Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 1910–1920
https://doi.org/10.54941/ahfe1008419

A Regulatory and Human Factors Approach to Overreliance in Legal Information Chatbots

Niki Georgiadou and Maria Rigou

Department of Management Science and Technology, University of Patras, GR 26334, Greece

ABSTRACT

Chatbots are rapidly transforming the way individuals from diverse social and economic backgrounds access legal information. This free access to legal knowledge can act as a corrective mechanism within society, reducing inequalities among social groups. Notwithstanding their benefits, the growing use of chatbots for legal matters raises significant concerns. One major issue is the accuracy of the information provided, since correctly applying a rule is a complex exercise, not a simple act of reading. Additional risks involve the processing of personal and sensitive data, which may be leaked or over-shared without users' deliberate consent or awareness. Concerns also arise regarding algorithmic bias, since AI systems may reproduce discriminatory patterns embedded in training or design data. Furthermore, chatbots may raise issues concerning the unauthorized practice of law. This article argues that legal question-answering chatbots, regardless of their pre-designed purpose, may endanger users' fundamental rights, and that classifying them as limited-risk systems undermines the goal of ethical, responsible and trustworthy AI. It further argues that the transparency framework under Article 50(1) of the AI Act is insufficient and proposes embedding ethics-by-design and ethics-by-default through a set of concrete interaction safeguards to actively support verification of chatbot-provided information before reliance: cognitive forcing functions, interaction friction, scaffolded verification, and calibrated uncertainty and limitation cues.

Keywords: Legal information chatbots, Automation bias, Trust, AI Act, Transparency, Ethics-by-design

INTRODUCTION

The widespread use of artificial intelligence technologies has had a significant impact across scientific professions. Internet users, from the comfort of their personal location, at no cost, address AI Large Language Models (LLMs) a variety of questions on scientific issues (for example on medicine, taxation, finance, commerce, employment, public service etc.) (Hasal et al., 2021). The legal field similarly experiences significant engagement, particularly in prompts requesting legal information or guidance (Setyawan, 2025). Chatbots (AI-driven systems capable of interacting with users through natural language) have emerged as an innovative solution which can provide immediate, scalable, cost-effective legal information, helping users navigate legal challenges. Upon initial consideration, this integration of

Received September 25, 2026; Revised October 30, 2026; Accepted November 6, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

chatbots into the legal field helps individuals comprehend legal problems and become aware of their rights and obligations and therefore come closer to experiencing a sense of justice, empowering especially social groups with a low economic or educational profile (Dahan & Liang, 2021). Chatbots, using natural language processing (NLP) and machine learning, can address these challenges in several ways and function as a corrective mechanism within society by reducing inequalities and mitigating disparities among social groups (Ogunsan & Johnson, 2025, pp. 2–12).

On the other hand, legal information or -even worse- guidance from chatbots involves risks, initially concerning the correctness of legal answers. The multifaceted nature of reality cannot be adequately captured through a prompt, particularly when prompts are formulated by individuals lacking familiarity with effective prompt design. In any case, the possibility of incorrect answers cannot be ruled out. Additional legal risks concern the protection of individuals' personal data and privacy, the prevention of biased outputs, and the preservation of legal and professional ethics within the legal profession (Tabrez et al., 2025).

This study investigates the deployment of chatbots in delivering free legal assistance and guidance, with particular attention to their limitations, ethical challenges, and broader implications for users and the legal system. The scope of this research is exclusively limited to individuals. It does not examine the use of specialized or general-purpose chatbots by legal professionals in the exercise of their duties, nor their use by public administrations or judicial authorities.

LIMITATIONS AND ETHICAL CHALLENGES

Information Accuracy

Legal language is notoriously **complex** and challenging to interpret (Katz et al., 2024). Terms in legal texts rarely have a single meaning, which ultimately often depends on the specific branch of law (e.g., torts vs. property law) or the surrounding legislative intent (Urbina, 2024, p. 1661). While polysemy deals with multiple defined meanings, vagueness involves terms which are not strictly defined in the law, such as **indefinite legal concepts** ("good faith", "good morals", "social manners and customs" etc.) or **principles** ("the principle of ultima ratio", "the principle of trust", "the principle of proportionality", "the principle of favoring the weaker party", etc.) or even **the law's intended purpose or spirit** that serves as the foundation for teleological interpretation. In the timelessness of juridical history, the effectiveness of legal approach directly depends on the interpretative practice of law (Strunskiy, 2021). Furthermore, the jurist must specify these concepts and principles according to the circumstances of each case, also considering emotional or mental parameters. Therefore, the correct implementation of a rule is not a simple act of reading but a sophisticated, multi-layered, analytical exercise. Attorneys routinely combine abstract reasoning and problem-solving skills in environments of legal and factual uncertainty (Surden, 2020). While chatbots can provide general legal information in simple questions, they are not equipped to handle the nuances and complexities of legal cases (Ogunsan & Johnson, 2025, pp. 2–12).

A further challenge concerns **jurisdictional diversity**. Each jurisdiction has its own national laws, usually written in its official language, which operate alongside and interact with European and international legal frameworks (Burchardt, 2019). The governing rule is typically a synthesis of these multi-level influences across various languages (Bahanov, 2021). An LLM trained predominantly on Anglocentric legal corpora may provide false answers when applied to users governed by non-English national statutes. Concurrently, it remains difficult for non-lawyers to ascertain whether the chatbot's outputs are accurate or up to date (Dahan & Liang, 2021).

Beyond the inherent challenges associated with interpreting legal texts, the accuracy of chatbots' responses is technically dependent on the prompt design (Pawlik, 2025). The greater the precision and specificity of the question posed, the more reliable the resulting answer is likely to be. The same applies to follow-up inquiries, which play a decisive role in determining the overall correctness and reliability of the response. Such requirements, however, are unlikely to be met by people who do not have the **digital literacy skills** to use the AI tools, or to use them effectively, thereby increasing the risk of inaccurate or misleading legal guidance (Pangrazio, 2020).

As a counterargument, it could be argued that a human lawyer may also make erroneous assessments or provide ineffective or inappropriate advice; however, the lawyer–client relationship constitutes an interpersonal legal relationship in which rights, contractual and ethical obligations and professional responsibilities are established (Xu et al., 2024). This clearly defined allocation of roles and duties represents a crucial distinction between legal guidance provided by a natural person and the algorithmically generated outputs of a chatbot.

Data Protection and Confidentiality

Given the sensitivity of legal issues, users may disclose personal information to chatbots, such as details about disputes, employment, taxation, income, or contracts, or criminal charges (Setyawan, 2025). The so-called "privacy paradox" is particularly evident in the context of chatbot interactions (Meng & Liu, 2025). Although individuals frequently express concerns regarding the protection of their personal data, they often disclose sensitive or extensive information to conversational AI systems in exchange for convenience, efficiency, or personalized assistance, while comparatively few users employ technical safeguards to protect their privacy online. Without strict data governance and encryption standards, this information could be exposed or misused, violating privacy laws such as the General Data Protection Regulation (GDPR).

At the same time, the growing demand for data (particularly text data) to train and enhance LLMs encourages developers to gather as much information as possible directly from users (Hasal et al., 2021). This can create tension with **key data protection principles** (Article 5 GDPR), such as data minimization (which requires data controllers to collect only the information necessary for a specific purpose), purpose limitation (which restricts the use of data to the purpose for which it was originally collected),

consent (which requires consumers to knowingly agree to the collection and use of their personal data), and storage limitation (which requires that personal data may be stored for no longer than is necessary for the purposes for which the personal data are processed). In confirmation of these concerns, major platforms that hold extensive data on billions of users have stated that, by default, they use users' inputs to further train their LLMs unless users actively opt out. This emerging practice is apparently inconsistent with the principles of data minimization and purpose limitation (King, 2025). Besides, the secure transfer of messages, voice, and images to a server on which the chatbot is hosted raises significant security concerns (Hasal, 2021).

Bias

Another concern from the use of chatbots is the possibility of bias, as the algorithmic result recycles **existing biases** and could reinforce inequalities and discrimination (Georgiadou et al., 2026). AI bias stems from the human bias (conscious or unconscious) hidden in AI algorithms throughout their development (Alkhaldi, 2024). Sources of potential bias are hidden either in datasets or in machine learning algorithms and other components of AI systems (Mehrabi et al., 2021). For example, if in the past courts had imposed harsher penalties on specific minorities for the same offense, AI can "learn" that this is the proper legal practice. The algorithm reinforces existing biases and fails to identify new patterns or trends. Discrimination can also occur due to biased sampling or other factors that may result in an unrepresentative dataset. If a model is trained on a dataset that includes only male employees, for example, it will not be able to accurately predict the performance of female employees. Therefore, the chatbot's output may also be biased.

There is also **linguistic bias** concerning the "English-centric" lens through which AI models view the world due to their training data. Since many AI models are built around English-language data, they may not always accurately capture the legal nuances of regions where English is not the primary language (Mukherjee et al., 2026). **Such bias in a legal QA chatbot** is even **more critical** than usual chatbots, as the consequences of a biased response may impact on a person's attitude towards his/her rights and legal situation.

Furthermore, chatbots are usually trained to learn from user reactions (e.g. thumbs up), enabling the AI models to provide more personalized responses that align with individual preferences and beliefs. So, they are often intentionally designed to **prioritize users' satisfaction, reflecting their prejudices back at them and reinforcing their biases**. This approach can unintentionally encourage AI systems to prioritize positive user feedback over providing accurate information (Williams et al., 2024).

Algorithms can also create a "**black box**" problem because they process and optimize multiple variables at once by finding geometric patterns in mathematically defined spaces that humans can't visualize (Bathaee, 2018, p. 897). In conclusion, there is a profound inability to understand the AI output, which brings about the risk of lack of transparency (Edwards & Veale, 2018). At the same time, algorithmic processing can have an air of

rationality, and people may blindly follow them (Borgesius, 2025). The tendency to believe computers or to follow their advice is called "**automation bias**", which typically leads to two types of errors: omission errors (missing events because the system failed to flag them) and commission errors (following incorrect automated advice over one's own judgment) (Romeo & Conti, 2026).

Unauthorized Practice of Law

Most jurisdictions restrict legal advice to licensed practitioners. No person—other than a licensee whose license is not suspended— is permitted to practice law. The delivery of legal services by non-lawyers is known as the unauthorized practice of law. This statutory prohibition was designed to protect the public from entrusting its legal affairs to untrained professionals (Dahan & Liang, 2021). State supreme courts and bar committees then discipline lawyers who fail to adhere to the rules of practice and ethical codes (Remus & Levy, 2016). However, while non-lawyers cannot deliver legal services, they can provide legal information (Dahan & Liang, 2021). Therefore, chatbots must be explicitly presented as **informational tools** rather than as substitutes for professional legal counsel (Surden, 2020). Failure to maintain this distinction could result in violations of legal ethics and professional conduct rules.

CHATBOTS' REGULATORY FRAMEWORK IN AI ACT

The EU AI Act (Regulation (EU) 2024/1689), the fundamental regulation of artificial intelligence in Europe, regulates AI according to a '**risk-based' escalating approach: the greater the risk, the stricter the rules** (Schuett, 2024). Standard conversational chatbots are initially considered to fall within the concept of 'limited risk', which includes AI systems intended for direct interaction with natural persons, while producing synthetic audio, image, video, or text content (Recital 132, AI Act). This includes virtual assistants (**rule-based chatbots**), which answer people's routine or procedural questions such as "*how many days of leave are left*", "*how do I apply for leave*", etc. and have **narrow functionality**. AI systems that fall into this category must be transparent, which means that people must be informed about their interaction with AI (unless this is obvious), important for generative AI systems and their content (Article 50(1,2) AI Act) (Block, 2026).

However, when a chatbot is based on a General Purpose AI (GPAI) model, it is capable of answering a wide range of questions and may influence users' decisions and attitudes concerning their fundamental rights, particularly when the interaction relates to legal matters. This raises the question of whether such functionality could qualify the chatbot as a high-risk AI system. Under the EU AI Act, the answer is generally negative. A chatbot that, for example, explains the scope of criminal rules, assists in drafting an out-of-court document or suggests possible legal options and defense strategies may indeed have a significant impact on a user's decisions regarding the exercise of their rights. However, this fact alone is not sufficient to place the system within the high-risk categories, since such individual use is not listed in Annex

III of the AI Act (Helberger & Diakopoulos, 2023). The chatbot would still be regarded as a limited-risk AI system. In this scenario, the relevant regulatory obligations would primarily arise from the GPAI model on which the chatbot is based (Revolidis, 2024). These obligations—including requirements relating to technical documentation, transparency, compliance with copyright law, and, where applicable, risk assessment and mitigation—are directed primarily at providers and deployers of the GPAI models and not at the providers or deployers of the AI systems built on top of those models.

In our view, this development may undermine the objective of ensuring the ethical, responsible and trustworthy use of AI. By excluding chatbots from the high-risk category, despite their potential impact on the exercise and protection of fundamental rights, deployers may not be subject to regulatory requirements that are designed to safeguard fundamental rights, such as fundamental rights impact assessments, enhanced human oversight, robust data governance, logging and traceability mechanisms. As a result, a potential regulatory gap may emerge in situations where AI systems significantly influence decisions affecting individuals' rights and legal interests (Block, 2026).

FROM TRANSPARENCY TO ETHICS-BY-DEFAULT AND ETHICS-BY-DESIGN

It has become so far evident that the common baseline rule for the individual use of an AI-powered legal information chatbot remains transparency under Article 50(1) AI Act. Products should include in their settings clear, prominent disclaimers stating that the user is interacting with an AI agent, not a person, for example through labels or banners stating, "*You are interacting with an AI system.*" (Felzmann et al., 2020). However, this minimum transparency requirement is insufficient to adequately protect users, who often place undue confidence in AI-generated outputs (Li, 2023) – the phenomenon already aforementioned as *automation bias*. While Article 50(1) of the AI Act requires providers to disclose the AI nature of the interaction, it does not require additional warnings, for example that the tool may provide inaccurate or incomplete or biased information and cannot replace advice or guidance from a licensed professional nor qualify critical decisions on legal matters.

Transparency in chatbot-based interaction should not be treated merely as a formal informative obligation, but as a core value corresponding to the Ethics Guidelines for Trustworthy Artificial Intelligence, developed by the European Parliament's independent High-Level Expert Group on Artificial Intelligence (European Parliament, 2020). Public trust in artificial intelligence can only be ensured through an **ethics-by-default** and **ethics-by-design** regulatory framework that guarantees full compliance with the Treaties, the Charter of Fundamental Rights, and secondary Union law. Within chatbot environments specifically, these ethics, strongly linked to the GDPR "privacy be design", mean that ethical safeguards should be embedded into the architecture and design process of the AI systems from the beginning

(e.g. opt-in rather than opt-out data sharing, minimal or anonymized data collection, clear disclaimers on the tool's limitations) (Brey & Dainow, 2024).

CHATBOT DESIGN SAFEGUARDS AGAINST OVERRELIANCE

The design of the user interface (UI) and the flow of interaction dialog between the general population and legal information chatbots can play a significant role to assuring ethics-by-design and ethics-by-default. This way the interaction with the chatbot can impose certain 'rules', require active user engagement and offer suggestions that communicate risks, potential inaccuracies in answers and reference to sources of information used resulting in a behavioral change on the part of the users. HCI research has consistently revealed that users over-rely on AI recommendations (adhering to a different level of the automation bias discussed above) demonstrating an increased unwarranted trust (Zhang et al., 2020). This overreliance is in fact a cost–benefit response to the low effort of accepting an answer relative to confirming it (Vasconcelos et al., 2023) and can therefore be affected by design. To this end, we propose four safeguards, as depicted in Fig. 1:

1. **Cognitive forcing functions.** The chatbot requires a deliberate step of active user engagement before disclosing the answer, thus reducing overreliance (Buçinca et al., 2021). In Fig. 1(a), the system withholds its conclusion until the user commits to an initial position.
2. **Interaction friction.** The user must take an action (e.g. confirmation or selection) before proceeding to a consequential step, such as to acknowledge that the information will be verified against the governing source or a qualified professional as depicted in Fig. 1(b). Friction should be handled cautiously and be reserved for cases of junctures of legal consequence to avoid purposeless user nuisance.
3. **Scaffolded verification**. A variant of the interaction friction safeguard; the chatbot guides the user through the verification rather than merely obstructing it, so that it remains usable for less expert users. Because interventions demanding effort tend to benefit most those users already high in literacy and motivation (Buçinca et al., 2021), undifferentiated friction risks widening the access gap such chatbots should close. As depicted in Fig. 1(c), the system isolates the single decisive fact, restates the applicable rule in plain language, and offers a low-effort route to free legal assistance, thereby lowering the cost of verifying rather than only raising the cost of accepting.
4. **Calibrated uncertainty and limitation cues.** The chatbot indicates a graded, content-sensitive reliability signal that reveals all substantive warnings: inaccuracy, incompleteness, bias, jurisdictional dependence, and the non-substitutability of professional advice. Since LLMs are systematically overconfident, stating incorrect answers and fabricated citations in fluent language and users overestimate their accuracy (Steyvers et al., 2025), a numeric confidence indicator would easily be considered as a proof of authority; instead, as depicted in Fig. 1(d), claims backed by a credible source are marked and linked, while unattributed assumptions are flagged as unverified.

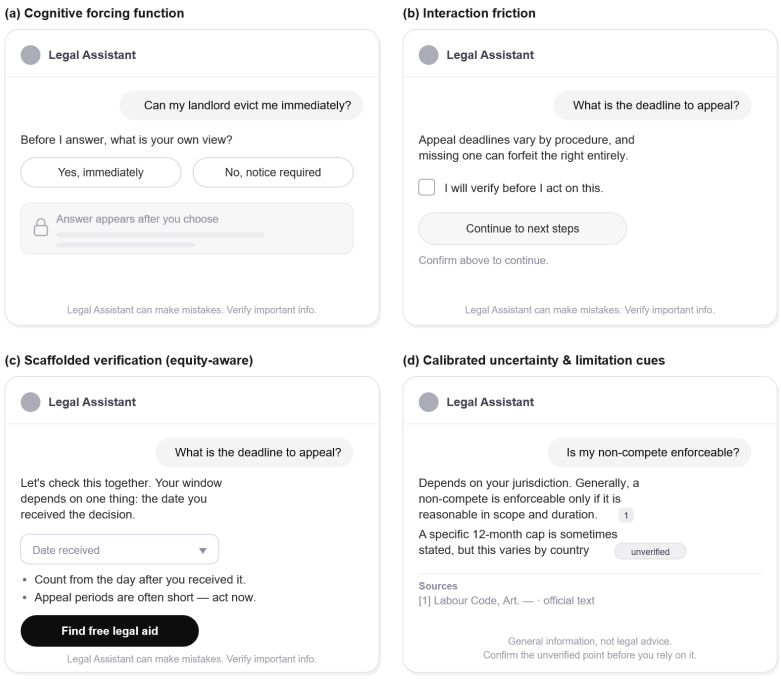


Figure 1: Examples of chatbot interaction using the proposed safeguards.

When built into the chatbot, there safeguards implement ethics-by-design shifting the objective from ensuring the user is informed of the system's nature to supporting verification before reliance. Moreover, when applied on all users as standard chatbot behaviour (and not as an opt-in feature), they also serve the ethics-by-default requirement. And since providers have little market incentive to add such safeguards on their own, these should be required by regulation.

CONCLUSION

GPAI chatbots represent a transformative development in the delivery of legal information. They promote accessibility and cost-efficiency, for meeting basic legal needs such as routine questions, legal document generation, or outcome prediction. These services may help users dissolve or divert disputes before they resort to the traditional adjudication service of the courts, which also benefits the justice system as it saves resources to focus on the most demanding, and complex mental processes. The current structure of the AI Act (Article 6 and Annex III) classifies AI systems as high-risk only when they are deployed in specific sectors, rather than based on their use by individuals, regardless of the likelihood of harm or the severity of the potential consequences. This deprives society of the benefits of a more proportionate, risk-based regulatory framework, one that is better suited to ensuring accuracy, accountability, and compliance with ethical standards. In our view, integrating ethics by design and ethics by default into the implementation process of a chatbot constitutes a preventive, ex ante approach, that is

essential for the responsible deployment of AI. Concretely, we proposed four interaction-level safeguards (cognitive forcing functions, interaction friction, scaffolded verification, and calibrated uncertainty and limitation cues) that operationalize these principles by shifting the design goal from merely informing users that they are interacting with an AI to actively supporting them in verifying its output before relying on it. As future work, we intend to test the proposed safeguards in exemplary chatbot designs through user studies, to evaluate their effectiveness.

ACKNOWLEDGMENT

The authors acknowledge the limited use of ChatGPT (version 4) and Claude (version Opus 5.5) solely for copy-editing purposes, including improvements to grammar, wording, and readability, as well as figure refinements. No generative AI was used for study design, data generation, analysis, interpretation, or the creation of original content.

REFERENCES

Ahmad, T., Maaz, A. S., Murtaza, S., & Jafri, A. (2025). Legal irregularities and bias in Deep Seek and ChatGPT AI: A comparative analysis. In *Proceedings of the International Conference on Law and Technology (ICLT 2025)* (pp. 42–57). Atlantis Press. https://doi.org/10.2991/978-2-38476-515-7_5

Alkhaldi, N. (2024). *What is AI bias really, and how can you combat it?* ITRex. https://itrexgroup.com/blog/ai-bias-definition-types-examples-debiasing-strategies/

Bahanov, B. (2021). Interpretation in EU multilingual law. *European Papers, 7*(1), 465–480. https://doi.org/10.15166/2499-8249/564

Bathaee, Y. (2018). The artificial intelligence black box and the failure of intent and causation. *Harvard Journal of Law & Technology, 31*, 889.

Block, M. J. (2026). Transparency in human–AI interaction: An analysis of Article 50(1) AI Act. *Computer Law & Security Review, 61*, 106327. https://doi.org/10.1016/j.clsr.2026.106327

Borgesius, F. Z. (2025). *Discrimination, artificial intelligence, and algorithmic decision-making* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2510.13465

Brey, P., & Dainow, B. (2024). Ethics by design for artificial intelligence. *AI and Ethics, 4*, 1265–1277. https://doi.org/10.1007/s43681-023-00330-4

Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW1), Article 188. https://doi.org/10.1145/3449287

Burchardt, D. (2019). The relationship between the law of the European Union and the law of its Member States: A norm-based conceptual framework. *European Constitutional Law Review, 15*(1), 73–103. https://doi.org/10.1017/S1574019618000512

Dahan, S., & Liang, D. (2021). The case for AI-powered legal aid. *Queen's Law Journal, 46*(2), 415–429. https://search.informit.org/doi/10.3316/agispt.20211105056491

Edwards, L., & Veale, M. (2018). Enslaving the algorithm: From a "right to an explanation" to a "right to better decisions"? *IEEE Security & Privacy, 16*(3), 46–54. https://doi.org/10.1109/MSP.2018.2701152

European Parliament. (2020). *Resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies* (2020/2012(INL)).

Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamò-Larrieux, A. (2020). Towards transparency by design for artificial intelligence. *Science and Engineering Ethics, 26*, 3333–3361. https://doi.org/10.1007/s11948-020-00276-4

Georgiadou, N., Thanasas, G., Theodoridis, K., & Mandilas, A. (2026). Algorithmic bias in automated decision-making: Hidden risks through the perspective of AI-driven tax compliance. The EU's response from GDPR to the AI Act. In J. P. Vieito et al. (Eds.), *Proceedings of the World Finance Conference*. Springer. https://doi.org/10.1007/978-3-032-19314-8_11

Hasal, M., Nowaková, J., Ahmed Saghair, K., Abdulla, H., Snášel, V., & Ogiela, L. (2021). Chatbots: Security, privacy, data protection, and social aspects. *Concurrency and Computation: Practice and Experience, 33*(19), e6426. https://doi.org/10.1002/cpe.6426

Helberger, N., & Diakopoulos, N. (2023). ChatGPT and the AI Act. *Internet Policy Review, 12*(1). https://doi.org/10.14763/2023.1.1682

Katz, D. M., Bommarito, M. J., Gao, S., & Arredondo, P. (2024). GPT-4 passes the bar exam. *Philosophical Transactions of the Royal Society A, 382*(2270), 20230254. https://doi.org/10.1098/rsta.2023.0254

King, J., Klyman, K., Capstick, E., Saade, T., & Hsieh, V. (2025). User privacy and large language models: An analysis of frontier developers' privacy policies. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 8*, 1465–1477. https://doi.org/10.1609/aies.v8i2.36646

Li, Z. (2023). Why the European AI Act transparency obligation is insufficient. *Nature Machine Intelligence, 5*. https://doi.org/10.1038/s42256-023-00672-y

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), Article 115, 1–35. https://doi.org/10.1145/3457607

Meng, X., & Liu, J. (2025). "Talk to me, I'm secure": Investigating information disclosure to AI chatbots in the context of privacy calculus. *Online Information Review, 49*. https://doi.org/10.1108/OIR-06-2024-0375

Mukherjee, A., Meng, C., & Anastasopoulos, A. (2026). *Lost in the Tower of Babel: The adverse effects of incidental multilingualism in LLMs* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2605.01224

Ogunsan, I., & Johnson, N. (2025). *The use of chatbots in providing free legal guidance: Benefits and limitations* [Unpublished manuscript]. ResearchGate. https://www.researchgate.net/publication/388179232

Pangrazio, L., Godhe, A.-L., & Ledesma, A. G. L. (2020). What is digital literacy? A comparative review of publications across three language contexts. *E-Learning and Digital Media, 17*(6), 442–459. https://doi.org/10.1177/2042753020946291

Pawlik, L. (2025). How the choice of LLM and prompt engineering affects chatbot effectiveness. *Electronics, 14*(5), 888. https://doi.org/10.3390/electronics14050888

Remus, D., & Levy, F. S. (2016). *Can robots be lawyers? Computers, lawyers, and the practice of law* [Preprint]. SSRN. https://doi.org/10.2139/ssrn.2701092

Revolidis, I. (2024). *Regulating general purpose artificial intelligence (GPAI) within the EU AI Act: Challenges and considerations* [Preprint]. SSRN. https://doi.org/10.2139/ssrn.5122935

Romeo, G., & Conti, D. (2026). Exploring automation bias in human–AI collaboration: A review and implications for explainable AI. *AI & Society, 41*, 259–278. https://doi.org/10.1007/s00146-025-02422-7

Ruschemeier, H., & Hondrich, L. (2024). *Automation bias in public administration: An interdisciplinary perspective from law and psychology* [Preprint]. SSRN. https://doi.org/10.2139/ssrn.4736646

Schuett, J. (2024). Risk management in the Artificial Intelligence Act. *European Journal of Risk Regulation, 15*(2), 367–385. https://doi.org/10.1017/err.2023.1

Setyawan, V. (2025). Robot lawyers and the ethical challenges of the legal profession in the era of artificial intelligence. *Balobe Law Journal, 5*(2), 95. https://doi.org/10.47268/balobe.v5i2.3116

Steyvers, M., Tejeda, H., Kumar, A., Belém, C., Karny, S., Hu, X., Mayer, L. W., & Smyth, P. (2025). What large language models know and what people think they know. *Nature Machine Intelligence, 7*(2), 221–231. https://doi.org/10.1038/s42256-024-00976-7

Strunskiy, A. D. (2021). Teleological interpretation in European legal tradition: The historical and comparative perspective. *ARSP: Archiv für Rechts- und Sozialphilosophie, 107*(4), 616–624. https://www.jstor.org/stable/48649833

Surden, H. (2020). Artificial intelligence and law: An overview. *Annual Review of Law and Social Science, 16*, 399–419.

Urbina, F. J. (2024). Reasons for interpretation. *Columbia Law Review, 124*(6), 1661. https://columbialawreview.org/content/reasons-for-interpretation/

Vasconcelos, H., Jörke, M., Grunde-McLaughlin, M., Gerstenberg, T., Bernstein, M. S., & Krishna, R. (2023). Explanations can reduce overreliance on AI systems during decision-making. *Proceedings of the ACM on Human-Computer Interaction, 7*(CSCW1), Article 129. https://doi.org/10.1145/3579605

Williams, M., Carroll, M., Narang, A., Weisser, C., Murphy, B., & Dragan, A. (2024). *Targeted manipulation and deception emerge when optimizing LLMs for user feedback* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2411.02306

Xu, J., Li, Z., & Li, S. (2024). Legal ethics and professional responsibility in the legal profession. *SHS Web of Conferences, 190*, 02006. https://doi.org/10.1051/shsconf/202419002006

Zhang, Y., Liao, Q. V., & Bellamy, R. K. E. (2020). Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 295–305). ACM. https://doi.org/10.1145/3351095.3372852