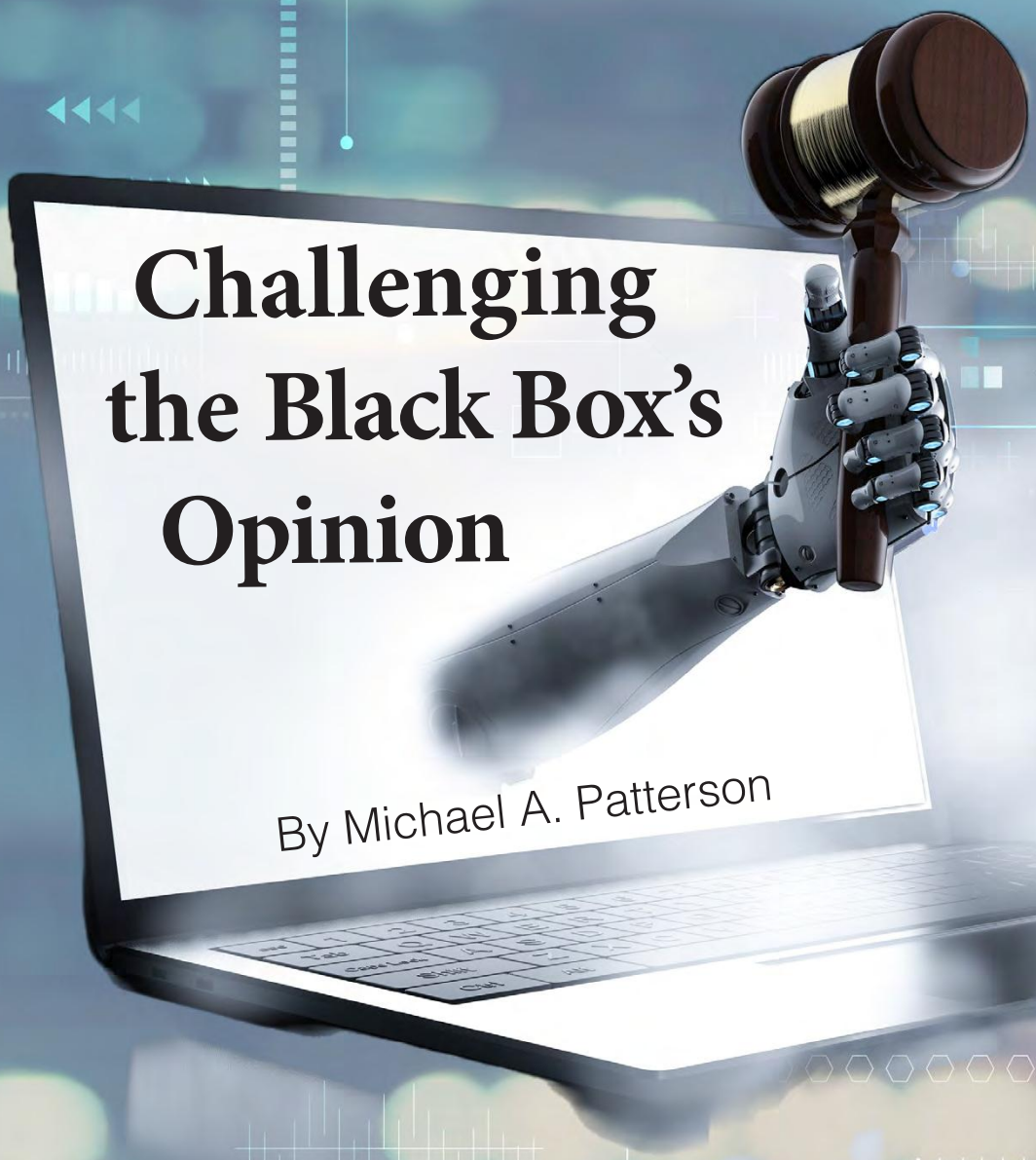
ARTIFICIAL INTELLIGENCE ISSUES:

Challenging the Black Box's Opinion



By Michael A. Patterson

Artificial intelligence algorithms are making high-stakes decisions in many areas of our lives. They diagnose and treat patients. They decide which employee is recommended for promotion or termination. They decide if your credit qualifies you for a credit card or a mortgage. They decide who gets admitted to college. They predict whether attorneys will win or lose a lawsuit. They identify offenders that are most likely to recidivate.

When a trial judge is deciding about release of an individual, he may now rely on systems that assess the likelihood of a defendant to re-offend when making bail decisions. The problem with these systems is that their predictions look to the past to make guesses about future events. "In a racially stratified world, any method of prediction will project the inequalities of the past into the future. This is true of the subjective prediction that has long pervaded criminal justice as it is of the algorithmic tools now replacing it."[1] Studies of these systems report that they underestimate the probability of white recidivism, while overestimating the probability of black recidivism.[2]

In the past, these opinions were only reached by humans, but now they are reached by artificial intelligence. The big difference is that if you are in a lawsuit over these opinions, and they were made by a human, you could cross examine that person about that opinion. You have no such opportunity if the opinions are made by an artificial intelligence algorithm.

Understanding the Terms

"Artificial Intelligence (AI) is the hypothetical ability of a computer to match or exceed a human's performance in tasks requiring synthesis, reasoning, creativity and emotion."[3]

"Machine learning (ML) enables computers to 'receive data and learn for themselves' from 'examples rather than a list of instructions.' ML uses statistical methods that incorporate algorithms to mimic human thought . . . ."[4]

"Deep learning is the most advanced part of the 'machine learning spectrum.' Deep learning 'refers to a set of highly intensive computational models' that 'allow an algorithm to program itself by learning from a large set of examples that demonstrate the desired behavior, removing the need [for humans] to specify the rules explicitly."[5]

". . . [B]ecause of how the AI Ecosystem operates, it may be impossible to reverse engineer the decision-making process to know on which data the AI system relied. This is the classic 'black box problem' that reflects the lack of transparency and explainability that may render the AI decision-making process impenetrable."[6]

How Will You Challenge the Black Box?

Like all evidence, the proponent must establish authenticity and admissibility.

The issue of authenticity has been addressed in federal court with the addition of Fed. R. Evid. 902(13) and (14). Louisiana has not yet adopted similar amendments to the Louisiana Rules of Evidence.

Continued next page

Fed. R. Evid. 902(13) provides:

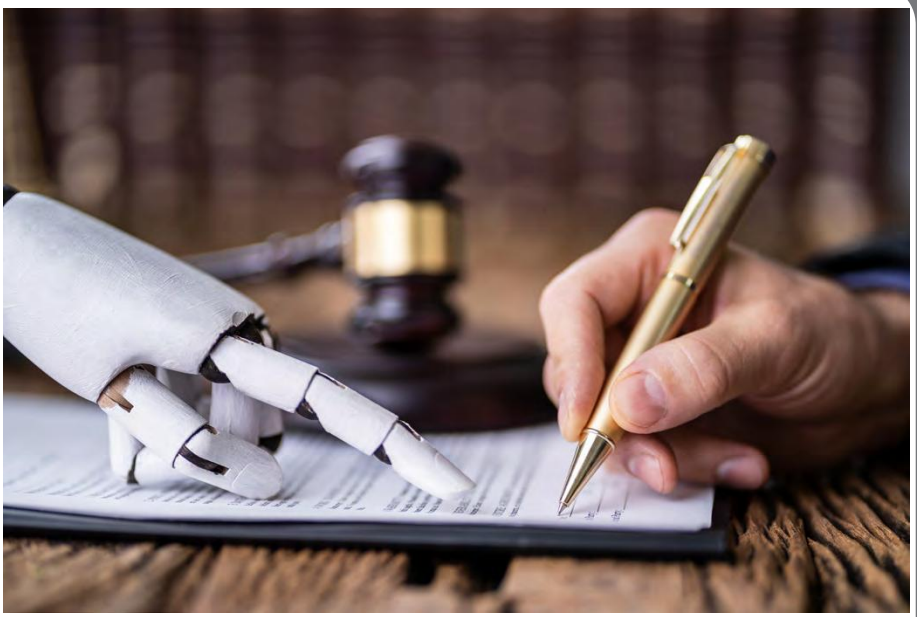
> A record generated by an electronic process or system that produces an accurate result, as shown by a certification of a qualified person that complies with the certification requirements of Rule 902(11) or (12), is self-authenticating. The proponent must also meet the notice requirements of Rule 902(11).

This rule eliminates the necessity of a foundational witness at trial. The person who makes the certification must meet Rule 602 (personal knowledge) and Rule 702 (scientific or specialized knowledge) and Rule 901(b)(9) requiring explanation of how the process or system that generated the electronic record produces reliable and accurate results.

With a human expert, if you want to challenge the opinion, you file a motion *in limine* which allows you to put on evidence prior to trial to contest the expert's qualifications to give the opinion or the expert's methodology.

The proponent of the AI opinion will have to establish that the AI algorithm produces accurate results. If the proponent cannot do that, then the evidence is unreliable. Unreliable evidence is not relevant.

Authentication requires the proponent to show the technology produces accurate and reliable results. "When the accuracy of technical evidence has been verified by testing; the methodology used to develop it has been published and subject to review by others in the same field of science or technology; when the error rate associated with its use is not unacceptably high; when standard testing methods and protocols have been followed; and when the methodology used is generally accepted within the field of similar scientists or technologists, then it can be established as authentic because it does what its proponents say it does."[7]



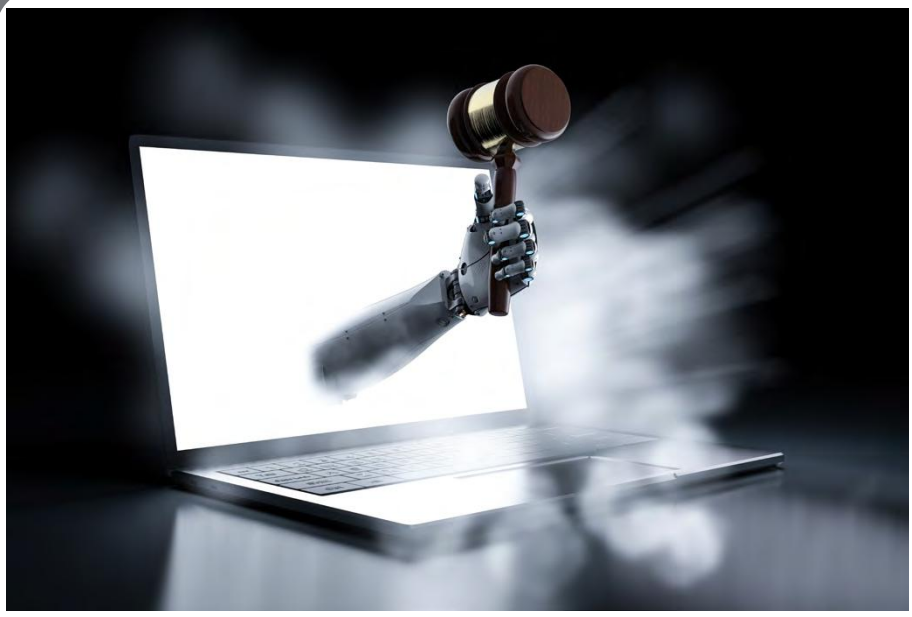
We know that a party planning to have an expert testify at trial and give opinions must comply with FRCP 26 to allow the other side to see the proposed opinions and what the expert relied upon and his methodology. A trial judge probably should require early disclosure of proposed AI opinion evidence with the same requirements of Rule 26 to allow the opponent time to challenge the evidence at a hearing prior to trial and to give the trial judge enough time to decide on the admissibility of the proposed evidence. This may be an area where we see trial judges making use of Fed. R. Evid. 706 to get expert assistance to help reach these highly technical decisions.

Another thing to consider is a special jury instruction which would make clear that the fact that the AI opinion was admitted and there was no cross examination does not mean the jurors should automatically accept the AI opinion but should consider it along with all other evidence.

CONCLUSION

At the end of the day, if the court finds the AI opinion authentic, then it will be admitted in evidence and the opponent will not be able to cross examine the opinion at trial.

Machine-derived evidence is only as unbiased and fair as it is designed to be. "The potential prejudicial effect upon jurors is that they see only the outputs generated by the AI, but they cannot 'peer into' the system generating those outputs. The danger is the presumption of reliability and credibility jurors may place on the 'testimony' provided by these systems without considering that, although faster and more efficient, algorithms are human-made and, therefore, can be flawed. Even with the advent of 'self-learning machines,' there is still no guarantee of a 'zero-error rate' because the genesis of even 'self-learning machines' are human beings who are flawed."[8]



FOOTNOTES

1. Sandra G. Mayson, "Bias In, Bias Out," 128 Yale L.J. 2218, 2218 (2019).

2. Yifat Nahmias, Maayan Peret, 58 Harv. J. on Legis. 145, (Winter 2021).

3. Paul. W. Grimm, Maura. R. Grossman and Gordon V. Cormack, *Artificial Intelligence as Evidence,* 19 Nw. J. Tech & Intell. Prop. 9, 14 (2021).

4. Frank Griffin, *Artificial Intelligence and Liability in Health Care*, 31 Health Matrix, 65, 73 (2021).

5. *Id.* at 73-74.

6. Ira Giuffrida, "Liability for AI Decision-Making: Some Legal and Ethical Considerations," 88 Fordham L. Rev. 439, 442 (2019).

7. The Sedona Conference, *Commenting on ESI Evidence & Admissibility*, 22 Sedona Conf. J. 83, 190 (2021); Paul W. Grimm, "Practical Considerations for the Admissibility of Artificial Intelligence Evidence," 2 No. 3 Md. B.J. 39, 40 (2021).

8. Victor Nicholas A. Metallo, "The Impact of AI Technology on Forensic Accounting Expert Testimony," 69 Emory L.J. Online, 2039, 2051 (2020).

Michael A. Patterson is a partner with Long Law Firm, LLP, in Baton Rouge and is a founding and managing member of The Patterson Resolution Group. He received his BA and JD degrees from Louisiana State University and a certificate and LLM in dispute resolution from Pepperdine University. He is an adjunct professor of trial advocacy and evidence at LSU Paul M. Hebert Law Center. He served as president of the Louisiana State Bar Association in 2010-11. (mpatterson@patterson-resolution.com; map@longlaw.com; 1800 City Farm Dr., Bldg. 6, Baton Rouge, LA 70806)