Jeffrey I. Ehrlich
EHRLICH LAW FIRM, APC

Courts will use AI to decide cases. Then what?

WE WANT THE MACHINE AND WE WANT THE CONSCIENCE

Here is a fact that should concentrate the mind of every trial lawyer: Generative AI can now read the moving and opposition papers in a motion or the briefs in an appeal, analyze the legal issues, and draft a tentative ruling or a draft opinion. It does this quickly, competently, and tirelessly. Given how overburdened our courts are – given the grinding pressure on judges and their staffs to move cases through the system – the pull to use this capability is likely irresistible. If it is not happening already, it will be soon, at scale.

This development invites legitimate concerns about AI taking over the legal system, about the dehumanization of justice, and "robots in robes." These reactions are understandable, but they can obscure what is genuinely interesting about this moment. The arrival of AI capable of competent legal analysis forces us to ask a question we have not had to reckon with before: What, exactly, is a judge *for*?

What AI can do today

According to a 1,700-lawyer survey taken in January and February 2025 by the Thomson Reuters Institute, 41% of respondents said they personally use publicly available tools such as ChatGPT and 17% said they used industry-specific

tools. Only 13% of respondents said that generative AI is central to their organization's workflow currently, but 95% think it will be central to their workflow within the next five years.

This suggests that readers of this article will fall into two camps: a substantial number who are already familiar with how well today's most powerful AI models can perform legal analysis, and an even larger number who will be surprised by how much these models can do.

If you have not experimented with these tools yourself, consider what Westlaw's most advanced product suite, marketed as "Westlaw Advantage," is currently offering. The marketing brochure explains that this version of Westlaw includes a tool called "Deep Research," which allows users to ask a legal question and receive within minutes a comprehensive research memo – complete with a detailed analysis of applicable law, arguments on both sides, and cases supporting each position.

Westlaw Advantage also includes a tool called the "Litigation Document Analyzer" that will examine briefs and memos to identify potential weaknesses, ensure citation accuracy, and flag vulnerabilities in both your work and your opponent's. Within this is a "Judicial Analysis" tool that promises to show you "the same analysis of briefs from both parties that judges can see," to "review the most relevant authority that neither party cited," and to "discover issues with the citations and quotations relied upon by the parties."

The ability for AI to do these tasks is not exclusive to Westlaw. Rather, Westlaw has harnessed the underlying capability of the ChatGPT 5 family for its product. There are competing legal-research products, such as Mid-Page.AI, which can do similar things, and may do them better than Westlaw does. Similarly, the underlying models themselves can do these tasks, and more.

The "Judicial Analysis" tool deserves a moment's pause. Westlaw is not speculating about what judges might someday do with AI. It is selling a product that appears to be *premised on* judges having access to AI-powered analysis of the briefs filed in their courts. Whether or not any particular judge is using such tools today, the capability exists, it is mainstream, and it is being marketed by the dominant legal-research company in the United States.

I think the important takeaway is that, as Westlaw and Lexis bundle highly capable AI models into their legal-research products, the notion of lawyers or judges "using AI" will seem quaint, and the concept of "legal research" will expand to incorporate the various outputs the models are capable of.

And what they are capable of – right now – is to provide thorough, accurate legal analysis, and to draft any type of legal document they are asked to produce. Their output is not perfect, or typically artful. But the same can be said for most outputs from overburdened courts or time-stressed lawyers. And the models become more capable with each upgrade.

For now, these tools work best as collaborators with skilled lawyers (or judges) who direct their analysis and verify their output. Remove the human from the process, and the quality of the work degrades significantly. But the models improve with each generation, and the balance of that collaboration is shifting.

These tools are currently available to lawyers and to courts. And given the strains that litigation places on lawyers and the judiciary alike, lawyers, judges, and their staffs have every incentive to use them. The question this article addresses is not whether AI will or should enter the judicial process – that is already happening – but what it means for how we think about judges and judging.

The components of judicial fairness

What we want from a judge

I think we would all agree, if asked what we want from judges, is that they be "fair." But when we examine what we mean by "fairness," the concept fractures into several distinct components, some of which exist in tension with each other.

First, we want *accuracy*: The ability to correctly identify the applicable statutes and legal rules, and to find the facts that trigger application of those rules. A judge who misreads a statute or misapprehends the facts has failed at the most basic level.

Second, we want *fidelity*: The judge should apply the law to the facts as the law directs, regardless of whether the judge personally likes the outcome. We do not want judges substituting their preferences for the commands of the legislature or the holdings of binding precedent.

Third, we want *disinterestedness*: The judge should have no personal stake in the outcome. A judge who owns stock in a corporate defendant, or who is related to a party, cannot be trusted to decide the case impartially.

Fourth, we want *emotional discipline*: The judge should set aside personal sympathies and antipathies so they do not distort the analysis. We do not want judges ruling for parties because they are likable or against parties because they are not.

Fifth, we want *productivity*: Courts must resolve cases and motions within reasonable timeframes.

So far, so good. These five components are mutually reinforcing, and they describe something that sounds much like an algorithm: take in the inputs (facts and law), process them according to established rules, and produce an output (the ruling). We expect judges to suppress their humanity – their preferences, their emotions, their sympathies – and function as neutral processors of legal information.

But there is a sixth component, one that exists in tension with the first five.

We also want *conscience*: The capacity to recognize when a mechanically correct outcome is nonetheless unacceptable, and to invoke available judicial tools to avoid it. We want a human being in the robe who can look at where the legal logic is pointing and say, "I cannot do this. There must be another way."

The relationship between these six components and the work judges do depends on what kind of case is before

the court. In most cases, the first five components do nearly all the work, and conscience rarely enters the picture. In some cases, conscience is the whole game – the case is fundamentally about which values should prevail. And sometimes, a case that appears to require only routine mechanical analysis can nevertheless generate a need for rulings that fall into the second category.

Understanding this relationship is essential to understanding where AI fits.

The tension within fairness

Notice the problem. Components two and four tell judges to suppress their feelings and follow the law wherever it leads. Component six requires judges to *feel* that something is wrong – to experience the discomfort of an unjust but seemingly correct outcome – and then to act on that feeling.

When we say we want judges to "set aside their emotions," we do not really mean all emotions. We mean the *distorting* emotions: bias, favoritism, irritation, unwarranted sympathy. But we want judges to retain the *corrective* emotions: the sense of justice that rebels against an unacceptable outcome even when the legal logic is impeccable.

This is a subtle distinction – one that I have not seen discussed much during my career. Rather, it seems that we have simply relied on human judges to manage both demands – to be emotionless rule-appliers most of the time while retaining the capacity for moral discomfort when it matters.

Two categories of cases

With this framework in hand, we can see that cases tend to fall into two categories – and that AI's usefulness depends heavily on which category a case falls into.

Resolution through established legal rules

The first category, and by far the larger, consists of cases that have answers derivable from applying established legal rules to found facts. The statute of limitations has run or it has not. The contract contains an arbitration clause or it does not. The defendant's conduct falls within the elements of the tort or it does not. These cases may be complex, but they are ultimately deterministic: a sufficiently skilled analyst, given the law and the facts, can work out the correct answer. In these cases, the first five components of fairness do nearly all the work; conscience is rarely activated. These are the cases that courts routinely decide using unpublished opinions.

This analysis assumes the facts are established or undisputed — the situation that is typically presented in motion practice and on appeal. Factfinding at trial, where credibility is contested, raises different questions I do not address here.

No resolution by applying established legal rules

The second category consists of cases that cannot be resolved by applying legal rules, because the rules themselves encode a clash of values whose resolution depends on what the decisionmaker holds dear. Constitutional cases frequently fall into this category. When does the state's interest in protecting potential life overcome a woman's liberty interest in bodily autonomy? When does public safety justify restrictions on the right to bear arms? When does equality require race-conscious remedies, and when does equality forbid them? These questions cannot be answered by legal analysis alone. They require choices among competing values, and different judges – applying the same legal materials with equal skill – will reach different answers because they hold different values. In these cases, conscience is the whole game.

This two-type frame is obviously an oversimplification. Some cases that fall into the first category may contain discrete issues that present a category-two issue. For example, in a straightforward breach-of-contract case, there can be evidentiary issues that turn on the trial judge's exercise of discretion, such as whether to exclude certain evidence as unduly prejudicial.

Nevertheless, for our purposes here, the two-types-of-cases rubric is helpful because it allows us to distinguish between the types of legal issues that AI is well suited to help decide and those where it is structurally incapable of providing much assistance.

The conscience function

There is a situation that complicates this tidy two-category scheme: The case that appears to fall into the first category – it seems to have a deterministic legal answer – but where that answer strikes the judge as unacceptable. These are the cases where conscience is unexpectedly activated, and they pose the hardest questions about AI-assisted judging.

For juries, we have a name for the response to this situation: nullification. A jury that believes a defendant is technically guilty but that conviction would be unjust can simply acquit. Jury nullification exists in a legal twilight; courts do not instruct juries that they have this power, but the system structurally cannot prevent its exercise. We tolerate the escape valve while maintaining a polite fiction that it does not exist.

Do judges have an equivalent power? Formally, no. Judges take oaths to follow the law. A judge who openly says, "the law requires X, but X is unjust, so I am ruling Y," faces possible discipline. For example, in Morrow v. Hood Communciations, Inc. (1997) 59 Cal.App.4th 924, Justice Anthony Kline filed a dissent in which he refused to follow what he considered to be an invalid precedent from an earlier California Supreme Court decision regarding stipulated reversals of appellate decisions, because he felt that the doctrine was "destructive of judicial institutions." The Commission on Judicial Performance initiated disciplinary proceedings against him for his refusal to follow binding precedent.

And yet, judges nevertheless exercise a kind of quiet, deniable resistance to unjust outcomes all the time. They find ambiguity in statutes that permits a more just reading. They distinguish precedent by emphasizing

factual differences. They exercise discretion to soften harsh outcomes. They invoke the absurd-results doctrine to avoid literal readings that produce unacceptable consequences.

None of this is called "nullification." It looks like conventional legal reasoning. But its function nevertheless fills some aspect of the nullification role: A human being looks at where the legal logic is pointing, feels that it is wrong, and finds a way to alter the outcome.

Whether we *want* judges to have this power is a harder question than it might appear. There are real costs to judicial resistance. It undermines predictability and the rule of law. It introduces the judge's personal values into what is supposed to be neutral adjudication. It is distributed unevenly, creating inconsistency. It can be unaccountable, hidden within "interpretation" where it cannot be openly debated.

A formalist would say that, if the law produces unjust results, the remedy is legislative amendment, not judicial subversion. But a realist would respond that the system has always depended on human judgment to sand down the rough edges of legal rules. Pretending otherwise is naive. We need the escape valve.

I raise this tension not to resolve it, but to observe that AI forces us to confront it. AI will give a "correct" answer. It will apply the rules to the facts and tell you what follows. It will not feel the discomfort that a human judge feels when the correct answer seems "wrong." It will not be motivated to find a different outcome.

AI and the judicial role

With these distinctions in hand, we can map AI capabilities onto the components of judicial fairness.

Accuracy: Very good, and improving. AI can identify applicable statutes, parse case law, and apply legal rules to facts with high reliability. Yes, AI can hallucinate cases. But these inaccuracies are easily spotted and corrected by lawyers or clerks exercising standard cite-checking techniques that should be incorporated into any document filed in or by a court.

Fidelity: Excellent. AI has no temptation to deviate from what the law requires.

Disinterestedness: Structurally superior to humans, though not perfect. AI has no personal stake in any outcome. But it may reflect biases embedded in its training data or design choices made by its developers – a different kind of interest than judicial self-dealing, but not nothing.

Emotional discipline: Again, structurally superior. AI has no personal preferences or sympathies to set aside. But it may have picked up tendencies embedded in its training material. For example, a legal analysis conducted by Google's Gemini and xAI's Grok might differ because Grok is trained on Elon Musk's posts on X.

Productivity: Extraordinary. AI can analyze a motion in minutes rather than hours or days.

Conscience: Absent. AI does not experience moral discomfort. It cannot feel that a correct answer is wrong.

A careful reader will notice a tension in this analysis. The very conditions that make AI attractive to courts – overwhelming caseloads, understaffed chambers, pressure to move cases – are the same conditions that create risk of insufficient human oversight. If a judge or clerk is too busy to draft a tentative ruling, are they likely to have time to carefully verify the AI's work?

The accuracy I attribute to AI assumes that someone is checking. The danger is that AI's competence becomes a reason to trust it without verification, and the verification that would catch errors gets squeezed out by the same time pressure that prompted AI adoption in the first place. Courts that adopt these tools will need to build verification protocols into their workflows, not treat AI output as presumptively correct.

The implication is clear: AI is well-equipped for five of the six components – arguably *better* equipped than human judges, who struggle with accuracy under time pressure, who sometimes let bias creep in, who have dockets they cannot keep up with. But AI entirely lacks the sixth component. And that sixth component may be the one that ultimately matters most.

This analysis suggests that AI is well-suited to category-one cases and unsuited to category-two cases. For the routine application of settled law to found facts, AI may actually *improve* the quality and consistency of judicial decisions. For cases requiring value judgments, human judges remain essential – not as calculators but as the conscience of the system.

The law clerk precedent

Before we react with abject horror to the idea of AI-assisted judging, we should bear in mind that judges have had assistance for a long time.

As explained on the website of the Federal Judicial Center, Justice Horace Gray of the United States Supreme Court hired the first law clerk – a recent Harvard Law graduate – and paid him out of his own pocket. Congress began appropriating funds for Supreme Court law clerks in 1886. By the early twentieth century, more justices had followed Gray's lead; Oliver Wendell Holmes Jr. and Louis Brandeis are often cited as among the first to systematically use recent law-school graduates as more than mere stenographers. After Congress funded "law clerks" in 1919, each justice had one clerk, later two. By the 1940s through 1960s, a Supreme Court clerkship had become a coveted career step, and essentially all justices employed clerks regularly.

With exploding caseloads in the 1960s through 1980s, the use of law clerks spread throughout the federal appellate and district courts, and many state courts followed. Today, federal district judges are authorized to hire two or three clerks, and appellate judges (and Justices) can hire four. State appellate courts commonly do the same or employ centralized staff attorneys.

The point is this: It's been a long time since most judges were personally researching and writing every word of every opinion. We have had 140 years

of judicial assistance being built into the system, and the profession adapted without crisis. Each expansion of the clerk's role – from stenographer to researcher to drafter – probably occasioned some hand-wringing, and each was eventually absorbed as normal.

A strong argument can be made that generative AI is simply the next step in this evolution: a particularly capable form of law clerk, one that can do in minutes what a human clerk does in days, and do it with tireless consistency. The question is not whether judges should have help – we answered that long ago. The question is what must remain with the human.

A model already exists

Division Two of California's Fourth Appellate District (the 4/2), sitting in Riverside, already operates under a model that could make a transition to acknowledged use of AI-drafted tentative opinions relatively seamless. Unlike other California appellate courts, the 4/2 sends the parties a tentative opinion in every appeal. The parties then argue the case from that tentative about a month or two after receiving it.

In my experience arguing before that court, the justices on the panel sometimes discuss the tentative as though none of them had any role in drafting it. I have observed justices referring to "what the tentative says" as though it were the product of an author who was not in the room. This is likely because the tentative was principally or entirely the product of a research attorney. In that event, the panel's role is to evaluate the analysis, stress-test it against the arguments of counsel, and ultimately decide whether to adopt, modify, or reject it.

This represents a separation of the drafting function from the judging function. And once that separation is made – once judges are comfortable saying, in effect, "this tentative came from somewhere in the court's apparatus and we are here to decide whether it is right" – the transition to AI-drafted tentatives becomes conceptually straightforward.

One can imagine a future in which courts simply acknowledge that tentative rulings are AI-drafted. The question for oral argument then becomes explicit: "The court's AI has analyzed this matter and proposes the following disposition. Counsel, what has the analysis gotten wrong?"

That framing is honest, and it might improve the quality of argument. While appellate lawyers are not of single mind on the pros and cons of receiving a tentative opinion from the 4/2 before argument, I love the process. I want to know what the court is thinking before I argue the case. If there is a problem in the analysis, it is far easier to address before the final opinion is filed than afterward.

Having the court's tentative opinion in hand before argument means lawyers no longer must shadowbox with what they imagine the court is thinking and instead engage directly with a concrete analytical document. And just as they do now in the 4/2, counsel would have a clear role: To expose the limits of mechanical legal analysis, to show why this case requires the human-conscience function, or to demonstrate that the AI has misunderstood something a human reader would catch.

Implications for practitioners: The machine in the room

As courts adopt AI-assisted analysis, lawyers will need to adapt.

Writing for an AI reader may differ from writing for a human one. AI responds well to clear structure, explicit mapping of facts to legal elements, and precise identification of the controlling legal standard. Emotional appeals and narrative flourishes that might move a human reader are unlikely to register with an AI reader.

And lawyers would seem to be well-served in flagging when a case or an issue does not involve a routine category-one application of law-to-fact. If a case involves a value clash, or turns on the exercise of discretion, or if the mechanically correct outcome seems unjust, lawyers will need to signal that to the human judge who will review the AI's draft. Lawyers must learn to find ways that are likely to activate the judge's conscience function – to look past the confident legal analysis and engage with why this case requires human judgment.

This applies at the trial level as well as on appeal. While my earlier discussion focused on appellate tentative opinions, trial courts face even greater volume pressure. Law-and-motion departments in busy courts may find AI-drafted tentatives irresistible. Trial lawyers should anticipate that the judge ruling on their motion may be reviewing an AI draft, and should write and argue accordingly.

Paradoxically, this might actually be empowering. The lawyer's job becomes, in part, to be the check on AI – to represent the human judgment that the system still requires. Lawyers who understand how AI analyzes legal problems, and who can identify where that analysis falls short, will be more effective advocates than those who ignore the machine in the room.

Conclusion: What judges are really for

The arrival of AI in the courts does not merely raise practical questions about efficiency and workload. It forces us to confront a tension we have managed to avoid for as long as we have had courts: We expect judges to behave like algorithms – to suppress their preferences, to mechanically apply law – while simultaneously wanting a human being available to ensure that the outcome is just. We want rule-following and we want an escape valve. We want the machine and we want the conscience.

AI can be the machine. It is, in many respects, better at being the machine than human judges are. It does not get tired, it does not get annoyed, it does not worry about a backlog or the impact of an unpopular decision on a future judicial-retention election.

What AI cannot be is the conscience. It cannot feel that a correct answer is wrong. It cannot be moved to find another way.

This suggests that AI's entry into the judicial system might ultimately clarify what human judges are really for. If the mechanical legal analysis can be handled by AI, the judge's core function becomes what it perhaps should have been all along: Not calculating legal outcomes, but ensuring that law serves justice.

The risk, of course, is that the AI does such a good job at legal analysis that human review becomes perfunctory – that the judge's signature becomes a rubber stamp and the conscience function atrophies from disuse. Guarding against that risk will be a challenge for courts as they adopt these tools.

But the opportunity is real as well. For too long, we have asked judges to be two contradictory things: algorithms and human beings. AI might finally allow us to separate those functions – to let the algorithm be an algorithm and to free the human being to be what only a human being can be: the keeper of justice in a system of rules.

Jeffrey I. Ehrlich is the principal of the Ehrlich Law Firm *in Claremont. He is a cum laude graduate of the Harvard Law School, an appellate specialist certified by the California Board of Legal Specialization, and an emeritus member of the CAALA Board of Governors. He is the editor-in-chief of Advocate magazine, a two-time recipient of the CAALA Appellate Attorney of the Year award, and in 2019 received CAOC's Streetfighter of the Year award. Jeff received the Orange County Trial Lawyer's Association Trial Lawyer of the Year award for "Distinguished Achievement" in 2023.*