

Analysis and Technology

Would an AI judge be able to efficiently dispense justice?

Judges are only human and can make mistakes, so could an artificial intelligence make better and more efficient decisions?

By [Matthew Sparkes](#)

📅 7 June 2024



▲ **The Statue of Justice at the Old Bailey in London**



Jamie Lorriman / Alamy Stock Photo

Should artificial intelligence be used in the justice system, and if so should it apply the letter or the spirit of the law? While there are no plans for AI judges yet, this is a question that the UK government is already wrestling with as it considers the [potential uses of AI in the English court system](#) <https://apnews.com/article/artificial-intelligence-ai-guidance-england-wales-judges-c2ab374237a563d3e4bbbb56876955f7>. Despite this, lawyers and computer scientists are warning that current systems can't handle the ambiguity and nuance often required in legal situations.


For example, [Viniya Jain](https://viniya.ai/) <https://viniya.ai/>, formerly at Stanford University and now working at Amazon, and her colleagues assessed how large language models (LLMs), of the type that powers ChatGPT, would judge people carrying out [ethically ambiguous actions driven by their circumstances](#) <https://arxiv.org/pdf/2403.14633.pdf>, such as stealing food to feed starving family members. While people may be split on whether or

not poverty or need was mitigation in these scenarios, the AIs had none of the understanding or empathy we might expect from human decision makers and almost always found the hypothetical lawbreakers guilty – which could be disastrous in the context of the legal system.


“The leeway that judges have where they can say ‘OK, this is wrong, but in this current situation this is justified’ is not something we can really expect an LLM to do as well as a human judge,” says Jain.

But the time-saving potential of AI is a tempting proposition for cash-strapped public services. [Youmna Hashem](https://www.turing.ac.uk/people/researchers/youmna-hashem)  <https://www.turing.ac.uk/people/researchers/youmna-hashem> at the UK’s AI think tank and research group, the Alan Turing Institute, and her colleagues, have put out a [report](https://arxiv.org/pdf/2403.14712.pdf)  <https://arxiv.org/pdf/2403.14712.pdf> quantifying how many interactions between citizens and the state – of which there are an estimated one billion each year in the UK – could be farmed out to AI. The group estimated that 84 per cent were “highly automatable”, although crime and justice was deemed to have a “low potential” for AI automation due to its complexity.

Hashem says that AI could be a huge timesaver in courts in limited roles, such as transcribing proceedings, but not to make decisions. “I think that even if the technology itself exists and performs relatively well, it’s still a question of whether we should or shouldn’t be implementing it in areas where the consequences are high,” she says.

[Burkhard Schafer](https://www.law.ed.ac.uk/people/professor-burkhard-schafer)  <https://www.law.ed.ac.uk/people/professor-burkhard-schafer> at the University of Edinburgh, UK, says that there is a concept in legal studies of Judge Hercules – a superhuman judge who knows all case law, all previous decisions and all the facts of the case, and can arrive at a perfect legal decision. This sounds like what we might want from an AI judge, but Schafer is sceptical that existing technology can step up to the job.

Schafer says human judges can display spontaneous acts of empathy, or a discretionary bending of the rules when it right or just, but that this door swings both ways. And while human judges have individual reputations for leniency or strictness, even for certain types of crime, an AI model would at least be uniform, he says. “We want cold logic, that is liberating,” says Schafer. “Rules liberate us, ultimately.”

But this cold logic brings its own problems, like the so-called [three strikes laws](https://en.wikipedia.org/wiki/Three-strikes_law)  https://en.wikipedia.org/wiki/Three-strikes_law in the US that can see a person with two

previous convictions caught stealing a loaf of bread face a mandatory life sentence. Schafer says that under that sort of system human judges essentially become computers, who could easily be replaced by AI. Neither option is desirable, he says.

“What we normally see in most legal systems is that both elements coexist in an uneasy tension. And as humans we are good at living with these ambiguities,” says Schafer.