 DOI: https://10.5281/zenodo.22903556

Artificial Intelligence and Legal Reasoning: Reconsidering Human Judgment and Accountability in the Digital Age

John W. Baker[1]

Abstract

The rapid development of artificial intelligence (AI), particularly generative AI and large language models, is transforming legal research, analysis, drafting, and decision-making. AI systems increasingly perform tasks traditionally associated with lawyers and other legal professionals, raising fundamental questions concerning human judgment and accountability within legal institutions. This article examines whether AI-assisted legal reasoning can be reconciled with the rule of law, procedural fairness, and human responsibility. The article argues that legal reasoning extends beyond the retrieval or probabilistic generation of legal information. It involves interpreting authoritative sources, evaluating facts, resolving competing principles, exercising judgment, and providing reasons capable of scrutiny and challenge. AI therefore presents significant risks, including hallucinated authorities, algorithmic bias, opacity, automation bias, and excessive reliance on machine-generated outputs. Through a comparative examination of emerging regulatory and professional approaches in the European Union, United Kingdom, United States, and India, the article assesses existing mechanisms for accountability in AI-assisted legal work. It proposes a framework centred on meaningful human oversight, independent verification, transparency, traceability, reason-giving, and contestability. The article concludes that AI should function as an assistive technological instrument that augments legal reasoning without displacing human judgment and responsibility necessary for legitimate and accountable legal decision-making.

Keywords: Artificial Intelligence, Generative AI, Law and Technology, Algorithmic Bias, Procedural Fairness

1. Introduction

Artificial intelligence (AI) has moved from the margins of legal technology into the ordinary infrastructure of legal work. Earlier generations of legal technology were predominantly concerned with information retrieval, document management, electronic discovery and predictive analytics. Contemporary generative AI systems, by contrast, can produce natural-language text, summarise legal materials, identify potentially relevant authorities, draft documents and formulate arguments in response to complex prompts. This development has intensified a longstanding debate about the extent to which technological systems can perform functions traditionally associated with legal

[1] **University of Chicago, Chicago, IL 60637, United States**

professionals. As Surden observes, AI has already affected the practice and administration of law across multiple domains, although its actual capabilities should be distinguished from speculative claims about technological autonomy (Surden, 2019).

The significance of this transformation lies not merely in the efficiency gains associated with automated legal research or drafting. It concerns the nature of legal reasoning itself. Law is not simply a body of information capable of being retrieved and reproduced. Legal reasoning ordinarily requires the identification and interpretation of authoritative sources, the characterisation of facts, the application of rules and principles, the reconciliation of competing authorities, the exercise of judgment in areas of legal uncertainty, and the articulation of reasons capable of public and institutional scrutiny. The increasing ability of AI systems to generate convincing legal language therefore raises a distinctive question: when an artificial system participates in the production of legal analysis, where should the boundary between technological assistance and legally significant human judgment be drawn?

This question has become more consequential with the emergence of large language models (LLMs). Unlike conventional legal databases, which principally assist lawyers in locating existing materials, generative AI can produce apparently reasoned answers to legal questions without necessarily revealing the evidentiary or inferential basis from which those answers have been generated. The distinction is legally important. A system may generate grammatically coherent and professionally persuasive legal prose while nevertheless misidentifying the governing law, misunderstanding a factual distinction, fabricating an authority, or presenting an uncertain proposition as settled law. Recent scholarship has consequently identified difficulties for generative AI in performing specifically legal reasoning tasks, including selecting the applicable legal framework, distinguishing *ratio decidendi* from *obiter dicta*, resolving conflicting legal provisions and applying open-textured standards such as reasonableness (see, e.g., recent analysis of generative AI and legal reasoning).

The problem is not confined to technological accuracy. An erroneous legal proposition produced by a search engine may be corrected by examining the underlying authority; an erroneous proposition generated by a sophisticated LLM may be more difficult to detect precisely because it is expressed with the appearance of legal competence. This creates what may be described as an epistemic asymmetry: the system can produce an output that appears sufficiently authoritative to influence a human decision-maker even where the system itself has no legal authority and no capacity to bear legal responsibility for the consequences of its output. The risk is therefore not simply that AI may occasionally be "wrong", but that the institutional mechanisms through which law distinguishes justified legal reasoning from unsupported assertion may become obscured by technological mediation.

These concerns are particularly significant within professional legal practice. Lawyers do not discharge their obligations merely by producing text that appears legally plausible. Their professional responsibilities include competence, confidentiality, communication with clients, supervision and candour towards tribunals. The American Bar Association's Formal Opinion 512, issued in 2024, expressly applies these existing professional obligations to lawyers using generative AI and emphasises the necessity of appropriate review and verification of AI-generated material

(American Bar Association, 2024). The significance of this approach is that the introduction of AI does not, in itself, transfer professional responsibility from the lawyer to the technological system. The lawyer remains the legally accountable actor.

The implications become still more profound when AI is introduced into adjudication. Judicial reasoning occupies a distinctive position within the legal order because judicial decisions do not merely provide predictions about legal outcomes; they exercise public authority and may determine rights, obligations and liabilities. A system used to assist a judge in identifying relevant law or analysing factual material may therefore affect the exercise of public power even if the formal decision remains that of the judge. The European Union's Artificial Intelligence Act recognises the heightened significance of this context. Regulation (EU) 2024/1689 classifies certain AI systems intended to assist judicial authorities in researching and interpreting facts and law, and in applying law to particular facts, as high-risk systems because of their potential implications for the rule of law, individual freedoms, effective remedies and fair trials. Crucially, the Regulation states that AI may support judicial decision-making but should not replace judicial authority, and that final decision-making must remain human-driven (European Union, 2024).

The same principle appears in international normative frameworks. UNESCO's Recommendation on the Ethics of Artificial Intelligence places human oversight, responsibility, accountability, transparency and explainability among its central principles. It provides that AI systems should not displace ultimate human responsibility and accountability and emphasises the importance of auditability and traceability where AI affects rights and freedoms (UNESCO, 2021). These principles are particularly relevant to legal reasoning because accountability requires more than the presence of a human somewhere within the decision-making process. It requires the identification of a human or institutional actor capable of understanding, scrutinising, explaining and, where necessary, correcting the basis of a legally consequential decision.

The resulting tension may be expressed as one between automation and accountability. Technological systems can potentially increase the speed, scale and consistency of legal analysis. They may assist lawyers in processing large bodies of material and may facilitate access to legal information that would otherwise require substantial time and resources. Earlier scholarship on technology and the legal profession anticipated precisely such transformations in the delivery of legal services (Susskind, 2008). Yet efficiency cannot, by itself, establish legal legitimacy. A legal system must also be capable of explaining why an outcome was reached, identifying who is responsible for it, allowing affected persons to challenge it, and ensuring that decisions remain within the bounds of applicable law.

This distinction is especially important because legal reasoning is simultaneously descriptive, interpretive and normative. It involves determining what the law provides, but also deciding how authoritative legal materials apply to particular circumstances and how competing principles should be reconciled. In difficult cases, the answer cannot always be derived through mechanical application of a predetermined rule. Even where AI can identify relevant authorities or generate plausible competing arguments, the legal institution must ultimately determine which interpretation is justified within the applicable constitutional, statutory and precedential framework.

The capacity to generate alternatives should therefore not be confused with the authority to select among them.

The central issue is consequently not whether AI can "think like a lawyer" in a philosophical or anthropomorphic sense. Such a formulation risks obscuring the legal question. The more important inquiry is whether AI-generated or AI-assisted outputs can be incorporated into legal processes while preserving the institutional conditions under which legal reasoning is recognised as authoritative. These conditions include the identification of applicable sources, methodological transparency, reason-giving, procedural fairness, independent judgment and the possibility of review. The question of AI and legal reasoning is thus ultimately a question about institutional design and responsibility, rather than merely technological capability.

This article examines that problem through the relationship between artificial intelligence, legal reasoning, human judgment and accountability. It proceeds from the proposition that the appropriate legal response should neither treat AI as inherently incompatible with legal decision-making nor assume that human involvement automatically resolves the risks created by AI. Instead, the decisive issue is the nature and quality of human involvement. A nominal "human-in-the-loop" may provide little protection if the human decision-maker merely accepts an AI-generated recommendation without meaningful scrutiny. Conversely, AI may be capable of valuable assistance where its outputs remain subject to independent verification, reasoned evaluation and accountable human judgment.

The article therefore advances a distinction between AI assistance and AI substitution. AI can assist legal actors by locating information, organising evidence, identifying patterns and generating preliminary analytical material. The legal difficulty begins when technological output becomes determinative without an adequate process through which a responsible human actor independently evaluates its accuracy, relevance and legal justification. The distinction is particularly important in adjudication, where the legitimacy of a decision depends not merely upon its outcome but upon the existence of legally intelligible reasons capable of being examined by the parties and, where appropriate, by an appellate or reviewing institution.

Against this background, the article asks three principal questions. First, what features distinguish legal reasoning from the computational processing of legal information? Secondly, what risks arise when generative AI participates in legal analysis and judicial decision-making, particularly in relation to accuracy, bias, opacity and accountability? Thirdly, what legal and institutional framework is necessary to ensure that AI augments rather than displaces accountable human judgment? These questions require consideration not only of technological limitations but also of professional ethics, procedural safeguards, judicial independence and regulatory design.

The article adopts a comparative legal perspective, examining emerging approaches in the European Union, United Kingdom, United States and India. The comparison is intended not merely to catalogue regulatory instruments but to identify the underlying principles through which different legal systems are attempting to reconcile technological innovation with existing legal values. The EU AI Act provides an especially significant example of risk-based regulation in the administration of justice, while professional guidance such as ABA Formal Opinion 512 illustrates how established duties may be applied to emerging generative AI technologies. International

principles articulated by UNESCO further demonstrate the growing recognition that transparency, accountability and meaningful human oversight are integral to responsible AI governance.

The article ultimately argues that human judgment should remain the locus of legal responsibility wherever AI materially contributes to a legally consequential determination. This does not require excluding AI from legal reasoning. Rather, it requires establishing institutional conditions under which AI-generated outputs remain contestable, verifiable and subordinate to accountable legal judgment. The appropriate objective is therefore not technological resistance, but accountable augmentation: a model in which AI expands the analytical capacity of legal institutions without becoming an unaccountable substitute for the human exercise of legal authority.

The argument has broader implications for the rule of law. If legal institutions increasingly depend upon systems whose outputs cannot readily be reconstructed, explained or challenged, technological efficiency could come at the expense of legal accountability. Conversely, if AI is deployed within structures requiring meaningful oversight, independent verification, transparency and reason-giving, it may enhance rather than undermine the capacity of legal institutions to perform their functions. The future of AI in law should therefore be assessed not by asking whether machines can reproduce the language of legal reasoning, but by asking whether their integration preserves the human responsibility, institutional justification and contestability upon which legitimate legal reasoning depends.

2. Artificial Intelligence and the Transformation of Legal Reasoning

The relationship between artificial intelligence (AI) and law is no longer confined to the theoretical question of whether machines might eventually perform tasks associated with lawyers and judges. AI is already being incorporated into legal research, document review, contract analysis, due diligence, litigation analytics, drafting and other forms of legal work. The contemporary transformation is particularly significant because generative AI and large language models (LLMs) do not merely retrieve or classify existing legal information; they can generate new text in response to legal questions and prompts. This changes the technological character of legal assistance from information retrieval towards the production of apparently reasoned legal analysis. Surden accordingly distinguishes the practical uses of AI in law from more speculative accounts of future machine intelligence and emphasises the importance of assessing AI according to its actual capabilities and limitations (Surden, 2019).

The transformation, however, should not be understood as a simple transition from human to machine reasoning. A more accurate description is the emergence of a new socio-technical architecture of legal reasoning, in which human legal actors increasingly interact with computational systems that search, classify, predict, summarise and generate legal material. The legal significance of this development depends upon what function AI performs within that architecture. There is an important difference between a system that helps a lawyer locate potentially relevant authorities and a system whose generated assessment materially influences the lawyer's conclusion. There is a further difference between a judicial research tool and an automated system that recommends or

determines the outcome of a dispute. The technological system may be identical or similar, but the legal consequences of its use are substantially different.

2.1 From Legal Information Retrieval to Generative Legal Analysis

Early applications of AI in law were principally concerned with structured forms of legal information processing. Legal expert systems attempted to represent legal rules and reasoning through computational structures, while later applications employed machine learning and statistical techniques to analyse large bodies of legal data. Contemporary legal analytics can identify patterns in judgments, classify documents, extract information from contracts and assist with predictions concerning litigation outcomes. Ashley describes the development of legal AI as involving computational approaches capable of connecting legal texts with models of reasoning, generating arguments, predicting outcomes and supporting human evaluation of those predictions (Ashley, 2017).

Generative AI represents a significant development because its principal output is not merely a classification or retrieval result but newly generated language. A user can ask an LLM to identify potentially applicable legal principles, compare competing authorities, draft submissions or explain a legal doctrine, and the system can produce an apparently coherent response. This capability has considerable practical value. The American Bar Association's Formal Opinion 512 recognises that generative AI may assist lawyers with legal research, contract review, due diligence, document review, regulatory compliance and the drafting of letters, contracts, briefs and other legal documents (American Bar Association, 2024).

The distinction between retrieval and generation is nevertheless crucial. A conventional database is generally designed to direct the user towards an existing source. A generative model instead predicts and constructs an output based upon patterns learned from its training data and the context supplied in the interaction. The resulting text can therefore possess the linguistic characteristics of legal reasoning without necessarily reproducing the methodological process by which a lawyer establishes the legal validity of a proposition. This distinction explains why linguistic fluency cannot be treated as evidence of legal correctness.

The problem is particularly acute because legal reasoning is authority-dependent. A proposition concerning the content of law ordinarily requires support from an identifiable legal source—legislation, judicial precedent, constitutional provision, regulation or another recognised source of law. The fact that an AI system produces a plausible proposition does not itself confer legal authority upon that proposition. AI therefore introduces a potentially important separation between the production of legal language and the justification of legal conclusions.

2.2 Legal Reasoning Is More Than Information Processing

The transformation of legal reasoning by AI becomes clearer when the nature of legal reasoning itself is examined. Legal reasoning involves multiple interconnected activities: identifying relevant legal materials; interpreting authoritative texts; determining the significance of precedent; characterising legally relevant facts; applying rules and principles to those facts; resolving conflicts

between authorities or principles; and articulating reasons capable of supporting the resulting conclusion. These activities do not always operate according to a single deterministic procedure.

This is particularly apparent in difficult cases involving open-textured concepts such as reasonableness, proportionality, fairness, good faith, negligence or legitimate expectation. Such concepts cannot necessarily be resolved through the mechanical identification of a matching phrase in a legal database. Their application depends upon legal context, precedent, factual characterisation and normative judgment. AI may assist with each component, but the availability of computational assistance does not eliminate the underlying interpretive problem.

Ashley's conception of computational legal reasoning is therefore particularly important. His work does not treat computational systems simply as replacements for lawyers; rather, it explores ways in which computational models can connect legal texts, arguments and predictions while enabling legal professionals to evaluate their outputs (Ashley, 2017). This conception provides a useful foundation for understanding AI-assisted reasoning as a form of human–machine collaboration, rather than as autonomous machine adjudication.

The distinction is also consistent with Surden's analysis of AI's practical capabilities. Contemporary AI systems can perform extremely well in particular, bounded tasks, but their capabilities should not be conflated with a general capacity to reason across all domains in the manner of a human legal professional (Surden, 2019). Legal reasoning frequently requires precisely the contextual flexibility that makes such generalisation difficult: a lawyer must determine not only what information is relevant but why it is relevant, what authority governs, which factual distinctions matter and how competing legal considerations should be reconciled.

Consequently, AI can substantially alter the process through which legal reasoning is performed without necessarily replacing the institutional function of legal reasoning. The lawyer remains responsible for determining whether the generated analysis is legally sustainable; the judge remains responsible for deciding the dispute; and the legal institution remains responsible for ensuring that the decision can be justified according to law.

2.3 AI-Assisted Legal Research and the Changing Role of the Lawyer

Legal research provides one of the clearest examples of this transformation. AI systems can assist lawyers in searching large bodies of case law and legislation, summarising authorities, identifying relationships between cases and generating preliminary research memoranda. The potential efficiency gains are considerable because legal professionals have historically devoted substantial time to locating, reading and organising legal material.

The ABA has expressly recognised this potential. Formal Opinion 512 notes that AI may improve efficiency and assist with legal research, while simultaneously stressing that lawyers must understand the technology's capabilities and limitations and must satisfy their existing professional obligations when relying upon it (American Bar Association, 2024).

The transformation consequently affects the lawyer's role. The lawyer may increasingly spend less time on the mechanical identification of potentially relevant material and more time evaluating the quality, relevance and legal significance of material produced or retrieved by AI. This

does not make legal expertise less important. It changes the point at which that expertise is exercised.

The distinction can be expressed as follows:

Traditional model:

Legal question → research → authorities → analysis → conclusion.

AI-assisted model:

Legal question → AI-assisted retrieval/generation → verification → legal evaluation → analysis → conclusion.

The insertion of AI between the legal question and the legal conclusion creates a new verification stage. The lawyer must assess not merely whether an output appears plausible but whether the authorities cited exist, whether they support the proposition attributed to them, whether the jurisdiction and temporal context are correct, and whether material contrary authorities have been omitted.

Empirical research demonstrates why this verification function is indispensable. A Stanford RegLab and Stanford HAI study evaluating leading AI-powered legal research tools found that, although specialised legal research systems performed better than general-purpose models, the systems tested still produced incorrect or inadequately grounded responses at material rates. The study found hallucination or error rates exceeding 17 per cent for Lexis+ AI and Ask Practical Law AI and exceeding 34 per cent for Westlaw's AI-Assisted Research in the researchers' benchmark dataset (Magesh et al., 2024).

Importantly, the study distinguishes between an answer that is substantively incorrect and one that reaches a correct proposition but cites a source that does not actually support it. The latter problem is particularly significant for legal reasoning because legal conclusions derive legitimacy not simply from their plausibility but from their relationship with authoritative legal sources (Magesh et al., 2024).

2.4 The Problem of Hallucination and the Authority of Legal Sources

The phenomenon commonly described as AI "hallucination" presents a distinctive challenge to legal reasoning. In ordinary conversational contexts, an incorrect factual statement may be inconvenient. In legal practice, however, a fabricated case, statutory provision or quotation can affect the validity of an argument, mislead a court or expose a lawyer to professional consequences.

The problem has already moved beyond hypothetical concern. The well-known *Mata v Avianca, Inc* litigation demonstrated the consequences of relying upon fictitious authorities generated by ChatGPT. The episode is significant not because it establishes that all generative AI is unreliable, but because it demonstrates the institutional consequences of treating generated legal language as though it were verified legal authority.

Empirical research reinforces the point. Stanford's evaluation of AI legal research systems found that even specialised systems could generate both incorrect legal information and "misgrounded" answers in which the cited source did not actually support the proposition for which it was offered (Magesh et al., 2024). The latter problem is especially important because conventional

methods of legal verification depend heavily upon the relationship between proposition and authority.

This produces an important conceptual distinction between syntactic legal plausibility and juridical validity. An AI-generated argument may be syntactically persuasive, doctrinally familiar and professionally expressed while remaining legally defective. The appearance of reasoning is therefore not equivalent to reasoned legal justification.

For this reason, AI should not be treated as an autonomous source of law. Its output may be evidence of what the system predicts or generates in response to a prompt, but it cannot itself establish what the law is. Legal authority remains external to the model and must be established through recognised legal sources.

2.5 Predictive Analytics and the Shift from Explanation to Prediction

Another important dimension of AI's transformation of legal reasoning is the growth of predictive analytics. Machine-learning systems can identify statistical patterns in judicial decisions and other legal datasets. Such systems may be used to estimate litigation outcomes, analyse judicial behaviour or identify patterns in case law (Surden, 2019).

Prediction, however, is conceptually different from legal justification. A system may accurately predict that a particular court is likely to reach a particular outcome without explaining why that outcome is legally justified. The distinction is fundamental to a legal system in which reasons matter independently of prediction.

A predictive model asks, in effect:

What outcome is likely?

Legal reasoning asks a different question:

What outcome is legally justified, and why?

The two questions may produce the same practical answer, but they are not interchangeable. A model that predicts judicial behaviour based upon historical data may reproduce patterns within previous decisions without necessarily identifying the legal principle that makes a new decision legitimate. This distinction becomes especially important when historical data contain inconsistencies, institutional biases or legally irrelevant correlations.

Consequently, predictive accuracy should not be treated as a substitute for legal reasoning. Prediction can be an important analytical input, but a legal conclusion requires a normative and institutional justification that statistical correlation alone cannot provide.

2.6 Generative AI and the Production of Legal Arguments

Generative AI also changes the nature of legal advocacy by lowering the cost of producing legal arguments. A lawyer can ask an LLM to formulate arguments for multiple sides of a dispute, identify counterarguments, restructure a submission or produce alternative interpretations of a statutory provision. This can enhance deliberation because legal professionals can examine a broader range of possible arguments before selecting and refining those that are legally sustainable.

The potential benefit is therefore not merely automation but argumentative expansion. AI can serve as a tool for generating hypotheses that a lawyer can test against primary authorities. Used in this manner, AI may function as a form of intellectual augmentation.

Yet the same capability creates a risk of argument proliferation without legal validation. The capacity to generate ten arguments does not mean that ten legally viable arguments exist. A system optimised to produce responsive language may generate arguments that are rhetorically plausible but doctrinally weak, distinguishable on the facts or unsupported by authority.

This again reinforces the distinction between generation and judgment. AI can increase the range of candidate arguments available to the lawyer, but the lawyer must determine which arguments satisfy the requirements of law, evidence and professional responsibility. The ABA's Formal Opinion 512 specifically warns that lawyers cannot rely uncritically on generated content and must satisfy their duties concerning competent representation and candour towards tribunals (American Bar Association, 2024).

2.7 From Human-in-the-Loop to Meaningful Human Judgment

The increasingly common expression "human-in-the-loop" is useful but potentially inadequate as a legal safeguard. The mere presence of a human somewhere in the decision-making process does not establish meaningful human control. If a lawyer or judge accepts an AI-generated conclusion without understanding its evidentiary basis, checking its authorities or considering plausible alternatives, human involvement may be formally present but substantively weak.

The relevant standard should therefore be meaningful human judgment, rather than merely human participation. Meaningful judgment requires the capacity and institutional authority to question, reject, modify and explain the AI output.

This distinction is already reflected in regulatory thinking concerning judicial AI. The EU AI Act classifies AI systems used by or on behalf of judicial authorities to assist in researching and interpreting facts and law and applying law to particular facts as high-risk systems. The legislation identifies the potential risks of bias, error and opacity and expressly states that AI may support judicial decision-making but should not replace it; final decision-making must remain human-driven (European Union, 2024, recital 61; Annex III, para 8(a)).

The significance of this approach extends beyond the European Union. It reflects a broader principle: where AI participates in the exercise of legal authority, the relevant question is not simply whether a human remains formally responsible but whether that human retains genuine capacity to exercise independent legal judgment.

2.8 The Transformation of Legal Reasoning as Institutional Transformation

The most significant consequence of AI may therefore lie not in the replacement of particular legal tasks but in the redistribution of cognitive labour within legal institutions. Tasks that once required substantial human time—searching, sorting, summarising and drafting—may increasingly be delegated to machines. Human effort may consequently shift towards verification, contextual interpretation, strategic judgment and responsibility.

This redistribution can be beneficial. It may allow lawyers and judges to devote greater attention to complex questions that require contextual and normative analysis. AI may also reduce the cost of certain forms of legal work and potentially broaden access to legal information and assistance. The ABA has recognised the possibility that AI-enabled efficiency may reduce costs and contribute to access to justice, while emphasising that professional obligations continue to apply (American Bar Association, 2024).

But redistribution also creates new dependencies. If lawyers become dependent upon AI systems for identifying relevant authorities or constructing initial analyses, the technological system can influence what the human decision-maker sees before the human begins independent evaluation. The problem is therefore not only whether an AI output is correct. It is also whether the system determines the informational environment within which human legal judgment occurs.

This is a subtler form of technological influence. A lawyer may remain formally free to reject an AI recommendation, yet the recommendation may frame the issue, identify the authorities considered relevant, suggest the available arguments and establish the apparent range of possible outcomes. AI can therefore influence legal reasoning without making the final decision itself.

2.9 From Automation to Augmentation

The appropriate conceptual framework is consequently one of augmentation rather than substitution. AI should be understood as a technological instrument capable of increasing the speed, breadth and analytical capacity of legal work while leaving legally consequential judgment within an accountable institutional structure.

Such augmentation requires at least four conditions.

First, source verification must remain central. AI-generated propositions should be tested against authoritative legal materials.

Secondly, contextual evaluation must remain human-led. The legal relevance of facts, distinctions between precedents and implications of competing principles cannot be assumed from generated text alone.

Thirdly, reason-giving must remain independent of the mere existence of an AI recommendation. A lawyer or judge must be able to explain why a conclusion follows from law and facts rather than merely state that an AI system produced it.

Fourthly, responsibility must remain attributable. The use of AI cannot create a gap in accountability in which the human actor attributes an erroneous conclusion to the machine while the developer, institution and user each deny responsibility.

The ABA's approach illustrates this principle in professional practice: lawyers remain subject to duties of competence, confidentiality, communication, supervision, meritorious advocacy and candour notwithstanding their use of generative AI (American Bar Association, 2024).

2.10 Reconstructing the Meaning of Legal Reasoning in the AI Age

AI therefore challenges legal scholarship to reconsider what is meant by "legal reasoning". If legal reasoning is defined merely as the ability to produce legally plausible propositions, advanced

language models may appear increasingly capable of performing it. If, however, legal reasoning is understood as an institutional practice involving authority, interpretation, justification, responsibility and contestability, AI's role becomes more limited and more complex.

This distinction is central to the argument of this article. The transformation of legal reasoning by AI should not be measured simply by asking how much legal work a machine can perform. The more important question is which elements of legal reasoning can be technologically assisted without compromising the conditions that make the resulting exercise recognisably legal.

The answer is unlikely to be an absolute division between tasks that are "human" and tasks that are "machine". Many aspects of legal work are already technologically mediated, and the historical development of legal information systems demonstrates that technology can change legal practice without necessarily undermining legal institutions. The critical distinction instead concerns the degree to which AI output influences legally consequential judgment and the safeguards surrounding that influence.

The transformation can therefore be understood across three stages:

First, AI as retrieval: the system helps identify potentially relevant legal information.

Secondly, AI as augmentation: the system analyses, summarises or generates possible arguments that a human legal actor independently evaluates.

Thirdly, AI as decision influence: the system's output materially shapes or recommends a legally consequential conclusion.

The legal risks increase as one moves from retrieval towards decision influence because the system moves closer to the exercise of legal authority. This is precisely why regulatory frameworks increasingly distinguish between low-impact administrative applications and AI used in the administration of justice. The EU AI Act, for example, excludes purely ancillary judicial administrative activities from the high-risk category while specifically identifying AI used to assist judicial authorities in researching and interpreting facts and law as high-risk (European Union, 2024, recital 61; Annex III).

The transformation of legal reasoning by AI should consequently be understood neither as the arrival of autonomous machine jurisprudence nor as a conventional efficiency improvement. It represents a restructuring of how legal information is encountered, how arguments are generated, how alternatives are evaluated and how professional judgment is exercised. The central legal challenge is to ensure that this restructuring does not silently transform assistance into delegation.

The future of AI-assisted legal reasoning will therefore depend less upon whether machines become capable of producing increasingly sophisticated legal language and more upon whether legal institutions develop mechanisms capable of distinguishing generated output from justified legal judgment. AI may become an extraordinarily powerful instrument for research, analysis and argumentation. But the transformation remains compatible with the rule of law only where the authority to determine what the law requires, and the responsibility for explaining and defending that determination, remain located within accountable legal institutions.

3. Human Judgment, Reliability and Accountability in AI-Assisted Law

The integration of artificial intelligence into legal practice does not eliminate the need for human judgment; rather, it changes the conditions under which that judgment is exercised. The central difficulty is that AI systems can produce outputs that are sufficiently fluent, structured and legally plausible to invite reliance even when the underlying analysis is incomplete or incorrect. This is particularly consequential in law because legal decisions and professional advice are not evaluated solely by their linguistic quality or predictive accuracy. They must be grounded in authoritative sources, responsive to the relevant facts, consistent with applicable legal principles and capable of being justified by an identifiable legal actor. The growing use of generative AI therefore requires a reconsideration of the relationship between technological assistance and human responsibility. The appropriate question is not whether a human remains somewhere within an AI-assisted process, but whether that human retains the knowledge, independence and institutional authority necessary to exercise meaningful judgment over the system's output.

Reliability presents the first difficulty. Large language models generate outputs probabilistically rather than by consulting a fixed body of law in the manner of a conventional legal database. Consequently, a response may be linguistically convincing without being legally accurate. This problem is particularly acute in legal contexts because an apparently minor error can materially alter the conclusion reached. A fabricated case, an incorrect statutory section, an outdated authority or an omitted qualification can transform an otherwise plausible argument into a legally defective one. Empirical evidence demonstrates that access to legal databases does not eliminate this problem. Magesh et al's preregistered evaluation of leading AI-powered legal research tools found that the specialised systems tested by the researchers hallucinated more than 17 per cent of the time, despite being designed specifically for legal research and despite using retrieval-augmented generation techniques intended to improve factual grounding (Magesh et al., 2025).

The significance of these findings extends beyond the familiar phenomenon of fabricated citations. The deeper concern is that legal AI may produce misgrounded reasoning: an answer may reach a legally plausible conclusion while the authority cited does not actually support the proposition for which it is offered. This distinction is critical because legal reasoning derives its legitimacy partly from the relationship between proposition and authority. A lawyer cannot establish that a proposition is law merely by presenting a persuasive explanation of it; the proposition must be supported by a recognised source of legal authority. Magesh et al therefore distinguish between different forms of error in AI-generated legal research, demonstrating that the problem cannot be reduced to obviously false statements. The reliability question must encompass accuracy, responsiveness, completeness and the relationship between an answer and its supporting authorities (Magesh et al., 2025).

Reliability should consequently be understood as context-dependent rather than absolute. An AI system might be sufficiently reliable for low-risk administrative tasks while being unsuitable for an unverified legal conclusion affecting a person's rights or substantial financial interests. The relevant inquiry is therefore not whether an AI system is “reliable” in the abstract, but whether its reliability is adequate for the particular legal function it performs. The greater the legal consequences of an output, the stronger the justification for verification, documentation and independent human review. This risk-sensitive approach is reflected in the EU Artificial Intelligence

Act, which classifies certain AI systems intended to assist judicial authorities in researching and interpreting facts and law and applying law to concrete facts as high-risk systems because of their potential effects on the rule of law, fundamental rights and fair trial guarantees (European Union, 2024, recital 61; Annex III, para 8(a)).

The reliability problem also interacts with a specifically human cognitive phenomenon: automation bias. Automation bias describes the tendency to give excessive weight to information or recommendations generated by automated systems, particularly where the system appears sophisticated or authoritative. The EU AI Act expressly requires human oversight for high-risk AI systems to address the risk that individuals may automatically or excessively rely upon system outputs. Article 14 requires persons responsible for oversight to understand the system's capabilities and limitations, remain aware of automation bias, correctly interpret outputs, and retain the ability to disregard, override or reverse an AI-generated output (European Union, 2024, art 14). This is an important regulatory insight because it recognises that the existence of a human reviewer is insufficient if the institutional environment encourages that reviewer simply to accept the machine's recommendation.

Accordingly, the concept of human-in-the-loop requires greater precision. If a lawyer receives an AI-generated memorandum and signs it without checking its authorities, the presence of a human does not necessarily amount to meaningful human oversight. Similarly, if a judge receives an AI-generated assessment and accepts its conclusion without independently evaluating the relevant law and facts, the formal retention of judicial authority does not resolve the underlying accountability problem. Meaningful human judgment requires more than the opportunity to disagree; it requires a realistic capacity to understand, interrogate and reject the system's output. The EU regulatory framework captures this distinction by requiring oversight mechanisms that enable the responsible person to understand relevant capabilities and limitations and to decide not to use, disregard, override or reverse the system's output (European Union, 2024, art 14).

Human judgment is particularly important because legal reasoning contains dimensions that cannot be reduced to the statistical likelihood of an answer. Legal professionals routinely distinguish between materially similar and materially different facts, determine the precedential weight of authorities, interpret ambiguous statutory language and reconcile competing principles. Such exercises involve contextual and institutional judgment. An AI system may identify cases containing similar language or generate competing interpretations, but similarity of language does not necessarily establish legal relevance. The significance of a precedent depends upon its jurisdiction, procedural posture, level of court, factual matrix, subsequent treatment and relationship with other authorities. The human legal actor must therefore determine not merely what the AI has found, but whether what it has found is legally significant.

This distinction becomes especially important where the law contains standards rather than mechanically applicable rules. Concepts such as reasonableness, proportionality, fairness, good faith and legitimate expectation require contextual application. An AI model may identify how courts have previously used these concepts, but historical patterns do not automatically determine the legally correct application in a new case. A system trained upon previous decisions may reproduce existing patterns, including inconsistencies within those decisions, but reproduction of historical

patterns is not synonymous with legal justification. This is one reason why predictive accuracy should not be treated as equivalent to legal validity. A system could potentially predict a court's decision while offering no adequate account of why that decision is legally justified.

The distinction between prediction and justification is central to accountability. A prediction asks what is likely to happen; legal reasoning asks what the law requires and why. The two may converge, but they remain conceptually distinct. If a legal institution were to rely upon a statistical prediction without an independently reasoned legal justification, the resulting decision could be difficult to reconcile with principles requiring reasoned adjudication. AI can therefore provide useful predictive information without itself supplying the normative justification necessary for the exercise of legal authority. The more consequential the decision, the stronger the case for maintaining this distinction.

Human judgment is also indispensable because legal accountability requires an identifiable decision-maker. A machine cannot presently occupy the institutional position of a lawyer, judge or public authority in the conventional sense. It cannot hold professional office, owe fiduciary duties in the manner of a lawyer, exercise judicial power or accept disciplinary responsibility for professional misconduct. Where AI contributes to a legal decision, responsibility must therefore remain attributable to the human or institution legally empowered to make that decision. UNESCO's Recommendation on the Ethics of Artificial Intelligence expressly adopts this principle, providing that AI systems should not displace ultimate human responsibility and accountability and calling for systems to be auditable and traceable (UNESCO, 2021).

This principle has significant consequences for professional legal practice. The use of AI does not create an exemption from existing duties of professional responsibility. The American Bar Association's Formal Opinion 512 makes this explicit. It states that lawyers using generative AI must consider their existing obligations concerning competence, confidentiality, communication, supervision, meritorious claims and contentions, candour toward tribunals and reasonable fees (American Bar Association, 2024). The significance of the Opinion lies in its regulatory methodology: rather than creating a wholly separate ethical regime for AI, it applies established professional duties to a new technological environment.

The duty of competence is particularly important. A lawyer cannot reasonably rely upon a technology that the lawyer does not understand sufficiently to evaluate. Formal Opinion 512 explains that competent use of generative AI requires an understanding of the technology's capabilities and limitations and appropriate review of its outputs (American Bar Association, 2024). Competence in the AI era therefore includes a technological dimension. The lawyer need not become a computer scientist, but must possess sufficient understanding to identify foreseeable risks, select an appropriate tool, configure its use responsibly and evaluate whether the resulting output can safely be relied upon.

The duty of competence also demonstrates why human oversight cannot be treated as a purely procedural formality. If a lawyer lacks the knowledge necessary to recognise that a generated case citation is fabricated, then requiring the lawyer to "review" the output does little to mitigate the underlying risk. Verification must be substantive rather than ceremonial. A meaningful review may require checking every cited authority, consulting the original judgment, confirming the current

status of legislation, considering contrary authorities and determining whether the AI has omitted legally significant facts or qualifications. The level of review should correspond to the importance and risk of the legal task.

The same principle applies to candour towards the tribunal. A lawyer remains responsible for the accuracy of material submitted to a court even when the material was initially generated by an AI system. Formal Opinion 512 specifically identifies the duty of candour among the professional obligations implicated by generative AI (American Bar Association, 2024). The use of AI therefore cannot become a mechanism through which responsibility for a false or unsupported submission is transferred from lawyer to machine. If the lawyer places the argument before the court, the lawyer remains responsible for its legal and factual integrity.

Confidentiality introduces another dimension of accountability. AI systems may process client information through external infrastructure, creating questions about data retention, access, security and the contractual terms governing the provider. Formal Opinion 512 accordingly treats confidentiality as a central consideration when lawyers use generative AI (American Bar Association, 2024). The importance of confidentiality is heightened because legal information frequently includes privileged communications, commercially sensitive material, personal information and litigation strategy. Responsible AI use must therefore consider not only whether the output is accurate but also whether the information supplied to obtain that output was lawfully and professionally handled.

The transformation of professional responsibility can thus be understood through a shift from delegation of tasks to delegation of judgment. Delegating a repetitive task does not necessarily undermine professional responsibility. A lawyer may appropriately use software to organise documents, search large datasets or identify potential clauses. Delegating judgment is different. When an AI system effectively determines which legal authority governs, which interpretation should be adopted or which litigation strategy should be pursued, the lawyer risks allowing the technology to occupy a role that traditionally belongs to an accountable professional. The critical question is not whether AI performed part of the work, but whether the human professional retained genuine control over the legally consequential conclusion.

This concern becomes even more pronounced in judicial decision-making. Judicial authority is exercised on behalf of the state and directly affects legal rights and obligations. The EU AI Act consequently treats AI used by judicial authorities to assist in researching and interpreting facts and law and applying law to concrete facts as high-risk. Its recital 61 expressly recognises risks of bias, errors and opacity and states that AI may support judicial decision-making but should not replace it; final decision-making must remain human-driven (European Union, 2024, recital 61). The provision reflects a broader constitutional principle: technological assistance cannot itself become a substitute for the institutional responsibility of the judge.

Judicial use of AI also raises questions of impartiality and fairness. The American Bar Association's judicial ethics guidance identifies several provisions of the Model Code of Judicial Conduct that may be implicated by generative AI, including impartiality and fairness, bias, external influences, competence, supervision and the handling of non-public information (American Bar Association, 2025). The significance of these principles is that AI may influence judicial reasoning without being visible to the parties. If a system assists with legal research, summarisation or

drafting, the parties may have no practical means of knowing which material influenced the decision unless appropriate disclosure or institutional safeguards exist. The resulting asymmetry can complicate the ability to challenge the basis of a decision.

This leads directly to the problem of traceability. Accountability requires that an institution be able to reconstruct, to a meaningful degree, how a legally consequential AI-assisted output was produced and relied upon. UNESCO's framework expressly links accountability with auditability and traceability (UNESCO, 2021). In the legal context, traceability may require records of the system used, the relevant version, the information supplied, the nature of the task, the output relied upon and the human verification undertaken. Such records are not merely technical documentation. They can become important evidence if an AI-assisted legal conclusion is subsequently challenged.

Transparency must, however, be distinguished from full disclosure of proprietary technical information. A legal system does not necessarily require every litigant or lawyer to understand the mathematical architecture of a model. What matters is whether the system's use can be sufficiently understood to permit meaningful evaluation of its role in the decision. The legal requirement should therefore focus upon functional transparency: what the system was used for, what material it generated or influenced, what limitations were known, what verification occurred and how the responsible human actor ultimately reached the conclusion.

The problem of accountability becomes more difficult when responsibility is distributed among several actors. An AI-assisted legal outcome may involve the developer of the model, the provider of the legal AI product, the law firm or court deploying it, the individual lawyer or judge using it, and potentially third-party providers of underlying legal databases. This creates the possibility of an accountability gap. If an AI-generated recommendation is wrong, the user may argue that the technology caused the error, while the provider may argue that the system was intended only as an assistive tool and that the user was responsible for verification. A robust legal framework must therefore distinguish between different forms of responsibility rather than allowing technological complexity to obscure responsibility altogether.

This does not mean that every AI-related error should automatically be attributed to the human user. Responsibility should depend upon the nature of the task, the foreseeable limitations of the system, the information available to the user, the safeguards provided and the degree of human reliance. A sophisticated legal AI product marketed for a specific legal purpose may reasonably generate different expectations from a general-purpose conversational model. The standard of reasonable verification must therefore be sensitive to the technological context and the foreseeable risks associated with the particular system.

Reliability also has an equality dimension. If AI systems perform differently across jurisdictions, legal fields, languages or categories of cases, their use may produce unequal access to accurate legal analysis. The Stanford research on legal AI demonstrates meaningful differences between systems in accuracy and responsiveness, illustrating why performance claims should be evaluated empirically rather than accepted solely on the basis of technological sophistication (Magesh et al., 2025). A legal institution should therefore avoid assuming that an AI system that performs well in one jurisdiction or task will perform equally well in another. The relevant

benchmark must correspond to the population, jurisdiction, language, legal materials and decision-making context in which the system will actually be used.

This is particularly relevant to jurisdictions with complex multilingual or plural legal environments. Legal reasoning depends heavily upon the availability and correct interpretation of authoritative sources. An AI system whose training or retrieval environment disproportionately represents one jurisdiction or legal tradition may produce outputs that appear generally applicable while overlooking locally controlling law. The problem is therefore not simply algorithmic "bias" in an abstract sense; it can manifest as jurisdictional incompleteness, temporal obsolescence or systematic under-representation of particular legal authorities.

The appropriate response is not to reject AI-assisted legal practice altogether. AI can provide substantial benefits where its role is clearly defined and appropriately supervised. It can reduce the time required to search large bodies of material, generate alternative formulations, identify potential inconsistencies and assist professionals in handling information-intensive tasks. Properly deployed, these capabilities may allow lawyers and judges to devote greater attention to questions requiring contextual evaluation and normative judgment. The objective should therefore be accountable augmentation, not technological exclusion.

Accountable augmentation requires a distinction between the source of an output and the source of legal authority. AI may generate a memorandum, identify a potentially relevant judgment or propose an interpretation, but the legal authority must continue to derive from recognised legal sources. Similarly, AI may recommend a course of action, but the legally responsible professional must independently determine whether that course is justified. This preserves the functional division between computational assistance and legal judgment.

A useful model can therefore be constructed around five requirements: verification, independence, explainability, traceability and responsibility. Verification requires that material legal propositions and authorities be checked against primary or otherwise authoritative sources. Independence requires that the human decision-maker retain genuine capacity to reject the system's output. Explainability requires that the legal conclusion can be articulated through reasons intelligible within the relevant legal framework. Traceability requires sufficient records to establish how AI contributed to the process. Responsibility requires an identifiable human or institution that remains answerable for the legally consequential decision. These requirements correspond closely with the principles of human oversight, accountability, transparency and auditability reflected in contemporary international and regulatory frameworks (UNESCO, 2021; European Union, 2024).

The central proposition is therefore that human judgment should not be understood as the mere final step in an AI-assisted process; it must remain an active and substantive component of that process. A human who simply approves an AI-generated recommendation has not necessarily exercised meaningful judgment. Conversely, a lawyer or judge who critically interrogates the output, verifies its legal foundations, considers alternatives and provides an independent justification can use AI without surrendering legal responsibility.

This distinction is ultimately important to the legitimacy of AI-assisted law. Legal institutions derive authority not simply from producing outcomes but from producing outcomes through procedures and reasons that can be scrutinised. If AI becomes an opaque intermediary

between legal materials and legal conclusions, it may weaken the connection between authority and justification. If, however, AI remains subject to meaningful human oversight, independent verification and institutional accountability, it can become an instrument through which legal professionals increase their analytical capacity without relinquishing responsibility.

The appropriate legal principle is therefore neither "human judgment at all costs" nor "AI wherever technically possible." It is a principle of proportionate human responsibility: the greater the potential impact of AI on rights, obligations, liberty, property or access to justice, the greater the requirement for independent human evaluation, transparency and accountability. The EU AI Act's risk-based approach provides one important regulatory expression of this principle, while UNESCO's emphasis on human responsibility and accountability provides a broader normative foundation (European Union, 2024; UNESCO, 2021).

Ultimately, the reliability of AI-assisted law cannot be established by the sophistication of the technology alone. It depends upon the relationship between the system, the legal professional, the institution and the affected person. AI may be capable of generating increasingly convincing legal analysis, but convincing output is not synonymous with legally justified judgment. The continuing legitimacy of AI-assisted law will therefore depend upon maintaining a clear institutional principle: AI may assist in producing legal analysis, but accountable human actors must remain responsible for determining, explaining and defending the legal conclusions upon which rights and obligations depend.

4. Regulating AI in Legal and Judicial Decision-Making: A Comparative Analysis

The regulation of artificial intelligence in legal and judicial decision-making presents a distinctive regulatory problem because the consequences of error extend beyond ordinary commercial or administrative activity. AI used in a commercial recommendation system may affect consumer choice; AI used within a legal institution may affect liberty, property, family relationships, procedural rights, access to justice and the determination of legal liability. The regulatory question is therefore not simply whether an AI system is technically accurate, but whether its deployment is compatible with the institutional principles that govern the exercise of legal authority. Across jurisdictions, the emerging regulatory landscape reflects different approaches to this problem. The European Union has adopted a comprehensive, risk-based legislative framework; the United Kingdom has relied principally upon judicial guidance and institutional governance; the United States has developed a combination of judicial guidance, professional standards and decentralised institutional rules; and India is moving towards a more explicit framework for judicial AI use, alongside emerging judicial decisions addressing AI-generated material. These approaches differ considerably in legal form and regulatory technique, but they increasingly converge upon several principles: human control, independent verification, protection of fundamental rights, transparency, accountability and preservation of judicial independence.

The European Union currently provides the most developed example of a statutory approach. Regulation (EU) 2024/1689, commonly known as the Artificial Intelligence Act ('EU AI Act'), establishes a risk-based regulatory architecture under which obligations vary according to the

potential risk posed by an AI system. This architecture is particularly significant for legal decision-making because Annex III identifies as high-risk certain AI systems intended to be used by or on behalf of judicial authorities to assist in researching and interpreting facts and law and in applying the law to a concrete set of facts. The legislation nevertheless distinguishes such systems from purely ancillary judicial activities, such as administrative tasks, provided that those activities do not materially affect judicial decision-making (European Union, 2024, art 6; Annex III, para 8(a)). This distinction demonstrates an important regulatory proposition: the legal significance of AI depends not merely upon the technology employed but upon the function it performs within the decision-making process.

The EU approach is therefore functionally differentiated. An AI system used to organise case files does not necessarily present the same risks as one that analyses legal authorities and recommends how law should be applied to the facts of a dispute. The latter directly approaches the exercise of adjudicative judgment and consequently attracts more demanding safeguards. This functional approach is preferable to a purely technological classification because the same underlying model could have substantially different legal consequences depending upon whether it is used for transcription, document management, legal research, predictive analysis or substantive decision support.

The EU AI Act's treatment of judicial AI is also significant because it does not prohibit the use of AI in adjudication as such. Instead, it seeks to establish conditions under which high-risk systems may be used consistently with fundamental rights and institutional safeguards. Article 14 requires high-risk AI systems to be designed so that natural persons can effectively oversee their operation. Human oversight must be proportionate to the risks, the level of autonomy and the context of use. Persons responsible for oversight must be able to understand the system's capabilities and limitations, recognise automation bias, correctly interpret outputs, disregard or override outputs where necessary, and intervene in or interrupt the system (European Union, 2024, art 14).

This is more demanding than the simple proposition that a human must remain "in the loop". The legislation requires the human actor to possess an effective capacity to question and override the machine. In the context of legal reasoning, this distinction is crucial. A judge who formally retains authority but routinely accepts an AI-generated recommendation without independently examining the law and facts would satisfy the formal condition of human presence but could fail the substantive objective of human oversight. Article 14 therefore provides an important conceptual foundation for distinguishing human presence from meaningful human control.

The EU framework also imposes broader obligations relevant to the reliability of judicial AI. High-risk systems must achieve an appropriate level of accuracy, robustness and cybersecurity and must perform consistently throughout their lifecycle (European Union, 2024, art 15). They are also subject to requirements concerning risk management, data governance, technical documentation, record-keeping, transparency and human oversight. These requirements recognise that reliability cannot be established merely through a one-time demonstration that a model performs adequately. AI systems can change through updates, data changes, altered deployment conditions and interactions with users. Regulatory reliability must therefore be understood as a continuing lifecycle obligation rather than a static property.

The EU approach is particularly important for the rule of law because it recognises that judicial AI can implicate fundamental rights. Recital 61 of the EU AI Act explains that AI systems used by judicial authorities to assist with research, interpretation and application of law may have potentially significant impacts on democracy, the rule of law, individual freedoms and the right to an effective remedy and fair trial. It therefore treats such applications as high-risk while making clear that AI may support judicial activity without replacing judicial authority (European Union, 2024, recital 61). The regulatory objective is consequently not technological neutrality in the sense of ignoring AI, but institutional neutrality in the sense of ensuring that technological assistance does not alter the fundamental allocation of legal authority.

The European framework also provides an important distinction between risk regulation and prohibition. Some uses of AI are prohibited because of their unacceptable risks, while other systems are subject to heightened requirements because their risks can potentially be controlled. Judicial decision-support AI falls predominantly into the latter category. This indicates that the EU legislature does not regard AI assistance and the rule of law as inherently incompatible. Instead, compatibility depends upon governance conditions capable of controlling foreseeable risks.

The United Kingdom has adopted a materially different regulatory technique. Rather than introducing a single comprehensive AI statute equivalent to the EU AI Act, the UK has relied upon a combination of existing legal frameworks, sector-specific regulation and guidance. The judiciary has developed specific guidance for judicial office holders, demonstrating a preference for institutional governance in the judicial context. The October 2025 *Artificial Intelligence (AI) – Judicial Guidance* applies to judicial office holders in England and Wales and was expressly designed to assist judges and associated judicial personnel in using AI responsibly. It emphasises the integrity of the administration of justice, the rule of law and the personal responsibility of judicial office holders for material produced in their name (Courts and Tribunals Judiciary, 2025).

The UK guidance is significant because it places professional responsibility at the centre of AI governance. Rather than treating AI as a separate technological domain detached from existing judicial ethics, it integrates AI into established principles of judicial conduct. The October 2025 guidance specifically expands its treatment of bias in training data and AI hallucinations, warns against entering private information into public AI tools, and addresses confidentiality and data incidents (Courts and Tribunals Judiciary, 2025). The underlying principle is that judicial office holders remain personally responsible for material produced in their name even when AI has been involved in its production.

This approach has an important institutional advantage. Judicial decision-making is already governed by principles of independence, impartiality, integrity and reasoned adjudication. AI governance can therefore be embedded within those existing norms rather than constructed entirely as a new regulatory field. The approach also recognises that the most immediate risks may arise not from the existence of AI but from how individual judicial officers use it. The emphasis on confidentiality, verification and personal responsibility consequently addresses the point at which technological risk becomes legal responsibility.

The UK approach nevertheless differs significantly from the EU model in legal form. Judicial guidance does not create the same comprehensive statutory framework as the EU AI Act,

nor does it necessarily impose the same provider-facing requirements concerning risk management, technical documentation and conformity obligations. Its primary regulatory subject is the judicial user and the institutional environment rather than the AI provider. The difference is therefore one between technology-centred regulation and institution-centred governance. The EU framework places substantial obligations upon providers and deployers of high-risk AI; the UK judicial model places considerable emphasis upon responsible human use within the judiciary.

This distinction illustrates a broader comparative question: whether regulation of legal AI should focus principally upon the characteristics of the technology or upon the legal consequences of its deployment. The EU AI Act combines both approaches by imposing technical and organisational requirements while classifying use cases according to risk. The UK judicial framework places greater weight on the professional responsibilities of those who use AI. Neither approach necessarily excludes the other. Indeed, effective governance of judicial AI may require both: reliable systems and accountable users.

The United States presents a further model characterised by institutional and professional guidance rather than a single federal statutory framework governing judicial AI. The American Bar Association has issued *Guidelines for Judicial Officers Regarding the Responsible Use of Artificial Intelligence*, which state as a foundational principle that judicial authority is vested in judicial officers rather than AI systems. The Guidelines recognise the potential utility of AI while emphasising that judges remain responsible for judicial decisions and must use appropriate safeguards when employing AI tools (American Bar Association, 2025).

The federal judiciary has likewise developed interim guidance concerning AI. The Administrative Office of the United States Courts reported in its 2025 Annual Report that an AI Task Force was established to address the opportunities and risks presented by AI. Its interim guidance cautioned against delegating core judicial functions, including decision-making and adjudication, to AI and recommended independent verification of AI-generated content. It further emphasised that judiciary users and those approving AI use remain accountable for work performed with AI assistance (Administrative Office of the United States Courts, 2025).

This approach is particularly revealing because it recognises that judicial AI can be beneficial without allowing the technology to displace judicial authority. AI can assist with administrative functions, document processing and other tasks while core adjudicative responsibility remains with judges. The United States model consequently resembles the UK's institutional approach more closely than the EU's comprehensive statutory model, although the precise legal mechanisms differ.

The decentralised American structure also means that AI governance may develop through multiple sources: judicial guidance, professional ethics rules, court-specific policies, state judicial standards and federal judicial policy. This can permit experimentation and adaptation to rapidly developing technologies, but it may also produce inconsistency. A lawyer or judge working across jurisdictions may encounter different expectations concerning disclosure, verification, permissible uses and data protection. The absence of a single comprehensive framework can therefore make the regulatory landscape more fragmented than the EU model.

Professional responsibility provides an additional layer of US regulation. The American Bar Association's Formal Opinion 512 establishes that existing professional duties apply when lawyers

use generative AI, including duties of competence, confidentiality, communication, supervision, candour toward tribunals and reasonable fees (American Bar Association, 2024). This is important because it demonstrates that technological innovation does not create a normative vacuum. Existing legal duties can continue to regulate emerging technologies where their underlying principles are sufficiently general.

The comparative significance of the United States approach is therefore its emphasis upon responsibility through existing legal institutions. Rather than treating AI as a wholly novel category requiring a separate legal profession, the approach asks how established obligations apply when lawyers and judges adopt new technological tools. This has considerable doctrinal value because professional duties are already designed around the protection of clients, courts and the integrity of legal proceedings.

India presents an especially important jurisdiction for comparative analysis because it combines rapid technological development with a constitutional framework in which judicial independence, equality, due process and access to justice occupy central positions. India's approach to AI in the justice system has historically developed through institutional initiatives and technological projects, but by 2026 the regulatory discussion has become considerably more explicit. The Supreme Court of India has published a *White Paper on Artificial Intelligence and Judiciary* and, importantly, in 2026 published a notice inviting comments and suggestions on draft *Regulations for Use of AI in Courts 2026*.

The proposed Indian framework is particularly relevant to the present article because its formulation expressly places AI beneath human judicial authority. The draft regulations provide that AI use in court processes must remain subservient to human judgment and judicial authority and that AI systems should operate in an assistive capacity rather than supplanting the independent exercise of judicial authority. They further state that ultimate authority to determine questions of law, fact and justice remains with judicial officers (Supreme Court of India, draft Regulations for Use of AI in Courts 2026). Because these are draft regulations, they should not be presented as settled law; nevertheless, their formulation provides important evidence of the direction of Indian institutional thinking.

India's developing framework has also been shaped by judicial experience with AI-generated material. In *Pooja Ramesh Singh v Jammu and Kashmir Bank Ltd* (2026 INSC 668), the Supreme Court considered a case in which the NCLT and NCLAT had relied upon six AI-generated citations that were subsequently found to be non-existent or to contain attributed paragraphs that did not exist. The Supreme Court treated the issue as one implicating the integrity of adjudication and emphasised the need for strict control over AI-generated material introduced into judicial reasoning (Supreme Court of India, 2026).

The case is particularly significant because it demonstrates that AI regulation in judicial decision-making is not merely a theoretical question. The reliability of AI-generated authorities can directly affect the legal validity and institutional legitimacy of judicial reasoning. The Supreme Court's approach illustrates a principle that appears across the jurisdictions considered here: AI-generated material cannot acquire the status of legal authority merely because it is presented in the language of law. Authorities must be independently verified before they are relied upon.

The Indian position is also noteworthy because the Supreme Court's approach distinguishes between legitimate AI assistance and reliance upon fabricated or hallucinated material. The Court's judgment does not establish that AI has no legitimate role in judicial administration. Rather, it addresses the particular danger of treating AI-generated material as genuine precedent without verification. The distinction is important for regulatory design: the objective is not necessarily to exclude AI from courts, but to ensure that AI-generated content does not bypass the verification mechanisms upon which adjudication depends.

The comparative experience of these jurisdictions reveals a substantial area of convergence. First, human judicial authority remains non-delegable. The EU AI Act permits AI assistance but requires human oversight; UK guidance emphasises the personal responsibility of judicial office holders; US judicial guidance states that judicial authority belongs to judicial officers; and India's emerging framework expressly subordinates AI to human judgment. (European Union, 2024; Courts and Tribunals Judiciary, 2025; American Bar Association, 2025; Supreme Court of India, 2026).

Secondly, independent verification has become a common regulatory response to hallucination and error. The US federal judiciary recommends independent verification of AI-generated content; the UK guidance highlights hallucinations and the need for careful use; the EU framework requires human oversight and appropriate accuracy and robustness; and the Indian Supreme Court's 2026 decision directly demonstrates the consequences of failing to verify AI-generated authorities. (Administrative Office of the United States Courts, 2025; Courts and Tribunals Judiciary, 2025; European Union, 2024; Supreme Court of India, 2026).

Thirdly, accountability remains attached to human and institutional actors rather than to the AI system itself. The EU requires meaningful human oversight; the UK places personal responsibility upon judicial office holders; the US judiciary states that users remain accountable for AI-assisted work; and India's draft framework reserves ultimate authority over law, fact and justice for judicial officers. These approaches differ in legal technique but share a common institutional premise: AI cannot itself become the bearer of judicial authority. (European Union, 2024; Courts and Tribunals Judiciary, 2025; Administrative Office of the United States Courts, 2025; Supreme Court of India, 2026).

Fourthly, the jurisdictions increasingly recognise that confidentiality and data protection are integral to responsible judicial AI. The UK's 2025 judicial guidance specifically warns judicial office holders against entering private information into public AI tools and addresses the reporting of inadvertent disclosures. The US judiciary has similarly identified protection of sensitive data and security of judicial information systems as significant concerns in its AI policy development. These concerns are particularly important because judicial records frequently contain highly sensitive personal and commercial information, and the consequences of inappropriate disclosure can extend beyond the immediate case.

Despite these common principles, substantial differences remain. The EU model is the most legislatively comprehensive and risk-based. It regulates the AI system through a detailed statutory framework and places obligations upon providers, deployers and users. The UK model is more principle-based and institutionally embedded, relying heavily upon judicial guidance and existing constitutional and professional obligations. The US model is comparatively decentralised, combining

federal judicial policy, professional guidance and local institutional rules. India's approach is developmental and institutionally evolving, with the Supreme Court playing a significant role in shaping safeguards while formal regulations for court AI remain under development.

These differences matter because each regulatory technique addresses a different part of the accountability chain. A provider-centred regime can impose technical requirements before a system reaches a court. A user-centred regime can ensure that judges and lawyers understand their responsibilities. An institution-centred regime can establish procurement, security, training and disclosure rules. A judicially developed framework can respond rapidly to new problems through case law and practice directions. No single technique necessarily addresses every dimension of AI risk.

A particularly important comparative lesson concerns the meaning of human oversight. The EU AI Act provides the most detailed statutory formulation: the human overseer must understand the system's capabilities and limitations, recognise automation bias, interpret its output and possess the ability to disregard or override it (European Union, 2024, art 14). The UK and US approaches similarly emphasise human responsibility, while India's emerging regulations place ultimate legal authority expressly with judges. (Courts and Tribunals Judiciary, 2025; American Bar Association, 2025; Supreme Court of India, 2026). The common weakness, however, is that "human oversight" can become purely formal unless institutions provide sufficient time, expertise, access to underlying sources and authority to reject the AI output.

The regulatory challenge is therefore to move from human-in-the-loop to human-in-control. A judge cannot meaningfully supervise an AI system if the judge cannot understand its limitations, identify errors or independently access the underlying authorities. Similarly, a lawyer cannot discharge professional responsibilities merely by reading an AI-generated memorandum if the lawyer does not verify the citations or understand the system's limitations. Human control requires institutional capacity, not simply formal responsibility.

The comparative analysis also demonstrates that regulation should distinguish between AI-assisted legal reasoning and AI-generated legal authority. There is considerable scope for AI to assist with administrative tasks, legal research, summarisation and preliminary analysis. The legal difficulty becomes substantially greater where AI output is treated as an authoritative determination of law or fact. The regulatory objective should consequently be to preserve a clear chain of legal justification: AI may generate information or analytical suggestions, but the human legal actor must establish the legal proposition through recognised sources and independently justify its application.

This principle is particularly important for judicial reasoning. Courts are not merely information-processing institutions. Their decisions constitute exercises of public authority. The legitimacy of an adjudicative decision therefore depends upon more than its statistical or practical accuracy. It depends upon lawful jurisdiction, procedural fairness, impartiality, reason-giving and the possibility of review. AI may contribute to the informational and analytical stages of this process, but it should not obscure the legal reasoning through which authority is exercised.

The comparative evidence consequently supports a model of regulated augmentation rather than either unrestricted automation or categorical technological exclusion. AI should be permitted to perform functions that improve efficiency and analytical capacity where adequate safeguards exist,

but the level of oversight should increase as the system approaches substantive legal judgment. Administrative assistance may require ordinary institutional controls; legal research may require verification; substantive recommendations may require documented human review; and any system capable of materially influencing adjudication should be subject to the strongest safeguards concerning accuracy, transparency, traceability, human control and accountability.

The emerging regulatory landscape therefore suggests a developing international principle: the closer AI comes to determining legal rights, obligations or judicial outcomes, the stronger the legal requirement for demonstrable human control and independent justification. This principle is visible, although expressed differently, in the EU's risk classification, the UK's judicial guidance, US judicial policies and India's emerging regulations and case law. (European Union, 2024; Courts and Tribunals Judiciary, 2025; Administrative Office of the United States Courts, 2025; Supreme Court of India, 2026).

The comparative analysis ultimately reveals that the most important regulatory question is not whether AI should be used in legal and judicial decision-making, but under what institutional conditions its use remains compatible with accountable legal judgment. The European Union supplies a detailed statutory model; the United Kingdom demonstrates the potential of judicial guidance; the United States illustrates the role of professional and institutional governance; and India illustrates how judicial experience and emerging regulation can respond directly to the evidentiary and adjudicative risks created by generative AI. Their differences are significant, but their convergence around human authority, verification and accountability provides an important foundation for developing a broader normative framework.

For the purposes of this article, the comparative evidence supports a principle of accountable human primacy. AI should remain subordinate to legally authorised human decision-makers whenever its outputs may materially affect legal rights or judicial outcomes. Human primacy, however, should not mean merely retaining a human signature at the end of an automated process. It should require genuine capacity to understand, interrogate, verify, reject and explain AI-assisted outputs. Such a model would preserve the efficiency and analytical benefits of AI while maintaining the institutional features—reasoned judgment, accountability, transparency and contestability—that distinguish the administration of law from automated prediction.

5. Conclusion

Artificial intelligence is fundamentally altering the environment in which legal reasoning is undertaken. The emergence of generative AI and large language models has moved the discussion beyond the traditional automation of administrative and information-retrieval functions towards technologies capable of generating legal propositions, synthesising authorities, constructing arguments and assisting with decision-support. The significance of this transformation, however, cannot be measured solely by the sophistication of the technology or by its capacity to reproduce the language of legal analysis. The central legal question is whether AI can be integrated into legal institutions without weakening the human judgment, accountability, transparency and institutional legitimacy upon which the rule of law depends.

The analysis demonstrates that legal reasoning cannot be reduced to the generation or retrieval of legally plausible information. Legal reasoning involves the identification and interpretation of authoritative sources, the assessment and characterisation of facts, the application of legal rules and principles, the reconciliation of competing authorities, and the articulation of reasons capable of scrutiny and challenge. These functions explain why linguistic fluency is not equivalent to legal validity. An AI system may produce a sophisticated and persuasive legal argument while citing a non-existent authority, mischaracterising a judgment, overlooking a controlling precedent or applying an apparently relevant principle to materially different facts. Empirical research concerning AI-powered legal research tools confirms that even specialised systems can generate hallucinated or misgrounded legal material, demonstrating that verification remains indispensable (Magesh et al., 2025).

The principal implication is that the use of AI cannot transfer legal responsibility from human actors to technological systems. AI systems do not become judges, lawyers or independent sources of legal authority merely because they can generate language associated with legal reasoning. The legal authority of a judicial decision continues to derive from the constitutionally and legally recognised institution exercising jurisdiction; the authority of a lawyer's submission continues to depend upon professional responsibility and its grounding in applicable law. AI may contribute to the process through which legal analysis is developed, but it cannot itself assume the institutional responsibility attached to the resulting legal determination. UNESCO's Recommendation on the Ethics of Artificial Intelligence similarly places human responsibility and accountability at the centre of AI governance and emphasises that AI systems should not displace ultimate human responsibility (UNESCO, 2021).

The comparative analysis reinforces this conclusion. Although the European Union, United Kingdom, United States and India employ different regulatory techniques, each increasingly recognises the importance of human control when AI is used in legally consequential contexts. The EU AI Act adopts a risk-based statutory framework and identifies certain AI systems used by or on behalf of judicial authorities to assist with researching and interpreting facts and law as high-risk systems. It requires appropriate human oversight and expressly provides that AI systems intended to assist judicial authorities should not replace judicial decision-making (European Union, 2024, recital 61; Annex III, para 8(a)). The United Kingdom's judicial guidance similarly emphasises that judicial office holders remain personally responsible for material produced in their name and highlights the risks associated with hallucinations, bias and confidential information (Courts and Tribunals Judiciary, 2025). In the United States, federal judicial guidance and professional standards likewise emphasise independent verification and prohibit the delegation of core judicial functions to AI (Administrative Office of the United States Courts, 2025). India's emerging judicial framework similarly places AI in an assistive role while retaining ultimate authority over questions of law, fact and justice with judicial officers, and recent Supreme Court experience illustrates the practical consequences of relying upon unverified AI-generated authorities (Supreme Court of India, 2026).

These developments suggest that the appropriate regulatory objective is neither unrestricted automation nor categorical rejection of AI. The more defensible approach is accountable augmentation. AI can legitimately increase the speed and breadth of legal research, assist with

document analysis, identify potentially relevant authorities, generate preliminary arguments and support administrative functions. Such applications may allow lawyers and judges to devote greater attention to complex questions requiring contextual assessment and normative judgment. The American Bar Association's Formal Opinion 512 illustrates this approach by recognising legitimate uses of generative AI while reaffirming lawyers' existing duties of competence, confidentiality, communication, supervision and candour towards tribunals (American Bar Association, 2024).

The critical boundary, therefore, is not between technology and humanity but between assistance and substitution. Assistance preserves the human actor's capacity to independently evaluate and reject an AI-generated proposition. Substitution occurs when the technological output becomes effectively determinative and the human actor merely confirms or adopts the machine's conclusion. The distinction is particularly important in adjudication. A judge may use AI to organise a large body of authorities or identify potentially relevant legal questions without delegating judicial authority. By contrast, a system that effectively determines the applicable law, evaluates disputed facts or recommends an outcome that the judge accepts without independent examination risks transforming judicial decision-making into technological decision adoption. The existence of a human signature at the end of that process would not, by itself, establish meaningful human judgment.

This distinction also requires a more rigorous understanding of the frequently invoked concept of human-in-the-loop. Human presence should not be treated as sufficient simply because a person formally remains responsible for the final decision. Meaningful human oversight requires substantive capacity to understand the system's capabilities and limitations, evaluate its output, identify errors, consider alternatives and reject the recommendation where necessary. The EU AI Act provides an especially developed expression of this principle by requiring human oversight mechanisms that enable the relevant person to understand limitations, recognise automation bias and disregard, override or reverse an AI-generated output where appropriate (European Union, 2024, art 14). The principle has wider relevance beyond the EU: human control must be real rather than merely procedural.

Reliability must consequently be treated as a continuing legal and institutional obligation. No AI system should be regarded as universally reliable simply because it performs well in a particular benchmark or legal domain. Reliability depends upon the jurisdiction, legal subject, quality of the underlying data, currency of legal materials, nature of the task and consequences of error. An AI system used to classify administrative documents presents a different risk profile from one used to assist with a judicial determination concerning liberty or fundamental rights. Regulation should therefore be proportionate to the potential legal consequences of error. This risk-based approach is consistent with the architecture of the EU AI Act and provides a useful foundation for the broader governance of legal AI (European Union, 2024).

The requirement of independent verification follows directly from this conclusion. Legal professionals must not treat AI-generated authorities as self-authenticating. Citations must be checked against authoritative sources; quotations must be verified against the original judgment or legislation; statutory provisions must be examined in their current form; and the relevance of precedents must be independently assessed. The problem is not confined to obvious hallucinations.

As empirical research has demonstrated, AI systems can also produce answers that appear correct while citing authorities that do not actually support the proposition stated (Magesh et al., 2025). In law, where the relationship between proposition and authority is itself constitutive of legal reasoning, this form of error is particularly serious.

Accountability must similarly extend beyond the immediate user. The increasing complexity of AI systems means that legal responsibility can potentially be distributed across developers, providers, deployers, institutions and professional users. This creates a risk that each participant may attribute an erroneous result to another actor. A meaningful regulatory framework must therefore establish a clear chain of responsibility. The lawyer who submits AI-generated material to a court cannot avoid professional responsibility by attributing the error to the model. A judicial institution cannot avoid institutional responsibility merely because a technology vendor supplied the system. At the same time, providers should not necessarily be insulated from responsibility where their own conduct creates foreseeable risks. Accountability should correspond to the functions, knowledge, control and obligations of each actor within the AI system's deployment environment.

Transparency and traceability are consequently essential components of responsible legal AI. Where AI materially contributes to a legally consequential decision, there should be sufficient information to establish what role the system played, what type of output it generated, what material informed that output and what human verification occurred. Complete disclosure of proprietary technical architecture will not necessarily be required in every case. What is essential is functional transparency sufficient to permit meaningful scrutiny of the system's role in the legal process. Without such traceability, it may become difficult for an affected person, appellate court, professional regulator or other oversight body to determine how an erroneous conclusion was produced.

This is particularly important because the legitimacy of legal decision-making depends upon more than the correctness of its final result. Courts and legal institutions operate within structures of reason-giving, procedural fairness and review. A decision must ordinarily be capable of being understood and challenged through recognised legal mechanisms. AI-assisted decision-making should therefore preserve, rather than weaken, the connection between a legal conclusion and the reasons supporting it. An AI-generated recommendation may form part of the analytical background to a decision, but it cannot itself substitute for a legally intelligible justification.

The implications for professional legal education are equally significant. Competence in the AI era increasingly requires lawyers to possess sufficient technological literacy to understand the capabilities and limitations of the systems they employ. This does not mean that lawyers must become programmers or AI engineers. It means that they must be able to identify foreseeable risks, determine when AI use is appropriate, verify outputs, protect confidential information and exercise independent professional judgment. The continuing application of professional duties to AI-assisted practice, as reflected in ABA Formal Opinion 512, demonstrates that technological change does not remove the need for professional standards; it changes the circumstances in which those standards must be applied (American Bar Association, 2024).

The same principle applies to judicial education. Judges need not understand every technical aspect of machine-learning architecture, but judicial institutions should ensure that judges using AI have sufficient knowledge to recognise hallucinations, automation bias, data limitations and other

foreseeable risks. Institutional training, procurement standards, audit mechanisms and clear policies concerning permissible uses are therefore not merely administrative matters. They are safeguards for judicial independence and the integrity of adjudication. The UK's judicial guidance and the United States judiciary's developing AI policies illustrate how such safeguards can be incorporated into existing judicial governance structures rather than treating AI as an issue external to judicial ethics (Courts and Tribunals Judiciary, 2025; Administrative Office of the United States Courts, 2025).

The comparative experience further demonstrates that there is no single regulatory model capable of addressing every dimension of legal AI. A comprehensive statutory framework such as the EU AI Act can establish binding obligations concerning risk management, human oversight, accuracy, transparency and accountability. Judicial guidance can translate broader principles into practical rules for judges and court personnel. Professional regulation can establish obligations for lawyers. Institutional policies can govern procurement, security, training and permitted uses. Judicial decisions can respond to concrete failures and develop principles through case-specific reasoning. Effective governance is therefore likely to require a layered regulatory architecture rather than reliance upon a single legal instrument.

Such an architecture should be guided by proportionality. The degree of human scrutiny should correspond to the potential impact of AI on legal rights and interests. Low-risk administrative applications may require relatively limited safeguards. AI-assisted legal research should require reliable source verification. AI systems that materially influence substantive legal analysis should require documented human review. Systems used in adjudication or other contexts involving fundamental rights should be subject to the strongest requirements concerning accuracy, transparency, human oversight, traceability and accountability. This graduated approach would avoid both excessive restrictions that prevent beneficial technological innovation and insufficient safeguards that permit consequential decisions to become effectively automated.

The article therefore proposes accountable human primacy as a governing principle for AI-assisted legal reasoning. Under this principle, AI may assist with the production, organisation and analysis of legal information, but legally consequential judgment must remain attributable to a human actor or institution possessing lawful authority to make that determination. Human primacy does not require humans to perform every analytical task themselves. Nor does it require the exclusion of technologically generated recommendations. It requires that the final legal judgment remain the product of meaningful human evaluation and that the responsible decision-maker be able to explain, defend and, where necessary, correct the conclusion reached.

This principle is also compatible with technological innovation. Artificial intelligence should not be viewed solely as a threat to legal institutions. Properly designed and responsibly deployed systems may increase access to legal information, reduce repetitive workloads, assist with large-scale document analysis and help legal professionals identify issues that might otherwise be overlooked. The relevant objective is therefore not to preserve every existing method of legal work unchanged. It is to ensure that technological efficiency does not become a justification for removing the institutional safeguards through which law derives its legitimacy.

The most important boundary is ultimately one of responsibility rather than capability. AI systems will continue to become more capable of generating sophisticated legal language, identifying relevant material and assisting with complex analytical tasks. The increasing technical ability of such systems does not, however, automatically confer upon them the institutional status of legal decision-makers. The law must maintain a distinction between a system capable of generating an answer and an institution authorised to determine what the law requires. That distinction is fundamental to preserving accountability.

The future of legal reasoning should therefore not be framed as a contest between human lawyers and artificial intelligence. The more consequential question is how legal institutions can structure human–AI interaction so that technological assistance enhances, rather than obscures, the exercise of legal judgment. AI should be capable of challenging assumptions, identifying alternative arguments and accelerating research; human legal actors should remain responsible for determining whether those outputs are legally sound. The resulting model is neither purely human nor purely automated. It is a system of technologically assisted but institutionally accountable legal reasoning.

Ultimately, the rule of law requires more than decisions that are efficient, predictable or statistically accurate. It requires decisions that can be attributed to lawful institutions, justified through recognised legal sources, subjected to appropriate scrutiny and challenged through established procedures. Artificial intelligence can contribute substantially to these processes, but it cannot replace the institutional responsibility through which they acquire legal legitimacy. The central challenge for contemporary legal systems is therefore not whether AI can reproduce the language of legal reasoning, but whether its integration can preserve the conditions under which legal reasoning remains accountable.

The appropriate response is consequently neither technological prohibition nor uncritical technological adoption. It is principled integration. AI should be incorporated into legal practice and adjudication where its benefits can be realised consistently with verification, transparency, confidentiality, human oversight, procedural fairness and institutional accountability. Where those safeguards cannot be adequately maintained, the legal consequences of AI deployment should be correspondingly restricted. Such an approach recognises both the transformative potential of artificial intelligence and the distinctive character of law as an institution of public authority.

The enduring principle should therefore be clear: artificial intelligence may assist legal judgment, but it should not become an unaccountable substitute for it. The legitimacy of AI-assisted law will ultimately depend not upon how convincingly machines can imitate legal reasoning, but upon whether legal institutions remain capable of identifying who exercised judgment, on what legal basis, through what process and with what responsibility. Preserving that chain of accountability is essential if the technological transformation of legal reasoning is to strengthen rather than diminish the rule of law.

References

Administrative Office of the United States Courts. 2025. Annual Report 2025: Court Operations. Washington, DC: Administrative Office of the United States Courts. [https://www.uscourts.gov/data-news/reports/annual-reports/directors-annual-report/annual-report-2025].

American Bar Association. 2024. Formal Opinion 512: Generative Artificial Intelligence Tools. Chicago: American Bar Association, Standing Committee on Ethics and Professional Responsibility, 29 July 2024. [https://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/ethics-opinions/aba-formal-opinion-512.pdf].

American Bar Association. 2025. "Generative Artificial Intelligence and Judicial Ethics: A Brief Overview." The Judges' Journal. Chicago: American Bar Association. [https://www.americanbar.org/groups/judicial/resources/jd-record/2025/generative-artificial-intelligence-judicial-ethics/].

American Bar Association. 2025. "Guidelines for U.S. Judicial Officers Regarding the Responsible Use of Artificial Intelligence." The Judges' Journal. Chicago: American Bar Association. [https://www.americanbar.org/groups/judicial/resources/judges-journal/2025-spring/guidelines-judicial-officers-responsible-use-artificial-intelligence/].

Ashley, Kevin D. 2017. Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age. Cambridge: Cambridge University Press.

Courts and Tribunals Judiciary. 2025. Artificial Intelligence (AI) – Judicial Guidance (October 2025). London: Courts and Tribunals Judiciary, 31 October 2025.

[https://www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance-october-2025/].

European Union. 2024. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). Official Journal of the European Union, L 2024/1689. [https://eur-lex.europa.eu/eli/reg/2024/1689/].

Magesh, Varun, Faiz Surani, Matthew Dahl, Mirac Suzgun, Christopher D. Manning, and Daniel E. Ho. 2025. "Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools." Journal of Empirical Legal Studies 22 (2): 216–242. [https://law.stanford.edu/publications/hallucination-free-assessing-the-reliability-of-leading-ai-legal-research-tools/].

Mata v. Avianca, Inc. 2023. 678 F. Supp. 3d 443 (S.D.N.Y.).

Pooja Ramesh Singh v. Jammu and Kashmir Bank Ltd. 2026. 2026 INSC 668. Supreme Court of India, 2 July 2026.

Supreme Court of India. 2026. Draft Regulations for Use of Artificial Intelligence (AI) in Courts, 2026. New Delhi: Artificial Intelligence Committee, Supreme Court of India, 3 June 2026. [https://www.sci.gov.in/].

Susskind, Richard. 2008. The End of Lawyers? Rethinking the Nature of Legal Services. Oxford: Oxford University Press.

Surden, Harry. 2019. "Artificial Intelligence and Law: An Overview." Georgia State University Law Review 35 (4): 1305–1338. [https://readingroom.law.gsu.edu/gsulr/vol35/iss4/8].

UNESCO. 2021. Recommendation on the Ethics of Artificial Intelligence. Paris: United Nations Educational, Scientific and Cultural Organization, adopted 23 November 2021. [https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence].

Publisher's Note