The Roles AI Is (and Should Be) Playing in Dispute Resolution:

On the Legitimacy of AI Legal Applications

Talia Schwartz-Maor¹ & Orna Rabinovich-Einy^{2†}

I. Introduction

While the transformation of legal processes and roles by new technologies has been the subject of a heated debate for some time,³ the rise of Large Language Models (LLMs) represents a 'game changer' in terms of the capabilities of Artificial Intelligence (AI) in the professional domain.⁴ ChatGPT-4 has demonstrated human-level performance on an array of academic and professional tasks, including the LSAT, the Medical Knowledge Self-Assessment Program, an advanced macroeconomics course designed for high school students, and passing a bar exam.⁵ Other LLMs were quick to follow, including Claude, Gemini, and Perplexity.⁶

¹ Post-doctoral fellow, Dispute Resolution Automation Lab, Faculty of Law, University of Haifa, Haifa, Israel. Orcid ID: https://orcid.org/0009-0008-2376-312X

² Professor, Faculty of Law, University of Haifa, Haifa, Israel. PI, Dispute Resolution Automation Lab at the University of Haifa. Orcid ID: https://orcid.org/0000-0002-9855-0072

[†] This study was supported by the European Union (ERC-2021-COG and 101044195). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

³ While some have long-envisioned dramatic changes to the legal professions (see Richard Susskind and Professions (Oxford Susskind, The Future the University Daniel of Press 2015) http://dx.doi.org/10.1093/oso/9780198713395.001.0001> accessed 3 November 2024: Herbert M Kritzer, 'The Professions Are Dead, Long Live the Professions: Legal Practice in a Postprofessional World' (1999) 33(3) L.& Soc'y Rev. 713 <<u>http://dx.doi.org/10.2307/3115110</u>> accessed 3 November 2024; M Ethan Katsh, Electronic Media and the Transformation of Law (Oxford University Press 1989), others were more skeptical of the idea that machines could substitute for the more challenging aspects of professional work. See Frank Pasquale and L.A. Review of Books, March 15, 2016, 'Book Review: Automating the Professions?' [2016] U of Maryland Legal Studies Research Paper No. 2016-21 <https://ssrn.com/abstract=2775397> accessed 3 November 2024; Dana Remus and Frank S Levy, 'Can Robots Be Lawyers? Computers, Lawyers, and the Practice of Law' [2015] SSRN Electronic Journal <http://dx.doi.org/10.2139/ssrn.2701092> accessed 3 November 2024.

⁴ Rishi Bommasani and others, 'On the Opportunities and Risks of Foundation Models' [2022] <<u>https://doi.org/10.48550/arXiv.2108.07258</u>> accessed 3 November 2024.

⁵ OpenAI and others, 'GPT-4 Technical Report' [2024] <<u>https://doi.org/10.48550/arXiv.2303.08774</u>> accessed 3 November 2024.

⁶ David Brown, '2024 AI Showdown: GPT-4o, Perplexity, Google Gemini, and Claude 3 Compared' (*perplexity*, 24 July 2024) <<u>www.perplexity.ai/page/2024-ai-showdown-gpt-4o-perple-OU.CI7U RxKW9NaE3WmEhQ</u>> accessed 9 November 2024.

Focusing on the legal domain, LLMs have shown competence on a wide range of tasks, including assistance in writing a legal paper,⁷ passing first-year law exams with a C+,⁸ preparing cross-examination questions for a case (performing well, though inferior to lawyers),⁹ drafting a complaint in a securities cryptocurrency class action lawsuit,¹⁰ determining employment status with over 90% accuracy,¹¹ and exhibiting fair or above-average performance in legal reasoning tasks¹². Specifically, one study found that GPT4 performed well, with an average score of 77.32 of 100 combined, on all legal reasoning tasks.¹³ Researchers have found that ChatGPT's strength lies in its performance on a procedural level, specifically its 'ability to understand and reason about users' needs.'¹⁴ LLMs have also been tested for their ability to resolve disputes, with some studies awarding them high ratings for their choice of intervention strategy and message wording as mediators, in comparison with humans.¹⁵ Another recent study, in which researches built an AI-based arbitration model, point to excellent performances.¹⁶ These capabilities are

⁷ Andrew Perlman, 'The Implications of OpenAI's Assistant for Legal Services and Society' [2022] Suffolk University Law School Research Paper No. 22-14 <<u>http://dx.doi.org/10.2139/ssrn.4294197</u>> accessed 3 November 2024.

⁸ Jonathan H Choi and others, 'ChatGPT Goes to Law School' (2023) 71(3) J.Leg.Ed. 387, 387-400 <<u>http://dx.doi.org/10.2139/ssrn.4335905</u>> accessed 3 November 2024.

⁹ Kwan Yuen Iu and Vanessa Man-Yi Wong, 'ChatGPT by OpenAI: The End of Litigation Lawyers?' [2023] SSRN Electronic Journal <<u>http://dx.doi.org/10.2139/ssrn.4339839</u>> accessed 3 November 2024.

¹⁰ Arianna Trozze, Toby Davies and Bennett Kleinberg, 'Large Language Models in Cryptocurrency Securities Cases: Can ChatGPT Replace Lawyers?' [2023] <<u>http://dx.doi.org/10.48550/arXiv.2308.06032</u>> accessed 4 November 2024.

¹¹ Cohen, M. C., Dahan, S., Khern-Am-Nuai, W., Shimao, H., & Touboul, J. (2023). The use of AI in legal systems: determining independent contractor vs. employee status. Artificial intelligence and law, 1-30.

¹² Martin, L., Whitehouse, N., Yiu, S., Catterson, L., & Perera, R. (2024). Better call gpt, comparing large language models against lawyers. arXiv preprint arXiv:2401.16212.

¹³ Neel Guha and others, 'LegalBench: A Collaboratively Built Benchmark for Measuring Legal Reasoning in Large Language Models' [2023] <<u>https://doi.org/10.48550/arXiv.2308.11462</u>> accessed 3 November 2024.

¹⁴ Jinzhe Tan, Hannes Westermann and Karim Benyekhlef, 'ChatGPT as an Artificial Lawyer?', *Workshop* on Artificial Intelligence for Access to Justice <<u>https://ceur-ws.org/Vol-3435/short2.pdf</u>> accessed 3 November 2024.

¹⁵ Jinzhe Tan and others, 'Robots in the Middle: Evaluating LLMs in Dispute Resolution' [2024] <<u>https://doi.org/10.48550/arXiv.2410.07053</u>> accessed 3 November 2024. See also Hannes Westermann, Jaromir Savelka and Karim Benyekhlef, 'LLMediator: GPT-4 Assisted Online Dispute Resolution' [2023] <<u>https://doi.org/10.48550/arXiv.2307.16732</u>> accessed 3 November 2024 (showing promising results for LLM mediator).

¹⁶ Kieffaber, Jack and Gandall, Kimo and McLaren, Kenny, We Built Judge. AI. And You Should Buy It (January 28, 2025). Available at SSRN: https://ssrn.com/abstract=5115184 or

http://dx.doi.org/10.2139/ssrn.5115184 (finding zero hallucinations and zero incomplete responses in an AI arbitration model)

particularly exciting as they bring to the fore what have traditionally been considered human 'soft skills' that have to do with emotional intelligence.

Alongside the excitement and hype, certain studies point to the underperformance of foundation models in the legal domain, including errors in statutory reasoning resulting from imperfect prior knowledge¹⁷ and weak legal reasoning capabilities in certain contexts, such as cryptocurrency securities¹⁸ and taxes.¹⁹ In addition, LLMs have been found to perform poorly in complex legal tasks such as legal research and reasoning in certain models (GPT3.5, PaLM2 and Llama 2)²⁰, or yield outputs that exhibit a lack of concision and citation.²¹ Such findings reinforce concerns regarding AI bias that predated LLMs, and were strongly associated with legal tools such as COMPAS (a recidivism prediction system widely used in the U.S.).²² Hallucinations have also become a concern, with LLMs occasionally producing outputs that are false, with high certainty.²³ At the same time, more recent studies on LLMs' performance in fulfilling complex tasks and making predictions in the legal domain demonstrate more promising results and greater accuracy,²⁴ which could explain why an ever growing number of professionals and lay people are relying on such tools for a wide array of uses.²⁵ In some cases, professionals choose to rely on AI

¹⁷ Andrew Blair-Stanek, Nils Holzenberger and Benjamin Van Durme, 'Can GPT-3 Perform Statutory Reasoning?', *ICAIL 2023: Nineteenth International Conference on Artificial Intelligence and Law* (ACM 2023) <<u>http://dx.doi.org/10.1145/3594536.3595163</u>> accessed 3 November 2024.

¹⁸ Trozze, Davies, and Bennett, Can ChatGPT Replace Lawyers? (n 10).

¹⁹ John Nay and others, 'Large Language Models as Tax Attorneys: A Case Study in Legal Capabilities Emergence' [2023] Minnesota Legal Studies Research Paper No. 23-15, Northwestern Public Law Research Paper No. 23-48 <<u>http://dx.doi.org/10.2139/ssrn.4476325</u>> accessed 3 November 2024 (findings LLMs to perform at high levels of accuracy but not as high as an expert tax lawyer would).

²⁰ Matthew Dahl and others, 'Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models' (2024) 16(1) J. Legal Anal. 64 <<u>http://dx.doi.org/10.1093/jla/laae003</u>> accessed 3 November 2024.
²¹ Bhambhoria, R., Dahan, S., Li, J., & Zhu, X. (2024). Evaluating ai for law: Bridging the gap with open-source solutions. arXiv preprint arXiv:2404.12349.

²² Julia Angwin and others, 'Machine Bias' (*ProPublica*, 23 May 2016) <<u>www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing</u>> accessed 3 November 2024.

²³ Dahl, 'Large Legal Fictions' (n 20).

²⁴ Daniel Martin Katz and others, 'GPT-4 Passes the Bar Exam' (2024) 382 Philosophical Transactions of the Royal Society A <<u>http://dx.doi.org/10.2139/ssrn.4389233</u>> accessed 3 November 2024; Nay and others, 'Large Language Models as Tax Attorneys (n 19). Wang, X., Zhang, X., Hoo, V., Shao, Z., & Zhang, X. (2024). LegalReasoner: A Multi-Stage Framework for Legal Judgment Prediction via Large Language Models and Knowledge Integration; See also Francesco Contini, 'Unboxing Generative AI for the Legal Professions: Functions, Impacts and Governance' (2024) 15(2) International Journal for Court Administration 1 <<u>https://doi.org/10.36745/ijca.604</u>> accessed 3 November 2024 (discussing some of the changes to Gen AI introduced as of 2023 that improved accuracy in domain-specific tasks);

²⁵ 'The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023' (*GOV.UK*, 1 November 2023) <<u>www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-</u>

tools trained, tweaked, and refined for different legal tasks and contexts using various methods, for example, Retrieval-Augmented Generation (RAG) (Wiratunga et al., 2024).

In addition to the use of LLMs and other legal tech products by professionals and lay people, a growing number of AI-based applications are being developed and implemented in the justice system. Uses of AI in these settings are expanding and can include legal information retrieval and assessment tools, case management, legal research, summary and translation, writing applications, document retrieval, analysis and drafting, trial preparation functions, case analytics, negotiation, predictions, and co-pilot tools in judicial decision-making or alternative dispute resolution (ADR) processes.²⁶ While most courts currently shy away from AI-based decision-making²⁷ and are still in the experimentation phase with many of the above-mentioned tools, a host of professional decision-making functions both in public-administrative²⁸ and private settings²⁹ are

declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>

accessed 3 November 2024, acknowledging that 'AI systems are already deployed across many domains of daily life including housing, employment, transport, education, health, accessibility, and justice, and their use is likely to increase'; Kathrin Eidenmueller, Conor McLaughlin and Horst GM Eidenmueller, 'Expanding the Shadow of the Law: Designing Efficient Judicial Dispute Resolution Systems in a Digital World — An Empirical Investigation' (2024) 29(2) Harvard Negotiation Law Review <<u>http://dx.doi.org/10.2139/ssrn.4686785</u>> accessed 3 November 2024 (findings that the usage of AI tools across different functions by UK legal practitioners has more than doubled within the period from 2019/2020 to 2022/2023).

²⁶ See chapter on 'law' and legal applications in Bommasani and others, 'On the Opportunities and Risks of Foundation Models' (n 4); Hadar Yoana Jabotinsky and Michal Lavi, 'AI in the Courtroom: The Boundaries of RoboLawyers and RoboJudges' (forthcoming) <<u>https://ssrn.com/abstract=4883326</u>> accessed 3 November 2024.

²⁷ But see Vikas Mahendra and Arunima Athavale, 'Use and Regulation of AI in Dispute Resolution: Focus on the United Kingdom, Singapore and India' (2024) 18(1) DRI 5 (describing the integration of AI into judiciary in several countries, specifically Singapore); Tania Sourdin, 'Judge v Robot? Artificial Intelligence and Judicial Decision-Making' (2018) 41(4) UNSWLJ 1114 http://dx.doi.org/10.53637/zgux2213 accessed 3 November 2024; 'CEPEJ European Ethical Charter on the Use of Artificial Intelligence (AI) in Judicial Systems and Their Environment - European Commission for the Efficiency of Justice (CEPEJ) www.coe.int' (European Commission for the Efficiency of Justice (CEPEJ)) <www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-injudicial-systems-and-their-environment> accessed 3 November 2024.

²⁸ Finding that nearly half of the studied federal agencies (45%) have experimented with AI and ML in a wide array of use types from enforcement tasks to analysis use-cases, and from provision of services to the public to internal management applications [David Freeman Engstrom and others, 'Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies' [2020] NYU School of Law, Public Law Research Paper No. 20-54 <<u>http://dx.doi.org/10.2139/ssrn.3551505</u>> accessed 3 November 2024].

²⁹ Jack Kelly, 'How Companies Are Hiring and Reportedly Firing With AI' (*Forbes*, 4 November 2023) <<u>www.forbes.com/sites/jackkelly/2023/11/04/how-companies-are-hiring-and-firing-with-ai/</u>> accessed 3 November 2024 (stating a 2023 survey that found 98% of surveyed HR executives said they plan to employ software and algorithms to reduce labor costs); Rachel Curry, 'Microsoft, Amazon Among the Companies Shaping AI-enabled Hiring Policy' (*CNBC*, 11 October 2023) <<u>www.cnbc.com/2023/10/11/microsoft-amazon-among-the-companies-shaping-ai-enabled-hiring-policy.html</u>> accessed 3 November 2024; Vijay

already performed by AI, with varying degrees of human involvement and/or oversight. In addition, there is evidence that some judges are using these tools on their own initiative, with some courts attempting to regulate such use.³⁰ Parties to a legal dispute (as well as their counsel) may also rely on their own AI tools in assessing the strength of their court case, developing their legal strategy, and in their resolution efforts.

The expansion in AI capabilities and uses in the legal domain raises pressing and difficult questions about the potential and the limits of automation.³¹ In this Chapter we look to the perceptions of lay people and professionals on these questions as a significant prism for guiding societal decisions on the desirability of AI use in the legal and dispute resolution settings. Such perceptions also constitute important means for assessing the legitimacy-related implications of the adoption of various AI-based legal tools and functions. We, therefore, begin with a review of legal AI tools and functions in Part II of this Chapter. In Part III, we draw on empirical research examining the perceptions of lay people and professionals regarding the legitimacy of AI use by courts and in alternative dispute resolution processes. Part IV offers concluding thoughts stemming from the juxtaposition of the review of AI capabilities and existing empirical research.

II. AI Legal Tools and Functions: An Overview

Two key developments lie at the heart of the current trend of AI-based legal tools: the evolution of the 'fourth party' as part of the rise of online dispute resolution (ODR) on

³⁰ Ralph Losey, 'Eleventh Circuit Judge Admits to Using ChatGPT to Help Decide a Case and Urges Other Judges and Lawyers to Follow Suit' (*JDSUPRA*, June 7, 2024)

https://www.jdsupra.com/legalnews/eleventh-circuit-judge-admits-to-using-1737989/ accessed January 13, 2025. Jennifer Kay. Delaware Judges, Law Clerks Cleared to Use Generative AI Tools. (October 22, 2024) <<u>https://news.bloomberglaw.com/artificial-intelligence/delaware-judges-law-clerks-cleared-to-use-generative-ai-tools</u>> accessed 1 February 2024

Pereira and others, 'A Systematic Literature Review on the Impact of Artificial Intelligence on Workplace Outcomes: A Multi-Process Perspective' [2021] Human Resource Management Review 100857 <<u>http://dx.doi.org/10.1016/j.hrmr.2021.100857</u>> accessed 3 November 2024.

³¹ Amy J Schmitz, 'Chapter 1. Responsible Use of AI in Civil Dispute Resolution' [2024] Ohio State Legal Studies Research Paper No. 870 <<u>https://dx.doi.org/10.2139/ssrn.4903238</u>> accessed 4 November 2024 (elaborating to the benefits and risks of AI in dispute resolution).

the one hand, ³² and the growth of a legal-tech industry on the other. ³³ ODR emerged in the mid-1990s in the e-commerce setting with the spread of internet communication. As remote transactions proliferated, it soon became clear that there was a need for online avenues of dispute resolution, with both courts and ADR avenues typically being unavailable or impractical for distant parties involved in small scale disputes.³⁴ eBay quickly became a pioneer, developing a first of its kind ODR system, involving a tiered process that made successful use of automation, offering disgruntled buyers and sellers an avenue of redress.³⁵ The role of technology in the resolution of disputes was termed by Katsh & Rifkin as the 'fourth party,' either supporting or supplanting the human 'third party.'³⁶ In the decades that followed, ODR processes expanded from private settings to courts and tribunals, and from online small scale disputes to a wide variety of claims, ranging from small claims and traffic cases to family and eviction claims.³⁷

In the years post-COVID-19, various mediums in- and outside the courts became increasingly common in addressing claims and disputes, ranging from asynchronous text-based processes to video and hybrid ones, and from processes involving human facilitators and decision-makers to automated resolution avenues.³⁸ With the advancements in AI, it was only a matter of time before such functions would be incorporated into an online algorithm-driven process.³⁹

³² Ethan Katsh and Janet Rifkin, Online Dispute Resolution: Resolving Conflicts in Cyberspace (Jossey-Bass 2001).

³³ David Freeman Engstrom and Jonah B Gelbach, 'Legal Tech, Civil Procedure, and the Future of Adversarialism' (forthcoming) <<u>https://ssrn.com/abstract=3551589</u>> accessed 4 November 2024; Richard Susskind, *Tomorrow's Lawyers: An Introduction to Your Future* (Oxford University Press 2023) 136; Robert Dale, 'Law and Word Order: NLP in Legal Tech' (2018) 25(1) Natural Language Engineering 211, 211-217 <<u>http://dx.doi.org/10.1017/s1351324918000475</u>> accessed 4 November 2024; Qian Hongdao and others, 'Legal Technologies in Action: The Future of the Legal Market in Light of Disruptive Innovations' (2019) 11(4) Sustainability 1015 <<u>http://dx.doi.org/10.3390/su11041015</u>> accessed 4 November 2024.

³⁴ Ethan Katsh and Orna Rabinovich-Einy, *Digital Justice: Technology and the Internet of Disputes* (Oxford University Press, Incorporated 2017) 35-36.

³⁵ ibid.

³⁶ Ethan Katsh and Janet Rifkin, *Online Dispute Resolution: Resolving Conflicts in Cyberspace* (John Wiley & Sons, Inc. 2001) 93.

³⁷ Orna Rabinovich-Einy, 'The Past, Present, and Future of Online Dispute Resolution' (2021) 74(1) Current Legal Problems 125, 125-148 <<u>http://dx.doi.org/10.1093/clp/cuab004</u>> accessed 4 November 2024.

³⁸ Orna Rabisnovich-Einy, 'Process Pluralism in the Post-Covid Dispute Resolution Landscape' (2022) 10(1) Texas A&M Law Review 55 <<u>http://dx.doi.org/10.37419/lr.v10.i1.4</u>> accessed 4 November 2024.

³⁹ See for example Amica. offering AI-supported negotiation for divorce cases in Austrlia: 'Amica - The Simple, Low Cost, Smart Way to Separate or Divorce Online' (*amica*) <<u>https://amica.gov.au/</u>> accessed 4 November 2024.

Another key development was the growth of the legal tech market, rooted in the spread of digital communication and the internet. These changes that took place towards the end of the 20th century were seen as both an opportunity and a threat to legal (and other) professionals.⁴⁰ As access barriers to professional information were lifted with the growth of the internet and the advent of search engines, lay people were exposed to knowledge that was previously available only through professional publications and required specialized expertise.⁴¹ In addition to access to legal information, tools were later developed to assist in the generation of legal documents, such as contracts and wills, and in the filing of claims.⁴² Similarly, AI-based tools for professional use by lawyers for contract drafting and review, due diligence, discovery, and more soon became available. Recently, law firms have turned to RAG systems, described above, in hope for enhanced performance and fewer errors and hallucinations.⁴³ The new capacities offered by these various AI tools have helped lawyers increase their efficiency and workload capacity, but also created pressures for reduced costs,⁴⁴ and strengthened the notion that, in certain cases, litigants could now manage some functions on their own.⁴⁵ While some highlighted that in the wake of these developments lawyers would need to transform their practice,⁴⁶ the

⁴⁰ Katsh, *Electronic Media and the Transformation of Law* (n 3) 224-226 (using the term 'Trojan horse' in this context). Orna Rabinovich-Einy and Ethan Katsh, 'The New New Courts' (2017) 67 Amer. U. L. Rev. 165, 213 <<u>https://ssrn.com/abstract=3508460</u>> accessed 4 November 2024.

⁴¹ Katsh, *Electronic Media and the Transformation of Law* (n 3) 224-226.

⁴² Engstrom and Gelbach, 'Legal Tech, Civil Procedure, and the Future of Adversarialism' (n 33) (referring to Legalmation, RocketLawyer, CaseText and Lawyaw in Document Assembly and Creation category).

⁴³ There is supporting evidence that RAG outperforms GenAI. Lakatos, R., Pollner, P., Hajdu, A., & Joo, T. (2024). Investigating the performance of Retrieval-Augmented Generation and fine-tuning for the development of AI-driven knowledge-based systems. arXiv preprint arXiv:2403.09727. However, RAG models have been also shown to suffer from errors and hallucinations. See Magesh, V., Surani, F., Dahl, M., Suzgun, M., Manning, C. D., & Ho, D. E. (2024). Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools. arXiv preprint arXiv:2405.20362 (finding Lexis +ai and westlaw ai to hallucinate at considerable rates).

⁴⁴ John O McGinnis and Russell G Pearce, 'The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services' (2014) 82(3041) Fordham Law Review <<u>https://ssrn.com/abstract=2436937</u>> accessed 4 November 2024 (discussing the effect of machine-led legal services on making services cost-effective as well as putting pressure on experts to cut costs).

⁴⁵ Amy J Schmitz and John Zeleznikow, 'Intelligent Legal Tech to Empower Self-Represented Litigants' (2022) 23 Columbia Science and Technology Law Review 142 <<u>http://dx.doi.org/10.2139/ssrn.4048335</u>> accessed 4 November 2024; JJ Prescott, 'Using ODR Platforms to Level the Playing Field', *Legal Tech and the Future of Civil Justice* (Cambridge University Press 2023) 286-304 <<u>http://dx.doi.org/10.1017/9781009255301.016</u>> accessed 4 November 2024.

⁴⁶ Susskind, *Tomorrow's Lawyers* (n 33) 136. See also Callister, Paul D., Generative AI and Finding the Law (December 8, 2023). 117 Law Library Journal 1 (2025).

profession has often resisted such calls on the grounds of 'unauthorized practice of law' related to the introduction of legal technologies.⁴⁷

Today, we find a myriad of AI-based legal tools and functions being developed for the public sector⁴⁸ (e.g., courts and tribunals), private settings (e.g., ADR service providers and law firms), and individual users – both professionals (e.g., lawyers, mediators, and arbitrators)⁴⁹ and lay people (typically those seeking redress or a defense).⁵⁰ Some of these tools have to do with the technical-procedural elements, such as case management,⁵¹ or document translation and transcription.⁵² Other functions have to do with document summaries or drafting capabilities.⁵³ Some tools offer legal research and analysis capabilities, such as jury evaluation⁵⁴ and displaying relevant past decisions.⁵⁵ Certain AI tools help parties explore their interests and options, either in the form of a tool guiding them in assessing their legal issue and strategies for addressing it, or as part of an ADR process,⁵⁶ and other tools offer outcome predictions or recommendations to parties, lawyers, or third parties (e.g., judges, arbitrators, or mediators).⁵⁷

Over the years, scholars and practitioners have offered various categorizations of AI usage in courts and ADR, distinguishing tools according to the scope of human

⁴⁷ See for example FTC's action and settlement in the matter of DoNotPay, a platform that offers among other things the filing of legal claims ('DoNotPay' (*Federal Trade Commission*) <<u>www.ftc.gov/legal-library/browse/cases-proceedings/donotpay</u>> accessed 4 November 2024).

⁴⁸ David Freeman Engstrom, *Legal Tech and the Future of Civil Justice* (Cambridge University Press 2023); Francesco Contini, 'Unboxing Generative AI for the Legal Professions: Functions, Impacts and Governance' (2024) 15(2) International Journal for Court Administration 1 <<u>https://doi.org/10.36745/ijca.604</u>> accessed 3 November 2024 (a possible reference to Chapter 22 Vermeys).

⁴⁹ Darla Wynon Kite-Jackson, '2023 Artificial Intelligence (AI) TechReport' (*ABA*, 15 January 2024) <<u>www.americanbar.org/groups/law_practice/resources/tech-report/2023/2023-artificial-intelligence-ai-</u>

<u>techreport/</u>> accessed 4 November 2024; Contini, Unboxing Generative AI for the Legal Professions (n 39); Katie Wolf, 'Current Use of Large Language Models (LLMs) in Law' (*Filevine*, 11 July 2024) <<u>www.filevine.com/blog/current-use-of-large-language-models-llms-in-law/</u>> accessed 4 November 2024.

⁵⁰ Tina Seabrooke and others, 'A Survey of Lay People's Willingness to Generate Legal Advice Using Large Language Models (LLMs)', *TAS '24: Second International Symposium on Trustworthy Autonomous Systems* (ACM 2024) <<u>http://dx.doi.org/10.1145/3686038.3686043</u>> accessed 4 November 2024. ⁵¹ Samuel Sahagian, 'ILawyer' (2021) 3 UCLA J. L. & TECH. DIGEST 24.

⁵² Jakub Harasta, Tereza Novotná and Jaromir Savelka, 'It Cannot Be Right if It Was Written by AI: On Lawyers' Preferences of Documents Perceived as Authored by an LLM vs a Human' [2024] 6 and references there <<u>https://arxiv.org/abs/2407.06798</u>> accessed 4 November 2024.

⁵³ ibid, 5-6 and references there.

⁵⁴ Sahagian, 'ILawyer' (n 51), 15-16.

⁵⁵ ibid, 10-12 on legal research.

⁵⁶ ibid, 20-24 on Legal Advice.

⁵⁷ ibid, 14-15 on predicting case outcomes. (a possible reference to Chapter 23 Giacalone).

involvement,⁵⁸ the stage of the process into which AI is introduced,⁵⁹ the types of legal functions performed,⁶⁰ as well as what's at stake (simple versus complex, technical versus emotional).⁶¹ Our focus is on the nature of the capabilities – those drawing on traditional AI-related strengths, such as efficiency and capacity, as opposed to those relating to 'soft skills,' shaping the quality of communication with disputants or their procedural experiences. Such 'soft skill' capabilities (such as empathy and creativity) have been seen, until not long ago, as the sole purview of human communication. AI was initially seen as a means of streamlining and automating high volumes of cases in an efficient and accurate manner by handling mass, simple, easily categorized disputes. We are now moving towards the handling of more complex, high-stake, sensitive, and emotional conflicts, with AI performing in ways that require more sophisticated language and emotional capabilities. This shift towards a more sophisticated use of AI in the legal domain raises a host of legal and normative questions: Can machines satisfy the needs created in these more complex arenas? Are people willing to have algorithms resolve such disputes? Do the more advanced machine capabilities we are now seeing in LLMs satisfy human expectations in procedural terms? Will people feel heard by such machines, or that they have been treated with respect and dignity? Will the outcomes of such processes seem fair and acceptable?

In the following section we draw on existing empirical research on perceptions of fairness and legitimacy of AI in courts and ADR, generating insights on the implications and limitations of introducing a growing array of AI tools. As we will see, generally, the research shows that AI is preferred based on qualities such as accuracy, while humans are valued for their discretion and empathy.⁶² Much of the research predates the widespread launch of advanced LLMs, leading to the questions whether resistance to the use of AI in

⁵⁸ Leah Wing and Chris Draper, 'Parameters of Online Dispute Resolution: Introducing a New Framework for ODR' (2022) 9(2) International Journal of Online Dispute Resolution 112 <<u>https://ssrn.com/abstract=4673536</u>> accessed 5 November 2024.

⁵⁹ Bommasani and others, 'On the Opportunities and Risks of Foundation Models' (n 4).

⁶⁰ Sahagian, 'ILawyer' (n 51).

⁶¹ Jabotinsky and Lavi, 'AI in the Courtroom: The Boundaries of RoboLawyers and RoboJudges' (n 26).

⁶² Outside the legal context, see Yochanan E Bigman and Kurt Gray, 'People Are Averse to Machines Making Moral Decisions' (2018) 181 Cognition 21 <<u>http://dx.doi.org/10.1016/j.cognition.2018.08.003</u>> accessed 5 November 2024 (finding overall machine aversion, i.e., preference towards a human agent, due to perceived lack of moral competence in machines, across decision domains).

certain domains will persist and whether the considerations that drive preference for AI over human decision-making will persevere.

III. Select Research Findings about AI and Legitimacy: An Empirical Perspective

In recent years, there has been a growing understanding that trust in and perceived legitimacy of AI are important measures of AI fairness and acceptability. Therefore, alongside the significance of the 'objective' evaluation of algorithms' performance, a substantial number of studies sought to capture how users of such systems (potential or actual, lay or professional) and the general public view the effectiveness, fairness, and desirability of AI use in various domains, including law and dispute resolution.⁶³ While many studies focus on using AI in various spheres of life, only a small portion relate to legal and dispute resolution settings.

Despite variations in research methods and in the types of perceptions they capture, and the often-vague description of AI tools provided to study participants, studies examining the fairness and legitimacy of AI in the legal and dispute resolution arena offer important insights into the role AI can and should play in these domains. An important lesson from these studies is that while many of them uncover algorithmic aversion, meaning that people perceive human decision-making as preferable and more fair or legitimate than that of AI,⁶⁴ a nuanced analysis shows that under certain conditions AI may

⁶³ Gefen, Karahanna and Straub, 'Trust and TAM in Online Shopping: An Integrated Model' (2003) 27(1) MIS Quarterly 51 <<u>http://dx.doi.org/10.2307/30036519</u>> accessed 5 November 2024 (defining trust as a key predictor of intendent use of technology in e-commerce settings); Ella Glikson and Anita Williams Woolley, 'Human Trust in Artificial Intelligence: Review of Empirical Research' (2020) 14(2) Academy of Management Annals 627 <<u>http://dx.doi.org/10.5465/annals.2018.0057</u>> accessed 5 November 2024 (meta-analysis on trust in AI across use-cases and domains); Christopher Starke and others, 'Fairness Perceptions of Algorithmic Decision-Making: A Systematic Review of the Empirical Literature' (2022) 9(2) Big Data & Society 205395172211151 <<u>http://dx.doi.org/10.1177/20539517221115189</u>> accessed 5 November 2024 (meta-analysis on perceptions of algorithms' fairness across decision-domains); Matthew Katsaros and others, 'Procedural Justice and Self Governance on Twitter' (2022) 1(3) Journal of Online Trust and Safety <<u>http://dx.doi.org/10.54501/jots.v1i3.38</u>> accessed 5 November 2024 (on the applicability of procedural justice to online realms); Morin-Martel, A. (2024). Machine learning in bail decisions and judges' trustworthiness. Ai & Society, 39(4), 2033-2044.

⁶⁴ See for example AJ Wang, 'Procedural Justice and Risk-Assessment Algorithms' [2018] SSRN Electronic Journal <<u>http://dx.doi.org/10.2139/ssrn.3170136</u>> accessed 9 November 2024 (where algorithms received the lowest score on preference, fairness and legitimacy compared with human decision-maker and guidelines-based risk assessment in bail hearings).

not only be as acceptable as human third parties,⁶⁵ but even be preferable to them (algorithm appreciation).⁶⁶

While various studies have offered explanations for factors that give rise to algorithm appreciation (or aversion), the overarching conclusion emerging from literature is that preferences for AI in the legal domain are highly contextual. Such explanations vary and include factors such as the domain in which AI is applied (e.g., criminal sentencing versus natural disaster planning),⁶⁷ the decision's complexity (e.g., monetary versus mental issues in a divorce case),⁶⁸ outcome control and decision-making authority (e.g., mediation versus arbitration, or having 'a human in the loop'),⁶⁹ or the stage of the legal process in which the algorithm is used (with AI being perceived more positively than human decision-making in earlier stages, a finding that is reversed in later stages).⁷⁰

⁶⁵ Ryan P Kennedy, Philip D Waggoner and Matthew Ward, 'Trust in Public Policy Algorithms' (2021) 84(2) The Journal of Politics 1132, 1132-1147 <<u>http://dx.doi.org/10.1086/716283</u>> accessed 5 November 2024.; Ric Simmons, 'Big Data, Machine Judges, and the Legitimacy of the Criminal Justice System' (2018) 52 University of California Davis Law Review 1067, 1067-1118 <<u>http://dx.doi.org/10.2139/ssrn.3156510</u>> accessed 5 November 2024.

⁶⁶ Theo Araujo and others, 'In AI We Trust? Perceptions About Automated Decision-Making by Artificial Intelligence' (2020) 35(3) AI & SOCIETY 611, 611-623 <<u>http://dx.doi.org/10.1007/s00146-019-00931-w</u>> accessed 5 November 2024; Eike Schneiders and others, 'Objection Overruled! Lay People Can Distinguish Large Language Models From Lawyers, but Still Favour Advice From an LLM' [2024] <<u>http://dx.doi.org/10.48550/arXiv.2409.07871</u>> accessed 5 November 2024.

⁶⁷ Michael C Horowitz and Lauren Kahn, 'What influences attitudes about artificial intelligence adoption: Evidence from U.S. local officials' (2021) 16(10) PLOS ONE e0257732 <<u>http://dx.doi.org/10.1371/journal.pone.0257732</u>> accessed 9 November 2024.

⁶⁸ Gizem Yalcin and others, 'Perceptions of Justice by Algorithms' (2023) 31 Artificial Intelligence and Law 269 <<u>http://dx.doi.org/10.1007/s10506-022-09312-z</u>> accessed 5 November 2024.

⁶⁹ Anthony Joseph Casey and Anthony Niblett, 'Will Robot Judges Change Litigation and Settlement Outcomes? A First Look at the Algorithmic Replication of Prior Cases' [2020] SSRN Electronic Journal <<u>http://dx.doi.org/10.2139/ssrn.3633037</u>> accessed 5 November 2024; Eidenmueller, Conor McLaughlin and Horst GM Eidenmueller, 'Expanding the Shadow of the Law (n 25) 31-32 (pointing to overall preference for online systems among dispute resolution professionals, while rejecting AI for decision-making: 'The following three proposals received the strongest support' (mean value): 'online platform for communication with the court and other parties' (3.94), 'online filing and case management' (3.89), and 'greater specialisation / subject matter expertise of judges' (3.84). In contrast, respondents (weakly) oppose the 'use of artificial intelligence to assist case handling and decision-making' (2.18); Ayelet Sela, 'Can Computers Be Fair? How Automated and Human-Powered Online Dispute Resolution Affect Procedural Justice in Mediation and Arbitration' (2018) 33 Ohio State Journal on Dispute Resolution 91 <<ht>https://ssrn.com/abstract=3074311> accessed 5 November 2024.</tt>

⁷⁰ Dovilė Barysė and Roee Sarel, 'Algorithms in the Court: Does It Matter Which Part of the Judicial Decision-Making Is Automated?' (2024) 32(1) Artificial Intelligence and Law 117 <<u>http://dx.doi.org/10.1007/s10506-022-09343-6</u> accessed 5 November 2024; Andreia Martinho, 'Surveying Judges About Artificial Intelligence: Profession, Judicial Adjudication, and Legal Principles' [2024] AI & SOCIETY <<u>http://dx.doi.org/10.1007/s00146-024-01869-4</u> accessed 5 November 2024.

A significant number of studies find that algorithms are preferred in the legal context for their consistency,⁷¹ accuracy,⁷² ability to perform complex calculations,⁷³ and impartiality.⁷⁴ One study found that AI's perceived accuracy was the main reason for algorithmic appreciation (and not human-related factors, such as 'voice,' the opportunity to tell one's side of the story).⁷⁵ Another study found that the vast majority (87.7%) of respondents (including those with a human preference) thought that introducing machinelearning would enhance legal outcomes' determinacy, contributing to greater certainty in judicial decisions.⁷⁶

Some studies find that humans (as opposed to AI) are appreciated for their 'soft skills,' such as empathy and providing an opportunity for voice,⁷⁷ and that the preference for human versus AI agents is grounded in ethical considerations, such as warmth, competence,⁷⁸ and voice.⁷⁹ One way of explaining such findings is that humans are perceived as superior to AI in certain 'emotional' qualities that are deemed essential to decision-making, such as reasoning and empathy.

⁷¹ Benjamin Minhao Chen and Zhiyu Li, 'How Will Technology Change the Face of Chinese Justice?' (2020) 34 Columbia Journal of Asian Law 1 < https://doi.org/10.7916/cjal.v34i1.7484 > accessed 5 November 2024. ⁷² Wang, 'Procedural Justice and Risk-Assessment Algorithms' (n 64).

⁷³ Glikson and Williams Woolley, 'Human Trust in Artificial Intelligence' (n 62) 627-660.

⁷⁴ Minhao Chen and Li, 'How Will Technology Change the Face of Chinese Justice?' (n 71).

⁷⁵ Benjamin Minhao Chen, Alexander Stremitzer and Kevin Tobia, 'Having Your Day in Robot Court' (2022) 36 Harvard Journal of Law & Technology 127 <<u>https://ssrn.com/abstract=3841534</u>> accessed 5 November 2024 (finding in mediation analysis that reduction in fairness ratings that came from having an algorithm rather than a human is most significantly attributed to accuracy (29%) vs. only 2% coming from voice.) ⁷⁶ Minhao Chen and Li, 'How Will Technology Change the Face of Chinese Justice?' (n 71).

⁷⁷ Eiichiro Watamura and others, 'Empathetic Robot Judge, We Trust You' (2023) 40(18) International Journal of Human-Computer Interaction 1, 1-10 < http://dx.doi.org/10.1080/10447318.2023.2232982> accessed 5 November 2024.

⁷⁸ Xusen Cheng and others, 'Human vs. AI: Understanding the Impact of Anthropomorphism on Consumer Response to Chatbots From the Perspective of Trust and Relationship Norms' (2022) 59(3) Information Processing & Management 102940 http://dx.doi.org/10.1016/j.jpm.2022.102940> accessed 5 November 2024 (examining the impact of Anthromorphism on consumer's trust in chatbots in e-commerce. Authors find that when consumers consider chatbots as having higher warmth or higher competence – they tend to trust them more).

⁷⁹ Reuben Binns, 'Fairness in Machine Learning: Lessons From Political Philosophy' (2021) 81 Journal of Machine Learning Research 1, 1-11 < https://doi.org/10.48550/arXiv.1712.03586 accessed 5 November 2024 (finding that humans are perceived as allowing more voice).

One study found that algorithmic aversion was anchored in concerns over lack of deliberation, voice, and procedural fairness.⁸⁰ Another found perceived empathy to be positively associated with trust in a judge (both human and AI), which in turn predicts positive judgment evaluation and overall acceptance of the judgment.⁸¹ That study further found that a human judge is perceived as significantly more empathetic than an 'AI-judge.' Much in the same vein, a study researching perceptions of human versus 'robot' lawyers found that the latter were perceived as not being able to fulfill tasks requiring empathy or operate with creativity, nor being able to engage in psychological warfare.⁸² Alternatively, the study showed that 'robot lawyers' were perceived as having the ability to collect and retrieve data and to analyze and predict cases.

As we can see, most of the features driving positive perceptions towards AI in these studies are the types that one would typically associate with AI strengths, such as accuracy or predictability. However, with advancements in LLM capabilities we may find that some of the distinctions between human and AI fortes are becoming less stark.

A recent study demonstrated LLMs' ability to mediate between people with divergent opinions on divisive issues and generate a consensus.⁸³ In that study, AI-generated statements reflecting common ground on divisive issues among group members were consistently preferred by participants over ones prepared by trained humans and were deemed by external experts as being comparatively more clear, informative, and fair. Most striking was the finding that the machine not only produced consensual statements that were preferred by group members over those prepared by humans, but was also successful

⁸⁰ Anna Fine, Stephanie Le and Monica K Miller, 'Content Analysis of Judges' Sentiments Toward Artificial Intelligence Risk Assessment Tools' (2023) 24(2) Journal of Criminology, Criminal Justice, Law & Society 31, 31-46 <<u>http://dx.doi.org/10.54555/ccjls.8169.84869</u>> accessed 5 November 2024.

⁸¹ Watamura and others, 'Empathetic Robot Judge, We Trust You' (n 77); Sela, 'Can Computers Be Fair?' (n 69).

⁸² Ni Xu and Kung-Jeng Wang, 'Adopting Robot Lawyer? The Extending Artificial Intelligence Robot Lawyer Technology Acceptance Model for Legal Industry by an Exploratory Study' (2019) 27(5) Journal of Management & Organization 867, 867-885 <<u>http://dx.doi.org/10.1017/jmo.2018.81</u>> accessed 5 November 2024.

⁸³ Michael Henry Tessler and others, 'AI Can Help Humans Find Common Ground in Democratic Deliberation' (2024) 386(6719) Science <<u>http://dx.doi.org/10.1126/science.adq2852</u>> accessed 5 November 2024.

in moving participants' views closer while showcasing minority opinions and providing them with the space in which they could sway others in their direction.

Along similar lines, a study comparing LLMs and human mediators found the LLM outperformed human mediators. Specifically, the LLM's choice of intervention strategies was rated equal or superior to human-selected interventions in 62% of cases. Moreover, in 84% of cases the messages crafted by the LLM were judged as better, or at least on par, with those written by human mediators (evaluation was guided by rubric items including understanding and contextualization, neutrality and impartiality, empathy awareness, and resolution quality).⁸⁴ While the evaluators in the study were not expert mediators, and therefore relied on comprehension and tone-assessment rather than professional criteria, the findings are striking. Another study explored the use of LLMs in the context of ODR, employing an experimental platform (LLMediator). Researchers demonstrated certain facilitative capabilities, including the LLMediator's ability to reformulate inflammatory messages to a more neutral tone, assist mediators by providing suggested interventions, and autonomously intervene in the negotiation process with the aim of guiding the parties toward an amicable resolution.⁸⁵

Clearly, with the advancement in LLM capabilities and the increased creativity they have garnered, these tools seem to be much more capable of building rapport, and conveying empathy and support. Indeed, it may very well be that LLMs' hallucinations are less of a 'bug' and more of a 'feature,' a cost one needs to pay if they are to enjoy these other extraordinary abilities.⁸⁶ Aside from the expected technological changes in upcoming years (model upgrades and technological advancements), habits and social norms⁸⁷

⁸⁴ Jinzhe, 'Robots in the Middle: Evaluating LLMs in Dispute Resolution' (n 15)

⁸⁵ Westermann 'LLMediator: GPT-4 Assisted Online Dispute Resolution' (n 15)

⁸⁶ Minhyeok Lee, 'A Mathematical Investigation of Hallucination and Creativity in GPT Models' (2023) 11(10) Mathematics 2320 <<u>http://dx.doi.org/10.3390/math11102320</u>> accessed 5 November 2024; Tanmoy Chakraborty and Sarah Masud, 'The Promethean Dilemma of AI at the Intersection of Hallucination and Creativity' [2024] Communications of the ACM <<u>http://dx.doi.org/10.1145/3652102</u>> accessed 5 November 2024.

⁸⁷ Schneiders and others, 'Objection Overruled!' (n 66) (Authors suggest that algorithmic aversion is shaped by social norms, or the expectation to trust human more that AI. Authors conclusion results from an interesting pattern in findings in a series of experiments on willingness to accept legal advice depending on its source: when the source of the advice was unknown to participants, they were more willing to act on the LLM-generated advice than the lawyer-generated advice. However, when source of advice was known, i.e., respondents were explicitly told it was a human lawyer, higher reliance on human was observed).

surrounding AI services are expected to transform. As we study this moving target, it remains to be seen how such changes will shape perceptions of human and AI capacities in the legal domain, as well as the underlying reasons driving people's preferences for human versus automated legal decision-making.

IV. Concluding Thoughts

With the dramatic leap in LLM capabilities, we are witnessing legal AI functions improve and expand from the technological automation domain to new frontiers. Research findings, perhaps due to the inherent gap between the pace of technology and the research publication process, demonstrate an overall algorithm aversion for certain legal functions, but an appreciation for algorithmic performance in particular contexts, frequently where traditional algorithmic qualities and strengths are most pronounced.

Will these findings persist despite the profound changes we are witnessing in the capabilities of LLMs? How will perceptions of AI shift now that AI capabilities in the legal arena expand beyond more traditional functions to include enhanced communication skills and emotional intelligence-like features that have long been considered hallmarks of the human skillset? Finally, even if algorithms *can* perform as well as humans in what have traditionally been areas of human strength, *should* we, as a society, fully delegate such functions to AI from a moral standpoint? The concern underlying this question has less to do with the outcomes that will be reached through such AI-led processes and more to do with what it means for us as individuals and as a society to undergo legal and dispute resolution processes without leaving some space in which we interact with human 'third parties.'