Pressing Questions for the Use of AI in ADR

Stephen Sullivan, Jennifer Shack & Jasmine Henry

Across the ADR field, everyone, it seems, is considering the potential ways AI tools might make things better. Courts are trying to make their processes more efficient by using chatbots to answer questions; administrators are seeing opportunities for AI tools to help people without attorneys participate more capably; mediators are using coaching tools to support their practice and AI assistants to summarize documents and generate settlement proposals. All these tools are being adopted in the hope that they will increase access to ADR processes and will lead to higher mediation and arbitration quality and greater neutrality in decision-making.

But there are risks to consider. The use of AI presents new ways for power imbalances and biases to be introduced. Overreliance on outsourced technology may lead to de-skilling of ADR neutrals. Mediation may become one-size-fits-all, with a value system created by the tech field rather than those in ADR. Finally, there are questions about whether AI will lead to a two-tiered system and conflict with societal values prevalent in the field.

How is AI being used in ADR?

Courts and dispute resolution organizations are using generative and agentic AI tools for different purposes. Generative AI tools, such as ChatGPT, produce text, images and other output types based on patterns learned from training data and user inputs. Within the legal and ADR contexts, generative AI is used for self-help chatbots that provide free legal information or help users to fight parking tickets, AI tools that draft demand and response letters or mediate issues, and AI coaches that help parties to prepare for mediation or help mediators complete mediation tasks.

Agentic AI systems operate autonomously to complete more complex workflows with a set of pre-determined objectives and with less supervision. This is often done with human review. Within the ADR context, American Arbitration Association uses agentic AI to review submissions and make determinations in paper-only construction arbitrations. The agentic AI determinations are then reviewed by a human.

How do we ensure that AI provides value to parties without perpetuating or creating issues?

AI in ADR can potentially make people and processes more time- and cost-efficient. Chatbots can answer questions that clerks and program staff generally field. Mediators can more efficiently review pre-mediation document submissions. AI tools can facilitate resolutions in online, text-based processes. As a result of these efficiencies, more people should have access to ADR services.

Another promised benefit of client-facing AI tools, from chatbot coaches to mediation assistance tools, is that they can help practitioners to be more impartial and consistent. By better identifying issues, needs and interests and by standardizing services, AI should help clients have clearer expectations and better experiences.

But AI does not come without costs. While AI assistance might increase program speed and volume, it may do so at the expense of equality and quality. Bias, power imbalances, process quality and the potential for a two-tiered system to develop are all issues that need to be addressed.

BIAS ISSUES

As we now know, AI adopts the biases that are "baked in" to the dataset it uses. Chatbots powered by AI have been found to generate covertly racist decisions based on a user's dialect. Dataset-driven suggestions for salary offers to a job candidate were found to change based on whether it categorized their name as male or female, Black or white.[1]

Biases also emerge at different stages of AI development. In addition to those in the training data itself, there are biases in the people selecting the data, prioritizing features and deciding how the tool is used. For example, researchers found that AFST (an AI tool to screen calls reporting child neglect) exhibited race and disability biases due to decisions developers made about data sourcing and feature design. Another legal AI tool, COMPAS (an AI algorithm to predict defendants' recidivism risk) disproportionately mislabeled Black defendants as "high risk" compared with white defendants.

POWER IMBALANCES

Every technological system that promises greater access for those with fewer resources raises concerns about how the new technology differentially affects those

[1] For example, a chatbot may say a job candidate with a name like Tamika should be offered a $79,375 salary as a lawyer, but switching the name to something like Todd boosts the suggested salary offer to $82,485.

who are tasked with using it. AI is no different. People without lawyers can use AI tools to create demand letters, fill out forms and construct opening statements with greater assuredness. But AI tools do not inherently address power imbalance. Aside from the general issue of some people not having access to these tools, inequities can also be caused by different levels of knowledge needed to assess convincing but incorrect information.

AI requires at least two areas of knowledge that everyone does not share equally: "prompt engineering" and understanding when AI outputs may not be accurate. Users who are more knowledgeable about AI may be able to more effectively prompt an LLM to perform tasks. AI experts note that prompt engineering is key to obtaining precise, accurate and higher-quality responses from AI systems. Thus, those who have more experience with AI and more understanding of prompt engineering may have an advantage.

Legal knowledge may also play a role when vetting AI responses. For example, does a particular arbitration demand cite the correct jurisdictional information? Does it present a sound legal claim? When one party has an advantage over the other when using AI to prepare documents, arguments and opening statements, a power imbalance is created or enhanced.

PROCESS QUALITY

Valorizing predictability and standardization comes with risks, such as the potential for de-skilling and for reducing opportunities to learn from others. Research suggests that increased AI integration risks reducing human cognitive and social capacities. AI also favors and reinforces dominant ideas — and marginalizes alternatives. Psychologists note that this feedback loop can be counterproductive to making thoughtful decisions. How will we ensure that AI tools do not hamper creative problem solving in ADR and efforts to improve human relationships?

TWO-TIERED APPROACH

Should courts move to adopt AI tools more widely, we should be asking as a field how to ensure that AI adoption does not result in those with fewer resources being shuttled to a lower-quality process. How do we avoid creating a system in which human-centered mediation is offered to those with resources and a cookie-cutter AI process is the only option open to those who cannot afford mediation?

Who decides what mediation is?

As a field, we must ask: What it is that we want to replicate, and what can we enhance? There is evidence that AI tools can help self-represented parties prepare for mediation in ways that mediators and mediation program staff cannot. Chatbots can help parties devise an opening statement or give advice on how to present their side of the dispute. AI tools may also help facilitate resolution with or in place of a neutral third party.

Within these functions, however, lie a number of potential questions tied to the

values instilled in AI models and their impact on any ADR process. For example, an AI tool that helps a party prepare opening statements or strategize about how to present one's story can significantly impact how a person processes their dispute and discusses it within the mediation. Also, the way a chatbot responds to a person's prompts may color the way the party views the dispute and the other party. While helping parties, will the chatbot focus solely on monetary issues, or will it look at other issues as well? Will it inform the parties of settlement options or create terms for them? Will it instruct them to keep their arguments free of emotions? What are the values and implications inherent in how these chatbots are programmed to help?

Similarly, we need to think about what values and emphases the AI tool is given. In human-mediated dispute resolution, mediators come with different emphases and values. How do we ensure that an AI facilitator will approach dispute resolution in a way that is compatible with the goals we want to achieve? For example, will the AI model focus primarily on settlement? Should it emphasize the relationship between parties? How much emotional expression should it encourage? Should it help the parties process emotions? For skilled mediators, this might shift from one case to the next. An AI trained to a specific model or value system may not do the same.

How does AI adoption by the field fit with societal values?

A final consideration is how the advent of AI and the decisions we make about adoption come to shape our concepts of "access" and "justice" themselves. How do we reconcile the documented harms of AI to marginalized communities with our values as a profession invested in justice? How do we reckon the social and environmental costs of AI with access to justice commitments? Who stands to benefit most from the ADR field's use of AI technologies? What does justice mean, particularly for those who are most proximate to the material and health-related harms of exponential and unregulated AI growth? What constraints will we establish to minimize user harm? How will we repair harm when it is experienced, particularly for SRLs and other marginalized users? Taking the needs and concerns that the public has about AI seriously will be imperative to ensuring that ADR continues to have a positive impact in people's lives.

What methods will we use to evaluate the impacts of AI tools?

Research finds that the current metrics used to determine LLM task efficacy in legal contexts are overwhelmingly quantitative and removed from actual experience. Given societal-level pressures to justify investment in AI, it is reasonable to expect that there will be an increased premium on AI-driven indicators over more humanistic assessment measures. This runs the risk of limiting the ability of evaluators to more

deeply assess ADR programs and the experiences of people they serve.

Qualitative assessments will become increasingly important for understanding how parties experience bias and power imbalance while using AI tools. As we discuss in our OPEN Project research report, prioritizing the needs of self-represented litigants with low digital literacy benefits all parties. Partnering with groups outside the tech world can also ensure that AI systems align with ADR values and ethics and the needs of court constituents.

As a field, we need to be asking now about how we will determine whether AI actually benefits parties. What measures will we use? What voices are driving demand for AI use and under what rationales and forms of evidence? How might we adjust or reconsider AI use if its anticipated benefits do not bear out?

CONCLUSION

Today's AI systems are not value neutral; they are a reflection of a particular worldview and set of culturally specific assumptions about what it means to be human and how problems are solved. Considerations about AI use in ADR also cannot be divorced from the context of AI hype narratives and the techno-solutionist worldviews driving the creation and marketing of these products. There is a draw to adopt these tools quickly and not "fall behind." However, it is important to ensure that we proceed with caution when building out any AI tools for use in ADR processes. Caution and care can help to keep expectations realistic and ensure products perform functionally.