Nuffield Family Justice Observatory

AI in the family justice system

Author

Aliya Saied-Tessier

May 2024

Briefing

This briefing is intended to prompt discussion about how artificial intelligence has the potential to improve experiences of families and professionals in the family justice system. It also outlines the challenges and risks involved, and the governance required for safe usage.

Recommended citation

Saied-Tessier, A. (2024). *AI in the family justice system.* Briefing. Nuffield Family Justice Observatory. https://www.nuffieldfjo.org.uk/news/briefing-paper-ai-in-the-family-justice-system

Disclaimer

The majority of the research for this paper was completed in December 2023 with some updates made to reflect new guidance published in 2024.

Nuffield FJO has funded this project, but the views expressed are those of the authors and not necessarily those of Nuffield FJO or the Foundation.

100 St John Street, London EC1M 4EH T: 020 7631 0566

Part of Nuffield Foundation: Registered charity 206601
nuffieldfjo.org.uk | @NuffieldFJO www.nuffieldfoundation.org | @NuffieldFound

What's in this briefing?

Background notes

What is the issue?

Discussions on the reach and scope of artificial intelligence are now commonplace. But how well prepared are we to understand its potential or to deal with its challenges and risks in relation to the family justice system? And could we be doing more to think about how it could improve families' experiences?

We know that some AI technologies are available, accessible and already in use – commercial law firms are increasingly using AI to perform tasks such as contract analysis, automatic routine tasks and legal research for example (The Lawyer Portal 2023, 23 October). Could families in public and private law proceedings, as well as professionals in the family justice system benefit from the safe and fair use of newer technologies such as generative AI?

What do we mean by 'artificial intelligence'?

There is no single accepted definition of 'AI' – but there is a consensus that the core concept refers to machines simulating human intelligence.

Human intelligence processes are extensive and include the ability to acquire and process information, learn from experience, reason and make decisions, and interact with the environment in a way that achieves specific goals. As such, AI tools are not homogenous. Uses include facial recognition, virtual assistants (e.g. chatbots) and assessing eligibility and risks (e.g. the Home Office visa-screening tool used from 2015–2020 (see Modhvadia 2023)).

What does this briefing paper do?

This briefing paper aims to prompt reflection and discussion in the family justice system. It highlights some potential uses for AI, centred around:

- improving families' experiences
- enabling administrative efficiencies
- supporting decision making.

It also raises some of the important questions about challenges and risks – such as the role of humans, fairness, accountability, privacy, transparency and accuracy – and ends with a discussion around regulation and governance.

Chat GPT was used for some initial ideas generation and to refine the drafting of some sections of this report.

The author would like to thank Laura Carter, Cliff Manning and Kayliegh Richardson for reviewing an earlier draft.

How to get in touch

For all questions on the research and our work in this area, please contact: contactfjo@nuffieldfoundation.org

For all media enquiries, please contact: mediafjo@nuffieldfoundation.org

Key terminology

Artificial intelligence (AI)

There is no single accepted definition of 'AI' – but there is a consensus that the core concept refers to machines simulating human intelligence.

Generative AI

AI that is used to create new text, video, audio or code (Central Digital & Data Office and Cabinet Office 2024).

Humans in the loop (HITL)

This is where humans are actively and directly involved in an AI system's decision-making process. HITL is often used in situations where human judgement and expertise are crucial, and AI is used as a tool to assist or augment human decision making. Examples include medical diagnosis systems where doctors review AI-generated recommendations or content moderation systems where human moderators review flagged content.

Human on/over the loop (HOTL)

This refers to situations where humans are not actively involved in real-time decisions but are responsible for monitoring and intervening when necessary. AI systems operate autonomously, but there is a human oversight function to ensure that the system behaves correctly and ethically. Anomaly detection systems in manufacturing, for instance, may operate autonomously but alert human operators when anomalies are detected.

Large language models (LLMs)

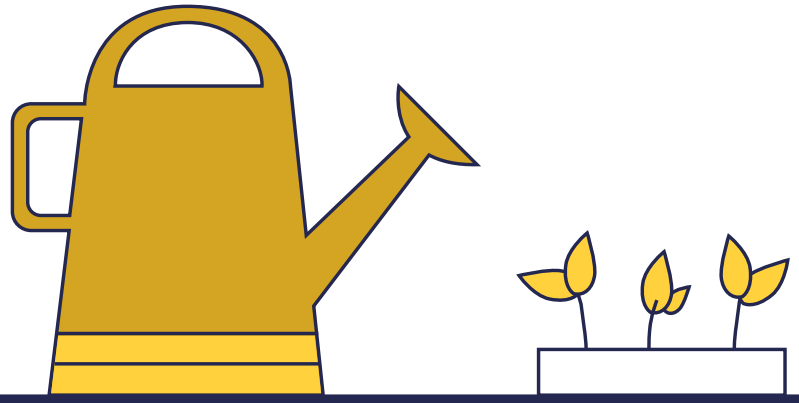
A type of generative AI that produce text output (Central Digital & Data Office and Cabinet Office 2024).

Machine learning

A subset of the tools that are currently being described and marketed as AI. While analysts building models to identify patterns and/or predict events is not new, the aim of machine learning is to enable computers to learn on their own, without being programmed by a human (Maini 2017, 19 August).

Key points for reflection

- How could AI be used to improve the experience of families in the family justice system?
- What additional guidance and legislation are required for responsible, legal, ethical AI use?
- Should there be a central record of where and how AI is being used in the family justice system in England and Wales?
- Given the importance of public trust in the family justice system, how and when should the public be engaged in debate around its use?



How might AI be used in the family justice system?

This section gives examples of possible uses of AI in the family justice system, including some examples of tools already being used and theoretical uses (the list is illustrative, not exhaustive). The uses of AI have been categorised as improving families' experiences of the family justice system, processing/administrative tasks leading to efficiencies or supporting decision making.

Outside of the applied examples, we have not discussed specific tools in this list, rather we have outlined areas with the potential to create new AI tools or adapt existing tools in the family justice system.

1. Improving family experiences	2. Efficiencies from processing or administrative tasks	3. Supporting decision making
Language translation	Document review	Online dispute resolution
Virtual assistants	Drafting (e.g. letters to children, child-friendly judgments, presentations and summaries	Predictive analysis
Legal advice	Case management	Risk assessment

1. Improving family experiences

Virtual assistants

Some courts incorporate virtual assistants or 'chatbots' on their websites to help users find information, complete forms and navigate the legal process.

Language 'translation'

Large language models (LLMs) could help improve understanding of family court proceedings by rewriting text containing legal jargon in plain English (or other languages), making court documents more accessible to all involved – including children.

There are examples of private companies offering these services in the US, such as Legalese Decoder.[1]

Legal advice

If generative AI had access to a sufficiently large set of existing case law information, it could theoretically support litigants in legal processes such as drafting submissions or citing relevant cases.

However, due to misunderstandings about how LLMs work, there are examples of people using ChatGPT to support them in court and getting incorrect information (known as 'hallucinations') (Lee 2024).

Applied example of a virtual assistant

"New Jersey state courts are handling an increasing number of public inquiries using a chatbot called JIA. JIA was initially released to the court's central office staff only. Staff manually entered all call centre inquiries (Human in the loop). On a daily basis, court staff reviewed a report of all inquiries and answers, adjusting JIA responses and adding question variations to train the system. When JIA was responding at an 80% accuracy level, the system was then released to more than 10,000 state court staff. Response accuracy again dropped to about 30% as state court staff asked questions the central office staff had not anticipated. Additional Q&A pathways were added until JIA was responding with 80% accuracy. The system was then added to the court's website making it live to the public (human on-the loop) without any kind of formal announcement."

– Joint Technology Committee 2020

[1] See: https://legalesedecoder.com/

2. Efficiencies from processing or administrative tasks

Applied example of supporting case management

"AI-empowered software is classifying and docketing e-filed documents at courts in Florida. The court started with three low risk/high volume case types, progressively expanding the variety and complexity of cases as they developed expertise with the robotic process automation (RPA) technology. The bots – each with its own name and user login – classify incoming e-filings, extract info from tagged fields, and docket them in the court's case management system. In 2020, 68 case types representing nearly a third of all Palm Beach County's e-filed documents are being docketed automatically. When the court first launched the system, humans double-checked 100% of the bots' work to verify accuracy (human in the loop). Today [2020], humans review 15% of all filings, whether docketed by a human or a bot."

– Joint Technology Committee 2020

Document review

AI tools could be used to scan and classify large volumes of documents for relevance to a case, reducing the time/cost associated with manual document review.

Case management

Case management systems could use AI tools to classify and route cases to appropriate teams.

Drafting

LLMs have been explicitly trained to create language and could be prompted to use specific writing styles (e.g. child-friendly or formal). LLMs can also learn new writing styles.

Guidance for the judiciary notes that AI could potentially assist in drafting summaries of large volumes of text (stressing the importance of checking accuracy), writing presentations (including suggesting topics to cover), and drafting emails /memoranda.[2]

Legal professionals could ask LLMs, such as ChatGPT, to redraft text to be understood by a child of a specific age so there is the potential to use LLMs to reduce the time taken to draft child-friendly documents such as letters and judgments.

[2] Courts and Tribunals Judiciary. (2023). *AI guidance for judicial office holders*. https://www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf

3. Supporting decision making (1)

Online dispute resolution

Some platforms use AI tools on online dispute resolution (ODR) platforms designed to facilitate the resolution of family disputes. AI can assist in mediation and settlement negotiations by suggesting potential compromises based on historic case data and legal precedents.

Applied example of online dispute resolution

Modria is an ODR system used by paying clients to give alternatives to litigation. It provides dispute resolution for a number of areas including divorce and separation, landlord-tenant disputes and employment disputes. According to Alexander 2019, the platform starts with a 'diagnosis module' that gathers relevant information. Next, a 'negotiation module' summarises areas of agreement/disagreement and makes suggestions for solving the issue. If these do not result in settlement, there is a 'mediation module', which takes place with a neutral third party. The final step is arbitration. Modria claims the 'vast majority' of claims are settled in the first two days without a human ever becoming involved.[3]

There are examples of the Modria platform having been used in Holland by the Dutch Legal Aid Board, though we could find no independent research about the results of the Modria platform, or information about it currently being used. In 2017, the Hague Institute for Innovation of Law discussed reasons why ODR is difficult to implement despite positive outcomes including resistance to public-private partnerships and a lack of resources from legal aid boards.[4]

[3] Alexander, N. (2019). *Ten trends in international mediation*. Singapore Academy of Law Journal, 31. https://journalsonline.academypublishing.org.sg/Journals/Singapore-Academy-of-Law-Journal-Special-Issue/Current-Issue/ctl/eFirstSALPDFJournalView/mid/503/ArticleId/1466/Citation/JournalsOnlinePDF

[4] Barendrecht, M. (2017, 21 June). *Rechtwijzer: Why online supported dispute resolution is hard to implement.* HiiL. https://www.hiil.org/news/rechtwijzer-why-online-supported-dispute-resolution-is-hard-to-implement/

Supporting decision making (2)

Predictive analysis

Some courts and children's social care departments use predictive analytics to assess the likelihood of various events – for example, predicting the likelihood of a child requiring social care interventions, or a young person's vulnerability to gang exploitation. This type of predictive analysis is well established when done without using AI tools (e.g. statistical analysis).

More recently, machine learning approaches have been adopted to analyse historical case data to predict various outcomes. Professionals can use these predictions to aid their decision making.

There is anecdotal evidence of local authority staff finding predictive analysis useful, though research published by What Works for Children's Social Care (WWCSC) in 2020 found that the machine learning models they built to identify children at risk using local authority social care data did not perform well in the sense that their measure of success was not met by any of the 24 models they tested.[5] The authors note that is it difficult to build models that predict outcomes well for children's social care.

Risk assessment

A subset of predictive analysis, AI can analyse historical data to make risk assessments, for example assessing the risk of a child being a victim of domestic violence or child abuse. There are examples of specific machine learning risk assessments performing better than existing approaches – for example, academics found a machine learning approach to domestic abuse risk assessment performed better than the Domestic Abuse, Stalking and Honour Based Violence Risk Identification (DASH) standard police protocol.[6] However, there are also examples of automated decision making errors, including inaccurate results and being overly punitive towards marginalised communities – for example, the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) system was designed to predict risks of recidivism and predicted higher risk values for black defendants and lower risk values for white defendants compared to actual risk.[7]

[5] Clayton, V., Sanders, M., Schoenwald, E., Surkis, L. and Gibbons, D. (2020). *Technical report: Machine learning in children's services: Does it work?*. What Works for Children's Social Care. https://whatworks-csc.org.uk/wp-content/uploads/WWCSC_technical-_report_machine_learning_in_childrens_services_does_it_work_Sep_2020.pdf

[6] Grogger, J., Ivandic, R. and Kirchmaier, T. (2020). *Comparing conventional and machine-learning approaches to risk assessment in domestic abuse cases.* Centre for Economic Performance. https://cep.lse.ac.uk/pubs/download/dp1676.pdf

[7] Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M.-E. et al. (2020). *Bias in data-driven artificial intelligence systems – An introductory survey.* Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. https://wires.onlinelibrary.wiley.com/doi/10.1002/widm.1356

Example of technology used in Australia's family justice system[8]

Information provision and automated drafting

- Document assembly applications that assist in drafting legal documents for both litigants and lawyers.
- Online guided interviews and guided form-filling software that walk users through the litigation process step by step and complete forms based on questions answered by the user.
- Do-it-yourself kits for different forms, including divorce.

Chatbots and expert systems

- Settify – a family law client intake system where potential clients respond to a series of questions before their first in-person meeting with a lawyer to help generate a list of instructions and save time.

Predictive analysis

- Big data can be used to provide lawyers and litigants with likely outcomes of a potential case based on similar cases (e.g. Split-Up).
- Machine text reading may enable automated programmes to analyse large datasets to identify patterns without help from a human (e.g. Lex Machina).

[8] Evans, J. and Ndegwa, A. (2022). *Use of Technology in the Family Justice System Annotated Bibliography.* Department of Justice Canada. https://www.justice.gc.ca/eng/rp-pr/jr/utfjsab-utsjfba/pdf/RSD2022_Use_of_Technology_in_the_Family_Justice_System_Report_EN.pdf

Challenges and risks

There are challenges and risks to using AI in any field, for example around fairness, accountability, safety and transparency. These issues, discussed below, are not specific to the family justice system. However, due to the nature of decisions made in family courts, our tolerance of any errors may be lower than other areas in which AI is being used.

- **Accountability**
- **Accuracy**
- **Fairness**
- **Human involvement**
- **Privacy and compliance**
- **Safety**
- **Transparency**

How much human involvement?

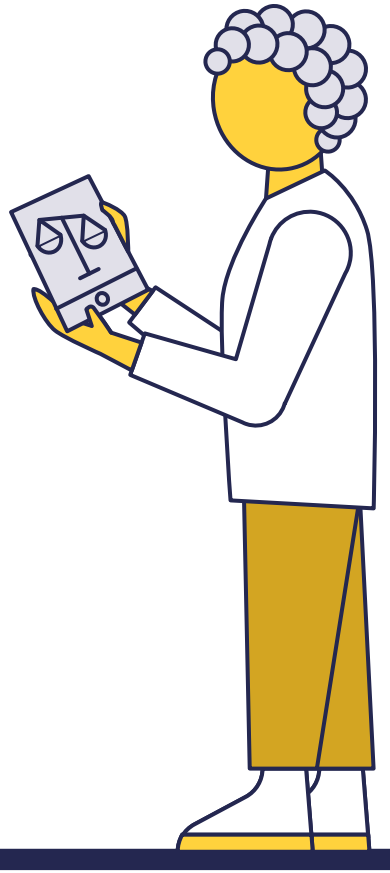
When thinking about how AI tools are designed and implemented in systems, it is useful to consider the appropriate level of human involvement.

This choice depends on the specific application, safety considerations, ethical concerns, and the level of trust that society and stakeholders place in the system.

In 'humans in the loop' (HITL) AI systems, humans are directly involved in the system's decision-making process. HITL is often used in situations where human judgement and expertise are crucial, and AI is used as a tool to assist human decision-making. Examples include medical diagnosis systems where doctors review AI-generated recommendations or content moderation systems where human moderators review flagged content.

'Humans on/over the loop' (HOTL) refers to situations where the AI systems operate autonomously but there is a human oversight function. For example, anomaly detection systems in manufacturing which scan without a human, but alert human operators when anomalies are detected.

There are other possibilities about the levels of human involvement such as human as a validator, human as an auditor and human in control.



Fairness

Bias

Bias in using AI can come from two sources – the data and how the AI tools are used. AI systems can inherit biases present in training data, potentially leading to discriminatory outcomes. Machine learning algorithms use data generated by humans which can cause them to reproduce or exacerbate existing biases such a racial bias (e.g. the COMPAS system predicted higher risk of reoffending for black defendants) or gender bias (e.g. Google Ads showing fewer adverts for high paying jobs to women).[9]

Addressing and mitigating algorithmic bias in AI systems is a complex challenge that requires ongoing monitoring and adjustment, and these tasks are not straightforward. For example, there are over 20 definitions of fairness used in the computer science literature with little consensus about the pros and cons of different definitions.[10]

Often different forms of bias must be traded-off against each other.[11]

There are also issues with bias around how algorithms are used. Okidegbe describes a three-pronged problem where organisations using ML algorithms are not transparent about their usage and adopt them without consulting marginalised communities. The same communities are often underrepresented – or not represented – in the process of creating those algorithms. Thirdly, even in areas where the public can share opinions about the use of algorithms, their views rarely lead to changes.[12]

Finally, people with insufficient means to fund quality human alternatives may turn to lower-quality and/or unregulated software tools in their place.

One example is people using unregulated AI tools that predict their likely financial settlements following family breakdown as an affordable alternative to professional legal advice. Another example is potentially unreliable AI translation in place of a human interpreter. This tilts the level-playing field away from those with fewer financial resources.

Digital exclusion

The adoption of AI technologies may not be uniform across all family justice institutions and may not be equally accessible across the population as some people do not have access to technology, reliable internet or training, potentially creating disparities in access to AI-assisted services and support.

[9] Ntoutsi et al. (2020).

[10] Ibid.

[11] Leavy, S., O'Sullivan, B. and Siapera, E. (2020). *Data, power and bias in artificial intelligence.* arXiv. https://arxiv.org/abs/2008.07341

[12] Okidegbe, N. (2022). *The democratizing potential of algorithms.* Connecticut Law Review, 53, 739. https://scholarship.law.bu.edu/faculty_scholarship/3138

The European Union Artificial Intelligence Act (2024) risk levels[13]

Unacceptable risk

Unacceptable risk AI systems are systems considered a threat to people and will be banned. They include:

- cognitive behavioural manipulation of people or specific vulnerable groups – e.g. voice-activated toys that encourage dangerous behaviour in children
- social scoring – classifying people based on behaviour, socioeconomic status or personal characteristics
- real-time and remote biometric identification systems, such as facial recognition – exceptions to this might include 'post' remote biometric identification systems where identification occurs after a significant delay to prosecute serious crimes, but only after court approval.

High risk

AI systems that negatively affect safety or fundamental rights will be considered high risk and will be divided into two categories:

- those that are used in products falling under the EU's product safety legislation (e.g. toys, aviation, cars, medical devices and lifts)
- those falling into one of eight specific areas that will have to be registered in an EU database
  - biometric identification and categorisation of natural persons
  - management and operation of critical infrastructure
  - education and vocational training
  - employment, worker management and access to self-employment
  - access to and enjoyment of essential private services and public services and benefits
  - law enforcement
  - migration, asylum and border control management
  - assistance in legal interpretation and application of the law.

All high-risk AI systems will be assessed before being put on the market and also throughout their lifecycle.

Generative AI

Generative AI, like ChatGPT, would have to comply with transparency requirements:

- disclosing that the content was generated by AI
- designing the model to prevent it from generating illegal content
- publishing summaries of copyrighted data used for training.

Limited risk

Limited risk AI systems should comply with minimal transparency requirements that would allow users to make informed decisions. After interacting with the applications, the user can then decide whether they want to continue using it. Users should be made aware when they are interacting with AI. This includes AI systems that generate or manipulate image, audio or video content, for example deepfakes.

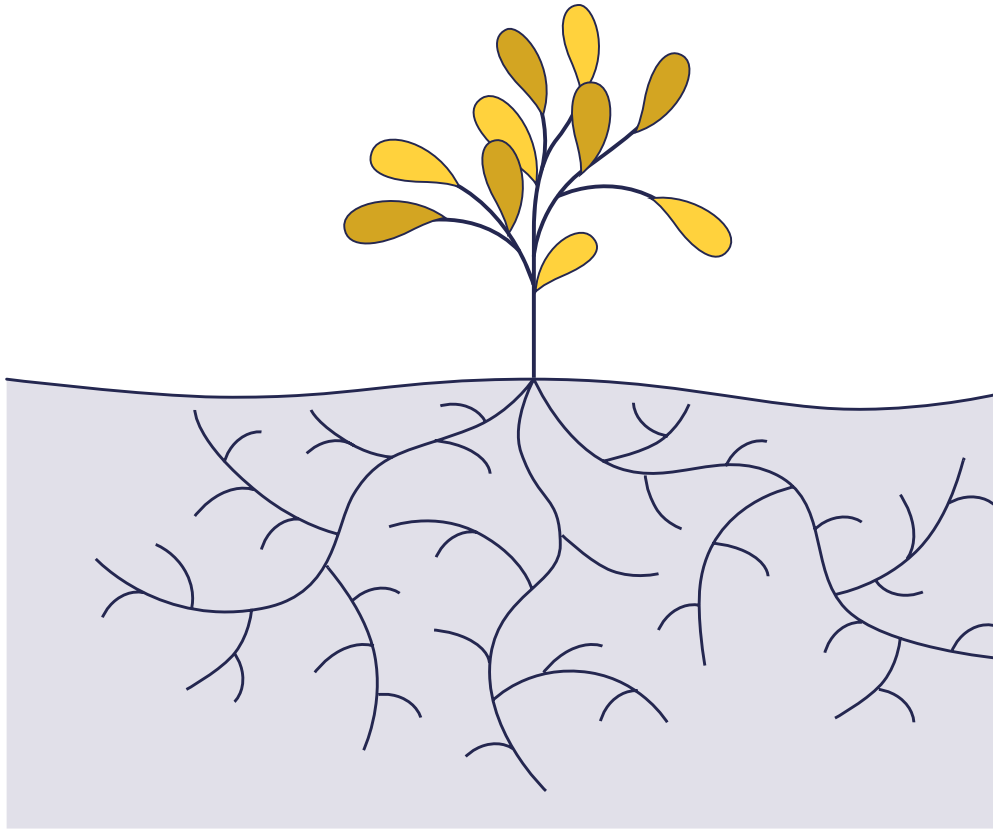
[13] European Parliament. (2023, December 19). *EU AI Act: First regulation on artificial intelligence.* https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Accountability

The Alan Turing Institute describes two challenges around AI projects being accountable:

- accountability gaps where models and hardware of AI systems are not responsible for the decisions they can produce
- the complexity of AI design where it is not straightforward to establish responsibility among the large number of people involved including technical experts, data teams, policy experts and users of the AI system.[14]

As well as AI projects, the accountability of AI users should also be considered. The Bar Council considerations on using ChatGPT and generative AI states that it is good to understand the underlying model before using a LLM, but in practice, there will be large variation in how well informed AI users are.[15]



[14] Leslie, D. (2019). *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector.* The Alan Turing Institute. https://doi.org/10.5281/zenodo.3240529

[15] Bar Council. (2024). *Considerations when using ChatGPT and generative AI software based on large language models.* https://www.barcouncilethics.co.uk/wp-content/uploads/2024/01/Considerations-when-using-ChatGPT-and-Generative-AI-Software-based-on-large-language-models-January-2024.pdf

Privacy and compliance

Data privacy

Handling confidential personal data in family law cases raises concerns about data privacy and security.[16]

AI is included in the scope of existing data protection laws (General Data Protection Regulation (GDPR) and the Data Protection Act 2018) if it involves personal data in training, testing or using models.[17]

There is currently only patchy, often anecdotal, information about how practitioners are using AI tools.

However, the Solicitors Regulation Authority Risk Outlook report notes that, at the end of 2022, 75% of the largest solicitors firms were using AI – nearly twice the number three years ago.[18] The same report suggests that over 60% of large law firms were exploring the potential of the new generative AI systems, as were a third of small firms.

Lawyers wanting to benefit from document review using AI tools will have to ensure they have permission to feed information into the tool (e.g. secure model purchased by an organisation as opposed to ChatGPT).

The Bar Council considerations acknowledges that confidential information or information subject to legal professional privilege should not be put into LLMs as it could be used as training data and repeated in future results.[19]

Legal standards and liability

Clarity in legal standards and regulations concerning the use of AI within the family justice system may be required. Governance is crucial.

For example, there are legal implications about the extent of automation in decision making systems. Article 22 of the GDPR concerns automated individual decision-making and gives people the 'right not to be subject to a decision based solely on automated processing' (though there are exceptions such as the individual giving their consent).[20] However, the picture is less clear when decision-making systems are semi-automated, as it often the case.[21]

[16] Samsung has restricted usage of ChatGPT following an engineer inputting confidential code into the model – see: The Economist. (2023, 27 November). *Generative AI generates tricky choices for managers.* https://www.economist.com/business/2023/11/27/generative-ai-generates-tricky-choices-for-managers

[17] See: Information Commissioner's Office and The Alan Turing Institute. (2022). *Explaining decisions made with AI.* https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/

[18] Solicitors Regulation Authority. (2023). *The use of artificial intelligence in the legal market.* https://www.sra.org.uk/sra/research-publications/artificial-intelligence-legal-market/

[19] Bar Council 2024.

[20] General Data Protection Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. (2016). Official Journal, L 119, 1-88. http://data.europa.eu/eli/reg/2016/679/oj

[21] For further details, see Palmiotto, F. (2024). When is a decision automated? A taxonomy for a fundamental rights analysis. *German Law Journal*, pp. 1-27. https://doi.org/10.1017/glj.2023.112

Transparency

Being transparent about when AI has been used is likely to be critical across a number of domains in order to secure public trust and allow for effective regulation, for example when drafting reports or using algorithms to support with decision-making. The Ada Lovelace Institute recommends 'transparency labelling' for AI-generated content.[22]

In order to understand the impact AI technologies, it is necessary to record when and how they are being used. Regarding the family justice system, researchers can only examine how cases vary if litigants in person are using AI, if the system collects information about when they are doing so.

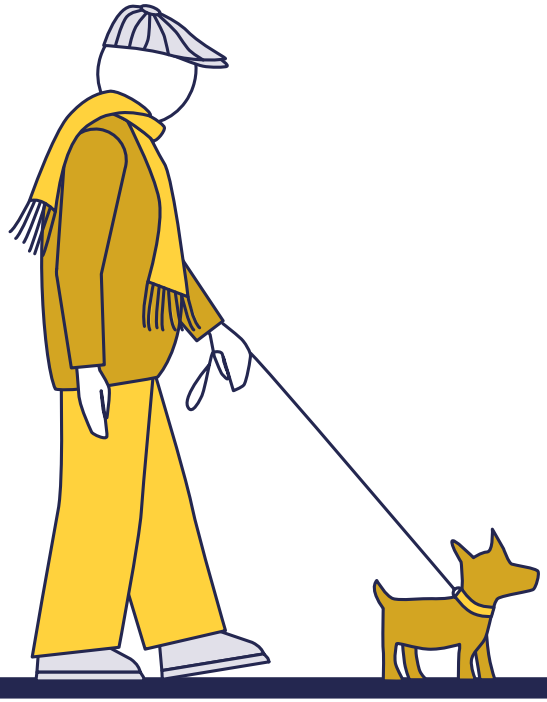
Interpretability

AI models used in risk assessment and decision support can be highly complex, making it difficult for both professionals and the public to understand how decisions are reached, which can be a barrier to effective use. Professionals responsible for decisions taken based on AI output will need to understand how results are reached to evidence their decisions.

Public trust

Ensuring that the public and legal professionals have trust in AI systems used in the family justice system is paramount. Negative experiences or perceptions of AI may erode this trust.

The Ada Lovelace Institute has done some research about public attitudes to AI in Britain and found that support for AI depended on how it was being used. People were broadly positive about the majority of uses covered by the survey – for example, using AI to detect cancer was seen as beneficial by 9 out of 10 people – but there were concerns about some uses, such as driverless cars and robotic weapons.[23] The research highlighted that the public believe that AI systems should be transparent.



[22] Davies, M. and Birtwistle, M. (2023, July 18). *Regulating AI in the UK.* Ada Lovelace Institute. https://www.adalovelaceinstitute.org/report/regulating-ai-in-the-uk/#_ftnref10

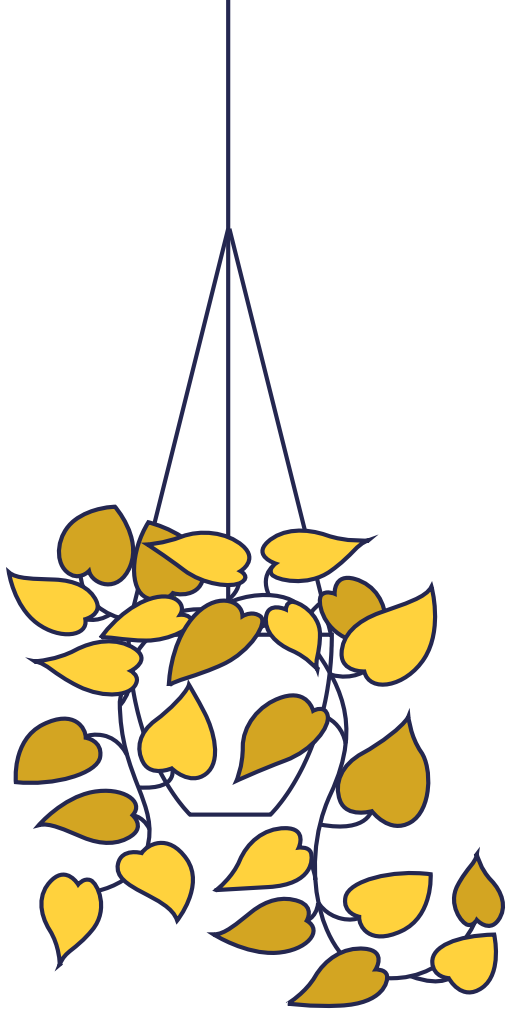
[23] Modhvadia, R. (2023, 6 June). *How do people feel about AI.* Ada Lovelace Institute and The Alan Turing Institute. https://www.adalovelaceinstitute.org/wp-content/uploads/2023/06/Ada-Lovelace-Institute-The-Alan-Turing-Institute-How-do-people-feel-about-AI.pdf

Accuracy

"Research by a firm called Vectara found that the main LLMs hallucinate or invent information between 3–30% of the time." – Metz 2023

Even when giving correct information, LLMs may provide advice that is not relevant. For example, legal information used to train LLMs has a US focus due to the nature of the data used to train the models – this may be irrelevant to people in England and Wales.[24]

People seeking legal support or advice from LLMs may not be aware of the accuracy and relevance of information from generative AI. An additional accuracy risk for the family justice system is AI tools being used to produce fake text, images and videos. Guidance for the judiciary notes that courts have always had to deal with forgeries and allegations of forgery, and advises judges to be aware of the new risks brought about by technology, which increases the ease of producing fake output.[25]



[24] Courts and Tribunals Judiciary 2023.

[25] Ibid.

What about governance?

Effective governance is required to ensure that challenges are successfully addressed, allowing any gains to be realised in a way that does not undermine public trust, professional accountability or access to justice.

As technologies and their uses continue to evolve, so the governance framework will need to adapt.

The following section outlines the current governance landscape in England, Wales and the European Union.

Box 3 of the Ada Lovelace Institute's paper *Regulating AI in the UK* has a summary of other countries' approaches to AI regulation.
(Davies and Birtwistle 2023)
https://www.adalovelaceinstitute.org/report/regulating-ai-in-the-uk/#_ftnref10

Principles v Rules

UK

In March 2023, the UK government published AI regulation: A pro-innovation approach, which outlined five principles:[26]

- safety, security and robustness
- appropriate transparency and explainability
- fairness
- accountability and governance
- contestability and redress.

The UK approach is based on existing regulators adapting their frameworks to incorporate the five principles with central functions in government supporting the regulators – for example by monitoring the overall framework's effectiveness.[27]

European Union

In contrast to the UK's principle-based approach, the EU has proposed a rules-based approach to AI regulation.

The Artificial Intelligence Act proposes that AI uses are categorised based on the level of risk they pose to users:

- unacceptable risks, such as social scoring based on a person's socioeconomic status
- high risk in sensitive sectors such as law enforcement or healthcare
- generative AI such as ChatGPT which needs to meet transparency requirements e.g. disclosing when content has been generated by AI
- limited risk.

In the justice context, using AI to assist with legal interpretation and application of the law is classified as high risk and one of eight uses that require registration in an EU database, as well as ongoing assessment.

[26] Department for Science, Innovation and Technology and Office for Artificial Intelligence. (2023). *AI regulation: A pro-innovation approach.* GOV.UK. https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

[27] See Davies and Birtwistle 2023 for discussion and recommendations about regulating AI in the UK.

Responsible AI application

The Alan Turing Institute's 2019 guide for the responsible design and implementation of AI systems in the public sector sets out values to underpin responsible delivery of AI system (respect, care, connect and protect).[28] It also includes principles for governing the ethical use of AI (fairness, accountability, sustainability, transparency) and a governance framework.

As the broad application of the principles and values above indicate, many aspects of AI regulation will apply across the public sector, though some issues will need specific thought regarding the family justice system. For example, guidance for civil servants (including lawyers) using generative AI was published in January 2024, emphasising that no sensitive or personal information should be entered into LLMs and to always treat outputs with caution.[29] The guidance mentions a number of specific uses, including using generative AI for research (which is contrary to guidance for the Judiciary that describes AI as a 'poor way' to conduct research,[30] summarising information, developing code and textual data analysis.

The family justice system may need its own guidelines/ethical principles – for example around the circumstances in which lawyers can use models such as ChatGPT.

Earlier this year, the Bar Council published considerations when using ChatGPT and generative AI states that barristers who use LLMs should do so 'responsibly' and weigh up the potential risks and challenges given their professional responsibilities.[31] The guidance advised barristers using LLMs to check and verify the models' outputs, respect legal privilege, confidential information and data protection compliance. The considerations end with a reminder to barristers to keep up to date with civil procedure rules, which they note may be updated to include parties disclosing where they have used generative AI in preparation of materials.

[28] Leslie 2019.

[29] Central Digital & Data Office and Cabinet Office. (2024). *Guidance to civil servants on use of generative AI.* GOV.UK. https://www.gov.uk/government/publications/guidance-to-civil-servants-on-use-of-generative-ai/guidance-to-civil-servants-on-use-of-generative-ai

[30] Courts and Tribunals Judiciary 2023, p. 7.

[31] Bar Council 2024.

Conclusion

There are numerous potential benefits of using various AI tools to support families and professionals in the family justice system, however these do not come without risks. There is also a question of how people in family proceedings and professionals working in the family justice system are prepared for AI. We know that AI technologies are available, accessible and already used by many. However, we do not know precisely how AI is being used by users of the family justice system, nor do we know the efficacy of the tools for dealing with requests made about family justice. At the moment there is not a vision for how those in the family justice system, both families and professionals, can safely harness the benefits of AI while being protected from the risks. Research shows that people are concerned about risks posed by AI and the majority welcomed regulation of AI to mitigate such risks.[32]

Before any regulation is agreed, given the potential challenges around public perceptions about AI, it will be important to include public engagement and collaboration with other parts of the justice system and public sector, such as academia, criminal justice, international partners.

[32] Modhvadia 2023.

References

Alexander, N. (2019). Ten trends in international mediation. *Singapore Academy of Law Journal*, 31. https://journalsonline.academypublishing.org.sg/Journals/Singapore-Academy-of-Law-Journal-Special-Issue/Current-Issue/ctl/eFirstSALPDFJournalView/mid/503/ArticleId/1466/Citation/JournalsOnlinePDF

Bar Council. (2024). *Considerations when using ChatGPT and generative AI software based on large language models.* https://www.barcouncilethics.co.uk/wp-content/uploads/2024/01/Considerations-when-using-ChatGPT-and-Generative-AI-Software-based-on-large-language-models-January-2024.pdf

Barendrecht, M. (2017, 21 June). *Rechtwijzer: Why online supported dispute resolution is hard to implement.* HiiL. https://www.hiil.org/news/rechtwijzer-why-online-supported-dispute-resolution-is-hard-to-implement/

Central Digital & Data Office and Cabinet Office. (2024). *Guidance to civil servants on use of generative AI.* GOV.UK. https://www.gov.uk/government/publications/guidance-to-civil-servants-on-use-of-generative-ai/

Clayton, V., Sanders, M., Schoenwald, E., Surkis, L. and Gibbons, D. (2020). *Technical report: Machine learning in children's services: Does it work?.* What Works for Children's Social Care. https://whatworks-csc.org.uk/wp-content/uploads/WWCSC_technical-_report_machine_learning_in_childrens_services_does_it_work_Sep_2020.pdf

Courts and Tribunals Judiciary. (2023). *AI guidance for judicial office holders.* https://www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf?utm_source=substack&utm_medium=email

Davies, M. and Birtwistle, M. (2023). *Regulating AI in the UK.* Ada Lovelace Institute https://www.adalovelaceinstitute.org/report/regulating-ai-in-the-uk/#_ftnref10

Department for Science, Innovation and Technology and Office for Artificial Intelligence. (2023). *AI regulation: A pro-innovation approach.* GOV.UK. https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

European Parliament. (2023, December 19). *EU AI Act: First regulation on artificial intelligence.* https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Evans, J. & Ndegwa, A. (2022). *Use of Technology in the family justice system annotated bibliography.* Department of Justice Canada. https://www.justice.gc.ca/eng/rp-pr/jr/utfjsab-utsjfba/pdf/RSD2022_Use_of_Technology_in_the_Family_Justice_System_Report_EN.pdf

General Data Protection Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. (2016). Official Journal, L 119, 1-88. http://data.europa.eu/eli/reg/2016/679/oj

Grogger, J., Ivandic, R. and Kirchmaier, T. (2020). *Comparing conventional and machine-learning approaches to risk assessment in domestic abuse cases.* Centre for Economic Performance. https://cep.lse.ac.uk/pubs/download/dp1676.pdf

Information Commissioner's Office and The Alan Turing Institute. (2022). *Explaining decisions made with AI.* ICO. Retrieved from https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/part-1-the-basics-of-explaining-ai/legal-framework/#legal_framework_2

Joint Technology Committee. (2020). *Introduction to AI for courts.* National Center for State Courts. https://www.ncsc.org/__data/assets/pdf_file/0013/20830/2020-04-02-intro-to-ai-for-courts_final.pdf

Leavy, S., O'Sullivan, B. and Siapera, E. (2020). *Data, power and bias in artificial intelligence.* arXiv. https://arxiv.org/abs/2008.07341

Lee, J. (2024). *Fabricated judicial decisions and 'hallucinations' – a salutary tale on the use of AI.* Financial Remedies Journal Blog https://financialremediesjournal.com/content/fabricatedjudicial-decisions-and-lsquo-hallucinations-rsquo-ndasha-salutary-tale-on-the-use-of-ai.1265d5deeb39450bbdc059ad5ae69818.htm

Legalese Decoder. (n.d.). Home. https://legalesedecoder.com/

Leslie, D. (2019). *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector.* The Alan Turing Institute. https://doi.org/10.5281/zenodo.3240529

Maini, V. (2017, 19 August). *Why machine learning matters. Machine Learning for Humans.* https://medium.com/machine-learning-for-humans/why-machine-learning-matters-6164faf1df12

References

Modhvadia, R. (2023, June 6). *How do people feel about AI.* Ada Lovelace Institute & The Alan Turing Institute. https://www.adalovelaceinstitute.org/wp-content/uploads/2023/06/Ada-Lovelace-Institute-The-Alan-Turing-Institute-How-do-people-feel-about-AI.pdf

Metz, C.. (2023, 6 November). *Chatbots learn to hallucinate.* The New York Times. https://www.nytimes.com/2023/11/06/technology/chatbots-hallucination-rates.html

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M.-E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., Broelemann, K., Kasneci, G., Tiropanis, T. and Staab, S. (2020). *Bias in data-driven artificial intelligence systems – An introductory survey.* Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. https://wires.onlinelibrary.wiley.com/doi/10.1002/widm.1356

Okidegbe, N. (2022). *The democratizing potential of algorithms.* Connecticut Law Review, 53, 739. https://scholarship.law.bu.edu/faculty_scholarship/3138

Palmiotto, F. (2024). When is a decision automated? A taxonomy for a fundamental rights analysis. *German Law Journal, 1-27.* https://doi.org/10.1017/glj.2023.112

Solicitors Regulation Authority. (2023). *The use of artificial intelligence in the legal market.* https://www.sra.org.uk/sra/research-publications/artificial-intelligence-legal-market/

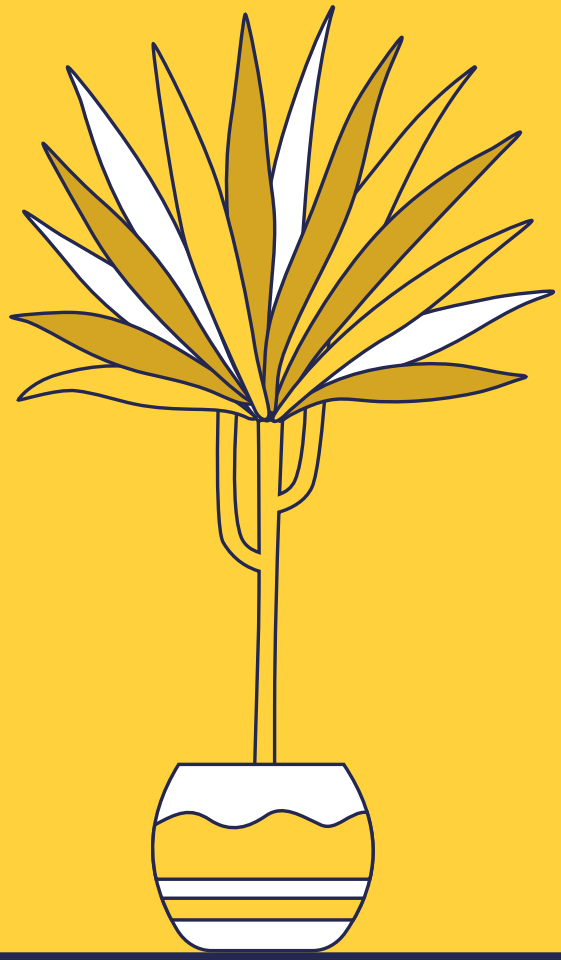
Stanley, D. (2024, 8 February). *What we don't talk about when we talk about AI.* Joseph Rowntree Foundation. https://www.jrf.org.uk/ai-for-public-good/what-we-dont-talk-about-when-we-talk-about-ai

The Lawyer Portal. (2023, 23 October). *Law firms that are leading the way with AI technology.* https://www.thelawyerportal.com/blog/law-firms-that-are-leading-the-way-with-ai-technology/

Nuffield Family Justice Observatory

Nuffield Family Justice Observatory (Nuffield FJO) aims to support the best possible decisions for children by improving the use of data and research evidence in the family justice system in England and Wales. Covering both public and private law, Nuffield FJO provides accessible analysis and research for professionals working in the family courts.

Nuffield FJO was established by the Nuffield Foundation, an independent charitable trust with a mission to advance social well-being. The Foundation funds research that informs social policy, primarily in education, welfare, and justice. It also funds student programmes for young people to develop skills and confidence in quantitative and scientific methods. The Nuffield Foundation is the founder and co-funder of the Ada Lovelace Institute and the Nuffield Council on Bioethics.



For further information or to get involved, please get in touch:
E contactfjo@nuffieldfoundation.org T +44 (0)20 7323 6242

100 St John St, London EC1M 4EH T: 020 7631 0566

Part of Nuffield Foundation: Registered charity 206601
nuffieldfjo.org.uk | @NuffieldFJO
www.nuffieldfoundation.org | @NuffieldFound

Nuffield
Family Justice
Observatory