AI-generated pleadings: separating facts from fictions

In the summer of 2023, stories began to emerge about cases concerning the use of pleadings generated by artificial intelligence (AI) before a court of law, raising questions both about a litigant's liability for submitting fictitious case law and also the court's approach when dealing with inaccurate submissions.

While these cases have generated some amusing headlines, the stakes are high: English legal precedent as a cornerstone of the legal system could suffer if inaccurate AI-generated submissions are accidentally adopted as true or accurate in judgments.

First case in the UK

In a December 2023 decision, *Felicity Harber v The Commissioners for HMRC*, a taxpayer, Mrs Harber, failed to notify HM Revenue & Customs (HMRC) of her liability to pay capital gains tax following a property sale (*[2023] UKFTT 1007 (TC)*). HMRC issued a penalty of £3,265.11, which she appealed on the basis of reasonable excuse, that being her mental health condition or that it was reasonable for her to be ignorant of the law.

Mrs Harber provided the First-tier Tribunal with the names, dates and summaries of nine alleged First-tier Tribunal decisions that supported her appeal that a reasonable excuse existed. She explained that these were provided by a friend in a solicitor's office. However, the tribunal discovered that none of these authorities were genuine but were instead AI-generated. Mrs Harber accepted that this was possible. The tribunal did not sanction Mrs Harber and proceeded with the case as if the submissions had not been made.

Attorneys penalised in the US

In June 2023, two New York lawyers were sanctioned $5,000 for submitting a legal brief that included six AI-generated fictitious case citations and making false and misleading statements to the court (*Roberto Mata v Avianca Inc, 22-cv-1461 (PKC), US DC Southern District of New York*).

A Colorado lawyer was suspended for one year and a day after he cited AI-generated case law in a motion submitted to a court in May 2023. The junior attorney failed to verify or even read the fictitious cases and he failed to alert the court about the existence of the incorrect cases or to withdraw the motion (*www.cbsnews.com/colorado/news/colorado-lawyer-artificial-intelligence-suspension/*).

Highly plausible but incorrect

These cases demonstrate the risk of AI "hallucinations", particularly any "highly plausible but incorrect results" as highlighted in a recent report published by the Solicitors Regulation Authority (SRA) (*www.sra.org.uk/sra/research-publications/artificial-intelligence-legal-market/*).

AI hallucination cases that alter existing authorities with incorrect facts can be more convincing and harder to spot than entirely fictitious cases. Counsel for HMRC in *Harber* described her extensive searches to challenge the veracity of these fictitious cases. "Baker v HMRC (2020)" cited by Mrs Harber was similar to the real case of *Richard Baker v HMRC* (*[2018] UKFTT 763 (TC)*). But while in the fictional case Mr Baker won his appeal by arguing that his depression was a reasonable excuse, in the real case he had lost.

Cases with some traits that are superficially consistent with actual judicial decisions are harder to spot, but the judiciary in the UK has recently issued guidance for judges in this area (*see box "Guidance on spotting AI hallucinations"*).

In *Mata*, the US court noted "stylistic and reasoning flaws that do not generally appear in decisions issued by United States Court of Appeals", demonstrating the nuances that are required to identify AI hallucinations. Another sign was the frequent repetition of identical phrases in different fictitious cases, as identified by the tribunal in *Harber*.

Impact on the legal profession and judicial system

The tribunal in *Harber* agreed with the US court in *Mata* that many harms may flow from the submission of fake opinions: the opposing party wastes time and money in exposing the deception, the court's time is taken from other important matters, the client may be deprived of arguments based on authentic judicial precedents, and there is potential harm to the reputation of judges and courts whose names are falsely invoked as authors of the bogus opinions and to the reputation of a party attributed with fictional conduct. It also promotes cynicism about the legal profession and the judicial system: a future litigant might be tempted to defy a judicial ruling by disingenuously claiming doubt about its authenticity.

Statistical errors by human witnesses are reported to have led to miscarriages of justice in the past. The SRA warns of the severe impact on affected people if legal proceedings are misinformed by AI hallucinations; a risk aggravated by research suggesting that people may place more trust in computers than in humans.

Guidance on spotting AI hallucinations

On 12 December 2023, the judiciary issued guidance for judicial officeholders that sets out some signs of AI-generated cases:

- References to cases that do not sound familiar or have unfamiliar citations, sometimes from the US.

- Parties citing different bodies of case law in relation to the same legal issues.

- Submissions that do not accord with the officeholder's general understanding of the law in the area.

- Submissions that use US spelling or refer to overseas cases.

- Content that, superficially at least, appears to be highly persuasive and well written, but on closer inspection contains obvious substantive errors (*www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf*).

Accountability

Some common themes in these cases include:

- Pressure of litigation or workload.

- Lack of skill and care or some level of recklessness in not being able to verify or not even attempting to verify the accuracy of AI-generated cases.

- Lack of transparency and an attempt to conceal the origin of the AI-generated case law, blaming a legal intern for inaccuracies, misunderstanding how the technology works or avoiding questions about it.

In *Harber*, the tribunal accepted that Mrs Harber was not aware that the cases were fabricated and that she did not know how to check authorities. However, there is notable disparity in the tribunal accepting ignorance as an excuse in relation to Mrs Harber's alleged active steps in generating the fictitious cases and submitting them to the tribunal, but not accepting ignorance of the law in relation to her failing to notify her tax liability within 30 days of her property sale (previously only notifiable at the end of the tax year), which her solicitor allegedly failed to advise on and which HMRC would likely have reminded her of anyway.

Lawyers will not be afforded the same leniency. Many law firms already use generative AI, including for legal research. However, lawyers remain responsible for any errors in that advice. Firms must have measures in place to meet these responsibilities and they should keep clients suitably informed of how AI is involved in their cases.

Perhaps in future, lawyers or litigants might be able to rely on AI-powered platforms designed for legal research. For example, under the AI Liability Directive proposed by the European Commission, damages may be claimed against the provider if AI-generated output is produced in breach of a duty of care or the law (*https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0496*). A reversal of the burden of proof will assist consumers, who will only need to prove damage and the provider will have to prove the lawfulness of its AI-generated outputs.

Interesting problems

The use of AI in the legal sector will continue bringing up interesting problems. At this time, the judiciary's message to the courts is that AI tools are a poor way of conducting research as they can make up fictitious cases, citations or quotes, or refer to legislation, articles or legal texts that do not exist. Their output cannot be independently verified and their analysis or reasoning is unconvincing. However, litigants are likely to continue making use of generative AI if they cannot afford representation and courts will continue facing the risk of wasted court time at the expense of other court users.

Measures will have to be adopted to ensure transparency by litigants about their use of generative AI, such as tick boxes on court forms confirming that no AI-generated content has been submitted or that any AI content has been verified. This will make it harder for unverified content to be submitted. It may also mean that litigants who attempt to conceal the use of generative AI in the future will likely be committing contempt of court.

However, the AI risk in litigation is wider than submitting fictitious pleadings. It also concerns any evidence that is prepared or processed through AI-powered systems. A December 2021 brief by the Centre for Security and Emerging Technology in the US suggested that courts should exclude AI evidence that is biased or otherwise unreliable and avoid the risk of a trial within a trial about the accuracy of AI content (*https://www.armfor.uscourts.gov/ConfHandout/Baker2021DecCenterForSecurityAndEmergingTechnology1.pdf*). Courts will have to develop frameworks for scrutinising the reliability of outputs delivered by legal technology; an industry which could grow from $1.3 billion in 2022 to upward of $8.7 billion by 2030 (*www.washingtonpost.com/technology/2023/11/16/chatgpt-lawyer-fired-ai/*).

Generative AI also enables fraud. While UK courts have always had to handle forgeries, and allegations of forgery, involving varying levels of sophistication, the courts must be aware of the convincing potential of deepfake technology (*see feature article "The rise of the deepfake: looking into a dystopian future?", www.practicallaw.com/w-026-8753*). While AI can be a useful tool for litigation, litigants must be held to a high standard of transparency and accuracy when relying on AI-processed content.

Alexander Dittel is a partner, and Scarlet Mitchell is a paralegal, in the Technology Practice at Wedlake Bell LLP.