Agentic AI in Judicial Decision-Making: Balancing AI, Ethics and Trust

John Keers BL
Director, Centre for Legal Technology, Ulster University, Northern Ireland

Dr. John McCord
Senior Lecturer, Law and Research Lead, Centre for Legal Technology, Ulster University, Northern Ireland

Abstract

Artificial Intelligence has permeated every layer of society—and the law is no exception. AI's integration into legal processes has accelerated from predictive caselaw analytics to emergent generative-AI tools adopted by practitioners and benches, however, the adoption of AI in courts varies dramatically: while some jurisdictions actively deploy AI to assist judges in case triage and summarisation, others remain cautious due to concerns over bias, opacity, reliability and data privacy. A lack of empirical evidence on real-world deployment has further hindered broader acceptance.

The integration of AI into judicial processes presents both transformative opportunities and profound challenges. Most of the existing deployment of AI into legal processes also relies on single-agent paradigms, limiting robustness, adaptability, and collaborative oversight. This study presents our ongoing research into the development of an Agentic AI prototype tailored to assist judicial decision making in the courts of the United Kingdom. Grounded in inter-disciplinary theories of legal reasoning and cognitive science, our approach seeks to explore a fine-tuned legal reasoner model on a curated UK case-law corpus to improve transparent and explainable reasoning in decision making in a legal setting. By harnessing distributed agent architectures, our work demonstrates how Agentic AI can elevate consistency, efficiency, and trust in judicial processes without compromising the rule of law.

We are developing and testing the prototype in Early Neutral Evaluation simulations and cases as a low-risk experimental setting to develop and test AI-assisted judgments, enabling an ethically safe examination of system behaviour without binding parties. Through simulations, practitioner and judicial feedback, we anticipate our findings will provide evaluation and empirical evidence of the prototypes performance to demonstrate how our Agentic AI can improve decision-issue identification consistency, improve legal reasoning and judgment drafting, while maintaining explainability through transparent reasoning. We also seek to uncover critical risks, including hallucinations, data privacy vulnerabilities, human–AI alignment failures, and potential erosion of judicial autonomy and public trust.

Keywords

Agentic Artificial Intelligence, Judicial Decision-Making, Explainable AI, Legal Reasoning, Trust, Ethics, Courts.