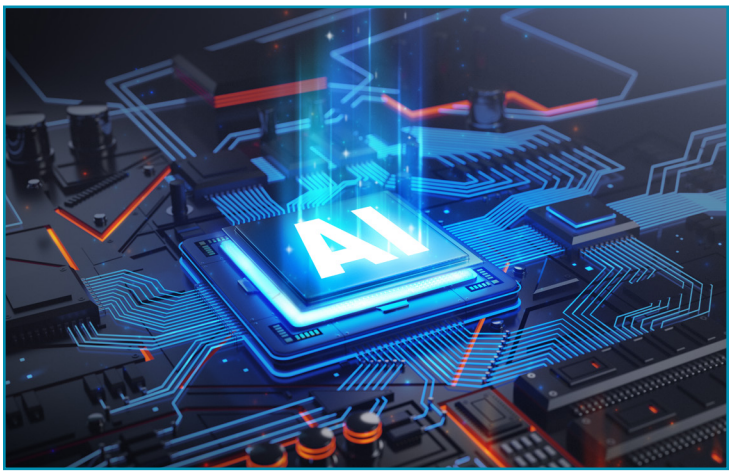
Judge Francis on Ediscovery, GenAI, and the Role of Special Masters

By Justin Smith



Hon. James "Jay" C. Francis IV (Ret.) has been a judge from practically the start of his career.

After graduating from law school in 1978, he practiced for just seven years before assuming his role as a U.S. Magistrate Judge for the Southern District of New York, and remaining in the judgeship until his retirement in 2017. In his current role as a mediator, arbitrator, and Special Master for JAMS, he's still taking an active role in the legal system.

During his career on the bench, Judge Francis saw ediscovery take shape and evolve into the practice it is today, making several consequential rulings on preservation and spoliation that still carry weight today.

And now, with the advent of generative AI as the next technological transformation facing the legal profession, he's excited about its potential for attorneys and judges alike.

Judge Francis spoke with Everlaw about becoming a judge early in his career, generative AI's use cases in the law, the role of Special Masters in ediscovery, and more.

You've had an interesting legal career, in part because you didn't practice for very long as an attorney before making your way to the bench. Was this a conscious decision to aim for a judgeship so quickly? Did you learn anything during your practice with The Legal Aid Society of New York that helped inform your early years on the bench?

It wasn't a conscious decision to aim for a judgeship at all.

When law students or young attorneys talk about aspiring to become a judge, it strikes me as a little naive because the path to the bench is subject to so much uncertainty and serendipity. Certainly, there are things that an attorney can do to put themselves in a better position to be selected, but given the number of positions available in either the federal or state systems, the idea of aiming for a judgeship is a little unrealistic.

In my case, it was more of a subconscious desire to become a judge. I clerked for a federal district judge, the Hon. Robert L. Carter, who was a true role model. He had been general counsel of the NAACP, and had litigated for *Brown* against the Board of Education, along with Thurgood Marshall.

The clerkship with Judge Carter reinforced my interest in legal decision-making and my passion for writing, so when the opportunity came to apply for positions as a magistrate judge, I jumped at the chance, but it was not something that I had planned out as any kind of anticipated trajectory.

As for my time at the Legal Aid Society, it helped prepare me for the bench in a couple of ways. Since my work there largely involved impact litigation in federal court, I was relatively comfortable with federal pretrial procedure.

It also taught me the importance of respecting litigants. Even when I did not agree with their position, I had to remember that the parties to any given case considered their own matter to be of the utmost importance, and that I should treat it that way.

You mentioned in a previous interview that "folks view electronic discovery as somebody else's problem." Now that a few years have passed since that interview and the amount of discoverable data has exploded in the way it has, do you still think legal professionals feel that way? What does the profession need to do in order to make ediscovery more of a priority in the eyes of both the courts and attorneys?

CONTINUED ▶

I think the view legal professionals have toward ediscovery has changed, and part of this is a function of its pervasiveness.

The partner who previously wouldn't have been interested in the details of ediscovery has come to recognize that the facts that will support or undermine the construction of their legal argument will likely come out of the ediscovery process.

There's now a better understanding of how central that process is to litigation generally.

The other thing that I think has changed is the popularization of technology generally. More lawyers are aware of things like ephemeral messaging and artificial intelligence, so they're becoming more comfortable using the technologies that are necessary for dealing with electronic data.

I don't think there's the same technophobia that existed 10 or 15 or 20 years ago. Folks are not only recognizing the importance of ediscovery, but they're feeling more comfortable with how it works.

Your ruling in the *Cat3* litigation that said courts do retain their inherent authority to bring spoliation sanctions in spite of the 2015 amendments to Rule 37(e) was notable in the legal community. With that in mind, what do you make of the Ninth Circuit's recent ruling in *Gregory v. State*, essentially saying that the district court in the case should not have used its inherent authority, and that the State's duty to preserve fell under Rule 37(e)?

To understand *Cat3* and my view of inherent authority, we need to take a step back and talk about the origins of Rule 37(e).

Before the 2015 amendments to the rules, federal courts generally applied common law principles to issues of spoliation. And since common law means judge-created, some differences arose in how the courts treated spoliation. In particular, some courts held that severe sanctions, such as dismissal or judgment by default or an adverse inference, could only be imposed where the court found that the spoliating party had acted with intent.

More lawyers are aware of things like ephemeral messaging and artificial intelligence, so they're becoming more comfortable using the technologies that are necessary for dealing with electronic data.

Other courts, including the Second Circuit in the case *Residential Funding Corp. v. DeGeorge Financial Corp*., held that in appropriate circumstances, even severe sanctions could be imposed to remedy negligent spoliation. The Rules Advisory Committee, in response to concerns that the threat of sanctions was causing parties to over-preserve ESI, explicitly rejected the *Residential Funding* decision and adopted the position of other circuits when it amended Rule 37(e). And the notes specifically say that the rule displaces inherent authority.

Gregory v. State of Montana, the Ninth Circuit case that you referred to, rejected the district court's use of inherent authority. In that case, the plaintiff had brought a civil rights action claiming the use of excessive force by a probation officer, and the state lost the videotape of the encounter. The court found that the state was only negligent, but nevertheless imposed an adverse inference pursuant to its inherent authority. The Ninth Circuit found that this was improper since Rule 37(e) applied.

Perhaps surprisingly, I agree with the Ninth Circuit. As it found, all the prerequisites for Rule 37(e) were present, and the district court had found that the police were negligent, not that they had destroyed the tape with the intent of depriving the plaintiff of the evidence. It's not my position that a court can always rely on inherent authority notwithstanding the rule. Rather, inherent authority fills the gaps where spoliation occurs, but for whatever reason, the conduct is outside of the rule.

For example, in *Cat3*, one of the parties manipulated ESI so that the evidence that was discoverable had in fact been fabricated. This circumstance is not addressed by the rule and, in my view, can therefore be sanctioned under inherent authority. In shorthand, my view is that the rule and inherent authority coexist but do not overlap.

At the same time, I think *Gregory* demonstrates why I believe that the amendment to Rule 37(e) was a bad idea, and that the position of *Residential Funding* was really the preferable law.

In *Gregory*, the police lost the evidence, and yet the cost of their negligence was imposed on the plaintiff, who's now deprived of potentially dispositive proof. I personally wouldn't have gone that way with the rule, but once the rule is in place, I agree with the court in *Gregory* that, in this case, the rule displaced inherent authority.

Can you speak a little bit more about why you feel the Rule 37(e) amendments were a bad idea, or what sorts of issues you have with them?

I wouldn't say that they're a bad idea across the board, but I think in the case of negligent versus intentional spoliation, the risk of spoliation should be on the spoliator, not on the innocent party.

CONTINUED ▶

I think the rule has shifted that risk to the innocent party, and I don't think the justification for that was empirically sound. That is, I don't think that parties were over-preserving ESI because of the threat of sanctions. They were certainly over-preserving ESI, but that tended to be, I believe, the result of a failure to have good data hygiene and also a function of big entities and big businesses having so much litigation and so many overlapping legal holds that it was impossible to reduce their preservation.

I think the problem existed, but not for the reasons that the Rules Advisory Committee had identified, and I believe they were solving the wrong problem.

Switching gears a bit to the emerging influence of generative AI on ediscovery, do you think generative AI has the potential to help with issues related to the ediscovery process? What do you see as its role in the discovery process overall?

I think it has a significant potential in discovery. And maybe we can go back and think about the history of discovery and how technology has impacted that, generally.

The first problem that technology had to confront was search. Some of us of a certain age remember when legions of associates or contract attorneys would have to go to the warehouse and pore through documents and try to sift out what was potentially relevant to a case. But with the advent of technology-assisted review, we've gone a long way toward mitigating that particular part of the problem. We've reduced the cost, increased the accuracy of the searches, and so forth.

Now, I think the big problem is determining how to craft a search that will be conducted in a manner that the requesting party can have faith in, because we end up with interminable fights over search terms and disputes over search protocols, and TAR protocols in particular. Generative AI has promise in this realm because, if you can imagine linking up the subject matter of the litigation with the process of search, it's eliminated many of those disputes.

If you train a tool with the pleadings, for example, and the request is for it to develop a search that can be utilized in a particular search tool, then you've gone a long way toward eliminating the disputes that occur in that interface. You could similarly ask the tool to modulate the search to reach particular recall numbers or particular precision numbers to satisfy counsel for whatever is feasible given the data set.

You could similarly ask the tool to modulate the search to reach particular recall numbers or particular precision numbers to satisfy counsel for whatever is feasible given the data set.

And potentially, at the end of the day, and with sufficient trust in the generative AI tool, you could even have counsel forgo human review for production. That is, they could agree that the generative AI tool is crafting a search that they are both satisfied with, and that they don't need to have a million attorneys looking at those documents that are produced by the search before the results are turned over to the other side.

If you train a tool with the pleadings, for example, and the request is for it to develop a search that can be utilized in a particular search tool, then you've gone a long way toward eliminating the disputes that occur in that interface.

Now, full disclosure, I made a prediction long ago that parties would use TAR in this way, and that they could agree that the results of a TAR review would be sufficient and they wouldn't need to do an eyes-on review. Frankly, I have not seen that come to fruition very frequently, so it may be that I'm over-predicting again, but I think it actually has more promise with generative AI. It just needs a track record so that counsel will become sufficiently comfortable with it.

You co-wrote a paper titled "Generative AI and Courts: How Are They Getting Along?", in which you discussed the rash of standing orders that have emerged from judges over the use of generative AI in their courtrooms, highlighting that the level of education isn't quite what it should be, and that some judges might not mean what they're saying in these orders. How can the judiciary better educate themselves about this technology? And, as a follow up, how should attorneys proceed when faced with a standing order that isn't particularly well-informed?

Let's talk about the genesis of those standing orders. There was a case in the Southern District of New York, *Mata v. Avianca Airlines*, where counsel submitted a brief to the court in which some of the content was created by a generative AI tool. Unfortunately, the tool hallucinated and produced a number of fabricated citations to cases, and counsel did not do a sufficient job of reviewing what the tool had done and submitted it to the court. Of course, not surprisingly, the court was quite upset at receiving a largely fabricated brief.

There were a number of cases thereafter where there was evidence that briefs had been submitted using generative AI-created content and had been less than accurate. A number of courts then imposed rules to try to limit this problem, and the rules ranged from ones that

CONTINUED ▶

required notice when counsel were utilizing a generative AI tool to one court that forbade any use of AI in the briefs.

In any event, these orders reflect the concern of the courts about receiving unreliable input from counsel. And I think the way that courts can be better educated about these kinds of issues is, number one, before you issue a rule, you vet it. That is, you talk to your colleagues, you talk to counsel, and you get input from experts to try and understand what the problem is and how you get inaccurate and unreliable information out of the generative AI tools, and then you can craft rules that are specifically directed toward what the problem is.

Additionally, as in any circumstance, counsel can attempt to educate the court. The judges are not hidebound. If they're shown why and how an individual rule is not entirely appropriate, they will generally adapt. The secret is to do it respectfully, and with the appropriate backup. That is, if you have an expert who can submit an affidavit that educates the court, I think that can be most effective.

What do you see as the main ethical issues facing legal professionals over the use of generative AI? How much should attorneys and law firms be thinking about ethics when selecting AI tools?

There are a range of ethical issues. The first and perhaps most obvious one is that counsel needs to conform to the requirements of competence in the rules of professional conduct, and that means getting up to speed and understanding what the limits and potential pitfalls of generative AI are.

The second is the issue of confidentiality, and this relates largely to the use of generative AI in discovery as opposed to in presentation of materials to the courts. But counsel have to understand that depending upon the nature of the tool, and the materials that they input to train the tool, those materials do not necessarily remain confidential. Once they are input, they may then be exposed to future users of the tool.

There are a range of ethical issues. The first and perhaps most obvious one is that counsel needs to conform to the requirements of competence in the rules of professional conduct, and that means getting up to speed and understanding what the limits and potential pitfalls of generative AI are.

A lot of folks understand that at one level, but then they forget that their adversaries may be using a generative AI tool as well. There needs to be some thought upfront, perhaps included in a protective order or in an ESI protocol, that addresses this.

If I turn my confidential material over to my adversary, I need to know that my adversary is not going to input it into a generative AI tool that is not self-contained.

You have fairly considerable experience working in the civil rights space. What sort of potential do you see for generative AI in helping level the playing field in these types of cases? What about its overall effect on addressing the access to justice issue?

I do think it has the potential to be quite impactful in the civil rights area, and in the area of access to justice by folks with limited resources.

It's going to depend on how the tools are priced, and how well-supervised they are. The danger is that if you tell a pro se litigant that a generative AI tool will generate a complaint for them that they can then file in federal court, they may input their problem, and the output that they then file in court, even if it's not hallucinated, is not sufficiently reviewed to be a legitimate legal filing.

It can surely assist impecunious litigants, but it still needs the kind of oversight that, say, a non-lawyer assistant who is helping a litigant would need from a lawyer. There still needs to be that overlay.

Having said that, it's still a major addition to the arsenal of tools that poor litigants can take advantage of. It has great promise, but it can't be used unsupervised.

Getting into your current role as a mediator and arbitrator with JAMS, has your attitude on the proliferation of legal technology changed since retiring from the bench and moving into the ADR space? How so?

It hasn't changed a lot, and the reason is that, notwithstanding the idea that arbitration is designed to be faster, less formal, and more streamlined, I think the parties in many cases still want the process to be fair. And to be fair, it often requires discovery, and discovery these days is often centered on the discovery of ESI.

As has often been noted, the arbitration process is beginning to look more and more like litigation, and in that respect, electronic discovery has become more a part of our arbitration practice. Many of the same disputes that occur in litigation over technology also appear in arbitration, so the distinctions have become blurred.

CONTINUED ▶

Acknowledging the release of the EDRM bench book, you wrote a blog about the use of Special Masters in ediscovery, highlighting the various benefits of using them. Do you wish they'd be used more widely? What do you see as their role in the current legal profession? What's the biggest barrier to their use?

I certainly think that there is more room for the use of Special Masters, and it's really a function of the caseload burden on courts. Courts simply do not have the time, and sometimes the inclination, to dig into issues like ediscovery, that are going to be very difficult to work out and may require a certain level of expertise that a judge doesn't have.

I think the problem is that if a judge doesn't use a Special Master in circumstances like that, you run the risk of getting relatively arbitrary outcomes. If I'm unfamiliar with ediscovery, I may make these blanket rulings that don't take into account the nuances of the case and the needs of the parties.

Just to take a simple example, when I'm confronted with disputes that have over 50 discovery requests, or 50 document demands, I can say to a party, "Choose your 10 best demands. You can have those, and we'll cut out the other 40." That's a little arbitrary. I haven't taken the time to dig in and try to figure out which of those requests are actually legitimate.

That's a role that Special Masters can really fill. They can get in and make those decisions or help the parties mediate the problem.

There are instances where I've been tasked as a Special Master, not with making rulings, but with trying to help the parties come to a resolution on ediscovery issues.

Special Masters bring different skills. Some have great technical abilities, and they are appropriate in certain kinds of cases where that skill is critical. Others are more familiar with decision-making, and they provide a different benefit.

Why do you think they aren't leveraged as much as they should be?

I think part of it is just tradition. If you're a judge and you see yourself as the decision-maker, you may not feel comfortable delegating that role, or at least you may not feel comfortable delegating it beyond the court. A district judge can always delegate it to the magistrate judge, and a state judge might delegate it to a referee. But they're not used to going outside the system. That's something that may take a little getting used to.

How do you think judges should be encouraged to utilize them more and reach outside the system?

I think the best way may be for counsel to ask for it because they have a stake in the matter. They're the ones who are going to be paying for it, and they understand where the judge's ability to deal with these problems begins and ends. Beyond that, I think advertising the effectiveness of Special Masters through judicial conferences and similar avenues is helpful.

What's something this next generation of judges and attorneys who are going to be dealing with generative AI and the continued development of legal technology can do to best educate themselves on its future impact, both in their roles and in the legal system at large?

One of the best things that lawyers and judges can do is get involved in situations where they're exposed to technologists, whether it's through conferences that have both lawyers and technologists or inviting technologists to do seminars in the courthouse or at the law firm.

It's really useful to hear directly from the people who can tell you what the positive attributes and the limits and the potential of the different technologies might be. When you hear it from another lawyer, it may not have the same impact, but getting the technologists in to talk to the lawyers and judges is really beneficial.

*JAMS arbitrator and mediator **Hon. James 'Jay' C. Francis IV (Ret.)** is a former U.S. Magistrate Judge for the Southern District of New York and a nationally recognized authority on electronic discovery. Learn more about him at jamsadr.com/francis.*

This article originally appeared in Everlaw.com and is reprinted with their permission.