Legal Expertise Meets Artificial Intelligence: A Critical Analysis of Large Language Models as Intelligent Assistance Technology

Frank Schilder[1]

[1]Thomson Reuters, TR Labs, 610 Opperman Drive, St. Paul, MN, 55123, USA

Abstract

This talk investigates an intelligent assistance (IA) approach to utilizing Large Language Models (LLMs) in the legal domain by addressing the risks associated with unchecked artificial intelligence (AI) applications. We emphasize the importance of understanding the distinctions between AI and IA, with the latter involving human-in-the-loop decision-making processes, which can help mitigate risks and ensure responsible use of this rapidly developing technology.

Using ChatGPT and GPT-4 as a prime example, we demonstrate its dual role as both an AI and IA application, showcasing its versatility in a variety of legal tasks. We look at recently reported explorations in particular in using very LLMs in addressing tasks such as multiple-choice question answering, legal reasoning, case outcome prediction, and summarization. We argue that to fully achieve "augmented intelligence," a reasoning and knowledge base component is required, allowing IA systems to effectively support human users in decision-making processes.

Keywords

artificial intelligence, augmented intelligence, intelligent assistance, large language models, chatGPT, GPT-4

1. Introduction

This extended abstract presents a comprehensive critique and analysis of the application of Large Language Models (LLMs), with a focus on GPT-4 and ChatGPT, within the legal sector. The talk unpacks the concept of Intelligent Assistance (IA), differentiating it from artificial intelligence (AI), and underscores its value within the context of human-in-the-loop decision-making, particularly in the legal domain. By doing so, we delve into the significant benefits and potential pitfalls associated with the unchecked use of these technological innovations in the legal sector.

The presentation initially emphasizes the critical distinctions between AI and IA. While both possess their strengths and unique features, IA is proposed as a more ethically responsible and practical solution in the legal sector due to its requirement for human involvement in the decision-making process. The key argument lies in the fact that IA, in contrast to pure AI, has a better potential to mitigate the risks associated with unsupervised technological applications, enhancing overall responsible use.

The talk uses the OpenAI-developed language models, ChatGPT and GPT-4, to highlight the dual capabilities of these models as both AI and IA applications. Through this demonstration, we seek to show how these models have been effectively deployed for a variety of tasks within the legal domain, including multiple-choice question answering, legal reasoning, case outcome prediction, and summarization.

Nevertheless, the crux of this presentation argues that the full realization of "augmented intelligence" necessitates not just the computational prowess of LLMs, but also a reasoning and knowledge base component. We posit that IA systems should be designed to augment rather than replace human users in decision-making processes. This means that while LLMs can process and generate human-like text based on vast amounts of data, they should also be built to collaborate with human users, enhance their decision-making capacities, and make their work more efficient and effective.

Via this talk, we aim to stimulate further discourse on the responsible and beneficial integration of LLMs in the legal sector, reinforcing the need for more sophisticated IA systems that can effectively balance the benefits of advanced AI technology with the invaluable expertise of legal professionals.

2. AI vs. IA

AI and IA can be distinguished along the following dimensions how the system would interact or be autonomous from human activity:

In: *Proceedings of the Third International Workshop on Artificial Intelligence and Intelligent Assistance for Legal Professionals in the Digital Workplace (LegalAIIA 2023),* held in conjunction with ICAIL 2023, June 19, 2023, Braga, Portugal.

frank.schilder@thomsonreuters.com (F. Schilder)

2.1. Decision Making

AI Artificial Intelligence makes decisions autonomously based on algorithms and data it has been trained on. An example of this would be AlphaGo, an AI developed by Google DeepMind, that defeated the South Korean professional Go player Lee Sedol, one of the best players at Go. It did this without any human input during the game, just based on its previous training data [1].

IA In contrast, Intelligent Assistance supports human users in their decision-making processes but does not make the final decision itself. For instance, IBM's Watson for Oncology helps doctors in diagnosing cancer and suggesting treatment plans, but the final decision is always made by the human doctor [2].

2.2. Error Correction and Learning

AI AI systems like the GPT-4 model learn from their mistakes autonomously by adjusting their algorithms based on the feedback from their output results, without any human intervention during pretraining [3]. Even though reinforcement learning with human feedback (RLHF) [4] relies heavily on human feedback, humans cannot directly control the behavior of the model when an error or an unacceptable response occurs nor is the chatGPT or GPT-4 correcting human errors or do these models support the human learning process.

IA On the other hand, Intelligent Assistance systems like Grammarly, a language-correction tool, aid humans in spotting and correcting errors, facilitating a learning process that is heavily reliant on human cognition [5].

2.3. Autonomy versus Collaboration

AI A seemingly prime example of an autonomous AI is Tesla's Autopilot feature, which can control the car's steering, acceleration, and braking within its lane without human input, albeit under supervision for safety reasons [6].

IA Conversely, a collaborative robot (cobot) in a manufacturing line, like those developed by Universal Robots, works alongside humans, assisting them in tasks that require heavy lifting or precision, but always under the control and supervision of human operators [7].

2.4. Predictive Capabilities

AI Artificial Intelligence models like Google's DeepMind's AlphaFold predict protein structures with remarkable accuracy, a task that has remained unsolved for decades. They do this by analyzing vast amounts of data without any human intervention [8].

IA On the other hand, Intelligent Assistance systems like predictive text features in email clients (such as Google's Smart Compose) assist users in writing emails by suggesting phrases but do not compose entire emails autonomously [9]

These examples elucidate the key difference between AI and IA: AI operates with relative autonomy, whereas IA operates in conjunction with and under the supervision of human users, enhancing their abilities rather than replacing them.

3. Risks

While Large Language Models (LLMs) present considerable potential when used in AI applications, it is important to acknowledge and address associated risks to ensure responsible and ethical use. These risks can range from reliability issues to ethical and legal concerns.

3.1. Reliability and Accuracy

LLMs, including an AI system like GPT-4, are trained on vast amounts of data, and while they can generate human-like text, they do not understand the content in the same way humans do. This can lead to potential errors or misinformation. For instance, if a legal AI system misinterprets a statute or case law, it could provide inaccurate advice or predictions [10].

3.2. Ethical and Bias Concerns

Since LLMs are trained on real-world data, they may perpetuate existing biases present in the training data [11, 12]. If unchecked, these biases can influence the advice or insights generated by the LLM, leading to potential discrimination or unfair treatment in a legal context.

3.3. Accountability and Transparency

As LLMs become more complex, the reasoning behind their outputs can become opaque, leading to a "black box" problem [13]. This lack of transparency makes it challenging to ascertain accountability if the AI system leads to incorrect or harmful decisions, especially in high-stakes legal settings.

3.4. Data Security and Privacy

LLMs used in legal AI applications will likely handle sensitive data. Ensuring the security and privacy of this data is crucial to protect client confidentiality and comply with legal requirements such as GDPR [14]. The misuse or breach of this data represents a significant risk.

3.5. Dependence on Technology

There's a risk of over-reliance on AI systems, leading to complacency and diminished critical thinking abilities among users. Legal professionals must continue to apply their expertise and judgment in conjunction with AI tools [15].

Addressing these risks requires a combination of technical solutions (like refining training techniques and improving transparency of AI decision-making processes), regulatory measures, and fostering user awareness about the strengths and limitations of LLM-based AI systems.

4. LLMs and Legal Interactive Assistants

The application of LLMs in legal settings, while promising, comes with an inherent limitation: these models lack an understanding of the semantic content they process. Despite their ability to generate human-like text, they do not grasp the implications or nuances of the content they generate or analyze in the same way a human user would [10].

In the context of the legal sector, where precision, understanding, and complex reasoning are paramount, this limitation is critical. Legal professionals need to reason about the law, apply it to specific cases, understand complex interdependencies, and navigate ambiguities. These tasks are not just about processing language, but about reasoning and understanding the underlying principles and consequences [16, 17].

To fully realize the potential of LLMs as IA in the legal sector, there is a need for a reasoning and knowledge base component that goes beyond mere language processing. Such a component would enable the IA system to support human users effectively in complex decision-making processes.

Existing LLMs, like GPT-4, are based on transformer models that excel in solving standardized tests but lack explicit reasoning capabilities. To fill this gap, we propose integrating these LLMs with knowledge graphs or similar structures that provide a contextual understanding of the data [18]. This way, the LLM could not only process text but also reason about it in a manner more aligned with human cognition.

References

[1] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, D. Hassabis, Mastering the game of Go without human knowledge, Nature 550 (2017) 354–359. URL: https://doi.org/10.1038%2Fnature24270. doi:10.1038/nature24270.

[2] Z. Jie, Z. Zhiying, L. Li, A meta-analysis of Watson for Oncology in clinical application, Scientific Reports 11 (2021). URL: https://doi.org/10.1038%2Fs41598-021-84973-5. doi:10.1038/s41598-021-84973-5.

[3] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, et al., Language models are unsupervised multitask learners, OpenAI blog 1 (2019) 9.

[4] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, et al., Training language models to follow instructions with human feedback, Advances in Neural Information Processing Systems 35 (2022) 27730–27744.

[5] H.-W. Huang, Z. Li, L. Taylor, The Effectiveness of Using Grammarly to Improve Students' Writing Skills, in: Proceedings of the 5th International Conference on Distance Education and Learning, ICDEL 2020, Association for Computing Machinery, New York, NY, USA, 2020, pp. 122–127. URL: https://doi.org/10.1145/3402569.3402594. doi:10.1145/3402569.3402594.

[6] S. Ingle, M. Phute, Tesla autopilot: semi autonomous driving, an uptick for future autonomy, International Research Journal of Engineering and Technology 3 (2016) 369–372.

[7] F. Sherwani, M. M. Asad, B. Ibrahim, Collaborative Robots and Industrial Revolution 4.0 (IR 4.0), in: 2020 International Conference on Emerging Trends in Smart Technologies (ICETST), 2020, pp. 1–5. doi:10.1109/ICETST49965.2020.9080724.

[8] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al., Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583–589.

[9] M. X. Chen, B. N. Lee, G. Bansal, Y. Cao, S. Zhang, J. Lu, J. Tsay, Y. Wang, A. M. Dai, Z. Chen, T. Sohn, Y. Wu, Gmail smart compose: Real-time assisted writing, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19, Association for Computing Machinery, New York, NY, USA, 2019, pp. 2287–2295. URL: https://doi.org/10.1145/3292500.3330723. doi:10.1145/3292500.3330723.

[10] E. M. Bender, T. Gebru, A. McMillan-Major, S. Shmitchell, On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?, in: Proceedings of the 2021 ACM conference on fairness, accountability, and transparency, 2021, pp. 610–623.

[11] S. Matthews, J. Hudzina, D. Sepehr, Gender and racial stereotype detection in legal opinion word embeddings, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 36, 2022, pp. 12026–12033.

[12] T. Bolukbasi, K.-W. Chang, J. Y. Zou, V. Saligrama, A. T. Kalai, Man is to computer programmer as woman is to homemaker? debiasing word embeddings, Advances in neural information processing systems 29 (2016).

[13] D. Castelvecchi, Can we open the black box of AI?, Nature 538 (2016) 20–23. URL: https://doi.org/10.1038%2F538020a. doi:10.1038/538020a.

[14] S. E. McGregor, H. Zylberberg, Understanding the General Data Protection Regulation: A Primer for Global Publishers (2018). URL: https://academiccommons.columbia.edu/doi/10.7916/D8K08GVB. doi:10.7916/D8K08GVB.

[15] R. E. Susskind, D. Susskind, The future of the professions: How technology will transform the work of human experts, Oxford University Press, USA, 2015.

[16] K. D. Ashley, Artificial intelligence and legal analytics: new tools for law practice in the digital age, Cambridge University Press, 2017.

[17] D. Zhang, F. Schilder, J. G. Conrad, M. Makrehchi, D. von Rickenbach, I. Moulinier, Making a computational attorney, in: Proceedings of the 2023 SIAM International Conference on Data Mining, SDM 2023, Minneapolis, MN, USA, April 27-29, 2023, SIAM, 2023.

[18] M. F. A. R. D. T. (FAIR)†, A. Bakhtin, N. Brown, E. Dinan, G. Farina, C. Flaherty, D. Fried, A. Goff, J. Gray, H. Hu, A. P. Jacob, M. Komeili, K. Konath, M. Kwon, A. Lerer, M. Lewis, A. H. Miller, S. Mitts, A. Renduchintala, S. Roller, D. Rowe, W. Shi, J. Spisak, A. Wei, D. Wu, H. Zhang, M. Zijlstra, Human-level play in the game of diplomacy by combining language models with strategic reasoning, Science 378 (2022) 1067–1074. URL: https://www.science.org/doi/abs/10.1126/science.ade9097. doi:10.1126/science.ade9097. arXiv:https://www.science.org/doi/pdf/10.1126/science.ade9097.

5. About the Author

Frank Schilder is a Senior Research Director at Thomson Reuters with TR Labs, leading a team of researchers to explore new machine learning and artificial intelligence techniques in order to create smart products for legal NLP problems. His research interests include summarization, question answering and information extraction, and natural language generation. Frank received the master's degree in computer science (Diplom-Informatik) from the University of Hamburg and the Ph.D. degree in cognitive science from the University of Edinburgh, Scotland. Before joining Thomson Reuters, he was an Assistant Professor at the Department for Informatics, University of Hamburg, Germany.