AI-Curiosity on the Bench: Judicial Boundaries, Opportunities, and Concerns from UK Focus Groups

Erin T. Solovey
Worcester Polytechnic Institute

Brian Flanagan
Maynooth University

Daniel Chen
Toulouse School of Economics

Abstract

The growing use of AI in courts is changing how judges and other legal professionals carry out their work. Systems based on AI have been suggested for tasks including legal research, decision support, evidence assessment, and case management. Yet there has been relatively limited empirical understanding of how such technologies align with judges' everyday professional practices, responsibilities, and roles. Drawing on a formative focus group study with 12 judges from the United Kingdom—including 5 justices of the UK Supreme Court—this chapter explores how AI is perceived across different judicial roles, from first-instance hearings to appellate decision-making. The findings reveal a nuanced picture of curiosity, caution, and conditional openness: judges identify areas of judicial work where AI support is viewed as potentially valuable, as well as functions that judges regard as fundamentally human and non-delegable. They also highlight key concerns relating to reliability, legitimacy, and professional skills. By foregrounding judges' own accounts, the chapter offers insights relevant to judicial training, court governance, and the responsible integration of AI into the administration of law in the UK.

1 Introduction

In his 1986 essay, 'Violence and the Word', Robert Cover famously reminded us that '[l]egal interpretation takes place in a field of pain and death' (1986, 1601). This is no less true of legal interpretation conducted or supported by artificial intelligence. Reform that would reengineer the practice of judicial reasoning in the ways now envisaged by legal technologists demand careful consideration from an array of perspectives. This chapter brings the bench's own perspective to those questions, drawing on qualitative focus groups with 12 United Kingdom judges, including five justices of the UK Supreme Court that aimed to explore judicial perceptions of AI's risks and opportunities for legal decision-making. These findings were originally reported in the Proceedings of the 4th Annual Symposium on Human-Computer Interaction for Work (CHIWORK '25): <https://doi.org/10.1145/3729176.3729192>[28]. We draw heavily on that paper here.

Justice systems worldwide have entered a period of crisis and opportunity. Judicial workers struggle with increasing case backlogs while ordinary people face the prospect of delays that impede their access to justice and undermine the rule of law. Conversely, with the arrival of generative AI, justice systems have been presented with a technology which might allow them to alleviate backlogs but which comes freighted with risks of its own for the legal system's legitimacy. Preliminary consideration of these risks

and opportunities has so far taken the form of philosophical analysis [12], benchmarking projects [14], and experiments on the comparative performance of human lawyers with and without chatbot assistance [9]. We advance the inquiry by reporting a qualitative study of judicial perceptions of how the integration of AI might transform the way judges and legal professionals work.

AI tools are increasingly used for tasks such as legal research, decision support, evidence analysis, and case management. However, in high-stakes contexts like judicial decision-making, these tools also raise concerns. Although there is extensive research on the technical capabilities of AI, less attention has been paid to understanding the human factors involved in their integration with judicial systems. Judges, clerks, and other judicial staff may have differing perceptions of AI's benefits and risks, influenced by their respective roles and tasks. Few studies explore the nuanced interplay between human and AI strengths for specific judicial tasks. The introduction of AI may shift traditional job roles, potentially automating some functions while creating new opportunities for specialized human involvement.

In this chapter, we report what judges across court levels say about where AI may enhance efficiency and access to justice and where human responsibility and presence must remain central. We explore the processes, beliefs, experiences, and needs of judicial professionals regarding the future of AI in their work and identify where judges themselves draw normative boundaries around acceptable AI assistance. We aimed to identify task-specific considerations for the appropriate use of AI in judicial work, particularly where such tools might support rather than replace human judgment and how they can respect the complexities of judicial reasoning and preserve critical human oversight. Exploring perceptions across different judicial roles highlights how acceptable forms of AI assistance vary by court level, task, and case type. These perspectives offer insights into how courts can integrate AI in ways that align with user needs and existing norms. Further, discussing the future of work in judicial contexts points to emerging implications for judicial training and professional development, as well as for how judicial roles may evolve in response to new AI capabilities.

2 Related Work

Legal theorists offer contrasting visions of AI's prospective contribution to judicial work. Alongside jurisprudence that emphasizes the risks of introducing such technologies [25], one finds philosophical analyses that highlight the efficiencies that they might bring [23]. Empirical work has broadly confirmed that the recent development of chatbots such as ChatGPT have brought us significantly closer to the creation of artificial legal intelligence. Whereas chatbots are prone to serious errors that humans are not, notably, the invention of fictitious legal authorities[11], they have also been shown to match human performance across a wide range of discrete legal tasks, such as writing [1], problem solving [21], and annotation [24]. There is even evidence that in hard cases chatbots simulate people's tendency to equivocate between the law's letter and the law's spirit [2].

It is perhaps unsurprising, therefore, that polling evidence suggests that 'judges, judicial support staff, prosecutors, and lawyers around the globe have started to use chatbots... to draft... judicial decisions, and elaborate arguments' [15]. In line with this trend, individual judges in several jurisdictions have sought to place the exploration of AI's potential on the public agenda, e.g., '[Judges] should consider whether

and how AI-powered large language models... might... inform the interpretive analysis' Judge Kevin Newsom, Court of Appeal 11th Circuit, USA, 2024. But empirical inquiry into judicial perceptions of legal AI is still developing.

One notable initial advance has been a representative survey of the Portuguese judiciary, in which judges' responses indicated a wariness of robot judges together with an enthusiasm for AI judicial clerks [18]. This combination of survey answers attests to the scope for nuance and invites the application of a qualitative method that might be better suited to 'elicit fine-grained, practice-informed insights' [8]. Indeed, in a recent speech, the former Deputy President of the UK Supreme Court, Lord Patrick Hodge suggested that 'collaboration between... judges... and academics (both legal scholars and computer scientists) offers the best prospect of facilitating and harnessing the new technology' [16]. This chapter seeks to start an AI research agenda in law that answers this call for collaborative research.

3 Formative Focus Group Study: Judges and AI

This chapter draws on a formative focus group study conducted in 2025 to understand judges' perception of the incorporation of artificial intelligence into their future work. The objectives of the study were the following:

(1) To learn about the way judges currently work without AI
(2) To understand the areas of their work that are particularly appropriate or necessary for the human to lead
(3) To identify the tasks that are likely going to integrate AI in the future and the benefits this could achieve
(4) Across the different job roles, to uncover concerns and important considerations for AI in judicial work

3.1 Participants

The study consisted of 3 focus groups with between 2 and 6 participants. One focus group was conducted online via Microsoft Teams and the other two were conducted in-person in London, in the UK Supreme Court Building and in the Royal Courts of Justice respectively. All participants were current judges in the United Kingdom legal system, and included 5 members of the UK Supreme Court, 1 member of the Court of Appeal, 5 members of the High Court, and 1 member of the County Court (Circuit judge). The study received ethics approval from the Harvard University Institutional Review Board (IRB24-1745) and Maynooth University Research Ethics Committee (SRESC-2025-39914), and all participants signed an informed consent form.

3.2 Procedure

During each 60-minute focus group session, two researchers co-moderated the discussion, guided by a set of pre-determined questions (see Appendix). Follow-up questions were asked when appropriate to delve deeper into topics that were discussed by the participants. Audio from the session was recorded.

The first part of the focus groups looked at current workflows without AI, aiming to gain a detailed understanding of how judges currently perform key tasks, including their processes and challenges, and to identify critical aspects that need to be preserved. The second phase aimed to understand perceptions

of AI, including its potential benefits and risks, and how these vary across roles. We then looked closely at specific tasks that came up in the discussion to determine the unique strengths of humans and AI in the context of a specific judicial task. Next, we looked at future implications of AI for judicial roles, assessing how participants envision their roles changing with AI integration, including the creation of new roles, the decline of certain tasks, and the potential for AI to improve or complicate their work. At the end, we allowed participants to share any additional thoughts or insights, summarized key discussion points and asked for any clarifications before closing.

3.3 Data Analysis

Audio recordings of the in-person sessions were transcribed using Open AI Whisper [19], while the online session was transcribed via Microsoft Teams. These auto-generated transcripts were then reviewed manually and errors were corrected and any identifiable comments (e.g., names) were removed. To uncover themes that emerged from the data, the focus group transcripts were analyzed using inductive thematic analysis based on the grounded theory framework [4]. One researcher coded the data using ATLAS.ti Version 9.1.3 for Mac [13] and ATLAS.ti Web, which facilitate qualitative analysis. This involved a *data familiarization* phase where the researcher reviewed the audio and transcripts, *initial coding*, *code refinement and recoding*, and finally, *theme identification*. The initial codes were based on the study objectives and included work tasks, AI perceptions, AI benefits and risks, human and AI strengths, and changing job roles. Further codes emerged from the data. To protect confidentiality, we do not attribute quotations to particular participants.

4 Results

This section discusses the focus group findings, starting with descriptions of current judicial roles, responsibilities and functions. We then set out findings related to human and AI strengths for various tasks as well as judges' perceptions of AI and its benefits and risks.

4.1 The Work of Judges

Task Category	Description and Subtasks
Legal Research & Pre-hearing Prep.	Reviewing case files, legal briefs, and relevant precedents before a hearing.
Courtroom Duties	Presiding over hearings, instructing jurors, ensuring fair hearing
Deliberation and Decision Making	Discussing cases, analyzing legal arguments, forming opinions
Judgment Writing	Summarize the facts, arguments, issue, and decision, drafting and revising
Writing for Other Audiences	Writing press briefs, explaining to public, creating child-friendly summaries

Sentencing	Reviewing any guidelines or precedent, identifying relevant considerations
Administrative and Managerial	Overseeing court staff, reviewing paperwork, creating orders, managing cases
Training and Development	Supporting judicial assistants, attending seminars
Public Engagement and Outreach	Writing and delivering speeches
Meetings	Administrative, leadership, outreach, committee meetings

Table 1: Core Functions of Contemporary Judicial Work

The first focus group began with a discussion of the day-to-day tasks and workflows of judges in different roles. In subsequent focus groups, we presented the tasks and phases of work that had emerged from the first focus group and asked for any additions or clarifications to ensure that we gained an inclusive set of task categories to describe the work of judges at the respective levels. The resulting categories can be found in Table 1.

Beyond this list of tasks, some critical aspects and values in judges' current work—without AI—were also discussed. For example, one participant stated "*judges are personally responsible for everything that goes out in their name and therefore they have to check everything that goes out in their name*". The application of this principle to the integration of AI seemed to be on the mind of most participants, and is stated clearly in the recently released guidance document for the use of AI in the UK judiciary [5]. Another value that came up several times concerns the writing of judgments in the highest courts. There, the judge fully writes the judgment and puts great effort into the language used, the reasoning, etc. In lower courts, the task of establishing the relevant facts is more central. It is a tenet of the Rule of Law that litigants must receive normatively acceptable reasons for a legal outcome (e.g., [20, 22]), but how this tenet is understood varies between jurisdictions.

4.2 The Value of Humans in Judicial Work

While AI has many potential uses in judicial work that will be discussed in Section 4.3, the focus groups revealed several areas where the human factor is critical and this should not be ignored when developing AI support tools. We begin with these important constraints.

4.2.1 Justice is rooted in human decision making and reasoning. At a fundamental level, the work of judges comes down to the "*evaluative judgment that you ultimately make... I don't think AI could do that. No, I hope not."* While there are strengths in AI's ability to analyze facts and potentially make logical decisions, one participant noted that *"law is not a matter of pure logic. It's a matter of practical reasoning"*. Thus, there is a sense that humans are deeply aware of human values and can identify situations when a line of reasoning will lead to an unjust decision: *"if logic is driving them to that end result, and the end result looks wrong, something's gone badly wrong... You're given the job, not just for intellectual ability, it's the judgment that you can see that logic is taking you in a direction that you shouldn't be going and you need a practical, humane result to a problem if it's humanly possible."* There was a feeling that AI would not have this capability.

We assumed that this hesitation about the quality of AI legal decision-making was due solely to the perception that AI would be unduly 'mechanical' in its application of legal rules, perhaps unable to identify intuitively exonerating circumstances. On reflecting again on the focus group transcript, however, another comment that we had not earlier reported also stands out: *"a critical part of writing the clever bit of a judgement is to think through the result in a human way. Whereas if what the AI is doing is working on a sort of predictive basis... it's just not reasoning and thinking in the way that is integral to the process of coming to a judgement on a case*." This further comment suggests a second jurisprudential concern about legal quality may also be at play.

Justice Oliver Wendell Holmes of the US Supreme Court famously prioritised the perspective of the 'bad man' who was satisfied with accurately predicting the behaviour of courts: "The prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law" [26 at 461]. Later, in The Concept of Law (1961), HLA Hart rejected this view as overly reductive: "Why should not law be equally if not more concerned with the 'puzzled man' or 'ignorant man' who is willing to do what is required [as it is with the 'bad man' who cares only for what a court is likely to do]" [27, 40]. The judge's perceived contrast between complex legal reasoning and superficially similar but prediction-powered AI outputs seems to echo this longstanding philosophical disagreement. It suggests a worry that the critical attitude inherent in legal judgment may always elude replication by techniques of next token prediction.

Some discussion considered current practices related to human assistants (clerks) to contextualize possible future assistance by AI. In contrast to the U.S. Supreme Court [10], UK Supreme Court judges typically write their judgments by themselves. One participant stated, "*There's no question that anything in a judgment that I hand down will be written by anyone other than me. My judicial assistant will do research for me and maybe give me an analysis of cases, but I will then go to the cases. Similarly, the judicial assistant might produce a chronology, but I will go to the individual documents when I'm writing the judgment. So the tradition in the UK is very much for the judge to write the document."* This was echoed by most of the judges at this level. In light of the existing emphasis on exclusive self-authorship, and the great care with which words are chosen, it may be less of a priority to have AI support for the drafting of legal reasoning. Another participant noted that "*I can't imagine it actually replacing most of the judgment writing that we do. Because each of us, I think, enjoys writing in possibly our own style. Now you could say, as I understand it, write this judgment in my style. [laughing]... We regard [the written judgment] as being an element of judgment that one might find difficult to replicate.*" Of AI's ability to articulate the reasoning behind a decision, a participant noted that "*AI isn't really undertaking that process."*

However, perceptions change at other court levels, where case volumes are much higher and the time that is currently devoted to judgment writing may be lower. For courts of first instance, which, unlike appellate courts, also perform the function of resolving disagreement over a case's facts, AI support for opinion drafting may make more sense, particularly to improve efficiency.

4.2.2 ***The human component in justice holds value***. Several participants noted that in some types of cases, a human judge is vital to providing emotional and psychological closure, and a sense of "dignity". A participant in Focus Group 1 noted "*people take comfort from having a human face, a human decision maker, listening to what they have to say, hearing them and making a human judgment based on the*

evidence. And I doubt whether AI will achieve that cathartic role that human justice does." A similar sentiment was expressed in the second focus group: *"You can't underestimate the catharsis that there is in a trial and the importance of that for peaceful dispute resolution so that the person who loses at the end of the day can say, well, I had a fair hearing. I understand why I lost. I don't agree with it, but I can move forward now."*

There also was a belief that people expect a human component in important legal decisions, particularly where a level of human empathy and understanding is important. *"I can imagine one might say that ultimately we would conclude that it's a matter of human rights to say the decisions about, for example, taking children away from families are decisions. It doesn't matter whether the AI would do it better. It doesn't matter if actually the AI would be less swayed by the well, you know, the human factors, because actually we want a decision as a matter of principle made by a human being."* However, there was an awareness from the judges that you could *"imagine a world in which society does accept that, you know, 15 years of everything else being done by [AI] decision making. That's going to affect things. I just say it's a moving picture there."*

4.2.3 ***The value of the human component is case-dependent and role-dependent.*** When evaluating the strengths of humans over AI for a particular task, participants noted that it can depend on the context. The expectation that a human judge would decide does not necessarily apply to all cases: "*It's one thing to have cheap and cheerful AI tool to resolve a £500 dispute over a second hand car sales contract. It's quite another if somebody's being sent to prison or somebody's having their children taken away from them and put into care. And I really struggle to, to see the role for AI in the decision making process in those sorts of cases, partly because everyone would expect there to be a human component, because it's an intensely human judgment that has to be made."* In another discussion, AI was noted as valuable for laying out the factual information for a case, to enable judges to focus on the judgment and legal questions. However, this was more true in appeals courts where the facts have already been determined. In first instance courts, a key role of the judge is to uncover the facts of the case, and this was an area that they would not delegate to AI: "*there's a big difference at what level we're talking about... one of the most important functions that first instance judges do is find the facts. And you couldn't rely on AI to set them out for you, because it may depend on who's telling the truth and not knowing the truth and all that. But when it comes on appeal, it is extremely tedious to set out the facts."* This was re-iterated by a judge that mostly handles such cases: *"that's the main part of our job, you know, in the county court, in the case, is finding the facts. And I would put that very much on the human side of the boundary."*

4.2.4 ***Decision making is collaborative with multiple points of view.*** Human collaboration and the social element were discussed as a critical aspect of judicial work at the highest level. One participant mentioned that *"there is a big social side to the job as well"* and proposed that if AI assists with other tasks, there might be more time for this human interaction. Another judge said *"the thing is the interaction between us, you know, before we go in and sit in a case, and when we come out of a case, we all meet together and discuss what we think about it and why. And that's, we can't have a room of robots doing that."* Notably, for the most difficult decisions, the judicial system has been designed to bring in more judges as a case ascends the levels of appeal. Typically, the Supreme Court has five judges on a panel and the Court of Appeals often has three; one participant noted that "*you are more likely to be right with five judges over one*." This led to a comment that *"it allows for plurality of opinions"* with a participant stating *"Now, in an AI world, I presume there's no point"* because the algorithm would be

likely to have a single line of reasoning. Another judge stated, "*if you went and looked up some decision in the Supreme Court where the judges are divided 3-2 with different reasoning and you came along, here's the AI answer. How are you ever going to say we're now doing it better? ... Conceptually, I just don't see how we'll ever say, certainly not in our legal careers, OK, AI can now do a complicated point of law with reasoning from lots of different traditions better.*"

4.3 AI Opportunities and Benefits

Alongside the emphasis on the human dimension of legal reasoning (Section 4.2), there was also a general sense of opportunity regarding the use of AI to improve efficiency and access to justice, across essentially all the tasks in Table 1. One judge summarized it as "*I mean, all of the these can be summarized as increasing productivity, reducing cost and reducing some of the drain on resources that we all have.*"

While primary legal decisions, and their written form, were considered a fundamentally human task, there are follow up tasks that might be well-supported by AI, such as creating a version that is written for the public or for a child, in cases involving children. This could also extend into the tools available to create other formats such as podcasts or videos that help to bring important topics from the judiciary to wider audiences. In addition, for high volume courts, the initial judgment drafting or summarizing of background was considered a potential boost to efficiency and better proofreading would also be helpful. There is also potential for "small claims" and some other types of cases to be fully resolved through AI, with a possible tradeoff between efficiency and quality of judgment.

Sentencing was identified as an area that AI could support, as AI could analyse the relevant background information, precedents and additional considerations and make a recommendation. There could be similar support for deciding what would qualify as a fair amount in settlement agreements (e.g. personal injury). Further, many "boring" or bulk administrative tasks were identified as areas in which AI could be beneficial.

Legal research and summarization of cases, disclosures, and bundles of documents was an area of much discussion. There is hope that AI could assist with these, but also currently a lack of trust that it would do so reliably without hallucination. Nevertheless, all judges agreed that there is sometimes an impossible amount of documents to read and also that AI may find things a human would require excessive time to discover. In addition, AI might remind judges or other legal professionals about aspects of the case that have been overlooked. As will be discussed below, however, there is also a concern that over-reliance on AI for such tasks could lead to de-skilling.

In relation to work within the courtroom, there was discussion of AI generating and presenting instructions to the jury for different types of cases and situations. From an administrative perspective, some judges saw exciting potential for AI to "interrogate" recently digitalized legal data and provide insights into the overall state of the judicial system.

Category	Supporting Quotes

Increase Consistency	"And that would have the advantage of ensuring consistency, of course." "We're kind of going in two different directions because the whole line of sentencing guidelines is to get rid of subjectivity and the huge variation."
Improve Information	"We can look up a case, and it'll tell you what other cases it's been referred to." "I think it can also stop going around down a kind of rabbit hole where actually AI can say, no, no, that provision was actually repealed by the time that this happened." "There are cases where it's not realistic to think that any human being genuinely can actually amalgamate, can assimilate that material. And the AI can at least... identify relevant material out of that mass."
Increase Efficiency	"Would save a huge amount of time." "A considerable time-saving device and something that would help increase productivity." "But one thing I'm absolutely certain AI could do that would save me 10 percent of my time is to take the paperwork and just make a suggestion, a short summary and suggestion of what the right answer is."
Increase Access to Justice	"There's no secret there are real resource constraints. There are backlogs in a number of courts. So anything that can improve efficiency and productivity whilst ensuring we don't lose the essence of what justice is, is exciting and to be welcomed." "We have a million and a half small money claims issued into the civil justice system every year. And the vast majority of them never get to a judge... And we could, you could imagine an AI resolving a lot of those. Now, not all of them... [but] most of them are very simple. Most of them, plenty are not."
Increase Understanding of Judicial Effectiveness	"The ability of AI to process the data, to give us evidence-based material on which sensible decisions can be made will be important in running the court service." "The aspect that I am most excited about is the opportunity with rich data to interrogate the data about how we do justice and to identify ways in which we might do it better, to identify whether there are inequalities in the way in which the justice system works. And I think there are opportunities to do that which we've just not had before, partly because we haven't had the data and partly because we haven't had the technology."
Reduce Bias	"AIs don't have other biases and the bias which we're talking about in this context, which might matter most is confirmation bias. Because when you're looking, when you're hunting through some material, you're immediately in a confirmation bias risk type situation, because you're already beginning to think you know what the answer is. The AI doesn't know what you think the answer is."

	"For the mass majority of sentencing decisions, for more minor offenses, I actually think having a human involved is negative, because it allows all of those individuals with their slightly subjective preferences to allow the decision to be affected by those.."
Reduce Cost	"Small value claims are uneconomic to pursue. Because if they require the intervention of a professional lawyer... AI seems to be one way in which Western countries could enhance access to justice because many people are never going to be able to afford a lawyer." "You could imagine an AI resolving a lot of those [small claims]."
Reduce Tedium	"There is a lot of boring stuff that we've still got to do manually." "AI can do all of the extensions of time... the mundane stuff.. if it goes wrong, it's reversible." "English judgments... they're too long because three-quarters of them are facts, what the statute may say, and then what the arguments were. I'd like to get to a system where I don't have to do that three quarters, where the facts, what the statute says, and what the case law says, what the arguments were, is done by AI. And then this is my decision."

Table 2: Potential Benefits of AI in the Judiciary by Category

AI's integration into the work of judges was thought to promise several benefits including the provision of *improved information* for legal decisions, increased *consistency*, *efficiency* and *access to justice* as well as reduced *bias*, *cost* and *tedium*. In addition, an increased understanding of *judicial effectiveness* was mentioned. These are listed in Table 2 along with direct quotes from judges related to these benefits. It is likely that more opportunities will arise as people become more familiar with the capabilities of AI.

4.4 Concerns and Considerations Around AI in Judicial Work

Along with the perceived opportunities for AI to improve the administration of justice (Section 4.3), many concerns and important considerations were expressed that would need to be addressed before wide adoption of AI in judicial work.

4.4.1 ***Reliability is currently insufficient for legal information.*** Reliability and trust in the AI came up frequently as a reason that judges could not currently use AI in certain parts of their tasks, calling it a "non-starter". Several participants described experiences they had while exploring AI where the results were factually incorrect, often also noting hallucinations, which are "confidence undermining". Several participants recalled asking AI about particular cases with which they were very familiar, and noted that, while the results were "plausible sounding", they were often glaringly incorrect: *"[it is] quite dangerous, because if you didn't know the answer was wrong, it looked very plausible. It made up citations, which didn't exist. It clearly knew the case in which the answer was to be found. But it summarised that case completely contrary to what the decision was. And it gave the opposite answer in language which it had*

obviously taken from the case." With several people having similar experiences, this was one of the highest priorities for enabling the use of AI in their work.

Related to this, several judges had experience with an AI-generated service that summarizes press articles, but that have inappropriate headlines and are also full of typographical errors, again reducing trust in the reliability and accuracy of AI for their work. Another noted that they had tried using AI to produce a transcript of a hearing but that it was not usable. Similarly, a participant reported their use of CoPilot to summarize a group of documents but found that "*it was very bad at that... I wouldn't be confident using it to summarize something that I didn't already know"*. Another issue is that general-purpose AI systems seem to be trained more on American law which can be problematic for use in UK law. These experiences led one judge to conclude "*At the moment, it just looks so risky that you wouldn't dream*", with another noting "*we need to make sure that it is trustworthy and reliable and that it won't undermine confidence in the judiciary if it's used.*" Optimistically, one judge remarked "*Personally, I think there may come a time when it will be of utility for legal research if it is carefully trained on authoritative sources. So collections of judgments and statutes and if you can ensure that it cross references everything so you can see immediately that it's a genuine source and that you can rule out confabulation in, in the results it gives you. But one has to be extremely careful about this.*"

4.4.2 ***Beyond reliability, judges are conscious of the importance of precision in their use of language***: "*I think also it being sensitive, to language and the importance of the choice of words and the nuance of words because so much of our job is reading and trying to understand what is meant by somebody else's words and then expressing it in our words. The level of sophistication of the use of language is very important.*" AI is known to be well equipped to adapt to linguistic and stylistic requirements. However, the legal context is specific and so more focus on this may be needed.

4.4.3 ***Privacy is a concern but not specific to AI.*** One judge noted that they were concerned that their search and conversation history may not remain private. Another agreed, but pointed out that it is a problem even in other areas such as doing a web search. Thus, it is not specifically an AI problem.

4.4.4 ***AI may develop misconceptions from uninformed user behavior.*** After discussing the way that current AI systems can confidently state incorrect facts, a concern emerged that these incorrect facts may be accepted by uninformed users. In addition to the risks to these initial users, this brought up an "*AI specific worry that it's learning from questions and answers that other people have asked. And they might be totally satisfied with the answers, but they may be completely wrong. And it [the model] is potentially drawing false inferences from the conversations that it's having with other people."*

4.4.5 ***AI bias needs to be understood.*** While gaining a better understanding of and mitigating human bias was brought up as a potential opportunity for AI, it also emerged as a potential concern. All judges were aware that AI systems are trained on human data which has biases and that algorithms can also introduce bias with one noting "*there is the issue, the discrimination point. [... There is] a lot of discussion as to whether the databases are discriminating themselves."*

4.4.6 ***AI could lead to de-skilling.*** The topic of de-skilling and training of new judges came up several times. For example, one judge noted that "*there's a risk in that as AI presumably increases, humans never go into the source documents. So you rely on what AI is telling you about the source documents. Whereas when you're browsing, whether it's through the case papers or the authorities, you come across something that doesn't fit or that does support something which you wouldn't have come across necessarily without browsing. And I think the risk of some of it is that you'd rely on AI's output without yourself ever going or properly going to the source documents or at least not in the same way.*" This was echoed in another group: "*I have a concern about skills... you've got really quite junior judges using it to generate model answers and so on. If people are doing that all the time, they're not actually acquiring the skills or practicing the skills that they need to engage critically with the answers that the AI may be giving them. So, it's kind of trying to strike that balance between people not just becoming de-skilled and pressing buttons and developing their skills, but also enabling the enhancement of their productivity and speed and so on.*" One participant noted that an aspect of having judicial assistants *"is to help them and to provide them with opportunity,"* and there was agreement in the room on this, indicating that this is something that should not be lost due to AI. In addition, there was recognition that *"there's a lot of people here [in the judiciary] who rather like hanging around law libraries [and that get] great satisfaction from the chase and following of the footnotes... the satisfaction of problem solving, you've put the puzzle together."* These practices, along with the discipline that goes into preparing for court and not taking "*shortcuts*" is something that should continue to be fostered in the training and mentoring of judges, and thus any work restructuring would need to take this important training into consideration.

*4.4.7. **AI might remove the intellectual challenge of legal reasoning.*** On reconsidering the focus group transcripts, however, we noted a comment we had not originally reported that discloses a further, non-instrumental aspect of judges' concern about the displacement of judicial work. Our participants also feared that the judicial career might lose an aspect of its intrinsic appeal: "*If the job becomes horrible to do, no one will do it. Everybody who takes a high court judgeship usually takes a very substantial cut in pay. And if you ...make it not as much fun as it is, then people will not follow that*". To those lawyers most attracted to the intellectual dimension of judging, the emergence of an AI-based method of adjudication that bypasses law's signature professional challenge might undercut the personal value of a judicial career. This development, in turn, might alter the demographics of judicial staffing at different court levels.

5 Discussion and Future Work

The focus groups revealed both opportunities and challenges for how judges interact with AI at work, particularly in the United Kingdom. Throughout the sessions, across most participants, there was acknowledgment that AI will change work practices, with many aspects discussed in the sections above. However, due to time constraints, we could not explore all topics in depth, which may have limited the breadth of themes uncovered. Additionally, participants' attitudes are shaped by their own levels of familiarity with AI and the current capabilities of AI tools–both of which are likely to evolve over time.

Much of the discussion, therefore, focused less on speculative or long-term transformations and more on how AI might support or augment existing judicial practices. For example, some discussions revolved around tasks that judges might already delegate to clerks or administrative staff. Such scenarios may

have been easier to imagine, given the existing norms around human delegation. In this capacity, AI would create efficiencies by expanding existing human support structures rather than by offering novel functionality. Equally, there was discussion of the suitability and unsuitability of AI in decision making and legal reasoning, which are key responsibilities of judges.

There were suggestions, however, of AI's potential to enable tasks that might not otherwise be performed, even if granted an abundance of specialist legal or non-specialist administrative human assistance. Participants identified several functions as possibly ideal for AI, such as the analysis of bulk sentencing data, court administrative data, the handling of small claims that often are not brought to court, and the ability to bring judicial topics to a wider audience through podcasts, videos, and the generation of audience-specific language (e.g., the drafting of child friendly summaries). In this way, participants entertained the prospect of AI transforming the delivery of justice beyond mere acceleration.

Further work could expand our study's reach by conducting a survey with broad participation to explore the topics covered in the focus groups to understand the similar and different perceptions and expectations of AI in the broader population of judges in the UK and around the world. Moving forward, it will be essential to determine the real-world impacts of these tools, for instance, whether they decrease or exacerbate existing human biases, preferably in early stages, before use is widespread. For example, in line with prior work [17], we suggest examining whether human decision makers become more or less biased in the presence of AI recommendations. Although AI systems can inherit biases from training data, ample evidence suggests that human actors may sometimes generate biases exceeding those of the very algorithms trained on their behavior [3]. This can occur, for example, when extraneous factors such as weather, sporting events, or time of day–variables absent from AI training–unduly influence judicial attention to legally relevant details [6, 7].

Several judges hypothesized that a rigorous, experimentally driven study could illuminate the true impact of AI on judicial decision-making. One judge mused, "*Suppose ... half the judges are given assistance by AI ... half were not. It's a randomized situation. And then you were to find that the ones with the AI assistance had decisions which went up the line of appeal ... much less frequently than the others.*" Others pointed out the practical barriers—"*You'd have to have 500 years to generate enough data*"—yet still considered this kind of controlled design as essential to testing whether AI indeed improves efficiency and fairness without eroding judges' capacity for critical reasoning. In other words, the bench is open to empirical evidence but insists that AI interventions be tested robustly, rather than simply be assumed to be beneficial.

In the same vein, the judges viewed the potential effects of AI through the three pillars of judicial state capacity—physical capital (the infrastructure of courts), human capital (the "skills" and "critical faculty" of judges), and normative capital (trust and legitimacy). While some participants endorsed using "*cheap and cheerful AI*" to handle small claims and to clear backlogs in the system, they also underscored the danger of reducing public confidence if human oversight were absent in serious cases. Concerns emerged that overreliance on AI could "*undermine confidence if the public suspects we're just rubber-stamping a machine's answer*," thereby depleting the judiciary's normative capital even as its physical and human

resources benefited. A well-designed, randomized control trial could thus measure not only "hard" outcomes—like case throughput or reversal rates—but also "soft" metrics such as litigant satisfaction, perceived fairness, and the system's overall legitimacy.

6 Conclusion

In summary, this chapter foregrounds judges' own accounts of how artificial intelligence may and may not be incorporated into judicial work. Drawing on qualitative evidence from across the UK court hierarchy, it shows that judges are neither resistant to technological change nor uncritical of its promise. Instead, they articulate task-specific boundaries that distinguish forms of assistance that may support efficiency, access to justice, and public understanding, and core judicial functions that must remain human, accountable, and visible.

These findings have implications that extend beyond technology design. They directly address questions of judicial training, institutional capacity, and public trust. By deeply engaging with judges, this chapter highlights the complexities of judicial processes and helps to enhance human-AI collaboration and prepare the judiciary for a future shaped by AI technologies. By investigating both algorithmic and human contributions to the administration of justice, we aim to inform more responsible design and implementation practices that safeguard fairness and rule of law.

Acknowledgments

We gratefully acknowledge the interest and practical support provided by members of the UK judiciary, without which this project would not have been possible.

References

[1] Armin Alimardani. 2024. Generative artificial intelligence vs. law students: an empirical study on criminal law exam performance. *Law, Innovation and Technology* 16, 2 (2024), 777–819.

[2] Guilherme FCF Almeida, José Luiz Nunes, Neele Engelmann, Alex Wiegmann, and Marcelo de Araújo. 2024. Exploring the psychology of LLMs' moral and legal reasoning. *Artificial Intelligence* 333 (2024), 104145.

[3] Victoria Angelova, Will S Dobbie, and Crystal Yang. 2023. *Algorithmic recommendations and human discretion*. Technical Report. National Bureau of Economic Research.

[4] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[5] Baroness Sue Carr, Sir Geoffrey Vos, Sir Keith Lindblom, and Lord Justice Colin Birss. 2023. Artificial Intelligence (AI) – Judicial Guidance. https://www.judiciary. uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance/ Accessed: 2025-02-05.

[6] Daniel L Chen. 2019. Judicial analytics and the great transformation of American Law. *Artificial Intelligence and Law* 27, 1 (2019), 15–42.

[7] Daniel L Chen. 2019. Machine learning and the rule of law. *Revista Forumul Judecatorilor* (2019), 19.

[8] Inyoung Cheong, King Xia, KJ Kevin Feng, Quan Ze Chen, and Amy X Zhang. 2024. (A) I Am Not a Lawyer, But...: Engaging Legal Experts towards Responsible LLM Policies for Legal Advice. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 2454–2469.

[9] Jonathan H Choi and Daniel Schwarcz. 2023. Ai assistance in legal analysis: An empirical study. *Available at SSRN 4539836* (2023).

[10] U.S. SupremeCourts. 2025. *Supreme Court Procedures*. https: //www.uscourts.gov/about-federal-courts/educational-resources/abouteducational-outreach/activity-resources/supreme-court-procedures

[11] Matthew Dahl, Varun Magesh, Mirac Suzgun, and Daniel E Ho. 2024. Large legal fictions: Profiling legal hallucinations in large language models. *Journal of Legal Analysis* 16, 1 (2024), 64–93.
[12] Timothy Endicott and Karen Yeung. 2022. The death of law? Computationally personalized norms and the rule of law. *University of Toronto Law Journal* 72, 4 (2022), 373–402.
[13] ATLAS.ti Scientific Software Development GmbH. 2021. ATLAS.ti: Qualitative Data Analysis Software. https://atlasti.com/
[14] Neel Guha, Julian Nyarko, Daniel Ho, Christopher Ré, Adam Chilton, Alex Chohlas-Wood, Austin Peters, Brandon Waldon, Daniel Rockmore, Diego Zambrano, et al. 2024. Legalbench: A collaboratively built benchmark for measuring legal reasoning in large language models. *Advances in Neural Information Processing Systems* 36 (2024).
[15] Juan David Guttierez et al. 2024. UNESCO Global Judges' Initiative: survey on the use of AI systems by judicial operators. (2024).
[16] Lord Patrick Hodge. 2023. The Law and AI - where are we going? Lecture at De Montfort University, Leicester. https://supremecourt.uk/speeches/the-law-andai-where-are-we-going Accessed: 2025-02-05.
[17] Jon Kleinberg, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2018. Human decisions and machine predictions. *The quarterly journal of economics* 133, 1 (2018), 237–293.
[18] Andreia Martinho. 2024. Surveying Judges about artificial intelligence: profession, judicial adjudication, and legal principles. *AI & SOCIETY* (2024), 1–16.
[19] Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. 2023. Robust speech recognition via large-scale weak supervision. In *International conference on machine learning*. PMLR, 28492–28518.
[20] Antonino Rotolo and Giovanni Sartor. 2023. Argumentation and explanation in the law. *Frontiers in Artificial Intelligence* 6 (2023), 1130559.
[21] Sascha Schweitzer and Markus Conrads. 2024. The digital transformation of jurisprudence: An evaluation of ChatGPT-4's applicability to solve cases in business law. *Artificial Intelligence and Law* (2024), 1–25.
[22] Lawrence B Solum. 2004. Procedural justice. *S. CAl. l. reV.* 78 (2004), 181.
[23] Eugene Volokh. 2019. Chief justice robots. *Duke law journal* 68, 6 (2019), 1135– 1192.
[24] Guillaume Zambrano. 2024. Case law as data: Prompt engineering strategies for case outcome extraction with large language models in a zero-shot setting. *Law, Technology and Humans* 6, 3 (2024), 80–101.
[25] Adrian Zuckerman. 2020. 291Artificial Intelligence in the Administration of Justice. In *The Civil Procedure Rules at 20*. Oxford University Press. doi:10.1093/oso/9780198863182.003.0020
[26] Oliver Wendell Holmes 1897, The Path of the Law 10 Harv. L. Rev. 457–478
[27] HLA Hart, 1961, *The Concept of Law* OUP.
[28] Erin Solovey, Brian Flanagan, and Daniel Chen. 2025. Interacting with AI at Work: Perceptions and Opportunities from the UK Judiciary. In Proceedings of the 4th Annual Symposium on Human-Computer Interaction for Work (CHIWORK '25). Association for Computing Machinery, New York, NY, USA, Article 5, 1–8. https://doi.org/10.1145/3729176.3729192

A Focus Group Discussion Questions

As noted above, two researchers co-moderated the discussions, guided by a set of guiding questions and topics listed below. The order and content varied, depending on the flow of conversation and time

constraints. Additional follow-up questions were asked when appropriate to delve deeper into topics that were discussed by the participants.

Understanding Current Workflows.

- Can you walk us through a typical day or week in your role?
- What is the process of doing X (today, without AI)?
- What are the most critical tasks you perform, and why are they important?
- If applicable: What factors do you consider when doing [a particular task]? What do you do if you are unsure? Where is there room for improvement in the process? When do you collaborate with others?
- What tools or systems do you currently use to support your work?
- Are there parts of your workflow that are particularly timeconsuming/repetitive/difficult/other?

Perceptions of AI by Role.

- Have you used AI in work-related tasks or non-work purposes? If so, what was the experience like?
- Could you list some of the benefits that would make it worthwhile to use AI?
- What concerns or risks come to mind?
- Do you think your role gives you a unique perspective on AI compared to other roles?

Human vs. AI Roles/Strengths.

- What aspects of your job do you feel it would be difficult for AI to replace or support?
- Are there tasks that you think AI might perform better than humans?
- Are there tasks where humans should always be involved, even if AI could technically perform them as well or better? • Are there aspects of the work that will be lost if an AI tool assists with or performs the task?
- Out of the many aspects of your job, are there ones that you are particularly concerned with or interested in?

Future Implications of AI on Judicial Roles.

- Do you see places where AI is being proposed in your work or where you imagine it coming soon in your work context? • How do you think AI might change your role in the next 5 years?
- Are there tasks and responsibilities that will no longer be needed because of AI?
- Thinking about your own day-to-day work, if you had the perfect AI assistant in the future, what would you want it to do for you?
- What training will be needed in the future?

Closing Reflections.

- What's one of the most important things to consider related to AI and the work of the judiciary?
- Do you have suggestions for computer scientists regarding tools that might be integrated into judicial systems?
- Is there anything we have not discussed that is important regarding the use of AI in judicial work?