI as threatening their custodianship of precedent while Chinese judges embrace it for standardization. This analysis has broad applicability beyond AI to other transformative technologies in law.

---

[8] Nature, "China is Leading the World on AI Governance: Other Countries Must Engage" (December 10, 2025), available at: https://www.nature.com/articles/d41586-025-03972-y. For example, China recently announced the Global AI Governance Plan in 2025 to promote industry collaboration and sustainable AI development while establishing standards for safety, governance, data protection, and ethical practices to ensure responsible AI use: Global AI Governance Action Plan, World Artificial Intelligence Conference & High-Level Meeting on Global AI Governance (July 26, 2025), available at: https://us.china-embassy.gov.cn/eng./zggs/202507/t20250730_11679377.htm. China also proposed to create the World Artificial Intelligence Corporation Organization to bring nations together to create a global AI governance system.

[9] Natalia Restrepo & Diego Socol, *Artificial intelligence at the bench: Legal and ethical challenges of informing—or misinforming—judicial decision-making through generative AI*, 6 DATA & POLICY, (Dec. 2024), available at: https://www.cambridge.org/core/journals/data-and-policy/article/artificial-intelligence-at-the-bench-legal-and-ethical-challenges-of-informingor-misinformingjudicial-decisionmaking-through-generative-ai/D1989AC5C81FB67A5FABB552D3831E46.
Jiang, Liwei, Yuanjun Chai, Margaret Li, Mickel Liu, Raymond Fok, Nouha Dziri, Yulia Tsvetkov, Maarten Sap, Alon Albalak & Yejin Choi, *Artificial Hivemind: The Open-Ended Homogeneity of Language Models (and Beyond),* arXiv preprint arXiv:2510.22954 (2025).

The final contribution is that our analysis has important practical implications for global AI governance. We demonstrate why international frameworks cannot succeed through harmonized terminology or technical standards alone. Due to the "discrepant frames"[10] that underpin the common law and Chinese perspectives on AI, we suggest that for a successful global AI instrument to exist, a frame-reflective translational approach should be adopted, whereby those involved in creating such regulatory frameworks "bring to consciousness"[11] how framing influences the formulation of questions concerning AI in each jurisculture. By explicitly recognizing the other, globalized AI instruments will, we hope, not only be able to address the threats posed by AI, but will also be viewed as being more than just an imposition of rules and assumptions from one jurisculture upon another.

One might object, however, that comparing common law with Chinese courts conflates *legal-systemic* differences (adversarial versus inquisitorial procedure) with *political-regime* differences (liberal democracy versus socialist authoritarianism), and thus, comparing common law with liberal European civil law jurisdictions would better isolate the former. We disagree for both methodological and substantive reasons. Methodologically, our thesis concerns the illusion of shared meaning created by superficially harmonized terminology. This illusion is most consequential, and most clearly exposed, when comparing common law and Chinese systems that appear to share global governance frameworks, while operating from fundamentally different premises. Thus, it is not illuminating for the purposes of this article to compare common law courts with leading European courts (such as those in France, Germany, Italy and Switzerland) as they generally share with common law systems the constitutional commitments to democracy, separation of powers, and individual rights[12], and thus, are likely to have overlapping understandings of what AI transparency and fairness mean in adjudication. Substantively, the legal/political distinction is itself analytically questionable here: as Part III suggests, AI governance in adjudication is political, presupposing and perpetuating particular distributions of authority. The procedural architecture of common law adversarialism is not merely a neutral "legal design" separable from liberal democratic commitments to individual autonomy, separation of powers, and constraints on state authority.[13] Similarly, Chinese "socialist rule of law with Chinese characteristics" is not an overlay on an otherwise neutral inquisitorial system; it

[10] Donald Schön & Martin Rein, FRAME REFLECTION: TOWARD THE RESOLUTION OF INTRACTABLE POLICY CONTROVERSIES 57 (Basic Books 1994).

[11] Ibid.

[12] Ran Hirschl, COMPARATIVE MATTERS: THE RENAISSANCE OF COMPARATIVE CONSTITUTIONAL LAW 112-150 (Oxford Univ. Press, 2014) (demonstrating the common intellectual genesis of comparative constitutional inquiry in Europe, the U.S., and Canada).

[13] On the constitutive relationship between procedural forms and political ideology, see Mirjan R. Damaška, THE FACES OF JUSTICE AND STATE AUTHORITY: A COMPARATIVE APPROACH TO THE LEGAL PROCESS 71-96, 181-239 (Yale Univ. Press, 1986) (arguing that adversarial procedure reflects the "reactive state" ideology of Anglo-American liberalism, where state activity is conceived as dispute resolution between autonomous individuals, while inquisitorial procedure reflects "activist state" ideology where proceedings serve policy implementation).

is constitutive of how that system conceptualizes the relationship between individual, court, and state.[14] Excluding China because its differences are political rather than juridical would precisely reproduce the formalist assumption our article questions—that legal concepts can be abstracted from their political-institutional-ideological embedding.

If the above potential objection pertains to comparative methodology ("you're mixing procedural differences with political differences", which we reject), another possible objection is the conceptual incommensurability critique, dressed in normative language: the comparison is inapt because common law systems have a different or superior conception of the rule of law or role of law as compared to that in China. The second objection assumes that valid comparative analysis requires similar underlying systems. But our thesis is that shared AI terminology masks divergent meanings. To demonstrate this, one must compare systems with different foundational premises. Further, global AI governance necessarily involves jurisdictions with different rule of law conceptions. Our comparison precisely calls into question one-size-fits-all standards. Finally, the objection conflates normative assessment with descriptive comparison. We do not argue or assume that common law and Chinese approaches are equally desirable or that Chinese courts have equivalent legitimacy. Rather, we argue that both systems use the same AI-related vocabulary while meaning and deploying fundamentally different things. This is a descriptive, analytical claim, not necessarily an endorsement. To reject the comparison on rule of law grounds is to refuse to understand how systems different from the common law, particularly one that has arguably the most advanced use of AI in adjudication, actually deploy technology. This is dangerous for AI governance: if we cannot analytically engage with how Chinese courts address issues related to AI "fairness", "bias" or "transparency", we cannot anticipate AI governance failures or negotiate effectively in international fora.

Rather than treating AI and related issues as universally established and understood concepts, our aim is to show that the meaning of these concepts is itself a product of the various forces that generate their manifestations. As Legrand observes, a comparatist must "understand why a particular legal culture has been attracted to a particular genre of cultural product"[15] – which, when applied to our examination of AI divergence, means exploring why certain manifestations of AI are introduced in and "are attractive to" some juriscultures but not others. Legrand's 'alterity' theory allows us to recognize and acknowledge that the otherness of each legal culture and interpretive community is largely irreducible, that the appeal of particular AI regulatory approaches are contingent upon a number of factors including historically, culturally and institutionally grounded patterns of thought, among others, and that only through sustained interpretive

[14] On socialist rule of law as constitutive rather than incidental to Chinese legal order, see Randall Peerenboom, CHINA'S LONG MARCH TOWARD RULE OF LAW 42-43, 63-65 (Cambridge Univ. Press, 2002) (distinguishing "rule of law" from "rule by law" and analyzing how the instrumentalist conception of law as governance tool shapes the individual-state relationship in Chinese courts).

[15] Pierre Legrand, *Alterity: About Rules, For Example*, *in* THEMES IN COMPARATIVE LAW: IN HONOUR OF BERNARD RUDDEN 21 (Peter Birks & Arianna Pretto eds., 2002).

engagement can we begin to grasp the trajectories that determine the reception of this technology. Thus, we should seek understanding and not convergence.

In Part II, we first define our use of AI. We then critically examine the different ways in which common law and Chinese courts interpret and implement core judicial values and concepts. In Part III, we analyse the extra-legal reasons for these divergences. In Part IV, we consider the implications of our analysis for global AI governance. Part V concludes.

II. Converging AI Policies, Diverging Interpretation and Implementation

A. Definition of AI

When attempting to define "artificial intelligence" in judicial contexts one is overwhelmed by the sheer number of AI tools and use cases across jurisdictions, all of which, despite being very different technically, are colloquially labeled "AI." These can include for example, technology assisted review, automated online dispute resolution, automated e-filing, risk assessment tools, automated decision-making, and prediction of litigation outcomes. In current discourse, the meaning of "AI" has thus expanded to encompass diverse technologies, approaches, and applications that differ substantially in capacities, constraints, and suitability for judicial deployment.[16]

As a working definition, we adopt a practical view: "AI in judicial contexts" are tools that assist courts with automation or semi-automation of judicial tasks. Judicial tasks can be grouped into "core judicial functions," "court management functions," and "interfacing with the public."[17] Of particular interest here are AI tools for "core judicial functions," namely judicial reasoning, decision-making, and adjudication. These include the use of controversial algorithmic risk-assessment tools in sentencing like COMPAS[18] and the use of large language models (LLMs) to assist with judicial decision-making[19], where a critical "core judicial function" is augmented, assisted, or potentially replaced by an AI tool.

---

[16] See e.g., Felicity Bell, Lyria Bennett Moses, Michael Legg, Jacob Silove & Monika Zalnieriute, *AI Decision-Making and the Courts: A Guide for Judges, Tribunal Members and Court Administrators*, 15-29 (Australasian Institute of Judicial Administration 2023), available at: https://mcusercontent.com/b49194247b86d30ad3cfe1745/files/42c3b454-e611-73e1-43b4-c2e67a1dd931/AI_DECISION_MAKING_AND_THE_COURTS_Report_V5_2022_06_20.pdf

[17] Colleen V. Chien, Isabella Ferrari & Niyati Narang, *JudgeGPT: When Progress Meets Precedent*, BERKELEY TECH. L.J. (forthcoming 2025).

[18] Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is a software program that US courts use to predict whether a defendant will commit future crimes: see State v. Loomis, 881 N.W.2d 749 (Wis. 2016).

[19] John Zhuang Liu & Xueyao Li, *supra* note 7.

To clarify, “algorithmic tools” and “AI” are often used interchangeably in judicial AI discourse, but they represent different points on a spectrum of computational complexity. Traditional algorithmic tools, such as actuarial risk assessment instruments like COMPAS, typically employ relatively well-established statistical methods: they apply predetermined formulas to structured input variables (age, prior convictions, employment status) to produce numerical scores. These systems are rule-based and, in theory, explicable: one can trace how inputs produce outputs, provided the developers are willing to disclose the formulae and variables. “AI” particularly machine learning systems and LLMs, operates differently. These systems learn patterns from vast training datasets rather than following predetermined rules, often developing internal representations and decision processes that resist straightforward human explanation, the so-called “black box” problem. LLMs like ChatGPT or specialized legal AI systems add another layer of complexity: they generate novel outputs (text, arguments, citations) rather than merely computing scores from fixed inputs. This distinction matters for governance because different tools present different risk profiles, raising different issues with respect to the judicial concerns relating to transparency and bias. Traditional algorithmic tools raise concerns about encoded biases in their formulas and input variables, but more advanced machine learning systems and LLMs raise additional concerns about emergent behaviors, hallucinations, and the difficulty of auditing learned representations. Throughout this article, we use 'AI' broadly to encompass this spectrum while being attentive to how different tool types generate different concerns in common law versus Chinese contexts.

A less explored aspect in computer science and legal scholarship is the social and political dimension of so-called AI tools. Technologies structure and transform social reality. If we are to critically consider AI in adjudication, these tools must be situated within power structures and relationships in a larger social matrix, and examined for their potential to transform authority and power.[20] Introducing AI in judicial settings is not a neutral automation effort; it can change the nature of judicial authority. Sensitivity to AI thus includes sensitivity to potential shifts in power, as will be discussed in Part III. This is evident in how common law and Chinese courts understand and deploy AI in light of judicial values and concerns.

While both traditions invoke similar principles such as reliability, bias prevention, transparency, and procedural fairness, they interpret and operationalize them in fundamentally divergent ways. Below, we proceed to discuss each of these principles in detail.

B. Reliability

[20] Langdon Winner, *Do Artifacts Have Politics?,* 109 DAEDALUS 121 (1980).

Courts in common law jurisdictions and Chinese courts[21] are deeply concerned about the reliability of generative AI outputs. Judicial practice notes and guidelines warn of inaccurate, irrelevant, or fictitious cases and analyses generated by LLMs. Judges in China[22], the U.S.[23], the U.K.[24], and Australia[25] have admonished and sanctioned litigants and lawyers for including fake cases in submissions. In China, providers are required to ensure the "accuracy" and "reliability" of generated content[26] . Yet it would be a mistake to assume the two systems understand and address reliability concerns regarding AI systems in similar ways.

A critical difference is that common law courts[27] prioritize individualized reliability or accuracy of AI outputs on a case-by-case basis within an adversarial framework. By

[21] Our references to "Chinese courts" necessarily generalize across significant institutional diversity, including specialized IP courts, internet courts (Hangzhou, Beijing, Guangzhou), financial courts, and the full hierarchical structure from Supreme People's Court to basic-level courts. We identify *structural* features distinguishing Chinese from common law approaches that hold across this variation. Importantly, vanguard innovations like the Shenzhen LLM experiment emerge and receive official endorsement within centrally-set parameters, confirming rather than undermining the systemic orientation we describe. Internal variation in China operates within the Supreme People's Court and the Chinese Communist Party-established boundaries, unlike the pluralistic, bottom-up variation characteristic of common law federalism in the U.S.

[22] There is at least one well-documented decision - Ren Moulou v Cao Mou and Beijing International Travel Agency Co Ltd (Beijing Tongzhou District People's Court, Civil Judgment, [2024] Jing 0112 Minchu No 19067, 10 April 2025) (任某璐诉曹某、北京某国际旅行社有限公司与公司有关的纠纷案 （北京市通州区人民法院（2024）京0112民初19067号， 2025年4月10日）. The Beijing court criticized but did not sanction the litigants for furnishing fake cases. Chinese newspapers have also reported other instances of litigants furnished fake AI generated legislative provisions, case law, and evidence in other courts in China: He Rui, "AI illusions" infiltrate courtrooms; courts in multiple locations explore governance mechanisms", Legal Daily, Safe China (January 8, 2026).

[23] In the U.S., there are more than 320 cases. The actions taken by judges include monetary fines, striking out of pleadings, public reprimands, state bar notification and warnings: see e.g. Mattox v Product Innovations Research Mattox v. Product Innovations Research, LLC, No. 6:24-cv-00235-JAR (E.D. Okla. Oct. 22, 2025) (the judge imposed a fine ranging from $1000 to $3000 on the lawyers). For all the cases, see generally: Damien Charlotin, *AI Hallucination Cases Database – United States*, Damien Charlotin.com, available at: https://www.damiencharlotin.com/hallucinations/?q=&sort_by=-date&states=USA&period_idx=0.

[24] In the U.K., there are more than 15 cases. The actions taken by judges include indemnity costs, striking out of claims, monetary fines, warnings, and notification of professional bodies: see eg Ayinde v. London Borough of Haringey, [2025] EWHC 1383 (Admin) (Eng.). where the judge fined the lawyer each GBP 2000.

[25] In Australia, there are more than 30 cases. The actions taken by judges include warnings and referral to the bar, order of costs against party and dismissal of application: see eg, Mark Khoury v. Nira Kooij, [2025] QSC 217 Sup. Ct. of Queensland (Aus.).

[26] *Interim Measures for the Management of Generative Artificial Intelligence Services* (Aug. 15, 2023), White & Case, AI Watch: Global Regulatory Tracker — China, available at: https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-china

[27] To be clear, we acknowledge variation within common law jurisdictions: US federal versus state court approaches differ, and the UK's judicial guidance differs in tone and specificity from Australian practice

contrast, Chinese judicial policy prioritizes the aggregate reliability of state-sanctioned AI systems, focusing on systemic control and proactive risk mitigation.

1. Individualized reliability

As of December 2025, there are more than 350 reported cases from courts and tribunals worldwide where AI produced hallucinated content.[28] Errors range from fictitious cases to incorrect legal arguments and summaries. Courts have criticized litigants and lawyers who furnish AI hallucinations.[29] In an egregious instance, a U.S. trial judge relied on fake cases in a divorce proceeding, resulting in reversal on appeal.[30] In such cases, courts focus on how failure to ensure reliability or accuracy has harmed the integrity of adjudication and the rights of the innocent party in that specific case. In short, the concern is individualized reliability tied to the facts and parties before the court.

To be clear, common law jurisdictions are not exclusively concerned with case-by-case reliability. Institutional guidance on structural AI deployment has emerged such as notably, the Victorian Law Reform Commission's report on AI in courts and tribunals[31] and the Australasian Institute of Judicial Administration's guide on AI and the courts.[32] Similar institutional guidelines exist in the U.K.[33] Nevertheless, the predominant frame through which reliability is operationalized in common law jurisdictions remains individualized: reliability is primarily tested through adversarial contestation of specific outputs in specific cases, enforced through professional responsibility sanctions on

---

notes. Our claim is that despite these variations, common law jurisdictions share *structural* features—adversarial procedure, party-driven process, individualized reliability focus, due process orientation—that distinguish them categorically from Chinese courts' approach. These commonalities justify treating "common law courts" as an analytically meaningful category for comparison with Chinese courts, while internal variation often concerns *degree* (how much disclosure of AI use to require, how severely to sanction hallucinations) rather than *kind* (the fundamental orientation toward AI in adjudication).

[28] See: Damien Charlotin, *AI Hallucination Cases Database*, Damien Charlotin.com available at: https://www.damiencharlotin.com/hallucinations/.

[29] See eg, Ayinde v. London Borough of Haringey, [2025] EWHC 1383 (Admin) (Eng.); Luck v. Commonwealth of Australia, [2025] FCA 68 (Fed. Ct. of Austl.).; Mata v. Avianca, Inc., No. 22-cv-1461 (PKC), 2023 WL 4114965 (S.D.N.Y. June 22, 2023).; Zhang v. Chen, [2024] BCSC 285 (Sup. Ct. of B.C.) (Can.).

[30] Shahid v. Esaam, 2025 Ga. App. LEXIS 299 (Ga. Ct. App. June 30, 2025, No. A25A0196).

[31] Victorian Law Reform Commission, *Artificial Intelligence in Victoria's Courts and Tribunals Report* (October 2025), available at https://media.lawreform.vic.gov.au/wp-content/uploads/2025/11/VLRC_AI_in_Victorias_Courts_and_Tribunals_Report.pdf .

[32] Bell et al., supra note 16.

[33] Courts & Tribunals Judiciary (UK), *Artificial Intelligence (AI) Guidance for Judicial Office Holders* (Oct 31, 2025), available at: https://www.judiciary.uk/wp-content/uploads/2025/10/Artificial-Intelligence-AI-Guidance-for-Judicial-Office-Holders-2.pdf

individual lawyers, and evaluated through case-by-case judicial assessment. The institutional dimension, while growing, has not displaced this individualized orientation as the primary regulatory mechanism.

2. Aggregate reliability

This is the hallmark of China's strategy. The Supreme People's Court has guided the nationwide "Smart Court" infrastructure, embedding government-owned and managed AI systems into core judicial processes.[34] Systems such as the Shanghai "Project 206", also known as "206 System", "[35] are not general-purpose AI; they are trained on restricted, curated datasets—statutes, regulations, and authenticated judgments (e.g., China Judgments Online).[36] Functions include Similar Case Push (suggesting authorized prior cases)[37] and the Uniform Sentencing Assistance System (empirically based sentencing guidance within legal bounds).[38] Unlike adversarial identification of hallucinations in common law submissions, the Chinese system aims to minimize aggregate unreliability ex ante.

The "Smart Court" approach keeps outputs within politically and legally permissible parameters that sustain systematic stability.[39] Its primary objective is not individualized "accuracy" in a single case but aggregate reliability in service of social stability.[40] Oversight of AI inputs and operational context is central.[41] The China Judgments Online

---

[34] Stern et al., *supra* note 6; Papagianneas & Junius, *supra* note 6.

[35] Ibid.

[36] Ibid. Note that the accuracy of the China Judgments Online platform depends on the cases that have been publicly disclosed, but scholars have found missing documents: see Stern et al., *supra* note 6, at 534, 538, 543-544. This will compromise the quality of the training data and hence the LLM outputs.

[37] Ibid.

[38] Ibid.

[39] This partially explains why China has removed cases that it considered as controversial after initially publishing them online: See: Benjamin Liebman, Rachel Stern, Xiaohan Wu & Margaret Roberts, *Rolling Back Transparency in China's Courts*, 123 COLUM. L. REV. 2407, 2447–62 (2023).

[40] Ibid.

[41] Matt Sheehan, *China's AI Regulations and How They Get Made*, 24 HORIZONS: J. INT'L REL. & SUSTAINABLE DEV. 112-113 (2023)., Noting that: "The [algorithm] registry is an online database of algorithms that have "public opinion properties or . . . social mobilization capabilities." Developers of these algorithms are required to submit information on how their algorithms are trained and deployed, including which datasets the algorithm is trained on. They are also required to complete an "algorithm security self-assessment report."

database is a principal dataset, but not every ruling is published; controversial or potentially disruptive cases may be omitted. The AI is trained on a screened subset endorsed and curated by governmental authority.[42] "Reliability" is thus predefined and limited by party-state decisions regarding training content, aligning outputs with officially sanctioned results. This prevents AI from "learning" or generating analysis conflicting with governmental policies.

The Uniform Sentencing Assistance System[43] illustrates this focus. It analyzes case circumstances and recommends sentencing parameters to minimize inconsistency across different courts. The emphasis is collective uniformity rather than the personalized accuracy that defense counsel might seek.[44] A sentence may be "correct" within the algorithmic framework while still arguably unsuitable for a particular defendant.

Two qualifications are warranted here. First, this does not necessarily imply that Chinese courts disregard individualized reliability because judges cannot disregard the law that applies to the specific facts. But validation resources are government-provided platforms, and incentives or ability to question the core reliability of these systems are limited. Challenging curated inputs lies outside litigants' or judges' designated roles. The system ensures individual case results do not undermine system-wide stability and uniformity. Second, the degree to which judges actually follow AI recommendations versus gaming compliance metrics such as by superficially aligning draft judgments with system parameters while exercising de facto discretion through other means, remains underexamined.[45] The risk of judges engaging in performative compliance, satisfying system metrics without genuine engagement with AI recommendations, warrants acknowledgment. Throughout our analysis, we describe the architectural design and stated objectives of these systems while recognizing that their actual operation may diverge from official accounts. Our analytical engagement with these systems should not be read as endorsing their efficacy or legitimacy.

One might ask whether China's approach to aggregate reliability reflects a genuinely different normative understanding of the concept or is simply a pragmatic response to the challenge of generating reliable systems at enormous scale. The answer is partly both, but the distinction matters. Scale undoubtedly drives the demand for centralized

[42] Ibid.

[43] The Supreme People's Procuratorate of the People's Republic of China, *智能量刑辅助系统—让认罪认罚从宽"看得见""落得实"[Intelligent Sentencing Assistance System — Making Leniency for Pleading Guilty Both Visible and Substantive]* (August 22, 2020), available at: https://www.spp.gov.cn/spp/llyj/202008/t20200822_477552.shtml.

[44] Ibid.

[45] On performative compliance in Chinese courts generally, see Kwai Hang Ng & Xin He, Embedded Courts: Judicial Decision-Making in China 45-72 (Cambridge Univ. Press, 2017); Carl F. Minzner, China's Turn Against Law, 59 Am. J. Comp. L. 935, 948-55 (2011).

curation: any system operating across thousands of courts with uneven judicial capacity requires standardized reference points.[46] However, scale alone does not explain the political filtering of training data--the exclusion of controversial rulings, the omission of cases inconsistent with party policy, and the screening of datasets by governmental authority.[47] A purely technical response to the scale challenge would involve comprehensive data inclusion with algorithmic quality controls, not politically determined content curation. The normative dimension is therefore irreducible: "reliability" in the Chinese context means not only technical accuracy at scale but conformity with state-sanctioned legal interpretations. This is a substantive value choice, not merely an engineering constraint.

In short, common law courts prioritize *individualized* reliability through the adversarial design (parties scrutinize facts and law and contest assertions, evidence, and citations, including AI outputs), procedural safeguards (civil procedure and professional responsibility rules requiring substantiation and forbidding misleading submissions), and technical measures (decentralized, private legal research LLMs scrutinized case-by-case). As Australia's Chief Justice Stephen Gageler observed, judges increasingly act as "human filters" of competing AI-generated arguments, an unsustainable burden.[48] By contrast, China prioritizes *aggregate* reliability through institutional design (curated datasets within the Smart Court system), procedural mechanisms (e.g., Uniform Sentencing Assistance System), and technological measures (controlling AI inputs through officially sanctioned materials).

C. Bias and Discrimination

Judicial concern with biased or discriminatory LLM outputs is another area where apparent convergence masks significant divergence. Common law guidance (e.g., the U.K.[49] and New South Wales[50]) warns of bias due to data limitations. The State Bar of

---

[46] See the discussion in Part III.A; Benjamin Minhao Chen & Zhiyu Li, How Will Technology Change the Face of Chinese Justice?, 34 COLUM. J. ASIAN L. 1, 21 (2020). George G. Zheng, *China's Grand Design of People's Smart Courts*, 7 ASIAN J.L. & SOC'Y 561, 562 (2020).

[47] Liebman et al., supra note 39; Sheehan, supra notes 41.

[48] AAP, *'Unsustainable': nation's leading judge sounds AI alarm,"* AAP (Nov. 21, 2025), https://aapnews.aap.com.au/news/unsustainable-nation-s-leading-judge-sounds-ai-alarm (last visited Dec. 8, 2025).

[49] Courts & Tribunals Judiciary (UK), *Artificial Intelligence (AI) Guidance for Judicial Office Holders* (Apr. 14, 2025) ¶ 4, available at: https://www.judiciary.uk/wp-content/uploads/2025/04/Refreshed-AI-Guidance-published-version.pdf

[50] Hon. A.S. Bell, Chief Justice of New South Wales, *Sup. Ct. of N.S.W., Practice Note SC Gen 23 — Use of Generative Artificial Intelligence (Gen AI)* ¶ 7(c) (Jan. 28, 2025).

California cautions that some generative AI is trained on biased information.[51] In China, while "bias" language is less prevalent, the Supreme People's Court requires AI products and services to be "free from discrimination and prejudice."[52]

Despite superficial similarity, the underlying frames differ. Common law courts overwhelmingly assess AI bias through individual rights, anti-discrimination norms, and judicial values of equality before the law and impartiality. In China, state-led AI policy aims to promote social stability, economic development, and national security; biased outputs are a threat to social harmony.[53] These frames accordingly lead to distinct mitigation methods.

In common law jurisdictions, if algorithms produce prejudicial outcomes targeting protected groups (such as in the contexts of incarceration or bail decisions), they jeopardize constitutional or fundamental commitments to non-discrimination.[54] In China, algorithmic bias is a primary risk because inequitable AI-influenced rulings (due to for example incomplete, biased or incorrect data) may trigger complaints, appeals, or unrest, undermining the administration's mission of social harmony.[55] Chinese law does not ignore justice or equality: statutes like the Personal Information Protection Law prohibit "unreasonable differential treatment," but "unreasonable" is assessed partly by its impact on social order, not solely on individual rights.[56]

Before proceeding, a conceptual clarification is warranted. "Bias" encompasses at least three analytically distinct meanings: (i) statistical bias--systematic deviation from a true value as a technical property of models; (ii) cognitive bias--systematic patterns in human judgment such as anchoring and automation bias; and (iii) discrimination--differential treatment based on protected characteristics or, more broadly, unjustified differential outcomes. Our analysis below touches on all three through the subcategories of system bias, data bias, and personal judicial bias. However, the cross-systemic divergence our article documents is most acute in the discrimination dimension: common law

---

[51] State Bar of Cal., Standing Comm. on Pro. Responsibility & Conduct, *Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law* 2, 4 (Nov. 16, 2023). https://www.calbar.ca.gov/Portals/0/documents/ethics/Generative-AI-Practical-Guidance.pdf

[52] The Supreme People's Court, *The Opinions on Regulating and Strengthening the Applications of Artificial Intelligence in the Judicial Fields* (Dec. 8, 2022) Principle 4 (China), https://www.chinajusticeobserver.com/law/x/the-supreme-people-s-court-the-opinions-on-regulating-and-strengthening-the-applications-of-artificial-intelligence-in-the-judicial-field-20221208

[53] Papagianneas & Junius, *supra* note 6. Mimi Zou & Lu Zhang, *Navigating China's Regulatory Approach to Generative Artificial Intelligence and Large Language Models*, 1 CAMBRIDGE FORUM ON AI: L. & GOV. e8, 2 (2025).

[54] See eg, State v. Loomis, 881 N.W.2d 749 (Wis. 2016); R (Bridges) v. Chief Constable of South Wales Police, [2020] EWCA Civ 1058 (Eng.).

[55] Stern et al., *supra* note 6, at 515, 525, 529, 543-545.

[56] *Personal Information Protection Law of the People's Republic of China* art. 24 (China).

jurisdictions frame discriminatory bias as an individual rights violation, while Chinese courts frame it as a threat to social harmony. The statistical and cognitive dimensions of bias may exhibit greater cross-systemic convergence as both systems would agree that a model producing systematically inaccurate predictions is technically deficient, and scholars have highlighted the risks that judges may be unduly influenced by algorithmic recommendations. The divergence that matters for global AI governance lies primarily in what each judicial system considers the relevant harm when AI produces differential outcomes, and whom the system is designed to protect against that harm.

1. Common Law Courts: Bias as Inconsistent with Individual Rights and Judicial Values of Equality before the Law and Impartiality

In common law jurisdictions like the U.S., the U.K., and Australia, bias is theoretically understood as violating individual rights to equal protection and freedom from discrimination based on protected characteristics.[57] Such bias is illegal when based on membership in legally defined groups, manifesting as system, data, or personal bias in AI use by judges or litigants.

a) *System Bias*

The COMPAS risk assessment system, used across U.S. states for bail, parole, and sentencing decisions[58], exemplifies system bias by incorporating variables like age, gender, and family structure that may enhance predictive accuracy but potentially violate equal protection principles, especially when harming historically disadvantaged groups.[59] Whether such differentiation is constitutionally permissible depends on adopting formal versus substantive equality approaches.[60]

Continuous learning systems create additional risks through feedback loops. A parole algorithm that updates based on reoffending statistics may initially assign higher risk

[57] In the US, see U.S. CONST. AMEND. XIV, § 1 (Equal Protection Clause). In the UK, see Equality Act 2010 and R (Bridges) v. Chief Constable of South Wales Police, [2020] EWCA Civ 1058 (Eng.). In Australia, see Racial Discrimination Act 1975 (Cth).

[58] Keith Kirkpatrick, *It's Not the Algorithm, It's the Data*, 60(2) COMMUN. ACM 21 (2017).

[59] Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, *Machine Bias*, ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.; Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, *Sample COMPAS Risk Assessment – COMPAS "CORE"*, ProPublica, https://assets.documentcloud.org/documents/2702103/Sample-Risk-Assessment-COMPAS-CORE.pdf.

[60] Derek Thompson, *Should We Be Afraid of AI in the Criminal-Justice System?*, *The Atlantic* (20 June 2019), https://www.theatlantic.com/ideas/archive/2019/06/should-we-be-afraid-of-ai-in-the-criminal-justice-system/592084 .

scores to certain demographic groups, leading to longer sentences. The resulting elevated reoffending rates from prolonged incarceration then become training data reinforcing the original bias, creating self-perpetuating algorithmic prejudice.[61] Though corrective measures like biased data sampling exist, they are inconsistently applied and of questionable effectiveness when the underlying system relies on discriminatory logic.[62]

b) *Data Bias*

The most persistent bias source is training data reflecting historical discrimination or skewed samples. ProPublica's 2016 COMPAS analysis revealed significant racial disparities: Black defendants received disproportionately higher false positive risk scores, while white defendants more often received false negatives.[63] When it comes to LLMs and word embeddings, these patterns of discrimination apply in many other contexts; it is well established for instance that gender-profession associations (for example, “woman-nurse”, “man-doctor”) tend to be preserved and perpetuated in the models that are trained on textual data.[64] These “garbage in-garbage out” patterns demonstrate how easy it is for historical data biases to become encoded and amplified in automated decision-making systems.

c) *Personal Judicial Bias*

Algorithmic deference occurs when judges defer to algorithmic recommendations perceived as objective or data-driven, effectively delegating moral responsibility to

---

[61] Danielle Ensign, Sorelle A. Friedler, Scott Neville, Carlos Scheidegger & Suresh Venkatasubramanian, *Runaway Feedback Loops in Predictive Policing*, in PROC. OF THE CONF. ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 160–171 (PMLR 2018); PROC. OF THE 2018 AAAI/ACM CONF. ON AI, ETHICS, & SOC. (AIES) (2018). See generally: Bernard E. Harcourt, *The Ratchet Effect: An Overlooked Social Cost*, in AGAINST PREDICTION: PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE 145, 145–71 (Univ. of Chi. Press, 2006).

[62] Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian & Janet Vertesi, *Fairness and Abstraction in Sociotechnical Systems*, in PROC. OF THE CONF. ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 59–68 (2019); Kenneth Holstein, Jennifer Wortman Vaughan, Hal Daumé III, Miro Dudik & Hanna Wallach, *Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?*, in PROC. OF THE 2019 CHI CONF. ON HUMAN FACTORS IN COMPUTING SYSTEMS 1–16 (2019).

[63] Ibid.

[64] Hadas Kotek, Rikker Dockum & David Sun, *Gender Bias and Stereotypes in Large Language Models*, in PROCEEDINGS OF THE ACM COLLECTIVE INTELLIGENCE CONFERENCE 12, 12–24 (2023). Tolga Bolukbasi, Kai-Wei Chang, James Y. Zou, Venkatesh Saligrama & Adam T. Kalai, *Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings*, in ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 29 (2016).

potentially biased technologies that appear neutral and authoritative.[65] Judge Babler's decision to extend a sentence based on COMPAS recommendations illustrates how algorithmic influence can compromise judicial independence.[66] His reliance on the algorithm's risk assessment shows how AI can transform from a tool to a de facto decision-maker, eroding the human judgment fundamental to the judicial process.

This over-reliance creates "anchoring effects" – which can be defined as "the tendency to focus too heavily on one piece of information when making decisions…" including being influenced "…by the order in which information is presented"[67], and additionally, the related "automation biases" – which are defined as "the tendency to use automated cues as a heuristic replacement for vigilant information seeking and processing"[68] – or in short, the tendency of users "to over-accept computer output."[69] Thus, AI predictions unconsciously influence human judgment, even as judges believe they are deciding independently.[70] Research on anchoring bias demonstrates that initial numerical suggestions—however arbitrary—significantly affect subsequent assessments.[71] In court, AI recommendations serve as powerful anchors that constrain the range of options judges consider.[72] Accordingly, it is unsurprising that AI systems displaying statistical patterns exert strong pressures toward conformity.[73] Even when judges believe cases

---

[65] Willem H. Gravett, *Judicial Decision-Making in the Age of Artificial Intelligence*, in MULTIDISCIPLINARY PERSPECTIVES ON ARTIFICIAL INTELLIGENCE AND THE LAW 281, 284 (Henrique Sousa Antunes et al. eds., 2024).

[66] State v. Loomis, 881 N.W.2d 749 (Wis. 2016); Case Comment, *State v. Loomis*, 130 HARV. L. REV. 1530 (2017).

[67] Isaac Cho, Ryan Wesslen, Alireza Karduni, Sashank Santhanam, Samira Shaikh & Wenwen Dou, *The Anchoring Effect in Decision-Making with Visual Analytics*, in 2017 IEEE CONF. ON VISUAL ANALYTICS SCI. & TECH. (VAST) 116, 116 (IEEE 2017).

[68] Kathleen L. Mosier & Linda J. Skitka, *Human Decision Makers and Automated Decision Aids: Made for Each Other?*, in AUTOMATION AND HUMAN PERFORMANCE 201, 205 (Raja Parasuraman & Mustapha Mouloua eds., CRC Press 1996).

[69] Kate Goddard, Abdul Roudsari & Jeremy C. Wyatt, *Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators*, 19 J. AM. MED. INFORMATICS ASS'N 121, 121 (2012).

[70] Rachlinski, Jeffrey J., Andrew J. Wistrich & Chris Guthrie, *Can Judges Make Reliable Numeric Judgments? Distorted Damages and Skewed Sentences*, 90 IND. L.J. 695 (2015);
Birte Englich, Thomas Mussweiler & Fritz Strack, *Playing Dice with Criminal Sentences: The Influence of Irrelevant Anchors on Experts' Judicial Decision Making*, 32 PERSONALITY & SOCIAL PSYCHOLOGY BULLETIN 188 (2006).

[71] Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision Under Risk*, 47 ECONOMETRICA 263 (1979).

[72] Anselmo Reyes & Adrian Mak, *Commercial Dispute Resolution and AI*, in THE CAMBRIDGE HANDBOOK OF PRIVATE LAW AND ARTIFICIAL INTELLIGENCE 515–16 (Ernest Lim & Phillip Morgan eds., 2024).

[73] Mari Sako & Teppo Felin, *Does AI Prediction Scale to Decision Making?*, COMM. OF THE ACM, Apr. 2025, at 18.

merit different outcomes, the numerical weight of precedent can discourage deviation from established patterns.

"Conformity bias" operates through multiple mechanisms: judges may follow established patterns to avoid criticism, while the legal principle of treating like cases alike reinforces uniformity. The challenge is distinguishing legitimate precedent adherence from uncritical adherence to statistical patterns. Conversely, awareness of AI over-reliance can trigger "contrarian bias," where judges may deliberately reject AI recommendations to demonstrate independence[74], causing arbitrary departures from accurate predictions and undermining consistency. Inadequate AI comprehension can foster wholesale rejection, preventing judges from leveraging valuable tools while maintaining an illusion of independence without addressing underlying bias issues.

d) *Potential Solutions to Address Bias*

Scholars have argued that judges should receive training on interpreting AI outputs and assessing their consistency with individual rights and judicial values like equality before the law.[75] Courts should disclose whether, how, and when AI is used by judges or litigants, and detail AI's influence on adjudication.[76] This disclosure should cover dataset selection, criteria and variable determination, acceptable error rates in algorithm application, and ongoing algorithm modifications. Additionally, independent oversight panels should be established to review judicial AI use, evaluate court AI systems, examine direct and indirect impacts (including indirect discrimination), and recommend remedial measures.[77]

To be clear, common law jurisdictions do not address algorithmic bias exclusively through individualized litigation. Institutional and structural responses relating to algorithmic accountability are increasingly being developed.[78] Proposed mechanisms to

---

[74] Anselmo Reyes & Adrian Mak, *supra* note 72, at 515-516.

[75] Tatiana Dancy & Monika Zalnieriute, *AI and Transparency in Judicial Decision Making*, OXFORD J. LEGAL STUD. (2025), gqaf030, https://doi.org/10.1093/ojls/gqaf030

[76] Ibid.

[77] Council of Europe – CEPEJ, European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment (2018).

[78] Ada Lovelace Institute, AI Now Institute & Open Government Partnership, Algorithmic Accountability for the Public Sector (2021). Available at: https://www.adalovelaceinstitute.org/report/algorithmic-accountability-public-sector/. Xukang Wang, Ying Cheng Wu, Xueliang Ji & Hongpeng Fu, Algorithmic Discrimination: Examining Its Types and Regulatory Measures with Emphasis on US Legal Practices, 7 FRONTIERS IN A.I. 1 art 1320277 (2024).

address bias include software procurement contracts[79], regulatory risk assessments[80], and institutional oversight[81]--managing bias structurally rather than case-by-case. That said, these remain largely decentralized, operating through contractual and regulatory channels rather than centralized state curation of permissible outputs.

2. China: Bias as a Threat to Social Harmony and Stability and as a Technical Flaw

In China's judicial system, "bias" primarily means threats to social harmony and stability rather than individual rights violations.[82] Courts and policy-makers thus view and frame bias as a collective problem that could generate public discontent and diminish faith in the Chinese Communist Party and state institutions.[83]

How this is reflected in practice is that Chinese courts employ two key systems: the "Same Case, Same Judgment" Database and the "Four Types of Cases" Supervision System to address bias.[84] The "Same Case, Same Judgment"[85] Database, part of China's "Smart Court"[86] initiative, enables judges to consult a nationwide AI-enhanced repository for cases.[87] The system analyzes cases, provides comparable historical rulings, suggests

---

[79] Cary Coglianese & Erik Lampmann, Contracting for Algorithmic Accountability, 6 ADMIN. L. REV. ACCORD 175 (2020). Elizabeth A. Rowe & Nyja Prior, Procuring Algorithmic Transparency, 74 ALA. L. REV. 303 (2022).

[80] Serena Oduro, Emanuel Moss & Jacob Metcalf, Obligations to Assess: Recent Trends in AI Accountability Regulations, 3 PATTERNS 100608 (2022).

[81] AI Now Institute, Algorithmic Accountability: Moving Beyond Audits (2023). Available at: https://ainowinstitute.org/publications/algorithmic-accountability.

[82] Papagianneas & Junius, *supra* note 6, at 9.

[83] Ibid; Xi Jinping, *Explanation of the "Decision of the Central Committee of the Chinese Communist Party on Several Major Issues of Comprehensively Advancing the Rule of Law"* (关于《中共中央关于全面推进依法治国若干重大问题的决定》的说明), Gov.cn (Oct. 28, 2014), https://www.gov.cn/xinwen/2014-10/28/content_2771717.htm.

[84] Supreme People's Court of the People's Republic of China, *Notice on Issuing the Guiding Opinions on Further Improving the Regulatory Mechanism for "Four Special Types of Cases"*, No. 30 (2021).

[85] Ibid.

[86] Supreme People's Court of the People's Republic of China, *'人工智能+审判' 深度融合 提升数字法院应用实效: 深圳法院启用人工智能辅助审判系统 (Integrating Artificial Intelligence with Judicial Adjudication: Shenzhen Courts Introduce an AI-Assisted Adjudication System to Enhance the Effectiveness of Digital Courts)* (2024), https://www.court.gov.cn/zixun/xiangqing/436981.html.

[87] HSF Kramer, *Asia Disputes Notes: Judicial Precedents Gain Prominence in PRC's Legal Landscape as the Chinese Supreme People's Court Issues a New Guiding Opinion on Similar Case Searches* (Aug. 11, 2020). Available at: https://www.hsfkramer.com/notes/asiadisputes/2020-08/judicial-precedents-gain-

applicable laws, and generates preliminary documents. Unlike Western systems, it uses government-curated databases pre-screened to eliminate anomalies, errors, or politically sensitive content, sanitizing training data to minimize biased human decisions. By compelling alignment with centralized "endorsed" precedents, it dramatically limits judicial discretion, establishing uniformity considered vital for stability.[88] However, it should be noted that the “sanitization” process is itself politically determined, and the exclusion of anomalies or sensitive content reflects state priorities rather than a neutral quality-control mechanism.[89]

The "Four Types of Cases" Supervision System mandates enhanced monitoring for high-stakes cases involving collective incidents (such as widespread employment terminations, environmental demonstrations), socialist principle violations (an expansive classification encompassing anything considered socially disruptive), or significant public attention.[90] AI recommendations receive additional scrutiny: judges cannot decide independently, requiring judicial panel evaluation and superior court notification.[91] Final rulings prioritizes social stability and political alignment.

However, it is clear that these systems create anchoring and conformity biases. The database presents historical cases and suggested rulings before judges thoroughly examine new cases, establishing precedent as the default baseline requiring deliberate effort to override. Heavy caseloads and efficiency pressures risk transforming AI outputs into practical standards difficult to reject. Furthermore, conformity bias persists through a dangerous feedback loop: AI learns from state-approved "correct" rulings, recommends similar verdicts, which then reinforce future patterns. This makes the system resistant to legal evolution outside approved patterns, systematically excluding dissenting reasoning and institutionalizing conformity bias. This is especially true considering that PRC judges face strong career incentives to conform to AI recommendations backed by thousands of historical cases. After all, judicial performance evaluation in China is structured around quantitative metrics including case closure rates, reversal rates on appeal, and consistency with higher-court guidance, creating direct professional consequences for deviation from established patterns.[92] When AI recommendations are drawn from the same centrally curated databases against

---

prominence-in-prcs-legal-landscape-as-the-chinese-supreme-peoples-court-issues-a-new-guiding-opinion-on-similar-case-searches

[88] Changqing Shi, Tania Sourdin & Bin Li, *The Smart Court—A New Pathway to Justice in China?,* 12 IJCA 1 (2021).

[89] Liebman et al., supra note 40; Sheehan, supra note 42.

[90] *Notice on Issuing the Guiding Opinions*, *supra* note 85.

[91] Ibid.

[92] On the structure and effects of judicial performance evaluation in China, see Carl F. Minzner, *China’s Turn Against Law*, 59 AM. J. COMP. L. 935, 948-55 (2011); Kwai Hang Ng & Xin He, EMBEDDED COURTS: JUDICIAL DECISION-MAKING IN CHINA 45-72 (Cambridge Univ. Press, 2017); Stern et al., supra note 6, at 529-30.

which judicial performance is evaluated, conformity may become a rational career strategy rather than merely a cognitive tendency.

In short, China's Smart Courts are engineered for predictable, uniform, politically secure results rather than independent and more discretionary jurisprudence. But what is important to emphasize is that these "biases" represent deliberate architectural features of the system as such. A clear trade-off becomes apparent - while reducing idiosyncratic human biases, such systems risk entrenching systemic bias prioritizing political compliance and algorithmic uniformity over flexible legal analysis.

Finally, taking a step back, the most revealing divergence may lie less in how bias is interpreted and more in where or how it is managed. In China, bias management is centralized: a state authority determines what data enters the system, which cases are permissible training material, and what counts as an acceptable outcome--deviation from which constitutes “bias.” In common law jurisdictions, bias management is distributed across software vendors, courts as procurement entities, contractual and regulatory mechanisms, and, as this article elucidates, case-by-case adversarial challenge. The Chinese model treats bias as deviation from a state-determined target; the common law model treats bias as deviation from a normatively grounded notion of neutrality or non-discrimination. This structural difference in the locus of bias governance may be the most consequential divergence for global AI standard-setting with respect to bias mitigation.

D. Transparency

Both common law and Chinese courts value "transparency" in judicial practice. For instance, China's Supreme People’s Court requires transparency in "technology development, product application and service operation."[93] Common law jurisdictions similarly emphasize transparency: Canada's Interim Principles allow external audits of AI-assisted data processing[94], while California's Judicial Council Rule 10.430 requires courts using generative AI to adopt transparent use policies.[95]

Despite this shared emphasis, the systems conceptualize transparency differently. In China's smart-court architecture, transparency means external visibility serving internal accountability. For example, this includes public-facing platforms and livestreams[96]

---

[93] The Supreme People’s Court, *The Opinions on Regulating and Strengthening the Applications of Artificial Intelligence in the Judicial Fields* (Dec. 8, 2022), Principle 6 (China), https://www.chinajusticeobserver.com/law/x/the-supreme-people-s-court-the-opinions-on-regulating-and-strengthening-the-applications-of-artificial-intelligence-in-the-judicial-field-20221208

[94] Federal Court of Canada, *Interim Principles and Guidelines on the Court’s Use of Artificial Intelligence* (Sept, 29 2025), available at: https://www.fct-cf.ca/en/pages/law-and-practice/artificial-intelligence.

[95] Judicial Council of California, *Report to the Judicial Council*, Item No 25-109 (June 16, 2025) available at: https://news.workcompacademy.com/2025/California-Court-Rules-on-AI.pdf

[96] Stern et al., *supra* note 6, at 523–24

which connect with unified back-end systems logging each procedural step, triggering deviation alerts, and mandating hierarchical review. This transparency renders adjudication legible and governable to advance legitimation and stability, not to enable adversarial challenge.[97]

In contrast, when it comes to common law courts, the idea of transparency is linked to open justice, reason-giving, and the ability to challenge decisions. It operates through disclosure of AI use[98] and procedural safeguards enabling oversight and appeal. These differences reflect deeper constitutional (and as we will show in Part III below, ideological) commitments and trade-offs, leading to different practices and risk tolerances for judicial AI.

1. How transparency is operationalized

a) Chinese Courts

Chinese courts have created multiple disclosure channels[99]—China Court Trial Online, China Judicial Process Information Online, China Judgments Online, and social media platforms like WeChat and Weibo—releasing visualized information[100] and expanding access through online litigation tools.[101] These channels may suggest that Chinese courts operationalize transparency similarly to common law jurisdictions. This is a mistake.

Consider the Faxin 2.0 system (launched in 2020), a national, Supreme People's Court-backed platform embedded in courts' electronic case-handling systems.[102] It aggregates laws, interpretations, prior decisions, and scholarly opinions, pushing similar cases,

---

[97] Straton Papagianneas, *Automating Intervention in Chinese Justice: Smart Courts and Supervision Reform*, 10 ASIAN J. L. & SOC'Y 463, 475-476 (2023).

[98] For examples of the variations of disclosure expectations or requirements in different courts, see Justice Jane Needham, *AI and the Courts in 2025: Where Are We, and How Did We Get Here?* (June 27, 2025), available at: https://www.fedcourt.gov.au/digital-law-library/judges-speeches/justice-needham/needham-j-20250627.

[99] Claire Cousineau, *Smart Courts and the Push for Technological Innovation in China's Judicial System*, New Perspectives on Asia (blog), CTR. FOR STRATEGIC & INT'L STUDIES (2021). available at: https://www.csis.org/blogs/new-perspectives-asia/smart-courts-and-push-technological-innovation-chinas-judicial-system

[100] Papagianneas & Junius, *supra* note 6, at 8.

[101] Stern et al., *supra* note 6, at 523-524.

[102] Cong Ma, "The Faxin 2.0 Smart Push System" is online! Fully connected with the National Court Electronic Case Handling System." Smart Court Times, 30 October 2020, available at: https://www.anpcn.com/hangyezixun/566.html. George G. Zheng, *supra* note 46, at 566.

applicable authorities, and sentencing frames to judges.[103] The Faxin 2.0 system comprises three elements. First, the Similar Case Push surfaces past decisions and consensus ranges (liability findings, sentence bands) within judges' workspaces.[104] Second, if a judge's draft judgment departs materially from the AI-generated recommendations, the system issues deviation warnings, which are recorded, tied to judges' accounts, and potentially triggering review by court supervisors.[105] Third, every search, acceptance, or override is time-stamped and user-attributed, feeding performance dashboards.[106] This creates permanent records of whether judges follow recommendations, becoming performance evaluation data visible to supervisors; in essence, transparency becomes a means of achieving algorithmic accountability where compliance is systematically monitored.

Accordingly, these aspects constitute "transparency as a means of supervision", and not "transparency as a means of adversarial explainability". Faxin exposes whether judges' decisions are aligned with AI-generated recommendations and whether judges heeded them. Importantly, Faxin does not provide explanations to litigants about why cases are deemed "similar," how features were weighted, or why the facts of this specific case fit the corpus of cases. Faxin's algorithms or recommendations are immune to challenge by the litigants. The tool remains a back-end governance mechanism; judges' written reasons remain the formal justification. Faxin's "transparent" artifacts are thus essentially compliance metrics (whether searches occurred, deviations happened, who reviewed), satisfying higher order standardization goals rather than external contestability.

For example, when judges draft civil judgments, Faxin auto-pushes "similar" decisions with damages ranges. If judges propose awards outside that range, the system issues deviation alerts, logs events, and queues cases for superior review. Parties see only final judgments and ordinary reasons, not deviation warnings, suggested ranges, or how internal review shaped outcomes. Via these means, Faxin makes adjudication governable and auditable *internally* within the judicial system while keeping algorithmic rationale outside of *external* scrutiny.

Similarly, Shanghai High Court's 206 System enforces criminal evidentiary procedures by automatically checking evidence against standards, flagging gaps or defects.[107] It

---

[103] Ibid. To be clear, China is not embracing stare decisis (as in the common law systems) - the algorithms in these systems are trained to find regular or similar patterns in the cases, not their ratio decidendi: Ibid, at 570.

[104] Supreme People's Court of the People's Republic of China, *Chinese Courts and Internet Judiciary* (White Paper, 2019).

[105] Ibid.

[106] Ibid.

[107] Jiang Wei, *China Uses AI Assistive Tech on Court Trial for First Time* (Jan. 24, 2019), available at: http://www.chinadaily.com.cn/a/201901/24/WS5c4959f9a3106c65c34e64ea.html.

transcribes audio/video, extracts "evidence factors," and maps relationships, turning heterogeneous materials into structured datapoints comparable across cases.[108] It quantifies case factors to assess social harm and pushes similar cases with sentencing references[109], centering uniformity by measuring outcomes against system-surfaced patterns.

By converting case files into standardized checklists, flags, and scores (evidence compliance, harm metrics), the system makes cases "legible"[110] in uniform format, but only for supervisors reviewing dashboards, and not for parties building arguments who mostly have to rely on statutes and the Supreme People's Court Guiding Cases. Audit trails monitor each judicial decision: overrides or deviations issue time-stamped alerts linked to judges' accounts, creating accountability to supervisors. The path of least resistance is aligning with recommendations, advancing uniformity by design.

The 206 system showcases transparency by demonstrating evidence was "checked," factors "extracted," and procedures "followed," projecting procedural orderliness publicly. Operative details (thresholds, weighting, internal alerts) remain within the supervisory remit, enabling control without subjecting algorithms to litigant scrutiny.

b) Common Law Courts

In common law jurisdictions, transparency measures target the public, including parties and legal representatives, to provide reasons, enable review and appeal, and reduce due process risks from opaque automation.

Courts have issued practice notes on AI disclosure expectations, with approaches differing across and within jurisdictions.[111] Beyond disclosure requirements for litigants' generative AI deployment, courts' own use of AI, such as predictive analytics, raises transparency issues. For instance, courts use predictive analytics in two contexts: AI-generated forensic evidence for conviction determinations,[112] and risk assessment tools

---

[108] Papagianneas & Junius, *supra* note 6, at 4.

[109] Ibid.

[110] On states' interest in uniformity and legibility see generally: James C. Scott, SEEING LIKE A STATE: HOW CERTAIN SCHEMES TO IMPROVE THE HUMAN CONDITION HAVE FAILED (Yale Univ. Press 1998).

[111] In Australia, New South Wales Supreme Court mandates disclosure if AI is used in preparatory steps for evidence and expert reports (PN SC Gen 23) and prohibits AI-generated affidavits, witness statements, or character references, though preparatory steps are acceptable. Conversely, Victoria has no mandatory disclosure but expects parties to disclose AI use to each other and the court when necessary (VSC Guidelines for Litigants – Responsible Use of AI in Litigation).

[112] See e.g., Commonwealth v Foley 71 A.3d 882, 38 Pa. D. & C.3d 31 (Pa. Super. Ct. 2012).

for sentencing and bail decisions.[113] In both contexts, courts require only minimum transparency: information on algorithmic performance, not how algorithms reached specific decisions or recommendations (the formulae or variables deployed to produce the output).[114]

For DNA evidence generated algorithmically to secure murder convictions, courts only require evidence of accuracy of algorithmic performance, not the algorithm itself.[115] Even if source code cannot be disclosed for proprietary reasons, this is irrelevant to admissibility of AI-generated forensic evidence if it demonstrates higher accuracy rates than other evidence types.

Similarly, COMPAS-generated risk and needs assessments (discussed in the bias section) consider offenders' prior offences and background factors (family, education, employment, among others) to predict recidivism. Courts do not require actual algorithm disclosure as statistical evidence of COMPAS's predictive accuracy is deemed sufficient.[116] Although evidence shows COMPAS has adverse discriminatory impact on African-Americans by classifying them as more susceptible to recidivism, this has not deterred courts from using COMPAS because they retain sentencing discretion.[117]

This raises the question of whether such transparency suffices for litigants to review and contest algorithmic outputs. As is typical with legal questions, the answer is: "it depends". First, it depends on whether cases are civil or criminal. Criminal sentencing involves liberty deprivation and the associated stigma, contrasting with, for example, contract cases that may use counterfactual AI[118] to evaluate whether the defendant's misrepresentation caused the claimant to enter into the contract and the losses sustained by it. Stronger justification exists for requiring information on how algorithms determine defendants' recidivism rates than in standard contract cases. COMPAS creators refused algorithm disclosure for proprietary reasons, relying on trade secret protections.[119] Justice Abrahamson criticized this opacity as serious, emphasizing courts

---

[113] See e.g., State v. Loomis, 881 N.W.2d 749 (Wis. 2016)

[114] Tatiana Dancy & Monika Zalnieriute, *supra* note 75, at 16.

[115] Ibid. at 14.

[116] Walsh Bradley J held that the overall predictive accuracy of COMPAS suffices as it had been established by "statistical validation studies", State v. Loomis, 881 N.W.2d 749 (Wis. 2016), ¶ 58.

[117] See e.g., State v. Headley, 926 N.W.2d 545 (Iowa 2019) and State v. Gordon, 921 N.W.2d 19 (Iowa 2018); People v. Younglove, No. 341901, 2019 WL. 846117, at *3 (Mich. Ct. App. 2019).

[118] On counterfactual, see Sandra Wachter, Brent Mittelstadt & Chris Russell, *Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR*, 31 HARV. J.L. & TECH. 841 (2017).

[119] State v. Loomis, 881 N.W.2d 749 (Wis. 2016), ¶ 51, Stating that: "Northpointe, Inc., the developer of COMPAS, considers COMPAS a proprietary instrument and a trade secret. Accordingly, it does not disclose how the risk scores are determined or how the factors are weighed. Loomis asserts that because COMPAS does not disclose this information, he has been denied information which the circuit court considered at sentencing."

must explain their consideration of such tools and limitations per longstanding discretionary sentencing decision requirements.[120] Yet what explanations satisfy due process remains uncertain, particularly when we question a step further, for instance, whether technical neural network explanations for more advanced tools are meaningful for litigants.

Further, transparency should depend on whether predictions are non-behavioral or behavioral.[121] For non-behavioral probabilistic evidence (AI-generated DNA), transparency can be satisfied by independent validation of overall performance and error distribution across sub-populations. However, for behavioral predictions like COMPAS recidivism rates, transparency demands extend beyond performance statistics to specific variables and their application. COMPAS accuracy statistics do not allow defendants to contest evidence without knowing the variables (education, employment instability, family background, offence nature and frequency) and weights producing recidivism scores. Thus, courts should be wary of systems whose developers refuse adequate explanations for commercial reasons.

In sum, both systems invoke "transparency" but pursue different objectives. In China, transparency functions as visible governance, reinforcing Chinese Communist Party-centered accountability, social stability, and user convenience. In common law systems, transparency serves justificatory openness—disclosure, reason-giving, and contestability sustaining open justice amidst technological opacity. These differences lead to different implementations: China emphasizes platforms and metrics for internal review; common law courts develop disclosure guidelines for AI use.

E. Procedural Fairness

Fairness is another judicial value understood and applied differently in China compared to common law jurisdictions regarding AI use. The Supreme People' Court Guiding Opinions' Principle 4 states that "fairness of the processes and outcomes of trials shall not be impaired by the introduction of technology."[122] Common law courts similarly emphasize fairness. The Federal Court of Canada ensures AI use does not "undermine fundamental rights, such as the right to a fair hearing before an impartial decision-maker."[123] Australia's Supreme Court of Victoria lists procedural fairness among "core

---

[120] Ibid. at ¶ 133, 141.

[121] Tatiana Dancy & Monika Zalnieriute, *supra* note 75, at 13.

[122] The Supreme People's Court, *The Opinions on Regulating and Strengthening the Applications of Artificial Intelligence in the Judicial Fields* (Dec. 8, 2022) Principle 4 (China), https://www.chinajusticeobserver.com/law/x/the-supreme-people-s-court-the-opinions-on-regulating-and-strengthening-the-applications-of-artificial-intelligence-in-the-judicial-field-20221208

[123] Federal Court of Canada, *Interim Principles and Guidelines on the Court's Use of Artificial Intelligence* (Sept, 29 2025), https://www.fct-cf.ca/en/pages/law-and-practice/artificial-intelligence

judicial values"[124] in its "Guidelines for Litigants: Responsible Use of AI in Litigation." The American Bar Association's Task Force on Law and Artificial Intelligence also highlights due process concerns from generative AI use.[125] Looking at these definitions, one might assume procedural fairness is conceived identically in China, the U.S., Canada, and Australia. However, upon closer investigation, the differences are stark, with profound implications for judicial AI use.

The fundamental distinction stems from underlying philosophical approaches. Common law courts (the U.S., the U.K., Australia) approach AI with wariness as a potential risk to individual procedural rights.[126] Within their adversarial framework, priority is given to litigants' rights to contest evidence, face accusers, and obtain reasoned decisions from human judges.[127] Opacity of these AI tools, as exemplified by the COMPAS algorithm, raises concerns about compromising openness and the right to fair trial.

By contrast, China's "Smart Court" systems embrace AI to strengthen government-determined fairness centered on communal objectives. Platforms like Shanghai's "206" system employ AI to standardize evidence handling, supervise judges, and ensure consistency, aligning outcomes with political and social stability goals.[128] Though providing procedural efficiency and openness, they emphasize Chinese Communist Party-state authority over individual representation. In short, while common law courts safeguard personal autonomy through procedural means, Chinese courts employ procedures to enhance governmental effectiveness and social stability.

1. Foundational Differences: Liberal Adversarialism vs. Socialist Legalism

---

[124] Supreme Court of Victoria, *Guidelines for Litigants: Responsible Use of Artificial Intelligence in Litigation*, Principle 11, https://www.supremecourt.vic.gov.au/forms-fees-and-services/forms-templates-and-guidelines/guideline-responsible-use-of-ai-in-litigation.

[125] Herbert B. Dixon Jr., Allison H. Goddard, Maura R. Grossman, Xavier Rodriguez, Scott U. Schlegel & Samuel A. Thumma, *Navigating AI in the Judiciary: New Guidelines for Judges and Their Chambers*, 26 SEDONA CONF. J. 1, 5 (2025) stating that: "Particularly when used as an aid to determine pretrial release decisions, consequences following a criminal conviction, and other significant events, how the AI or GenAI tool has been trained and tested for validity, reliability, and potential bias is critically important. Users of AI or GenAI tools for these foregoing purposes should exercise great caution."

[126] See e.g., Frank Pasquale, *Inalienable Due Process in an Age of AI: Limiting the Contractual Creep toward Automated Adjudication*, chap. 3, in CONSTITUTIONAL CHALLENGES IN THE ALGORITHMIC SOCIETY 42–56 (Hans-W. Micklitz et al. eds., Cambridge Univ. Press 2021).; Giulia G. Cusenza, *Litigating Governmental Use of AI*, 21 OHIO ST. TECH. L.J. 1 (2024). Felicity Bell et al., *supra* note 16, at ¶ 4.

[127] On the judicial duty to give reasons, see eg, Flannery v Halifax Estate Agencies Ltd. [2000] 1 W.L.R 377 (Eng.).; English v Emery Reimbold & Strick Ltd. & Orrs [2002] 1 WLR 2409 (Eng.)., [2002] EWCA Civ 605 ¶118 (Eng.).Hock L. Ho, *The Judicial Duty to Give Reasons*, 20 LEGAL STUD. 42 (2000).

[128] Stern et al., *supra* note 6, at 540-541.

Common law procedural fairness is a fundamental principle developed from the English Magna Carta and the notion of "natural justice."[129] It connects to the rule of law, which demands governmental power be exercised within a framework of known, predictable, and impartial rules. As articulated by scholars like A.V. Dicey[130] and Lon Fuller[131], the rule of law shields individuals from arbitrary state power. From this perspective, procedural fairness constitutes a pathway to justice, including rights to advance notice, impartial tribunal hearings, opportunities to submit evidence, and to challenge accusers.[132] This reflects a liberal democratic order where individuals are the fundamental legal and moral unit, with judicial proceedings as adversarial competitions between substantially equal parties and the state acting as impartial mediator through the judiciary.[133]

In contrast, China's legal framework is based on "socialist rule of law with Chinese characteristics."[134] This framework is essentially instrumentalist and monist: not limiting Chinese Communist Party authority but enabling its guidance and policy objectives of national revival and social harmony.[135] In such a framework, law serves as a governance instrument. This perspective rejects liberal democracy's pluralism, that is, competing values and interests require balancing. Only one substantive standard exists: the state's success in accomplishing its goals.[136]

Fairness and justice are conceived as belonging to the community and defined by the state, not grounded in the primacy of individual rights.[137] This draws on China's imperial past, where Legalism treated law as a state power tool.[138] In essence, the Chinese

---

[129] See e.g., Catherine Donnelly SC, Joanna Bell & Ivan Hare KC, DE SMITH'S JUDICIAL REVIEW (9th ed., Sweet & Maxwell 2024).

[130] Albert. V. Dicey, INTRODUCTION TO THE STUDY OF THE LAW OF THE CONSTITUTION (10th ed., Macmillan, 1959), Chap. IV.

[131] Lon L. Fuller, THE MORALITY OF LAW (rev. ed., Yale University Press, 1969), Chap. 2.

[132] See e.g., Henry J. Friendly, *Some Kind of Hearing*, 123 U. PA. L. REV. 1267, 1283 (1974).

[133] See e.g., Sward, Ellen E. *Values, ideology, and the evolution of the adversary system*. IND. LJ 64 (1988): 301, 308-313.

[134] The State Council, *The Socialist System of Laws with Chinese Characteristics* (2011), available at: https://english.www.gov.cn/archive/white_paper/2014/09/09/content_281474986284659.htm.

[135] Ibid.

[136] Communiqué of the Fourth Plenary Session of the 18th Central Committee of the Communist Party of China (Oct. 23, 2014), available at: http://www.china.org.cn/china/fourth_plenary_session/2014-12/02/content_34208801.htm.

[137] Eric W. Orts, *The Rule of Law in China*, 34 VAND. J. TRANSNAT'L L. 43, 86 (2001). Noting that: "the Hobbes' method of "possessive individualism" may strike many Chinese theorists as odd, as would most Western views that begin with the interests of the individual rather than the collective"

[138] Ibid; Albert H.Y. Chen, *China*, in ELGAR ENCYCLOPEDIA OF COMPARATIVE LAW 229–35 (Jan M. Smits, Jaakko Husa, Catherine Valcke & Madalena Narciso eds., Edward Elgar Publ'g 2023).

approach openly prioritizes the collective over the individual, as articulated by the Chinese Communist Party.

2. The Meaning of "Being Heard": Influence versus Participation

A crucial aspect of common law procedural fairness is the “right to be heard”: individuals should receive proper hearings before judges capable of genuinely comprehending, analyzing, and deciding their cases.[139] Sound judicial decision-making requires more than legal expertise; it necessitates skilled courtroom navigation, conflict resolution abilities, capacity to identify the root causes of disputes, and sensitivity to parties' emotional and psychological dimensions. Such competence develops through judges' interpersonal experiences and decades of practice—wisdom AI systems cannot yet acquire from legal databases alone. As the U.S. Supreme Court Chief Justice John Roberts observed, "Nuance matters: Much can turn on a shaking hand, a quivering voice, a change of inflection, a bead of sweat, a moment's hesitation, a fleeting break in eye contact. And most people still trust humans more than machines to perceive and draw the right inferences from these clues."[140] This view is consistent with a recent study of 12 UK judges (including 5 Supreme Court members) viewing law as practical reasoning rooted in human decision-making, not pure logic.[141] For instance, one respondent pertinently noted that AI cannot replace the "cathartic role that human justice does."[142] However, individuals need not have the right to human judges in all cases. As one judge observes, "It's one thing to have cheap and cheerful AI tool to resolve a £500 dispute over a second hand car sales contract. It's quite another if somebody's being sent to prison or somebody's having their children taken away."[143]

In contrast, in China's Smart Court ecosystem, being "heard" often means participating in a streamlined state-managed workflow.[144] Litigants can file materials, join proceedings

---

[139] See e.g. US, Mathews v. Eldridge, 424 U.S. 319, 333 (1976); South Buckinghamshire DC v Porter (No 2) [2004] UKHL 33 (Eng.); [2004] 1 WLR 1953, para 36; Ridge v Baldwin [1964] AC 40 (HL), 80-81; Kioa v West (1985) 159 CLR 550, 585-586 (Austl.)

[140] Chief Justice Roberts, *2023 Year-End Report on the Federal Judiciary* (Dec. 31 2023), 6, available at: https://www.supremecourt.gov/publicinfo/year-end/2023year-endreport.pdf

[141] Erin Solovey, Brian Flanagan & Daniel Chen, *Interacting with AI at Work: Perceptions and Opportunities from the UK Judiciary*, in *Proceedings of the 4th Annual Symposium on Human-Computer Interaction for Work* 1–8, ¶ 4.2.1 (Jun. 23-25, 2025). available at: https://dl.acm.org/doi/pdf/10.1145/3729176.3729192

[142] Ibid, ¶ 4.2.2.

[143] Ibid, ¶ 4.2.3.

[144] Supreme People’s Court, *Online Litigation Rules of the People’s Courts* (effective 1 Aug 2021), art. 2, https://cicc.court.gov.cn/html/1/219/199/201/2208.html,

via video, and track case status through extensive, technologically sophisticated interfaces.[145] However, the system frequently funnels submissions into preset fields for standardized handling. This participation mode confers outcome legitimacy for participants and the public, fostering a sense that cases received fair attention, even though the core evaluative process remains opaque.[146]

Due to these starkly different understandings of procedural fairness, common law and Chinese courts have developed different responses to AI use.

3. The Role of Technology in Procedural Fairness: Threat versus Tool

a) *Common Law: AI as a Potential Threat to Due Process*

The "right to a reason" is a fundamental procedural fairness aspect that AI potentially undermines due to its black box nature.[147] One of the reasons is that AI reasoning may be incomprehensible even to its developers.[148] When judges depend on predictive analytics like algorithmic or AI risk assessments or sentencing suggestions, defendants find mounting meaningful objections challenging.

State v. Loomis[149] illustrates this. In 2013, Eric Loomis received a sentence influenced by COMPAS algorithm risk evaluation. While the Wisconsin Supreme Court affirmed the sentence, it recognized significant due process concerns: Loomis could not contest the algorithm's proprietary methodology, factor weighting, or potential racial bias. The case demonstrates how a tool designed to introduce "rationality" undermined the defendant's ability to challenge COMPAS output used to determine his sentencing outcome, a fundamental procedural fairness aspect.

Given threats posed by predictive analytics to procedural fairness, many common law courts restrict its use to administrative functions presenting least risk to procedural

---

[145] Sophia Z. Tang, *Virtual Hearing in China's Smart Court* (2021), available at: https://conflictoflaws.net/2021/virtual-hearing-in-chinas-smart-court/.

[146] Straton Papagianneas, *supra* note 97, at 485.

[147] See e.g., Cynthia Rudin, *Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead*, 1 NATURE MACHINE INTELLIGENCE 206 (13 May 2019); Paul W. Grimm, Maura R. Grossman & Gordon V. Cormack, *Artificial Intelligence as Evidence* 19 NW. J. TECH. & INTELL. PROP. 9 (2021).; State v. Pickett, 466 N.J. Super. 316 (App. Div. 2020).

[148] Anna Rogers, Olga Kovaleva & Anna Rumshisky*, A Primer in BERTology: What We Know About How BERT Works*, 8 TRANS. ASS'N FOR COMPUT. LINGUISTICS 842 (2020). Stating that: "Although it is clear that BERT works remarkably well, it is less clear why, which limits further hypothesis-driven improvement of the architecture."

[149] State v. Loomis, 881 N.W.2d 749 (Wis. 2016) (see specially Justice Abrahamson's concurring judgment in ¶ 133,141-2).

rights.[150] For example, case scheduling, document filing, legal research assistance, and initial document review can leverage AI efficiency without displacing human judgment.

But going beyond the critical problems of predictive analytics, we note that generative AI also poses due process threats: undermining party autonomy, weakening adversarial structure, and compromising confidential case preparation.

Due process guarantees parties, not external actors, control case presentation. This ensures individuals and entities bear responsibility for legal positions and make strategic choices reflecting their unique circumstances, values, and risk assessments. When LLMs generate legal arguments, frame issues, or select authorities, critical aspects of the case may be determined by AI's architectural design and training data patterns rather than parties' legal strategies. This fundamentally departs from traditional legal research tools. While lawyers have always consulted secondary sources, databases, and precedent collections, these resources provide raw material counsel must actively analyze, synthesize, and deploy according to client objectives and case-specific strategy. Generative AI operates differently: it often produces complete legal arguments, pre-formulated reasoning chains, and ready-made analytical frameworks potentially embedding aggregate patterns, statistical correlations, or training data regularities not necessarily related to individual party circumstances or chosen approaches.[151] The procedural fairness concern becomes acute when parties' substantive legal positions are materially shaped by algorithmic outputs rather than human strategic judgment. If arguments advanced, authorities cited, or analytical frameworks employed essentially reflect what the AI system learned from its training corpus rather than what the party and counsel determined best serves their interests, there is a risk the party allows AI to materially influence its judgment and discretion.

The adversarial system's legitimacy depends on outcomes resulting from each party having *genuine* opportunity to advance their own case theory, marshalling evidence and argument serving their chosen strategy. When algorithmic systems substantially shape these fundamental choices, procedural fairness may be compromised regardless of whether AI-generated positions are legally sophisticated or factually sound.

Furthermore, regarding weakening adversarial structure, the adversarial system fulfils its due process function only when opposing parties genuinely contest positions, subjecting

---

[150] See e.g. on the judicial approval of the use of Technology Assisted Review (TAR), Pyrrho Investments Ltd v MWB Property Ltd [2016] EWHC 256 (Ch) (Eng.); Brown v BCA Trading [2016] EWHC 1464 (Ch). (Eng.)

[151] See eg, Emily M. Bender, Timnit Gebru, Angelina McMillan-Major & Shmargaret Shmitchell, *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?*, in PROCEEDINGS OF THE 2021 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 610–23 (2021). available at: https://doi.org/10.1145/3442188.3445922; American Bar Association, Formal Opinion 512: Generative Artificial Intelligence Tools (29 July 2024), available at: https://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/ethics-opinions/aba-formal-opinion-512.pdf.; Courts & Tribunals Judiciary (UK), *Artificial Intelligence (AI) Guidance for Judicial Office Holders* (Apr. 14, 2025), available at: https://www.judiciary.uk/wp-content/uploads/2025/04/Refreshed-AI-Guidance-published-version.pdf.

factual claims and legal arguments to rigorous examination. This contestation serves an epistemic purpose essential to justice, not merely form. Judicial decision-making is most reliable when competing advocates identify weaknesses in opposing arguments, expose evidentiary gaps, and advance alternative interpretations. Generative AI deployment threatens this core function when similar systems trained on overlapping datasets generate arguments for opposing parties. The risk is not that AI produces inferior or inaccurate legal analysis (although it does), but that it may produce convergent analysis.[152] In fact, recent findings prove this convergence in open-ended questions – questions which arguably lie at the heart of legal practice. Researchers have called this phenomenon whereby "*different* models produce strikingly similar outputs" as "*inter-model homogeneity*," which, when looked at collectively, can also be described as an "Artificial Hivemind."[153]

When both plaintiff's and defendant's briefs emerge from systems drawing on substantially similar training data and employing comparable analytical architectures, resulting arguments may present only surface-level opposition while sharing fundamental premises, analytical approaches, and rhetorical structures. Courts then confront what appears to be adversarial advocacy but may actually represent algorithmic variations on common themes: both parties' positions shaped by the same underlying patterns and both arguments bounded by similar architectural constraints. This threatens procedural fairness subtly but profoundly: parties retain formal adversarial representation trappings while losing its substantive benefit. Courts may render decisions without benefit of the searching adversarial examination that common law procedure assumes will expose flawed reasoning, weak evidence, or overlooked considerations.

Regarding compromising confidential case preparation, procedural fairness requires parties prepare cases confidentially without premature disclosure. This protection, embodied in attorney-client privilege and work product doctrine, ensures litigants can candidly assess weaknesses and make informed tactical decisions. When lawyers use generative AI to draft pleadings, analyze evidence, or develop arguments, due process concerns arise. If AI providers retain query data or if AI outputs can be forensically traced to prompts, confidentiality essential to case preparation may be compromised.[154] Parties cannot fully prepare cases[155] as they must consider whether AI use will reveal strategies, undermine privilege claims, or expose preliminary theories. The result is procedural

---

[152] Natalia Restrepo & Diego Socol, *supra* note 9.

[153] Jiang, Liwei, et al, *supra* note 9, at 1.

[154] Brandon K. Sorensen, *Generating . . . Client Confidentiality Concerns in the Use of Generative AI Technology,* N. ENGL. L. REV. (published online May 16, 2024), available at: https://newenglrev.com/2024/05/16/generating-client-confidentiality-concerns-in-the-use-of-generative-ai-technology/.

[155] David Levine, *Avoiding Ethical Pitfalls as Generative Artificial Intelligence Transforms the Practice of Litigation*, NAT'L L. REV. (May 6, 2025). available at: https://natlawreview.com/article/avoiding-ethical-pitfalls-generative-artificial-intelligence-transforms-practice.

disadvantage: litigants face difficult choices between using AI tools at confidentiality costs or forgoing tools to preserve attorney-client privilege. Either undermines the level playing field procedural fairness demands. Large law firms can work with LLM developers to create bespoke generative AI ensuring data stays within the model's four walls, preserving confidentiality. However, small or medium-sized firms are unlikely to have resources to do so.

b) *China: AI as a Tool for Enhanced Governance and Control to Promote Fairness*

China's Smart Courts view AI not as a due process threat but as a superior tool for achieving the state's conception of fairness: uniformity, control, and stability.[156] Two AI systems seek to promote fairness.

First, the Shanghai 206 System for criminal matters systematizes evidentiary procedure.[157] Once a case is logged, AI directs law enforcement and prosecutors to necessary conviction evidence, examines procedural violations, and identifies contradictory or inadequate evidence.[158] This establishes "internal accountability," ensuring lower-level officials strictly follow centrally established standards. The objective is to reduce individual judicial discretion, which the Chinese government regards as a source of arbitrariness rather than justice.[159] The goal is ensuring judges apply law consistently according to state-approved interpretations, not empowering parties to autonomously shape litigation strategy.

Second, the Faxin 2.0 system provides comparable cases and rulings to judges.[160] Perhaps more significantly, it also continuously monitors judges' rulings. For instance, when a judgment differs from algorithmic suggestion, the system generates an alert and the case may be marked for senior judge examination. In effect, this transfers discretion from individual judges to the system's creators and the Chinese Communist Party overseers. Similarly, Jiangxi Provincial High Court's Trial e-Management Platform measures judges' decisions against peers.[161]

---

[156] Papagianneas & Junius, supra note 6; Papagianneas, supra note 97.

[157] Stern et al., *supra* note 6, at 541.

[158] Ibid;

[159] Yiming Yin, *Analysis on the Problems and Solutions in Artificial Intelligence-Assisted Sentencing*, in PROCEEDINGS OF THE 2024 3RD INTERNATIONAL CONFERENCE ON SOCIAL SCIENCES AND HUMANITIES AND ARTS (SSHA 2024) 866 (M.F.B. Sedon et al. eds., 2024).

[160] Stern et al., supra note 6, at 541-2.

[161] Papagianneas & Junius, *supra* note 6, at 6.

Unlike common law jurisdictions expressing grave reservations about using AI to detect judicial biases,[162] Chinese courts use predictive analytics to correlate judges' identities with their decisions, ensuring actual or predicted decisions comport with state-sanctioned algorithmic databases.

Further, neither the Shanghai 206 System nor Faxin 2.0 undermines party autonomy because protecting party autonomy is not the procedural goal of China's inquisitorial system where judges actively participate in fact-finding. The 206 System enhances the court's investigative capacity, precisely what Chinese procedural law envisions.[163] When the system identifies evidentiary gaps or contradictions, it assists judges in fulfilling their affirmative duty to investigate thoroughly. As for Faxin 2.0, the "threat" to party autonomy is actually a feature, not a bug: by channelling judicial consideration toward algorithmically identified patterns in curated datasets, Faxin disincentivizes parties from advancing arguments deviating from state-sanctioned interpretations. In a system where law serves as a governance instrument, this paradoxically represents procedural fairness, ensuring individual cases receive treatment aligned with collective standards rather than idiosyncratic party strategies.

Finally, because China does not operate under the adversarial model, algorithmic convergence concerns in common law systems, in actuality, become a positive feature. Convergent AI systems prevent emergence of genuinely adversarial arguments that might destabilize proceedings. When both prosecution and defense arguments emerge from systems drawing on substantially similar training data, both parties' positions are bounded by state-approved interpretations. This prevents defense lawyers from advancing novel legal theories or challenging fundamental prosecution premises in ways that might slow proceedings or generate instability.

Consider how this operates in the 206 System and Faxin 2.0 systems. Both draw on China Judgments Online, with datasets screened by judicial administrators to eliminate anomalies or politically sensitive rulings. When prosecutors and defense lawyers both use systems trained on these sanitized datasets, their arguments naturally converge around state-sanctioned interpretations. This convergence enhances efficiency. The 206 system has reduced average trial times by approximately 20-30% by standardizing evidentiary presentations and channelling arguments within narrow pre-approved boundaries.[164]

---

[162] For example, in Australia, courts have held that statistical evidence which shows the bias of judges does not meet the standard for proving actual or apparent bias and hence lacks probative value: see BDS17 v Minister for Immigration and Border Protection [2018] FCA 1683. (Austl.)

[163] Wanqiang Wu & Xifen Lin, *Access to technology, access to justice: China's artificial intelligence application in criminal proceedings*, 81 INT'L J. L., CRIME & JUST. 100741, 6-7 (2025).

[164] Jinhua Cheng *人工、智能与法院大转型[Artificial, Intelligence, and the Great Transformation of Courts],* 27 J. SHANGHAI JIAO TONG UNIV. (PHIL. & SOC. SCI.) (**上海交通大学学**报（哲学社会科学版）) 33, 39 (2019) (China).

AI systems operate as constitutional technologies that either threaten or advance procedural fairness depending on the constitutional order they serve. In common law systems, AI threatens procedural fairness when it undermines party autonomy, weakens adversarial contestation, or compromises confidential preparation, which are mechanisms protecting individuals from state overreach. In Chinese judicial systems, AI advances procedural fairness when it standardizes outcomes, channels arguments toward consensus, and creates comprehensive audit trails, mechanisms ensuring consistency, stability, and supervisory oversight serving collective welfare.

The same technological features constitute threats in one system and advantages in the other because systems pursue different goals through different procedural mechanisms. Procedural fairness has no universal content; it is always procedural fairness for a particular constitutional order. AI systems threaten or advance procedural fairness based not on intrinsic technical features, but on compatibility with the constitutional commitments and procedural structures they serve.

III. Extra-legal perspectives

The preceding analysis has demonstrated that even when common law and Chinese courts employ ostensibly identical terms such as reliability, bias, transparency, and procedural fairness, these concepts operate within fundamentally divergent semantic and institutional fields. Yet while these doctrinal and jurisprudential legal-theoretic perspectives illuminate much about how AI is framed and deployed differently across jurisdictions, they cannot fully account for the profound disparities we observe. The question remains: why do these divergences persist, and what underlying forces shape the receptivity or resistance to AI in adjudication?

To answer this question requires moving beyond traditional legal theory to examine the broader socio-technical, ideological, and epistemic infrastructures that condition how AI is received and operationalized in judicial contexts. The legal frameworks and institutional structures discussed in Part II are themselves embedded within larger systems of meaning-making, power relations, and knowledge production that determine not merely how AI principles are interpreted, but whether the fundamental prerequisites for AI adoption exist at all.

Our first hint regarding the extra-legal dimensions of AI reception emerges from a close reading of legal discourse surrounding AI hallucinations. In common law decisions, the arguably alarmist language, stands in stark contrast with the focus on technical solutions to generative AI's problems that we observe in the legal discourse in China.[165] For

[165] John Zhuang Liu & Xueyao Li, *supra* note 8. Gao Yandong & Jia Xinyu, ***大模型技****术下，****AI****幻觉的风险防范 [Risk Prevention for AI Hallucinations under Large-Model Technology]*, Sup. People's Procuratorate Daily, Sept. 10, 2025, https://www.spp.gov.cn/llyj/202509/t20250910_706137.shtml (China).

instance, judicial discourse in common law cases regarding AI hallucinations frames the problem as a systemic concern to the entire legal edifice - for example, classifying these actions as "abuse of the judicial system"[166] and an "abuse of process"[167], and voicing a concern that fake precedents promote "...cynicism about the legal profession and the American judicial system"[168], and amplifying these statements by acknowledging that that AI's use will have "serious implications for the administration of justice and public confidence in the justice system",[169] among other similar statements - without a doubt, this is a sensitive topic for common law judges regardless of jurisdictional affiliation.

For instance, in the *Mata* case, LLMs were still a novelty, and few institutional frameworks or procedural safeguards were known. The judges' statements therefore are valuable because they reflect a gut reaction upon a first encounter - a reflexive response motivated by deeper considerations regarding the legal system as such. In this sense, AI use and the resulting hallucinations are seen as a serious threat and a risk that needs to be carefully managed and addressed as it is seen to have the potential to undermine the foundation of entire legal order.. This stance effectively precludes an unquestioning and uncritical deployment of AI in litigation in common law jurisdictions.

However, such a restrictive approach is not as clearly visible in Chinese legal discourse, where it is well-documented that generative AI has been actively utilized in court systems for a considerable time, and as we noted Chinese judges already use LLMs to assist with the writing of their decisions.[170] Supporting this point is the fact that Chinese legal discourse on hallucinations focuses on technical and computer science-based solutions. For example, statements from Chinese legal scholars that state that "the procuratorial organs should build an AI hallucination prevention system in the procuratorial context on the basis of understanding the mechanism and specific performance of AI hallucinations, and specifically prevent the adverse effects of AI hallucination on procuratorial work"[171] while recognizing the problems of hallucinations, clearly frame the problem as one that can be managed on a technical level, addressing the LLM technology directly. Liu and Li also confirm this approach, stating that "special attention was paid to mitigate the hallucination issue" with courts spearheading "the

---

[166] Mata v. Avianca, Inc., No. 22-cv-1461 (PKC), 2023 WL 4114965 (S.D.N.Y. June 22, 2023).

[167] Bandla v. Solicitors Regulation Authority, [2025] EWHC 1167 (Admin) (Eng.).;
Zhang v. Chen, [2024] BCSC 285 (Sup. Ct. of B.C.) (Can.).

[168] Mata v. Avianca, Inc., No. 22-cv-1461 (PKC), 2023 WL 4114965, at 1 (S.D.N.Y. June 22, 2023). (opinion and order on sanctions) - emphasizing the "gatekeeping role" of attorneys "to ensure the accuracy of their filings".

[169] Ayinde v. London Borough of Haringey, [2025] EWHC 1383 (Admin) (Eng.) ¶ 9.

[170] John Zhuang Liu & Xueyao Li, *supra* note 8; Stern et al., *supra* note 6. Chen & Li, supra note 47.

[171] Gao Yandong & Jia Xinyu, ***大模型技****术下，AI****幻****觉的风险防范[Risk Prevention for AI Hallucinations under Large-Model Technology]*, Sup. People's Procuratorate Daily, Sept. 10, 2025, https://www.spp.gov.cn/llyj/202509/t20250910_706137.shtml (China).

development of the model, even bringing full-time judges in to participate in the product development, ensuring that the model would meet the needs of judges' work."[172]

These contrasting approaches underscore why extra-legal explanations are necessary as complements to doctrinal or legal-theoretic analyses. As we see, the alarmist and reflexive response of common law judges cannot be fully understood through legislation or formal procedural rules alone, just as the pragmatic, technically oriented approach in China cannot be explained solely by the law.

We present three levels of extra-legal explanation for the observable differences in the structural fitness of legal systems to "assimilate" AI technologies. This typology is not intended to be exhaustive, but it clarifies the progressively widening lenses through which such differences can be understood. Our discussion proceeds along three increasingly abstract levels: the *efficiency* perspective, the *ideological* perspective, and the *archival-epistemic* perspective. Each level represents a further zooming out: from the immediate, day-to-day uses of AI in adjudication, to the broader institutional and socio-technical motivations that shape adoption, and finally to a critical examination of the epistemic and archival power relations that underpin how judicial knowledge is constituted and legitimated across different juriscultures.

A. Efficiency Perspectives

At this first level, we focus on one of the most basic non-legal rationales for explaining the differences in framing of AI in adjudication, namely efficiency considerations. While the overarching logic of efficiency is shared across jurisdictions we study, the manifestations and forms that efficiency arguments in legal and judicial discourse take seem to differ.

We firstly observe that in China, the emphasis seems to be on the development of a more efficient legal system from the ground up, hence the embrace of technologies to address classical judicial problems of speed, uniformity and consistency. Indeed, China's embrace of AI in many areas of society is well-documented,[173] and in the context of law, we can see that already in 2020, even before LLMs came onto the scene, Stern et al. have shown that Chinese courts were significantly ahead of the rest of the world in AI adoption into their legal system, stating that "Chinese courts are plainly leapfrogging efforts elsewhere, moving rapidly to a world where computers suggest legal outcomes to judges, either by analyzing millions of past cases or through a decision-tree designed to match the fact pattern in the case with the correct legal solution."[174]

---

[172] John Zhuang Liu & Xueyao Li, *supra* note 8.

[173] Huw Roberts, Josh Cowls, Jessica Morley, Mariarosaria Taddeo, Vincent Wang & Luciano Floridi, *The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation*, in ETHICS, GOVERNANCE, AND POLICIES IN ARTIFICIAL INTELLIGENCE 47 (Springer Int'l Publ'g 2021).

[174] Stern et al., *supra* note 6, at 519.

However, to be clear, current Chinese practices primarily standardize and constrain judicial discretion through structured recommendations rather than displacing judicial reasoning altogether. Existing systems predominantly operate as structured recommendation tools rather than outcome-generating engines. The Uniform Sentencing Assistance System, discussed above, analyzes case circumstances such as offense type, aggravating and mitigating factors, defendant characteristics, against a database of prior sentences to recommend parameter ranges, but does not generate verdicts from unstructured factual narratives. Similarly, Faxin 2.0's Similar Case Push, discussed earlier, surfaces prior decisions based on structured queries and coded features, requiring judges to assess relevance and applicability. These systems thus augment judicial decision-making within predefined parameters rather than replacing it. The more recent Shenzhen province experiment with LLMs for judgment drafting represents a potential shift toward more autonomous generation[175], though even here the AI assists with drafting rather than determining outcomes independently.

The adoption of AI solutions in China's legal system, from the earlier "Smart Court" initiatives[176] to the recent generative AI-assisted adjudication efforts[177] can be seen as a logical response aimed at managing the immense and growing caseload, particularly in the context of a scarcity of legally qualified judges.[178] Commenting on this relatively quick adoption, Zheng points out that indeed due to a lack of an "established "legal nobility" (Max Weber's "legal honoratiories"[179]), the application of cutting-edge technology to judicial works encountered little resistance."[180] In effect, the shortage of judges both generates the demand for AI in adjudication while ensuring little opposition to its deployment - a happy coincidence which creates an unusually conducive environment for the rapid integration of AI tools in adjudication in the pursuit of systemic efficiency. This provides further bases for the differences in framing we discussed previously.

It is worth mentioning that this scarcity of legally qualified judges is not merely quantitative but qualitative and geographic. While first-tier cities have access to well-trained judicial personnel, basic-level courts in less developed regions often lack judges with formal legal education. This uneven distribution creates risks of inconsistent, arbitrary, or legally questionable decisions—a serious concern for a system prioritizing

---

[175] John Zhuang Liu & Xueyao Li, *supra* note 8.

[176] Changqing Shi, Tania Sourdin & Bin Li, *supra* note 88.

[177] John Zhuang Liu & Xueyao Li, *supra* note 8.

[178] Benjamin Minhao Chen & Zhiyu Li, *supra* note 46, at 21; Noting that "Remarkably, legally qualified judges are still a relatively new development in China."

[179] Max Weber, ECONOMY AND SOCIETY: AN OUTLINE OF INTERPRETIVE SOCIOLOGY 784 (Guenther Roth & Claus Wittich eds., Univ. of Cal. Press 1978).

[180] George G. Zheng, *supra* note 46, at 562.

uniformity and social stability. AI-driven standardization through systems like the Similar Case Push and Uniform Sentencing Assistance System directly addresses these concerns by providing judges, regardless of their training or location, with consistent reference points derived from centrally curated databases. In effect, AI operates as a mechanism for quality control and judicial capacity-building, substituting technological standardization for the professional socialization and legal education that underwrite consistency in common law systems. The attractiveness of such tools is thus inseparable from the developmental context of China's judiciary: where established legal training infrastructure cannot yet produce sufficient numbers of qualified judges, AI offers an alternative pathway to reducing regional disparity and ensuring baseline consistency in the application of law.

A subtler point is that the quick and efficient resolution of legal disputes with the help of AI may inhibit the emergence of community-based networks that assist in resolving said disputes. If legal issues can be quickly resolved individually and within minutes over a platform like WeChat, there is very little incentive to form support and solidarity networks within one's community to address legal grievances[181]. Thus, rapid, efficient, and individualized legal resolution leads to more effective social control, channeling disputes through state-mediated platforms, which prevents legal mobilization and the creation of social movements at the grassroots level.[182] In this way, we can see that the drive for legal systemic efficiency serves multiple purposes, satisfying administrative efficiency imperatives but also social control considerations. And in a society as large and as rapidly developing as the Chinese one, these issues are arguably of paramount concern that determine the adoption of AI in legal contexts.

Conversely, in the U.K., the U.S., and Australia, where the law is considered to be more "uncommonly puzzling"[183] when compared to the civil law, efficiency can be seen as a matter of preservation and maintenance of the existing legal edifice, ensuring that technology does not disrupt established processes. We see this especially clearly in the language used in the AI hallucination opinions where, for instance, in *Mata*, Judge Castel states that the harms that flow from "submission of fake opinions" are many, including, inter alia, the fact that "the opposing party wastes time and money in exposing the

[181] Benjamin Minhao Chen & Zhiyu Li, supra note 46, at 45; Noting that: "Technology alleviates this tension in authoritarian legalism by bringing law—and courts—closer to the people, and in so doing reduce the need for individuals to rely on others, be they friends or colleagues, volunteers or professionals. By helping citizens know and claim their rights, the legal system broadens popular access to justice. At the same time, however, it inhibits the coalescence of grievances and dampens the centrality of rights advocates. Villagers able to contest an administrative fine on their mobile phones will be less likely to consult neighbors like Zhou Guangli. Women able to sue employers for gender discrimination through Wechat will be less inclined to turn to non-governmental organizations like the Zhongze Center."

[182] Frank K. Upham, *Litigation and Moral Consciousness in Japan: An Interpretive Analysis of Four Japanese Pollution Suits*, 10 LAW & SOC'Y REV. 579 (1976).

[183] Frederick Schauer, *Is the Common Law Law?*, 77 CALIF. L. REV., 455, 455 (1989).

deception [and] [t]he Court's time is taken from other important endeavors."[184] Similarly, in *Ayinde* one of the relevant factors that ought to determine a court's response in hallucination cases is "the time and expense incurred by other parties to the case, and the resources used by the court in addressing the matter."[185] The question facing common law courts is thus posed not so much in terms of whether AI innovations in adjudication could improve overall legal system's macro-level efficiency, but more in terms of whether such technologies could potentially reduce efficiency by creating new and unforeseen costs and burdens for judges and litigants in the day-to-day administration of justice. Thus, skepticism of AI in adjudication is actually a means of conserving the legal system's established efficiency - in other words, for common law judges, conservatism is likely a reflection of a "if it isn't broken, don't fix it" mentality.

Indeed, as Dame Sharp in *Ayinde* made clear, analogizing hallucinated citations to "the mislabelling of a tin where the tin, in fact, contains the correct product…entirely misses the point and shows a worrying lack of insight."[186] In other words, an AI providing a correct rule but with a wrong citation is a very serious problem "showing a worrying lack of insight" precisely because it can disrupt the established and familiar common law ecosystem for the exploration of case provenance, lineage, and chain of doctrinal authority. From the common law system's perspective, what at first appears as an efficiency gain, at the systemic level, thus becomes a potential source of pollution and contamination, increasing the maintenance costs of interacting with the legal system, particularly the so-called legal archive. In this sense, a narrative of unintended consequences becomes even more apparent, and "empirically, complex systems do not have obvious one-dimensional cause-and-effect mechanisms, and that under opacity, you do not mess with such a system."[187] Hence, the "efficient" thing to do is to preserve the system in its current state and prevent further potential unverifiability of sources and thus destabilization of the entire structure, at least until AI stops hallucinating - and from the perspective of common law's critical dependence on categorically verifiable precedent, this could indeed be the right call, especially considering that we have mathematical evidence showing the difficulties in completely eliminating the hallucination problem.[188] But it is important to note that while some degree of

---

[184] Mata v. Avianca, Inc., No. 22-cv-1461 (PKC), 2023 WL 4114965, at 1-2 (S.D.N.Y. June 22, 2023) *(opinion & order on sanctions)*.

[185] Ayinde v. London Borough of Haringey, [2025] EWHC 1383 (Admin) (Eng.) ¶ 24.

[186] Ibid. ¶ 67.

[187] Nassim Nicholas Taleb, SKIN IN THE GAME: HIDDEN ASYMMETRIES IN DAILY LIFE, 9 (Random House 2018).

[188] Ziwei Xu, Sanjay Jain & Mohan Kankanhalli, *Hallucination Is Inevitable: An Innate Limitation of Large Language Models*, arXiv:2401.11817, at 10-11 (2024). Noting that: "…hallucination is defined as inconsistencies between computable LLMs and a computable ground truth function. By utilizing results in learning theory, we show that hallucination is inevitable for computable LLMs if the ground truth function is any computable function. Since the formal world is a part of the real world, we further conclude that it is impossible to eliminate hallucination in the real world LLMs… We emphasize that since hallucination is inevitable, rigorous study of the safety of LLMs is critical and urgent."

hallucination may remain an inherent property of LLMs, significant research is currently dedicated to developing sophisticated hallucination mitigation frameworks and hallucination-reduction techniques.[189]

However, it would not be entirely correct to state that all aspects of "efficiency" are ignored in the common law systems. If we consider the acceptance of algorithmic decision-making in the name of "efficiency,"[190] particularly in the U.S.,[191] but opposition to AI adjudication despite possible efficiency gains, we are faced with a paradox. In fact, the judiciary's attitude towards integration of algorithmic tools versus AI tools into the common law reveals a fundamental contradiction in common law's attitude towards "efficiency." On the one hand, the judiciary, as demonstrated by the *State v Loomis* case, are willing to conditionally accept automated decision-making systems, for instance, risk assessment tools like COMPAS, to inform correctional and sentencing decisions, citing the aforementioned goals of "efficiency" and "evidence-based practice."[192] Furthermore, In *Loomis*, Judge Bradley was much more willing to give a chance to these opaque "evidence-based" risk-assessment tools, stating that "the concerns we address today may very well be alleviated in the future"[193] - in other words, time should be given, and the outright rejection of these opaque, flawed and discriminatory tools would be premature. On the other hand, *Mata, Ayinde*, and other cases with AI hallucinated precedent do not show such welcoming deference to the possibility of developing a more "efficient" system - here, the introduction of generative AI into the adversarial process meets strong judicial resistance, and from the perspective of the attorneys, their "gatekeeping role" for the purpose of "ensur[ing] the accuracy of their filings" in AI hallucination contexts is emphasized - which shows a clear system level thinking about AI's role in the legal edifice.[194]

---

[189] For a recent survey on hallucinations and hallucination mitigation methods, see generally: Lei Huang et al., *A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions*, 43 ACM Trans. on Info. Sys. 1, 1–55 (2025).

[190] State v. Loomis, 881 N.W.2d 749 (Wis. 2016).

[191] Sarah Brayne & Audrey Christin, *Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts*, 68 SOC. PROBS. 608 (2021). Sarah Brayne, *Predict and Surveil: Data, Discretion, and the Future of Policing* (Oxford Univ. Press 2021). For instance, the proliferation of privatized AI surveillance through companies like Palantir exposes a fundamental tension in the AI ethics discourse, namely that the widespread deployment of these technologies contradicts the narrative that innovation in this space should be handled with extreme caution.

[192] State v. Loomis, 881 N.W.2d 749 (Wis. 2016), stating that "the judiciary "has a vital role to play in ensuring that criminal justice systems work effectively and efficiently to protect the public by reducing recidivism and holding offenders accountable." ¶ 1

[193] Ibid.

[194] Mata v. Avianca, Inc., No. 22-cv-1461 (PKC), 2023 WL 4114965, at 1 (S.D.N.Y. June 22, 2023). *(opinion & order on sanctions)*.

A peculiar question emerges - if *Loomis* was decided today, but, for the sake of argument, the parties' submissions contained AI generated precedent, would the judges find more cause for concern on the point of fake citations rather than with the possibility that harsher sentences would be administered as a result of relying on automated risk assessment tools? In short, while algorithmic and AI tools directly affecting underprivileged populations via sentencing outcomes, such as the case in *Loomis*, are tolerated by the courts in the name of effectiveness and efficiency, cases involving fake citations such as in *Ayinde* and others, are rebuked as having "serious implications for the administration of justice and public confidence in the justice system"[195], which shows selective sensitivity. It thus seems that the common law is more forgiving of algorithmic tools that have the potential to directly affect the liberty of individuals than of AI tools that might distort legal reasoning and disrupt the historical accuracy of the legal precedential record. Thus, the efficiency question transforms itself into a question of "socio-technical division of work", or "the distribution of acting functions among humans and machines."[196] In other words, the deeper question lies in exploring which judicial functions in the common law are tolerated for optimization and efficiency gains and which ones are not. Algorithms and AI for sentencing, surveillance and policing are tolerated, and in fact, encouraged, whereas AI for adjudication is met with opposition.

Thus, in our view, efficiency rationales are necessary but not sufficient as rationales, precisely because efficiency considerations are non-uniform across different legal tasks and contexts and even within the same jurisdiction. Legal systems may theoretically pursue AI-driven efficiency, yet institutional and systemic constraints determine where change is permitted and encouraged and where it is resisted.

B. Ideological Perspectives

What can be called the common law's paradox of efficiency—its ready acceptance of algorithmic and AI tools in certain legal sub-domains in the name of efficiency, but a marked resistance and anxiety about uncritical AI deployment in adjudication and litigation—reveals the selective nature of first order efficiency rationales. To explain this paradox, it is necessary to examine AI innovations in adjudication not only for how they improve efficiency but "…look behind technical things to notice the social circumstances of their development, deployment and use."[197]

To explore the ideological dimension of the reception of AI in adjudication, it may be helpful to turn to Langdon Winner's work on the politics of technology where he introduces the concept of "inherently political technologies" - which are technologies

[195] Ayinde v. London Borough of Haringey, [2025] EWHC 1383 (Admin) (Eng.) ¶ 9.

[196] Günter Ropohl, *Philosophy of Socio-Technical Systems*, 4 SOC'Y FOR PHIL. & TECH. Q. ELEC. J. 186 (1999), 191.

[197] Langdon Winner, *supra* note 20, at 122.

which either "require…or are strongly compatible…with the creation and maintenance of a particular set of social conditions."[198] Winner provides Plato's famous "ship of state" analogy as an example, whereby a ship, by its very nature as a piece of technology, requires a particular kind of political arrangement for its effective use—there can only be one captain, and every other sailor must obey the captain for "no reasonable person believes that ships can be run democratically."[199] Nuclear technologies are another illustrative example, whereby centralized authoritarian and hierarchical institutional arrangements are necessarily required in order to deal with things like theft of hazardous nuclear material.[200] Conversely, solar power is compatible with a more "democratic, egalitarian society" in that it is preferable to "build solar systems in a disaggregated, widely distributed manner."[201] Applying the notion of an "inherently political technology" to the present discussion, we can ask whether AI technologies required for adjudication possess features that similarly predispose them toward particular institutional arrangements. In other words, what are AI's technical characteristics that may be more compatible with certain legal and organisational forms?

For instance, consider the basic requirements for developing a high-quality legal LLM AI system. At a minimum, such models require enormous quantities of electrical power, specialised computational infrastructure, and large datasets and storage. In fact, the energy and storage demands involved in AI development have become so significant that companies like Google are now making deals with nuclear energy companies to build nuclear power plants to support their compute needs.[202] All of this is notwithstanding the substantial cadre of highly specialised engineers and computer scientists required to design, maintain, and validate the systems on a consistent basis. Clearly, the introduction of AI in adjudication necessarily exhibits centralizing political tendencies by concentrating research and development in institutions capable of mobilising significant capital, AI infrastructure, and expertise, namely large corporations or states. And as Winner reminds us, technological change "affects the relative distribution of power, authority, and privilege in a community,"[203] and innovations in AI adjudication, being no exception to this rule, carry a distinctly political dimension in the common law in that they could potentially change the traditional domain of judicial authority and discretion of common law judges, which helps explain why they are more sceptical of the use of AI in adjudication as compared to its use in sentencing or bail contexts.

---

[198] Ibid. at 123.

[199] Ibid. at 129.

[200] Ibid. at 134.

[201] Ibid. at 130.

[202] Google turns to nuclear to power AI data centres, BBC News (Nov. 12, 2024), https://www.bbc.com/news/articles/c748gn94k95o.

[203] Langdon Winner, *supra* note 20, at 127.

If AI adjudication is an "inherently political technology" which requires certain institutional arrangements to be implemented, it is thus unsurprising that common law judges almost instinctively perceive AI technology as a direct threat to the entire legal system, as seen in their rhetoric; they are acutely aware that even modest technological influence could adversely alter the delicate balance of existing instruments of power at their disposal. Furthermore, if "inherently political technologies" are more compatible with certain institutional arrangements, such as centralized and hierarchical governance structures, it is also unsurprising why Chinese judges do not see AI adjudication as a threat to the legal edifice. As we have discussed in Part II, judges in China serve a comparatively more bureaucratic role and are inherently subordinate to state's policy and thus do not have the same professional mandate as we see in common law judges. The comparison can be better illustrated by an analogy between interrelated disciplines of architecture and civil engineering: common law judges see themselves as more of architects, responsible for designing and updating the legal edifice, whereas Chinese judges see themselves as more of civil engineers, focused on implementing practical solutions within a centrally designed plan.

Connected to the distinct perceptions that common law and Chinese judges have of their roles is the fact that AI may pose a threat to the instruments of power at the disposal of the common law judiciary that need to be preserved, potentially even at the expense of speed and efficiency that AI seems to offer. There are two such key instruments of power. The first is the judges' discretionary power and the second lies in the power of the judge to interpret the legal archive. We elaborate on each of them below.

On the utility and significance of discretionary power, 18th century English criminal law offers a timeless lesson. One of its strange puzzles was the fact that although there was a significant increase in capital statutes, known as the "Bloody code", surprisingly, the proportion of death sentences that were carried out was, in actuality, declining.[204] The role of the pardon -discretionary, inefficient and even absurd -played a key role in this. Poor offenders were frequently spared, while on some occasions even the members of the elite class, such as the "wicked aristocrat" Lord Ferrers, were executed.[205] This selective and strategically deployed discretion created a powerful ideological effect. It sustained the appearance of a just and responsive legal order, even though the machinery of criminal law remained firmly in the hands of the ruling class in addition to the fact that the entire pardoning process was deeply paternalistic. Discretion was, in essence, an ideological instrument that allowed judges and elites to carefully calibrate punishment "with delicacy and circumspection" to social pressures, political contingencies, and prevailing moral sentiments. The result was a system that preserved its legitimacy, a feat all the more impressive considering the fact that England at the time was "without a police force and without a large army."[206]

---

[204] Douglas Hay, *Property, Authority and the Criminal Law*, in ALBION'S FATAL TREE: CRIME AND SOCIETY IN EIGHTEENTH-CENTURY ENGLAND 17, 23 (Douglas Hay et al. eds., Pantheon Books 1975).

[205] Ibid. at 34

[206] Ibid. at 56.

The common law's ideological effectiveness "lay first in its very elasticity... the ability to sacrifice punishment when necessary to preserve the belief in justice"[207] and indeed, "the law's very inefficiency was part of its strength as an ideology."[208] In short, the systemic inefficiency and discretionary fluidity of the common law was not a bug, but a feature: they were mechanisms that stabilised authority by maintaining the appearance of justice while retaining ultimate control in the hands of the ruling classes.

Seen through this lens, modern common law judges retain a structurally and ideologically similar role. Their power can be interpreted as depending on maintaining ideological flexibility and elasticity, the ability to interpret, distinguish, soften, or expand doctrine while presenting such moves as neutral, principled adjudication. Any technology that constrains or standardises this flexibility—automating reasoning, homogenising interpretations, or limiting the narrative space in which discretion operates—threatens a key instrument of judicial authority in the common law. By contrast, Chinese courts seek to eliminate judicial discretion because it threatens the authority of the Chinese Communist Party.[209]

Incidentally, the potential "constraints on flexibility" were the same grounds upon which more "rational" reforms of the criminal law were "instinctively rejected" in eighteenth-century England, for a "complete rationalization of the criminal law would remove those very elements of discretion, such as the pardon, which contributed so much to the maintenance of order and deference."[210] This is why we see such sensitivity, why the seemingly "inefficient" processes are jealously guarded, for they preserve the critical zones of discretion through which legitimacy is continually produced and maintained. And this leads directly to the second essential instrument of common law's judicial power threatened by AI adjudication, namely the control over the legal archive.

---

[207] Ibid. at 55.

[208] Ibid. at 33.

[209] A related consideration is that reliance on government-approved AI systems may function as a mechanism of responsibility-shifting within China's judicial hierarchy. When judicial decisions are supported by officially endorsed AI tools—the Similar Case Push, the Uniform Sentencing Assistance System, or Faxin 2.0 recommendations—individual judges gain a measure of insulation from personal blame for outcomes that generate dissatisfaction. Public criticism that might traditionally target an individual judge can instead be redirected toward 'the system' or 'the algorithm.' This diffusion of responsibility serves multiple functions: it reduces the personal stakes for judges following AI recommendations, thereby encouraging compliance with centrally determined standards; it provides a bureaucratic defence against accusations of corruption or favouritism; and it channels grievances away from individual state agents toward impersonal technological processes. Paradoxically, AI thus strengthens both uniformity and individual judges' protection simultaneously; conforming to algorithmic recommendations becomes the safest path both for producing state-sanctioned outcomes and for avoiding personal accountability. This responsibility-diffusing function represents an underappreciated appeal of AI standardisation within China's judicial bureaucracy, complementing the state's interest in eliminating discretion with individual judges' interest in self-protection.

[210] Douglas Hay, *supra* note 204, at 57.

C. Archival Perspectives

We now turn to a critical archival perspective to explain why common law jurisdictions often resist AI adjudication while readily embracing algorithmic risk assessment and other AI-enabled decision-support tools. We argue that the judiciary seeks to preserve a key instrument of authority, discretion and power, namely the power to determine, curate, and interpret the "legal archive."[211] From this critical archival perspective, the so-called "archival turn", AI "hallucinations" cannot be seen as mere technical flaws that can be "fixed", but in fact, are arrows that strike at the very foundations of legal power.

To better understand the "archival turn" and the power relations embedded in the archive, our first clue comes from the word "archive" itself, which is directly derived from the ancient Greek word "archon", meaning "those who commanded."[212] For example, the power of the Hellenistic archons has been linked to the power to interpret documents and records in the archive:

> "...the archons are first of all the documents' guardians. They do not only ensure the physical security of what is deposited and of the substrate. They are also accorded the hermeneutic right and competence. They have the power to interpret the archives. Entrusted to such archons, these documents in effect state the law: they recall the law and call on or impose the law."[213]

In other words, all authority, including judicial authority, since times immemorial, has been linked to the archive, at the very least on an etymological and epistemic level. The French anthropologist, Claude Lévi-Strauss, provides perhaps the most convincing evidence of this epistemic link between the archive and the archon. In a famous episode recounting his interactions with the Nambikwara of the Brazilian Amazon, aptly titled "a Writing lesson", Lévi-Strauss recounts a story of the leader of the Nambikwara community (a community which had no written language) who started scribbling "wavy lines" on a piece of paper "and pretended to read from it."[214] Of course, the leader of the Nambikwara did not suddenly learn how to write in the brief hours of interacting with Lévi-Strauss - he merely saw Lévi-Strauss doing it and started mimicking it. But nevertheless, despite his lack of knowledge of writing, the leader immediately and intuitively used this

---

[211] Renisa Mawani, *Law's Archive*, 8 ANNU. REV. L. & SOC. SCI. 337 (2012).

[212] Jacques Derrida, Archive Fever: A Freudian Impression, 25 *Diacritics* 9, 9–63 (Summer 1995). Alex McAuley, Officials and Office-Holding, in A COMPANION TO ANCIENT GREEK GOVERNMENT [Hans Beck, ed.] (John Wiley and Sons 2013).p. 179, stating that: "The Greek vocabulary of office-holding is derived from the verb *archein*, generally translated as "to rule, govern, or command." Of course, the more familiar words relating to state authority like "monarch" (single "archon"), and "anarchy" (no "archon") stem from the same root.

[213] Jacques Derrida, *Archive Fever: A Freudian Impression*, 25 DIACRITICS 9, 10 (Summer 1995).

[214] Claude Lévi-Strauss, TRISTES TROPIQUES, trans. John Russell (Atheneum 1961). 289.

newly acquired "skill" to arrange exchanges of presents between Lévi-Strauss and the Nambikwara peoples, looking "up and down his 'list' for the objects to be given in exchanges" in order to "persuade [his companions] that *his intermediacy* was responsible for the exchanges."[215] Upon reflecting on this peculiar incident, Lévi-Strauss hypothesises that writing "was not a question of knowing specific things, or understanding them, or keeping them in mind, but merely of enhancing the prestige and authority of one individual - or one function - at the expense of the rest of the party."[216] Lévi-Strauss notes that this is a universal phenomenon that can be observed "from Egypt right across to China, at the moment when writing makes its debuts... "[217] The written word, records, documents, archives - and the whole apparatus of producing, interpreting, and storing them - functions, in Lévi-Strauss's view and in keeping with the etymology of "archive", not simply as a passive and neutral vehicle of knowledge but as an instrument and source of power first and foremost. The archive, in essence, is the manifestation of the asymmetry of knowledge and power, a manifestation of epistemic gatekeeping - and indeed, "without asymmetries in knowledge, there cannot be differentials in power."[218]

What is aptly called the "archival turn"[219] - which is defined as the move from "archive-as-source to archive-as-subject" where the "focus on the *politics of knowledge* is a methodological commitment to how history's exclusions are secured and made" – allows us to critically question the archive itself, and allows us to use "the terminology or image of the archive to describe a set of ideas about the structural power of history and its ideological and technological conditions."[220] In this vein, the archive and its compositionality have been directly linked to both how legal issues are framed and their legal significance, where decisions about excluding or including certain records (including what is cited and footnoted) serve to legitimize and set precedent for narratives aligned with state objectives.[221] In this sense, 'the archive' does not merely preserve the law (as in the classical juridical view) but determines what is deemed relevant or irrelevant, legally significant or insignificant, and what enters or does not enter the

[215] Ibid., emphasis added.

[216] Ibid., 290

[217] Ibid.

[218] Konstantin Offer, Zoe Rahwan & Ralph Hertwig, Foucault's Error: The Power of Not Knowing, EUR. REV. SOC. PSYCHOL. 1, 1–36 (2024). 2

[219] For a historical tracing and a brief genealogy of the "archival turn" see: Ann Laura Stoler, ALONG THE ARCHIVAL GRAIN: EPISTEMIC ANXIETIES AND COLONIAL COMMON SENSE (Princeton Univ. Press 2008). 44

[220] Samantha Callahan, *When the Dust Has Settled: What Was the Archival Turn, and Is It Still Turning?*, 83 ART J. 74 (2024).

[221] Ann Laura Stoler, *supra* note 219, at 50.

archive. Through these elaborate and countless means and permutations of influence, the custodians of 'the archive' actively participate in the making of the law and in determining what counts as law to begin with.[222]

Armed with this understanding of the archive, we can state that under this conception the judge-archon is not simply the one who "applies the law", but is the one who also has the power to guard, modify, define and interpret the archive and the record. The archival understanding allows us to state that the "legal archive" is not a mere neutral repository of legal knowledge but is "the product of ongoing struggles over the production, politicization, and institutionalization of knowledge", and "a politically charged site of contestation… a site of force and command."[223] Law's archive operates on the "double logic of violence, as both the violence of law *and* the violence of the archive"[224]; all of us are familiar with the former, but in legal scholarship, there is relatively little focus on the latter. Could this give another explanation to the question why the issue of AI hallucinations when discussed by common law judges has "serious implications for the administration of justice and public confidence in the justice system?"[225]

Of course, one might counter these claims about "the archive" by stating that everybody knows (or at least implicitly understands) the enormous power that politicians, judges and lawyers, and legal professionals have over the determination of the "law's archive" - but here, we should be reminded that "[ideology] presents itself as an 'everybody knows that' - a kind of anonymous universal truth."[226] Upon explicitly recognizing the ideology of the archive as an essential element of the ideology of power, violence and domination by sources "determined by the upper class"[227] , we can come to a better understanding of the common law judiciary's visceral and almost reflexive sensitivity and skepticism with regards to AI's use in adjudication: it is clear that, at least from the perspective of legal professionals, their power of archival sovereignty is being undermined. As we established, decisions of what to include or exclude in a law's archive are not mere chronicles that reflect "legal reality", but should be read as a history of strategic choices that lie at the very core of archival sovereignty which constantly (re)shape the legal reality of the past, the present and the future.

Here, we gain another perspective on the question of why there is reflexive sensitivity to AI hallucinations in the common law, and relatively less concern about it in Chinese judicial discourse. From the perspective of the archive, AI systems that generate false

---

[222] Renisa Mawani, *supra* note 211, at 340.

[223] Ibid. at 351.

[224] Ibid. at 360 (emphasis added).

[225] Ayinde v. London Borough of Haringey, [2025] EWHC 1383 (Admin) (Eng.) ¶ 9.

[226] Terry Eagleton, IDEOLOGY: AN INTRODUCTION, 3 (Verso 1991)

[227] Ann Laura Stoler, *supra* note 219, at 47.

citations and legal reasoning intrude precisely into the core of archival sovereignty of the judiciary, which, as we established, is perhaps the most important element of judicial power. The narratives through which law legitimates its own power are a unique province of the judge-archon, where intrusion and interlopers are to be destroyed. Importantly, what is at stake, then, is not technical accuracy or efficiency of the legal system, but a destabilisation of the institutional understanding and arrangement that only the judiciary may speak for the archive.

But going a step further, the consequences might extend beyond institutional arrangements - what is also implicated is the legal professional identity and the social legitimacy that flow from the judiciary's exclusive claim and power to produce legal meaning. In this sense, AI systems have the potential to disrupt the social arrangement that reserves the interpretation, curation, and articulation of the law's archive to a social class of professional legal actors, thus, exerting pressure not only on institutional authority of the judiciary, but on the very identity of legal actors and the epistemic hierarchy that sustains the legal order as such. At a deeper level, the resistance to AI suggests that the common law's sensitivity to AI may be less about concerns over errors – for biased risk assessment tools are widely developed and used - and more about the preservation of specific power arrangements that define who is permitted to determine and speak for the law's archive. Thus, the radical challenge of AI adjudication in the common law goes beyond mere problems with hallucinations, and raises question regarding judiciary's epistemic gatekeeping and monopoly over the law's archive as such, bringing the underlying power dynamics and of the legal order into sharp relief.[228]

In short, the archival perspective helps us answer the question why common law jurisdictions tolerate algorithmic tools that operate as advisory or statistical instruments, while reacting defensively to fake AI generated citations; the former augment judicial power while the latter compete with it. Thus, the sensitivity about AI hallucinations in common law judicial discourse is not only a technical issue about the errors in the outputs, but is in fact, at its core, a political, ideological and epistemic rejection. The archival perspective also helps explain China's relative lack of anxiety over AI hallucinations namely, that the Chinese judiciary does not occupy the same position of archival sovereignty as its common law counterparts. Unlike the common law which "derives its command from its lengthy history of precedent and its mythical claims to 'time immemorial'"[229], the Chinese judge derives its authority as a bureaucratic agent of the state; the ultimate authority of the archive does not (yet) lie in the hands of the Chinese judiciary but with the state, as we have discussed in Part II of this article. And

[228] Samuel I. Becher & Benjamin Alarie, Legal Order in the Age of AI Agents, U. TORONTO L.J. (forthcoming 2026) (manuscript at 13), available at: https://ssrn.com/abstract=6001277. This epistemic archival perspective on judicial power is further solidified by recent scholarship on the potentials of AI Agents in the legal domain, stating that "AI legal agents would do more than distribute legal knowledge—they could redistribute interpretive authority. This, in turn, would empower users to engage in legal meaning-making, re-shaping how legal systems function and evolve."

[229] Renisa Mawani, *supra* note 211.

because the Chinese judge is not the ultimate sovereign of the law's archive, the challenge of AI is not met with such resistance.

IV. Implications for Global AI Governance

The foregoing analysis in Part III reveals that divergences in AI governance extend beyond institutional differences or policy choices. They reflect deep structural incompatibilities: efficiency serving different ends, technology embodying political arrangements, and archival sovereignty defining judicial power itself. These insights fundamentally raise questions of how we should approach global AI governance.

Three implications arise from the fundamentally incompatible meanings. First, attempts at universalizing or harmonizing global AI governance must be treated with caution. Second, ensuring AI legitimacy or alignment with public interest cannot be based on technical standards alone. Finally, the civil law/common law divide may facilitate or undermine how AI is adopted or its risks addressed.

A. The Limits of Harmonized Terminology

International AI governance frameworks (like OECD AI Principles or EU AI Act) relying on terms like "trustworthy," "transparent," or "fair" AI are likely to fail in creating meaningful global judicial standards. AI systems presuppose particular institutional arrangements. Common law systems, where judicial power derives from archival sovereignty and interpretive discretion, cannot simply adopt Chinese approaches premised on centralized state control, and vice versa, because the underlying technologies are compatible with fundamentally different political orders.

These terms function as homonyms: same words, entirely different operational meanings. China can claim "transparency" compliance while implementing systems (Shanghai 206, Faxin 2.0) that make judicial decisions legible to supervisors but opaque to litigants. Global governance efforts must acknowledge this semantic divergence rather than assume shared understanding.

Both common law and Chinese courts endorse human oversight of AI, but this produces opposite power dynamics. Under the common law, judges serve as checks on algorithmic error, overriding AI when it threatens individual rights. By contrast, judicial human oversight in China means supervisors reviewing judges who deviate from AI, a mechanism for enforcing algorithmic conformity rather than checking algorithmic error.
Global standard-setting must incorporate definitional clarity and minimum common denominators, avoiding assumed shared meanings. For example, "transparency" in

global standards should specify whether it targets party-facing contestability (common law) or internal supervisory monitoring (China), and "fairness" should indicate whether it is measured by individual harm or social harmony impacts.

B. The Inadequacy of Technical Standards Alone

Common law systems derive legitimacy from adversarial process, allowing erroneous or questionable AI outputs or decisions to be discovered, contested, and corrected. Chinese systems derive legitimacy from state control where erroneous outputs or decisions are detected and prevented through pre-emptive data curation and supervision of judicial deviations from algorithmic recommendations.

International standards often call for high-quality training data or removal of erroneous data.[230] China's Supreme People's Court can truthfully claim it ensures data quality by excluding anomalies or rulings inconsistent with the Chinese Communist Party policy. Data undergoes training not on the complete spectrum of Chinese jurisprudence but on a screened subset endorsed by governmental authority. By contrast, common law systems likely use comprehensive case law databases without pre-screening for political correctness. Differences in training models and data curation underscore that focusing on technical "data quality" requirements alone misses a key point: these technical requirements are vehicles for embedding political values.

Certification regimes[231] will likely diverge. An AI system certified as "safe" in China (because it outputs only Chinese Communist Party-approved legal interpretations) would fail common law standards (because parties cannot meaningfully contest its reasoning). Conversely, Western commercial AI trained on uncurated or non-state pre-approved data would be considered unacceptable by Chinese courts; nor would giving judges discretion to deviate from algorithmic recommendations based on state-sanctioned databases. This divergence calls into question whether technical standards alone suffice for ensuring AI systems are "safe" or aligned with public values.[232]

---

[230] See eg, EU AI Act: Article 10(3) (Data and data governance for high-risk AI systems). It states that training, validation, and testing datasets must be "relevant, sufficiently representative, free of errors and complete as far as possible," taking into account the intended purpose of the system. ISO/IEC 5259-2:2023 — Data quality for analytics and AI — Part 2: Data quality assessment.

[231] See eg, ISO/IEC 42001:2023Information technology — Artificial intelligence — Management system https://www.iso.org/standard/42001; OECD AI Policy Observatory – Catalogue of Tools & Metrics for Trustworthy AI, AI certifications and quality marks, https://oecd.ai/en/catalogue/tools?approachIds=3&approachIds=2&toolTypeIds=20&toolTypeIds=21&orderBy=dateDesc

[232] see eg, NIST AI Risk Management Framework (SP 1270), https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf. NIST explicitly frames AI risk and trustworthiness. as socio-technical, noting that standards, tests, and metrics are necessary but not sufficient on their own; effective assurance requires governance, context-specific practices, and broader stakeholder

International AI governance bodies should abandon one-size-fits-all technical standards. Instead, they should create parallel tracks with system-specific benchmarks, acknowledging that "trustworthy AI" means different things in different legal contexts, and recognising a plurality of governance toolkits including decentralized adversarial safeguards and centralized technocratic controls.

C. The Common Law/Civil Law Variable

Parts II and III address the semantic divergence of ostensibly shared AI judicial concepts in common law and Chinese courts and their underlying reasons. The comparison between AI deployment in these two kinds of courts raises a distinct, broader issue for global AI judicial deployment - whether a system is common law or civil law may influence how AI is used and what guardrails are established.

In civil law jurisdictions—inquisitorial systems—judges lead in identifying and evaluating law, facts, and all relevant evidence[233], unlike common law systems where judges generally depend on litigants to put forward arguments and evidence.[234] Judges in civil law systems generally do not depend on parties to bring evidence and submit arguments, unlike in common law adversarial systems. The common law adversarial system has been described as a "sporting theory of justice"[235] where judges act as neutral arbiters, not systemic problem solvers, adjudicating one case at a time.[236]

An important implication of this is that courts in civil law jurisdictions minded to use AI may be more inclined to use it in different aspects and stages of legal proceedings, given that civilian judges can provide more direction of the process and have more control of the outcome of a case. Compared to judges in most common law systems[237], judges in certain civil law countries appear to have fewer reservations about using AI, especially in determining whether decisions will be heard on appeal and in writing judgments.

---

considerations. This implies that relying solely on technical standards cannot guarantee safety or value alignment, especially across varied regimes.

[233] See John H. Langbein, *The German Advantage in Civil Procedure*, 52 UNIV. CHI. L.REV. 823, 831–32 (1985).

[234] Robert A. Kagan, ADVERSARIAL LEGALISM: THE AMERICAN WAY OF LAW 8 (2nd ed. 2019).

[235] Roscoe Pound, *The Causes of Popular Dissatisfaction with the Administration of Justice*, 14 AM. LAW. 445, 447 (1906).

[236] Colleen F. Shanahan et al., *The Institutional Mismatch of State Civil Courts*, 122 COLUM. L. REV. 1471, 1524, n.235 (2022).

[237] Leo You Li, *Digitization, Adversarial Legalism, and Access to Justice Reforms*, 76 SCL REV. 883 (2024). 901-2

Beyond the Chinese example, consider Latin America, a major civil law continent. In Argentina, courts use generative AI extensively in case management and for legal research and analysis.[238] AymurAI operates across 7 provinces to systematically compile and release anonymized judicial records on gender violence, detect potential feminicide risk patterns, and enhance justice system transparency.[239] Another example is a custom AI platform for handling tax debt litigation that automates document validation, identifies connected proceedings, and creates digital dockets.[240] This system cut average case resolution from 6.5 to 2.86 days[241] and generated approximately 6,000 decisions by September 2023.[242] In Brazil, the highest court employs VICTOR, an AI platform processing the massive number of submitted cases by automating appeals review.[243] Specifically, it determines which cases meet the repercussão geral criteria—a mandatory threshold before the Supreme Court can hear appeals.[244] Another Brazilian court AI system is ASSIS (Assistente de Inteligência Artificial Generativa), which produces preliminary versions of court rulings, judgments, and legal memoranda through GPT-4-powered generative technology.[245] It customizes content to reflect individual judges' linguistic preferences and jurisprudential history by analyzing their previous decisions. By contrast, judges in numerous common law jurisdictions have expressed grave reservations about using AI to assist with judgment writing.[246]

---

[238] Decisión histórica Poder Judicial argentino: por qué la IA puede ser clave para su funcionamiento (*Historic Decision: Argentine Judiciary – Why AI Could Be Key to Its Functioning),* (Oct. 1, 2024), available at: https://www.innovaciondigital360.com/i-a/poder-judicial-argentino-por-que-la-ia-puede-ser-clave-para-su-funcionamiento/

[239] *Aymur AI: Measuring Gender-Based Violence in Latin America)*, A+ Alliance (July 19, 2023), available at: https://aplusalliance.org/aymur-ai-measuring-gender-based-violence-in-latin-america/

[240] Río Negro: un fuero dictó casi 6.000 fallos con inteligencia artificial Nacionales (*Río Negro: A Court Issued Nearly 6,000 Rulings Using Artificial Intelligence)*, Opinorte (Sept. 28, 2023), available at: https://opinorte.com.ar/rio-negro-un-fuero-dicto-casi-6-000-fallos-con-inteligencia-artificial/

[241] La IA reduce los plazos (*AI Reduces Deadlines)*, Diario Judicial (Apr. 15, 2024), available at: https://www.diariojudicial.com/news-97536-la-ia-reduce-los-plazos

[242] Río Negro: la Justicia hizo 6.000 sentencias con IA (*Río Negro: The Judiciary Issued 6,000 Rulings Using AI)*, LM Cipolletti, available at: https://www.lmcipolletti.com/pais/rio-negro-la-justicia-hizo-6000-sentencias-ia-n1061370

[243] Salomao L. F., Braga R*The role of the Judiciary in the realization of the UN 2030 Agenda* (2020). available at: https://www.conjur.com.br/2021-jul-09/salomao-braga-judiciario-agenda-2030-onu. See also: https://portal.fgv.br/en/news/artificial-intelligence-Judiciary-and-itsrole-implementing-un-agenda-2030; https://sifocc.org/app/uploads/2020/06/Victor-Beauty-or-the-Beast.pdf

[244] Ibid.

[245] Tribunal de Justiça do Estado do Rio de Janeiro (TJRJ), *O Projeto – ASSIS* (launch page for AI-assisted judicial assistant) (2025), available at: https://www.tjrj.jus.br/magistrado/servicos/assis/o-projeto

[246] For UK see: Courts & Tribunals Judiciary (UK), *Artificial Intelligence (AI) Guidance for Judicial Office Holders* (Apr. 14, 2025), available at: https://www.judiciary.uk/wp-content/uploads/2025/04/Refreshed-AI-Guidance-published-version-website-version.pdf; Erin Solovey, Brian Flanagan & Daniel Chen, *Interacting with AI at Work: Perceptions and Opportunities from the UK Judiciary*, in *Proceedings of the 4th Annual Symposium on Human-Computer Interaction for Work* 1–8, ¶4.2.1 (Jun. 23-25 2025).; For Canada

A word of caution is warranted. While certain[247] civil law systems may be better positioned to adopt AI than common law systems, and evidence suggests their courts have already begun doing so, there is an important paradox. Courts pioneering AI adoption, particularly when judges allow algorithms to substitute for their own reasoning and discretion, may find their autonomy diminished rather than strengthened by technology. Extensive AI incorporation could result in judicial power being curtailed by the very tools meant to assist it. As discussed in Part III above, this potential curtailment of power is perceived as a profound threat to judges' authority in common law systems.

The potential curtailment of judges' power in civil law compared to common law systems must be counterbalanced by the access to justice problem, arguably worse in common law systems. If the judge, as the most important figure in an inquisitorial system, takes charge of proceedings, the judge in a common law adversarial system depends on two equal parties putting forward opposing arguments advancing their client interests.[248] The implication: groundwork must be furnished by the parties—affidavits, bundle of authorities, legal arguments. In cases where one party has more resources, the other party with fewer resources is more likely to rely on free foundation models (as opposed to law-specific LLMs like Harvey[249]) or has lesser capability to ascertain AI output veracity. This is nowhere truer than in cases involving self-represented litigants, where most AI hallucinatory cases can be found.[250] Thus, paradoxically, contrary to the notion that AI has democratizing or equalizing benefits, it may exacerbate access to justice problems in common law courts.[251]

Another aspect of the common law/civil law distinction: in the latter, cases are not decisive; interpretation of statutory provisions is critical. By contrast, cases in common

---

see: Federal Court of Canada, *Interim Principles and Guidelines on the Court's Use of Artificial Intelligence* (Sept, 29 2025), available at: https://www.fct-cf.ca/en/pages/law-and-practice/artificial-intelligence; For Australia see: *Supreme Court Practice Note SC Gen 23: Use of Generative Artificial Intelligence (Gen AI),* available at: https://supremecourt.nsw.gov.au/documents/Practice-and-Procedure/Practice-Notes/general/current/PN_Generative_AI_21112024.pdf?fbclid=IwY2xjawG5qUNleHRuA2FlbQIxMAABHdBdpXjl6wrlRAbxyqFq9FUUtbKV2GvKMZ7gUpyCCxbLoqFIHlbZ1-Ia-A_aem_v065CZgaA6Ji2gB32NRdag

[247] To be clear, not all civil law systems per se will take the lead, but the nature of the legal system in itself, all factors being equal, will better facilitate or undermine the use of AI.

[248] For e.g., Judicial ethics rules, such as Canon 3B(7) of the California Code of Judicial Ethics, prohibit judges from independently investigating facts in a proceeding, even when such investigations could illuminate systemic patterns or abusive behaviors" CAL. CODE. OF JUD. ETHICS Canon 3B(7) (2020).

[249] Harvey AI, https://www.harvey.ai (accessed Dec. 10, 2025).

[250] Michael Legg, *More People Are Using AI in Court, Not a Lawyer. It Could Cost You Money — and Your Case,* The Conversation (Sept. 28, 2025), available at: https://theconversation.com/more-people-are-using-ai-in-court-not-a-lawyer-it-could-cost-you-money-and-your-case-264340.

[251] Drew Simshaw, *Access to A.I. Justice: Avoiding an Inequitable Two-Tiered System of Legal Services*, 24 YALE J.L. & TECH. 150 150 (2022).

law tradition are legally binding and relevant authority can make decisive outcome differences. Thus, citing and relying on persuasive authorities is critical in common law adjudication. Litigants have every incentive to find the right cases. Increasingly, AI has been used to locate cases supporting claimant's or defendant's positions. Unsurprisingly, the bulk of cases on fake citations or precedents stem from common law courts.[252] Thus, AI governance should be legal system-specific. For example, a judicial governance framework designed to prevent citation fraud (common law risk) may be less relevant or insufficient for civil law jurisdictions facing evidence fabrication (such as spurious audio and visual recordings).

D. A Way Forward: Frame Reflective Reciprocal Translation

The divergent ways common law and Chinese courts understand and operationalize judicial concerns, rooted in competing political arrangements, efficiency logics, and archival sovereignties, suggest differences run deep and may be intractable. Nevertheless, while global AI governance standards for courts warrant caution, finding common ground on how we build, deploy, and evaluate AI systems that enhance justice without compromising what makes justice meaningful in different contexts remains important.

As Schön and Rein remind us, even intractable frame conflicts need not lead to stalemate: they call for frame-reflective dialogue rather than higher-order appeals to universal principles.[253] They argue that many enduring policy battles are intractable controversies rooted in competing frames—the implicit assumptions, values, causal stories, and metaphors through which actors "name and frame" situations. Frames, which are defined as “generic story lines that underlie the particular problem-setting stories one finds in any particular policy controversy”[254], can be used to both diagnose and prescribe; they make sense of complexity and simultaneously point to solutions that feel self-evident within that frame. Because different actors highlight different facts and interpret the same facts differently, no neutral, frame-free position exists from which one can definitively validate one frame over another.[255]

---

[252] In a worldwide record of hallucination cases from common law and civil law countries, approximately 89% of the cases come from common law countries, of which approximately 75% come from the US: Damien Charlotin, *AI Hallucination Cases Database*, Damien Charlotin.com, available at: https://www.damiencharlotin.com/hallucinations/

[253] Donald Schön & Martin Rein, *supra* note 10.

[254] Martin Rein & Donald Schön, *Frame-Critical Policy Analysis and Frame-Reflective Policy Practice*, 9 KNOWLEDGE & POL’Y 85, 89 (1996).

[255] Donald Schön & Martin Rein, *supra* note 10, at 44–45.

Schön and Rein recommend a frame-reflective approach centering on reciprocal translation, with each side "reflecting on the action frames"[256] and learning to restate the other's meanings in its own language, to build understanding where higher-order appeals rarely settle matters. Or, in the words of Schön and Rein "participants in such a [reflective policy] conversation must be able to put themselves in the shoes of other actors in the environment, and they must have a complementary ability to consider how their own action frames may contribute to the problematic situations in which they find themselves."[257] This approach treats policymaking as a design process under uncertainty in which problems and solutions co-evolve, intentions shift with experience, and feedback is interpreted and used to iteratively redefine and reframe problems. We can apply Schön and Rein's frame reflective reciprocal translation theory to bridge divergent understandings of reliability, bias, transparency, and procedural fairness.

However, it might be queried whether Schön and Rein's frame-reflective approach, developed for domestic policy controversies, can be applied to inter-state AI governance negotiations. We believe it can, for two reasons. First, the core insight--that intractable disputes arise from competing frames rather than factual disagreements, and that progress requires reflecting on these frames and making them visible rather than appealing to supposedly neutral higher-order principles[258]--applies with equal force to international negotiations where parties operate from fundamentally different institutional and ideological premises. Second, Schön and Rein themselves emphasize that frame conflicts are characteristic of situations where multiple legitimate perspectives coexist and no single frame can claim objective superiority--a condition that describes the global AI governance landscape precisely. The alternative--pursuing harmonized terminology while ignoring the divergent meanings documented throughout this article--is not merely ineffective but counterproductive, as it produces the illusion of consensus that our analysis has shown to be misleading.

1. Reliability

Common law courts could learn from China's focus on aggregate reliability, while China can allow for individualized reliability. As both systems are concerned about unreliable LLM outputs, common law courts should invest in aggregate reliability mechanisms such as curated case databases that litigants and lawyers rely on and cite, while continuing to require parties to ensure submission accuracy. For China, individualized reliability can

[256] Ibid. at 187

[257] Ibid.

[258] Ibid. at 23, stating that: "We see policy controversies as disputes in which the contending parties hold conflicting frames. Such disputes are resistant to resolution by appeal to facts or reasoned argumentation because the parties' conflicting frames determine what counts as a fact and what arguments are taken to be relevant and compelling. Moreover, the frames that shape policy positions and underlie controversy are usually tacit, which means that they are exempt from conscious attention and reasoning."

be integrated via narrow, rule-governed exceptions allowing judges to depart from AI-set ranges when unusual facts justify it, with brief judge-authored reasons captured for supervisory review so individualized justice does not undermine overall uniformity.

2. Bias and discrimination

Common law courts can view bias in AI systems as including not only differential treatment of protected groups, but also systemic patterns generating arbitrary variation across comparable cases. Where algorithmic tools yield inconsistent outcomes without principled justification, such inconsistency can constitute indirect discrimination or arbitrary decision-making. Thus, common law courts can consider adopting variants of Chinese courts' uniformity tools, such as curated case repositories and AI decision-support systems surfacing relevant precedents, to mitigate idiosyncratic judicial bias and prevent entrenchment of patterns with disparate impact.

Chinese courts can view common law courts' emphasis on individual rights and anti-discrimination as aligning with their core concern: unequal treatment without principled justification aligns with PRC law's prohibition on discrimination. In PRC terms, differential treatment lacking reasonable basis and resulting in public dissatisfaction constitutes discrimination endangering social stability. Thus, Chinese courts should explicitly put safeguards in place allowing judges to recalibrate or suspend AI tools if they reveal patterns of unreasonable differential treatment or if following AI recommendations would result in unequal treatment disturbing social order.

3. Transparency

From common law judges' perspective, both systems recognize transparency serves legitimation, ensuring judicial authority sustains public confidence. While common law emphasis is on enabling parties to challenge algorithmic outputs, common law judges may view Chinese courts' systematic oversight as achieving this goal more efficiently.

Hence, rather than relying on resource-intensive adversarial challenges case-by-case, common law courts can consider developing a variation of the Chinese Faxin system that aggregates laws and prior cases with recommendations as to the reasons and outcome when the facts and evidence of a particular case are inputted into the court’s system, and then comparing parties' submissions to the results produced by the court’s AI system. Courts can then disclose the result of the comparison to the parties. This may enable an earlier and quicker detection of erroneous arguments or fake citations, ultimately enhancing the administration of justice.

From Chinese judges' perspective, both systems understand transparency as legibility, making adjudication comprehensible and accountable within respective frameworks. Thus, Chinese courts could require decisions generated by the Similar Case Push platform and Uniform Sentencing Assistance System to be disclosed to parties under certain circumstances, enabling them to understand why their case is distinguishable or why judicial deviation is justified, promoting judicial accountability in AI use.

4. Procedural fairness

From the perspective of frame reflective reciprocal translation, common law judges can recognize that although they do not subordinate individual autonomy to state-determined collective goals, they can acknowledge the Chinese system addresses a genuine concern: inconsistent application of legal standards across cases can itself constitute unfairness. The common law adversarial system addresses this through appellate review and precedent; the Chinese inquisitorial system through algorithmic standardization and institutional supervision. Both reflect the principle that procedures should constrain arbitrary power exercise.

Thus, common law courts can consider incorporating elements of Chinese courts' conception prioritizing systemic consistency and institutional accountability, instead of relying solely on individualized party control and adversarial contestation. Common law courts can learn how to use AI to ensure judges exercise discretion consistently and predictably across all similar cases where appropriate. Hence, common law courts can consider adopting variants of Chinese courts' AI audit trails and the 206 System creating comprehensive records of how decisions were reached, without stifling judicial discretion.

Conversely, while Chinese courts do not adopt a conception prioritizing litigants' control over collective consistency and social stability, PRC courts can recognize the common law system addresses a genuine concern: overly rigid standardization ignoring case-specific circumstances can itself constitute unfairness. Thus, Chinese courts should consider whether fairness can be augmented through party-driven adversarial processes with judicial discretion, instead of relying exclusively on official-driven inquisitorial processes. For example, Chinese courts could allow parties to submit targeted objections to AI-generated recommendations before final decisions, within bounded parameters. After the 206 System or Faxin 2.0 generates recommendations, parties may receive summaries of key algorithmic factors, then submit written objections arguing: factual circumstances not captured by standardized inputs make their case distinguishable; the system's checklist did not anticipate certain relevant evidence; and applying standard recommendations would produce outcomes inconsistent with stated policy goals.

These frame reflective reciprocal translations acknowledge what the ideological and the archival perspectives make clear: technology is unlikely to be neutral, and meaningful

governance requires engaging with the political orders and power structures that AI systems both presuppose and perpetuate.

5. Operationalizing frame reflective reciprocal translations

Finally, in order for meaningful frame reflective reciprocal translations to take place, we propose five concrete governance mechanisms that can be put in place as preparatory, facilitative steps. First, interpretive annexes should accompany any global AI governance instrument, requiring signatory jurisdictions to file "local meaning statements" specifying how they operationalize key terms--"transparency," "fairness," "bias," "reliability"--within their legal systems. This mechanism draws on precedents in international treaty practice, where reservations and interpretive declarations already allow states to specify how they understand shared obligations.[259] The crucial difference is that local meaning statements would be mandatory rather than optional, ensuring that terminological convergence does not mask substantive divergence. For instance, a jurisdiction filing a local meaning statement for "transparency" would be required to specify whether it targets party-facing contestability, internal supervisory monitoring, or both, and what institutional mechanisms give effect to the stated meaning.

Second, glossary protocols should be developed within existing frameworks. The OECD AI Principles[260] currently deploy terms like "transparency" and "accountability" without jurisdiction-specific definitions. These could be supplemented with interpretive guidance requiring signatories to disclose whether "transparency" as implemented targets adversarial contestability by litigants, hierarchical supervisory review, public-facing disclosure, or some combination thereof. The OECD's mandate to monitor implementation of the AI Principles[261], combined with its established peer review methodology[262], provides an institutional vehicle for scrutinizing whether stated meanings correspond to actual practice

Third, implementation reports may be modeled on the Universal Periodic Review[263] mechanism, which could require periodic disclosure of how AI governance principles are operationalized in practice, enabling cross-jurisdictional scrutiny of the gap between

---

[259] Vienna Convention on the Law of Treaties, arts. 2(1)(d), 19–23, May 23, 1969, 1155 U.N.T.S. 331.Article 2(1)(d) defines a "reservation" as "a unilateral statement, however phrased or named, made by a state, when signing, ratifying, accepting, approving or acceding to a treaty, whereby it purports to exclude or to modify the legal effect of certain provisions of the treaty in their application to that state." Articles 19–23 set out the framework governing the formulation, acceptance, and effects of reservations.

[260] OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449 (2019, revised 2024), available at: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

[261] OECD (2023), "The state of implementation of the OECD AI Principles four years on", *OECD Artificial Intelligence Papers*, No. 3, OECD Publishing, Paris, https://doi.org/10.1787/835641c9-en.

[262] OECD, Peer Review: An OECD Tool for Co-operation and Change (2003), OECD Publishing, Paris,

[263] G.A. Res. 60/251, ¶ 5(e) (Mar. 15, 2006) (establishing the Universal Periodic Review mechanism within the Human Rights Council). For a description of the mechanism, see OHCHR, Universal Periodic Review, https://www.ohchr.org/en/hr-bodies/upr/upr-home

terminological commitments and actual implementation. Such reports would move beyond self-assessment to include structured questions derived from the divergences this article identifies such as asking whether "reliability" is assessed at the individual case level or the aggregate system level, and whether "fairness" is measured by individual harm or by social stability impacts.

Fourth, these mechanisms should be embedded within existing institutional initiatives rather than creating new governance structures. The Global AI Governance Action Plan proposed by China[264], the UNESCO Recommendation on the Ethics of AI[265], and ISO/IEC 42001 certification ecosystems[266] each provide institutional platforms within which steps can be taken in order to implement the frame-reflective translation. For example, ISO certification processes could be required to include jurisdiction-specific assessments of how "trustworthy AI" is defined and measured, ensuring that a system certified as "transparent" in one jurisdiction is not assumed to meet transparency requirements in another without explicit comparison of what each system means by the term. The UNESCO framework, with its global reach and explicit attention to judicial contexts, may offer a promising venue for developing the interpretive annexes and glossary protocols proposed above.

Finally, bilateral or plurilateral "judicial AI translation dialogues" between courts or judicial training bodies, building on existing judicial exchange programs, could facilitate direct engagement with how practitioners in different systems understand shared terminology. These dialogues would differ from existing judicial exchanges in a crucial respect: rather than assuming shared professional understanding, they would begin from the premise that identical terms mask divergent meanings, and would structure discussion around surfacing and articulating those divergences. Participants would engage in the frame reflective reciprocal translation exercise our analysis describes: restating each system's understanding of "fairness" or "reliability" in the other system's terms, identifying where translation succeeds and where irreducible differences remain.[267] These proposals do not resolve the underlying frame conflicts our analysis identifies; rather, they make those conflicts visible and manageable within institutional processes, which is what frame-reflective governance demands.

V. Conclusion

We have demonstrated that when common law and Chinese courts speak of "transparency," "bias," "reliability," or "fairness" in AI deployment, they are not merely

[264] Global AI Governance Action Plan, supra note 8.

[265] UNESCO, Recommendation on the Ethics of Artificial Intelligence, adopted 2021, available at: https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence.

[266] ISO/IEC 42001:2023, supra note 231.

emphasizing different aspects of shared concepts; rather they are operating within fundamentally incommensurable semantic universes. The divergences run deeper than doctrinal preference or regulatory design; they reflect competing epistemologies, archival sovereignties, and visions of judicial authority that no amount of technical standardization can reconcile.

Yet this incommensurability need not condemn us to mutual incomprehension or a retreat into jurisdictional silos. Instead, it demands that we reconsider the need for universal AI governance principles and embrace the recognition that our own frameworks for understanding judicial AI are no less contingent, no less shaped by hidden power relations and archival constraints, than those we study. The future of AI in courts lies not in harmonization but in sophisticated frame reflective reciprocal translation practices that make visible the unspoken assumptions each system brings to the table.

Indeed, one avenue for developing such translation practices may lie in AI itself. Computational methods for cross-jurisdictional semantic analysis, including natural language processing techniques for mapping how shared terms are operationalized differently across legal corpora, and LLMs trained on multilingual legal texts to identify divergent usage patterns, could make the translation process more systematic and less dependent on individual interpretive effort.[268] AI-assisted semantic mapping could systematically identify when “transparency” in Chinese judicial documents correlates with supervisory metrics while the same term in common law judgments correlates with party-facing disclosure requirements, rendering visible the divergences our analysis has traced through traditional comparative methods. There is a productive irony here: the very technology whose governance challenges this article examines may also offer tools for diagnosing and structuring the cross-jurisdictional semantic divergences that make governance so difficult. We flag this as an important direction for future research, while noting that such computational approaches would themselves require careful attention to the framing effects our analysis identifies--the risk that AI-assisted translation reproduces rather than reveals the assumptions embedded in its training data is precisely the kind of challenge our article illuminates.[269]

Finally, we want to ask whether either system can survive AI's inexorable advance, while preserving its conception of what adjudication is. To answer this question, courts have to be critical interrogators of the epistemologies AI embodies. Judges in both common law and Chinese systems should develop not just technical competence in evaluating AI outputs, but a deep understanding of how AI systems encode particular visions of legal reasoning, authority, and justice.

[268] On computational approaches to cross-jurisdictional legal analysis, see Michael A. Livermore & Daniel N. Rockmore eds., LAW AS DATA: COMPUTATION, TEXT, AND THE FUTURE OF LEGAL ANALYSIS (2019).

[269] Jiang et al., supra note 9.