

Artificial intelligence & justice:

What do we know from the evidence?

Policy brief

Dr Holli Sargeant, University of Cambridge

July 2026



Artificial intelligence & justice: What do we know from the evidence?

This briefing document accompanies a systematic scoping review commissioned by the Nuffield Foundation's *Public right to justice* programme to examine the evidence on artificial intelligence (AI) and data-driven technologies in civil justice settings in England and Wales.

The review maps 45 tools in use or in development across courts and tribunals, government departments, advice services, and private providers. General-purpose tools such as ChatGPT and Claude are increasingly used informally by practitioners, judges, and litigants in person. This brief highlights conclusions about the state of the evidence rather than about the net effect of AI itself.

Background and context

AI is being rapidly deployed across the justice system. The Ministry of Justice has set out an *AI Action Plan for Justice* (2025), established a dedicated Justice AI Unit, and extended judicial access to Microsoft Copilot through eJudiciary. AI tools are also used by litigants in person, practitioners, and advice services, in ways that already shape how justice processes are encountered. AI is increasingly placed at the centre of plans for swifter, fairer, and more accessible justice.

The case for that ambition is well-motivated, and it is strongest where the system is under most pressure. Court and tribunal backlogs are creating delays, the advice sector is capacity-constrained, and unmet legal need is high. For people facing everyday legal problems, public-facing information, triage, navigation, and intake tools that are low-cost, available at any time, and operate in plain English and multiple languages may reach those whom conventional provision does not, and may help them recognise a problem as legal and actionable.

Against this sits the central feature of the current landscape: a gap between adoption and independent evaluation. Only 7 of the 45 mapped tools have any publicly available evaluation.

Deployment is materially outpacing the production of evidence about what these tools do to the people the system serves. The UK evidence base is modest given the breadth of deployment, although the available evidence supports a small number of claims.

The findings of this review suggest that decisions about usage, research, and procurement of AI in justice settings cannot reasonably be deferred. AI is already shaping the conditions under which people encounter the justice system. There is reasonable ground to think that AI, designed and used well, can extend access and reduce friction for users who would otherwise be poorly served. Whether it will, in fact, advance or undermine the public's right to justice is, at present, an open question, and one the public is entitled to have answered. It is an empirical question, and one the current evidence is not well placed to answer.

The four structural gaps

The review revealed four structural gaps in our understanding and deployment of AI in justice: measurement, deployment, legitimacy, and transparency.

1. **A measurement gap: evaluation is measuring narrow performance metrics, not the outcomes that matter most to people using the justice system.**

Current UK evaluation measures the wrong things for the questions that matter most to people facing everyday legal problems. It captures the experience of system operators rather than the people whose cases are processed; it relies on self-reported time savings and satisfaction rather than independent assessment; and it tests narrow tasks in controlled conditions rather than the real settings in which legal problems arise.

Part of the gap is structural. Many legal needs never reach formal advice or adjudication, legal data rarely contains stable ground truth, and the people most likely to be poorly served are the least visible to system analysis. International evidence reveals that certain demographic groups may experience worse outcomes from AI in legal aid triage, and court digitisation shows uneven uptake across user groups. No UK justice-specific study has measured how AI-assisted processes affect individuals, or how effects differ across groups.

AI tools are also typically evaluated in isolation. The cumulative effect of multiple AI tools acting on a single case is largely unstudied.

2. A deployment gap: real-world use is not usually evaluated or studied.

The systems studied in the literature are increasingly not the systems used in practice. UK deployment is dominated by general-purpose proprietary models: Microsoft Copilot across the Ministry of Justice and HM Courts & Tribunals Service, GPT-based systems for Home Office, and Claude underlying Citizens Advice's Caddy advice tool. Proprietary systems cannot be inspected at the level of weights, training data, or architecture, so behavioural evaluation must do the work that structural evaluation cannot.

Evaluative research indicates that legal-domain and smaller non-generative models can outperform general-purpose ones for some UK legal tasks; procurement does not yet reflect that finding. These general-purpose proprietary models also change frequently with new versions and system instructions, so point-in-time assurance does not persist.

These issues also extend beyond formal deployment: publicly accessible general-purpose models (ChatGPT, Claude, etc) are increasingly used by litigants in person and practitioners, introducing challenges that currently sit beyond the purview of formal justice processes but already shape how people encounter them.

The decisions being taken now about which systems to embed in justice processes will be costly to reverse, and they are being taken faster than the evidence is being produced.

3. A legitimacy gap: users perceive and experience AI in terms that current evaluations do not measure.

How a system performs and how it is experienced require materially different evaluation, and the second is largely under-measured. Limited UK evidence suggests that even modest assistive uses of AI reduce the public's sense that a process is fair and transparent. In contrast, some international evidence indicates that groups already distrustful of the existing system may regard AI as a chance to improve fairness rather than a threat to it.

Human oversight is perceived as vital to the procedural fairness of AI use, and is treated in policy and judicial guidance as the principal safeguard. Yet no UK deployment evaluation examines whether

that oversight operates effectively in practice. Its real costs and limits, including time demands, psychological feasibility under workload pressure, and sensitivity to automation bias, are largely invisible in current evaluations.

While senior judicial guidance and engagement has been active and positive towards the benefits of AI in justice, judges and legal practitioners in survey evidence reveal principled concerns about the risk of deskilling, the normative importance of human judgment, and reliability of AI outputs.

The legitimacy question is therefore not whether AI use is acceptable in the abstract, but whether the balance between procedural protection and accessible resolution is being struck in the right place, with the right safeguards, for the right users.

4. A transparency gap: AI is deployed faster than it can be openly scrutinised.

Where evaluation exists, it is often internal, partial, or methodologically thin. The dominance of proprietary models makes external scrutiny harder. Internal and vendor-led evaluation cannot easily escape actual or perceived bias, and apparent effectiveness may reflect selective reporting rather than demonstrated performance.

The structure of the legal data environment shapes both accountability and innovation. Efforts to improve access to judgments and legal data create important conditions for research, scrutiny, and development. However, uneven access and the dominance of proprietary systems risk reinforcing existing asymmetries in accessibility, limiting who can evaluate or build legal AI systems.

The justice system has, historically, a less developed tradition of independent and empirical evaluation than other parts of the public sector, and building that capacity for interdisciplinary, sociotechnical, and more adaptive evaluations is itself a precondition for meaningful scrutiny.

Priorities for policy, practice, and research

AI is already shaping how people encounter the justice system, but the evidence needed to know whether it is advancing or undermining the public's right to justice does not yet exist, because what has been measured is not what matters most to the people the system serves. The findings of this review translate into several priorities across three dimensions.

Policy

- Resist framing AI as a remedy for entrenched problems it cannot solve alone; presenting it as an answer to unmet legal need risks substituting a tool for the investment the system requires.
- Commission projects based on people-centred outcomes from the outset where possible, rather than an afterthought.
- Treat cost transparency, auditability, and the ability to change provider as procurement requirements, because dependency on proprietary systems is the most consequential and least reversible choice being made now.

Practice

- Orient deployment towards people facing everyday legal problems who are underserved by the current system, not only those already inside it.
- Embed public-facing tools for sensitive domains within human support rather than offering them as standalone self-service products.
- Design and test human oversight as a real safeguard, with attention to the conditions under which it fails.
- Prioritise disclosure and transparency on the use of AI, especially where it bears on a person's case, so that transparency is built into the justice system.

Research

- Measure outcomes for the people whose problems are processed, disaggregated by group, rather than the experience of operators alone.
- Study the experience of AI-mediated justice in real settings rather than in vignettes.
- Examine the informal use of public-facing tools by litigants in person and practitioners, in order to prioritise design and deployment opportunities for improving accessibility of legal information and advice.
- Alongside conventional independent evaluation, design and adopt faster, structured test-and-learn methods so that evidence is generated closer to the pace of deployment.

About the research

This brief summarises a systematic evidence scoping review on AI and data-driven technologies in UK justice settings, with international evidence where relevant or transferable. The review mapped 45 AI tools in use or in development and synthesised the academic, policy, and grey literature against research questions on individual and group experiences and outcomes, societal factors, and system effectiveness. The review was conducted by Dr Holli Sargeant, commissioned by the Nuffield Foundation for the *Public right to justice* programme. The full report is available on the project page.

This report was commissioned by the Nuffield Foundation as part of its [Public right to justice](#) programme. The views expressed are those of the author and not necessarily the Foundation. The report forms part of a series of evidence reviews and briefing papers critically examining different dimensions of the justice system and access to justice in England and Wales. The programme aims to develop an interconnected body of research, providing policymakers, practitioners and the public with a better understanding of how the justice system operates, where it falls short, and how it could better meet the needs of those it serves.

As an independent charitable trust with a mission to advance social well-being, the Foundation funds and undertakes rigorous research, encourages innovation and supports the use of sound evidence to inform social and economic policy, and improve people's lives. It is the founder and co-funder of the Nuffield Council on Bioethics, the Ada Lovelace Institute, and the Nuffield Family Justice Observatory.

Find out more at: nuffieldfoundation.org

Bluesky: @nuffieldfoundation.org

LinkedIn: Nuffield Foundation