

---

# RESPONSE

## AI AND ADVERSARIALISM<sup>†</sup>

ANDREA ROTH\*

### CONTENTS

|                                                                                                                                     |      |
|-------------------------------------------------------------------------------------------------------------------------------------|------|
| INTRODUCTION .....                                                                                                                  | 2208 |
| I. SOME AI-ENHANCED EVIDENCE IS AKIN TO A HUMAN ASSERTION<br>AND DESERVES <i>DAUBERT/FRYE</i> SCRUTINY .....                        | 2211 |
| II. SUBJECTING SOME AI-“ENHANCED” EVIDENCE TO <i>DAUBERT/FRYE</i><br>SCRUTINY WILL NOT EXACERBATE THE “INEQUALITY OF<br>ARMS” ..... | 2215 |
| CONCLUSION .....                                                                                                                    | 2219 |

---

<sup>†</sup> An invited Response to Mary D. Fan, *AI-Enhanced Evidence*, 105 B.U. L. REV. 2149 (2025).

\* Professor of Law and Barry Tarlow Chancellor’s Chair in Criminal Justice, U.C. Berkeley School of Law. Thanks to Cameryn Bryant for research assistance.

## INTRODUCTION

Mary Fan's fascinating and important article, *AI-Enhanced Evidence*,<sup>1</sup> starts with a tale of two cases, both involving what she calls "AI-enhanced evidence," or evidence resulting from an artificial intelligence ("AI") program's "enhancement" or "clean up" of grainy video or hard-to-hear audio recordings.<sup>2</sup> One, a Washington case, was offered by the defense as evidence of innocence and yet was excluded by the court as not yet generally accepted by the scientific community as reliable.<sup>3</sup> The other, a case in Florida, was offered by the prosecution as evidence of guilt and was admitted by the court as being sufficiently reliable.<sup>4</sup>

Professor Fan's conclusion from these cases is that the threshold for admitting such AI-enhanced evidence should be lowered, to eliminate the "inequality of arms" between parties, and to allow the defense to offer such evidence more easily.<sup>5</sup> She further argues that the threshold for this sort of evidence has in fact traditionally been low, through lenient rules governing authentication and demonstrative evidence,<sup>6</sup> and that the tendency of courts in cases like *State v. Puloka*<sup>7</sup> and *State v. Magbanua*<sup>8</sup> to subject such evidence instead to *Daubert/Frye* reliability requirements<sup>9</sup> is both new and inappropriate.<sup>10</sup> Fan surmises that this newfound skepticism of AI-enhanced evidence is an overcorrection by courts based on legitimate concerns about an entirely different modern problem—"AI-generated" evidence like deepfake images and videos.<sup>11</sup> Still, acknowledging potential reliability issues with AI-enhanced evidence, Fan argues that a more reasonable, modest alternative to *Daubert/Frye* scrutiny would be judicial standing orders, capable of adapting more quickly to

---

<sup>1</sup> Mary D.M. Fan, *AI-Enhanced Evidence*, 105 B.U. L. REV. 2149 (2025).

<sup>2</sup> *See id.* at 2183.

<sup>3</sup> *See id.* at 2152-53 (discussing denial of evidence of self-defense enhanced by AI offered by defendant).

<sup>4</sup> *See id.* at 2153-56 (discussing conviction of one defendant after state introduced evidence enhanced by AI which captured inculpatory dialogue).

<sup>5</sup> *Id.* at 2157.

<sup>6</sup> *Id.* at 2160.

<sup>7</sup> No. 21-1-04851-2 (Wash. Super. Ct. Mar. 29, 2024).

<sup>8</sup> No. 2016CF3036 (Fla. Cir. Ct. Apr. 22, 2022).

<sup>9</sup> By "*Daubert/Frye*" reliability requirements, I mean admissibility requirements for expert testimony used around the country. *Daubert* jurisdictions, like the federal system, require a judge to find that the expert's method is reliable and reliably applied. *See* FED. R. EVID. 702; *Daubert v. Merrell Dow Pharms.*, 509 U.S. 579, 592-96 (1993). Alternatively, in *Frye* jurisdictions, the proponent of a novel scientific method must show the method has "gained general acceptance" in the relevant scientific community. *See Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923).

<sup>10</sup> *See* Fan, *supra* note 1, at 2184-85.

<sup>11</sup> *See id.* at 2185 ("Haunting judicial resistance to AI-enhanced evidence is anxiety over generative AI and the risk of manipulated, altered, or fake evidence.").

technological change than the laborious rulemaking process, that would require disclosure of the program and how it works.<sup>12</sup>

Professor Fan's article makes three critical points, with which I completely agree, that need to be better understood by judges and practitioners. First, Fan rightly points out that the legal community often conflates *authentication* of an image, video, or other software output, on the one hand, with establishing the *reliability* or *accuracy* of such output, on the other.<sup>13</sup> The difference matters. The question of whether an image is a deepfake is one of authentication: Is this image what it claims to be—the Ukrainian president telling his troops to surrender? On the other hand, the question of whether a facial recognition algorithm has correctly identified the person in a video is one of *reliability* or *accuracy*: Is the algorithm's claim—that this is the person in the video—true or false? Perhaps this conflation is natural given the complexity of the subject matter. But it might stem from the “original sin” of the drafters of the Federal Rules of Evidence (“FRE”) in the late 1960s choosing to deal with the burgeoning issue of “computerized business records” through a quick addition to FRE 901(b)(9).<sup>14</sup> The drafter chose to deal with authentication, which requires that the proponent of “evidence about a process or system” produce evidence sufficient to support a finding that the process or system “produces an accurate result,”<sup>15</sup> rather than a true reliability requirement for software results akin to FRE 702 for human experts.<sup>16</sup>

Second, Professor Fan persuasively argues that some recent proposals to deal with AI-related evidence would place a nearly impossible burden on parties presenting evidence, especially those that are under-resourced.<sup>17</sup> One such proposal, she notes, would apparently require an expert affidavit as a condition of admitting any video in any trial.<sup>18</sup> Luckily, as Fan acknowledges, these proposals do not seem to be gaining steam.<sup>19</sup> Additionally, as Ed Cheng and Rebecca Wexler have recently argued,<sup>20</sup> even the rules governing authenticity

---

<sup>12</sup> See *id.* at 113-14.

<sup>13</sup> See *id.* at 115-16.

<sup>14</sup> See Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972, 2012-13 (2017) (discussing original reasons drafters inserted the phrase “accurate result” in FRE 901(b)(9)); Edward J. Imwinkelried, *The Challenge that the Advent of Artificial Intelligence (AI) Tools Poses to the Procedures for Determining the Existence of the Preliminary Facts that Condition the Admissibility of Items of Evidence*, 108 MARQ. L. REV. 621, 642-43 (2025) (arguing FRE 901(b)(9) is for authentication and not substitute for reliability).

<sup>15</sup> See FED. R. EVID. 901(b)(9).

<sup>16</sup> See FED. R. EVID. 702.

<sup>17</sup> See Fan, *supra* note 1, at 2191-92.

<sup>18</sup> See *id.*

<sup>19</sup> See *id.*

<sup>20</sup> See Edward K. Cheng, *Deepfakes, Photographs, and Trust in Evidence*, 112 VA. L. REV. ONLINE (forthcoming 2026) (manuscript at 7-8) (on file with author) (arguing against newer more restrictive rules to address deepfakes because jurors weigh evidence based on trust in source of information, which current rules of authentication are well-equipped to handle);

arguably should not be further tightened to address deepfakes. To change the rules for all types of media that could theoretically be doctored or tainted by AI would largely upend the system of proof we know.

Third, Professor Fan rightly reminds the legal community that judicial standing orders are an underused, nimble way of addressing some of these issues, especially given how long the FRE amendment process takes.<sup>21</sup> Although there are now 310 AI-related standing orders around the country as of this writing,<sup>22</sup> they appear so far to be mostly related to the use of AI in legal drafting rather than imposing robust reliability or disclosure requirements for evidence.<sup>23</sup> To be sure, the FRE Advisory Committee is currently discussing potential amendments to FRE 909 that would address deepfakes,<sup>24</sup> and there is a proposed new federal rule of evidence currently open for public comment that would treat computer-generated conclusions akin to expert testimony uttered by a human expert.<sup>25</sup>

I disagree with Professor Fan, however, on two points that I discuss in greater depth in the remainder of this Response. First, I disagree with Fan about whether AI-enhanced evidence is ever deserving of *Daubert/Frye* scrutiny (I think it sometimes is). And that disagreement relates to a change I would make to Fan's description of the "spectrum" of AI evidence, from AI-"enhanced" evidence (where AI has "cleaned up" an existing image, video, or audio recording) to AI-"generated" evidence (essentially, a deepfake).<sup>26</sup> In my view, this spectrum both

---

ADVISORY COMM. ON EVIDENCE RULES, REPORTER'S MEMO ON AI AND MACHINE LEARNING 9-10 (2024), <https://www.uscourts.gov/forms-rules/records-rules-committees/agenda-books/advisory-committee-evidence-rules-april-2024> [<https://perma.cc/R6V5-PPQV>] (describing Rebecca Wexler's opposition to proposed amendment to 901 to deal with deepfakes).

<sup>21</sup> See Fan, *supra* note 1, at 2200-02.

<sup>22</sup> *Standing Orders, Local Rules, and Decisions on the Use of AI*, ROPES & GRAY, <https://www.ropesgray.com/en/sites/artificial-intelligence-court-order-tracker> [<https://perma.cc/QU87-E7RK>] (last visited Dec. 14, 2025).

<sup>23</sup> See Audrey Mitchell, *Governing Generative AI in Litigation Through Standing Orders* 20 (Apr. 24, 2025) (unpublished note) (on file with author).

<sup>24</sup> See Daniel J. Capra, *Deepfakes Reach the Advisory Committee on Evidence Rules*, 92 FORDHAM L. REV. 2491, 2492 (2024); ADVISORY COMM. ON EVIDENCE RULES, MINUTES OF THE MEETING OF NOVEMBER 8, 2024, at 8-13 (2025), <https://www.uscourts.gov/forms-rules/records-rules-committees/agenda-books/advisory-committee-evidence-rules-may-2025> [<https://perma.cc/LA8C-9RWY>]; AI Windham, *AI-Generated Deepfakes in Court: An Emerging Threat to Evidence Authenticity?*, NAT'L L. REV. (June 11, 2025), <https://natlawreview.com/article/ai-generated-deepfakes-court-emerging-threat-evidence-authenticity> [<https://perma.cc/NQ56-EB8C>].

<sup>25</sup> See ADVISORY COMM. ON EVIDENCE RULES, *supra* note 20, at 5-10 (discussing proposed amendments to FRE 901); JUD. CONF. OF THE U.S., PRELIMINARY DRAFT: PROPOSED AMENDMENTS TO THE FEDERAL RULES OF APPELLATE, BANKRUPTCY, CIVIL, AND CRIMINAL PROCEDURE, AND THE FEDERAL RULES OF EVIDENCE 109 (2025), [https://www.uscourts.gov/sites/default/files/document/preliminary-draft-of-proposed-amendments-to-federal-rules\\_august2025.pdf](https://www.uscourts.gov/sites/default/files/document/preliminary-draft-of-proposed-amendments-to-federal-rules_august2025.pdf) [<https://perma.cc/FX64-YA2P>] (proposing new Federal Rule of Evidence 707).

<sup>26</sup> See Fan, *supra* note 1, 2157-58.

omits AI-generated independent analysis that is akin to expert testimony and fails to acknowledge that AI-enhanced evidence can, in some cases, bleed into AI-generated independent analysis.

In turn, while I agree with Fan that the impact on the defense's "inequality of arms" is typically underappreciated in debates about admissibility of evidence,<sup>27</sup> I think this context is one in which the defense would on balance benefit from robust reliability scrutiny of some AI-"enhanced" evidence. Of course, such a reliability requirement could come in the form of a standing order rather than a new rule of evidence. Ultimately, I agree with Fan on the need for both parties to understand how AI output works, how it is used, and its track record.

I. SOME AI-ENHANCED EVIDENCE IS AKIN TO A HUMAN ASSERTION AND DESERVES *DAUBERT*/*Frye* SCRUTINY

Though Professor Fan does not directly define AI-enhanced evidence, she describes that side of the spectrum (from AI-enhanced to AI-generated evidence) as involving "useful, truth-revealing clarification of data" by AI, rather than AI-generated fake content.<sup>28</sup> Yet as she later briefly mentions, there is a third category of AI-generated evidence, which she refers to as "[a]nalytical AI—using techniques and algorithms associated with AI to analyze and interpret data."<sup>29</sup>

What should be further acknowledged, though, is that the line between using AI merely to clarify data versus to "analyze and interpret" data will be blurry at the margins, meaning that some of what Professor Fan labels AI-enhanced evidence is more akin to analytical AI.<sup>30</sup> As it turns out, the question of whether AI-enhanced evidence merely clarifies data or further analyzes it is analogous to a lower-tech problem that has plagued courts for years: whether a human court reporter or interpreter adds another layer to a hearsay analysis.<sup>31</sup> Some courts view a court reporter or interpreter as merely a "language conduit" who simply regurgitates or translates a declarant's statement.<sup>32</sup> Others treat court reporters and interpreters as additional declarants, who themselves are making independent reflective assessments about what someone else said.<sup>33</sup> The question, for these purposes, is whether the additional layer (of

---

<sup>27</sup> See *id.* at 2159-60.

<sup>28</sup> *Id.* at 2158.

<sup>29</sup> *Id.* at 2186.

<sup>30</sup> See *id.* at 2157-58 (explaining AI-enhanced evidence); *id.* at 2186 (explaining analytical AI).

<sup>31</sup> See Casen B. Ross, *Clogged Conduits: A Defendant's Right to Confront His Translated Statements*, 81 U. CHI. L. REV. 1931, 1931-33 (2014) (outlining whether interpreter is declarant for hearsay purposes).

<sup>32</sup> See *id.* at 1948-49.

<sup>33</sup> See *id.* at 1933.

relater/interpreter) is adding an “independent analysis” rather than merely a conduit.<sup>34</sup>

Some AI output is obviously analytical (conducting what is akin to “independent analysis”) and would presumably not be labeled mere “enhancement” in Professor Fan’s taxonomy.<sup>35</sup> An obvious example might be a Large Language Model (“LLM”) trained to analyze a videotaped interview and predict whether the subject is malingering a mental illness.<sup>36</sup> As a side note, and in line with Fan’s point about courts conflating authentication and reliability,<sup>37</sup> it might surprise some readers to know that such “analytical AI” is subject to *Daubert* only if it is deemed the method used by the expert in reaching their own independent opinion.<sup>38</sup> Otherwise, it is subject only to FRE 901(b)(9)’s authentication requirement.<sup>39</sup> A new proposed FRE 707 would extend FRE 702 to computer-generated opinions so that these opinions would be subject to the same FRE 702 analysis as if uttered by a human expert, but this new proposed rule is in early stages of consideration.<sup>40</sup>

On the “conduit” side of the spectrum, some uses of AI are clearly mere recitations of, or a “platform”<sup>41</sup> for, human assertions. For example, basic tape recorders or dictographs, text message platforms, social media sites that contain a person’s posts, databases listing entries made by employees, or other type of “electronically stored information,” would all be in the same category.<sup>42</sup> Indeed, FRE 902(13) and (14) already reflect the difference between a record “generated

---

<sup>34</sup> See *id.*

<sup>35</sup> See Fan, *supra* note 1, 2158-59 (differentiating between AI-enhanced evidence and AI-generated evidence).

<sup>36</sup> See Jacob Salinsky, Raymond Finzel, Martin Michalowski & Sergui Pakhomovet, *Automated Neural Nursing Assistant (ANNA): An Over-The-Phone System for Cognitive Monitoring*, 2023 INTERSPEECH 684, 684 (describing development of AI which determines, based on series of conversations, whether individual is or is not mentally deteriorating).

<sup>37</sup> See, e.g., Fan, *supra* note 1, at 2161-62 (discussing efforts by courts to not admit AI-enhanced evidence due to fear of deepfakes).

<sup>38</sup> See FED. R. EVID. 702(d).

<sup>39</sup> See *United States v. Lizzaraga-Tirado*, 789 F.3d 1107, 1109-10 (9th Cir. 2015) (explaining only FRE governing Google Earth location estimates are 901(b)(9) and 201 on judicial notice).

<sup>40</sup> See JUD. CONF. OF THE U.S., *supra* note 25, at 109-11 (proposed FRE 707 and advisory committee note).

<sup>41</sup> See Roth, *supra* note 14, at 2001-02 (discussing difference between “platforms” and testimony-like machine conveyances of information).

<sup>42</sup> See George Cornell, Note, *The Evidentiary Value of Automatically Transcribed Voicemail Messages*, 17 B.U. J. SCI. & TECH. L. 259, 277-78 (2011) (discussing how “electronically stored communications” like automatically transcribed voicemails may merely memorialize human statements).

by” an electronic process and one “copied from” an electronic device or medium.<sup>43</sup>

Of course, even the most basic “cleaning up” of an image, video, or audio recording, even a basic change in volume, could be seen as different from the original. But if the difference is a matter of changed inputs that a human witness could explain and demonstrate quite simply—such as changing the treble or bass setting on a played recording, or changing the lighting on a photograph—that seems to be the type of change that a lay witness could explain, without the need for a *Daubert* hearing or any expert at all.<sup>44</sup> Perhaps if the inputs were more complicated and difficult for lay people to understand, an expert might be needed to explain those inputs or changes, in which case whatever method the expert used would (and should) be subject to *Daubert* or *Frye*.<sup>45</sup> Not surprisingly, then, the cases cited by Professor Fan to support the claim that courts before deepfakes took a lenient approach to technological “enhancement” of audio and visual evidence were cases where an expert fully explained the basic input changes they themselves made to alter an image or audio recording (e.g., changes to lighting).<sup>46</sup>

In contrast, cases like *Puloka* and *Magbanua* seem much closer to the “independent analysis” side, where the AI tool itself is controlling what changes to make, in ways that are proprietary and inscrutable. For example, in *Puloka*, the State alleged that the Topaz Labs proprietary algorithm, which is powered by machine learning, changed (for reasons unexplainable by the testifying expert) colors of participants’ clothes and shoes and showed a sharp object protruding from a man’s leg that in the original was only a shadow.<sup>47</sup> Likewise, the proprietary algorithm in *Magbanua*, GLIMPSE, also uses machine learning.<sup>48</sup> Some commentators have noted that such audio enhancement algorithms might go beyond simply making a voice easier to discern and instead

---

<sup>43</sup> See FED. R. EVID. 902(13)-(14). The ability of email, texts, social media posts, or a tape recording to be manipulated, however, does not render the resulting product an independent analysis rather than a conduit. See Roth, *supra* note 14, at 2002-04. Rather, the doctored information would be akin to a deepfake. The admission of such evidence may turn on authenticating whether the human declarant actually made the statement, but it raises no novel issue of machine credibility. See *id.* at 2002.

<sup>44</sup> See Roth, *supra* note 14, at 2014 (noting some courts have admitted evidence produced by nonscientific technical process without *Daubert* hearing).

<sup>45</sup> See *id.* at 2021-22 (explaining more complex technical processes are subject to *Daubert* and other evidence rules to determine reliability, but as technology becomes more sophisticated, legal system may need more than *Daubert*).

<sup>46</sup> See Fan, *supra* note 1, at 2164 n.89 (collecting cases exemplifying previous more flexible approaches to admitting evidence).

<sup>47</sup> See Findings of Fact and Conclusions of Law Re: *Frye* Hearing on Admissibility of Videos Enhanced by Artificial Intelligence at ¶¶ 16-18, State v. Puloka, No. 21-1-04852-2 (Wash. Sup. Ct. Mar. 29, 2024) [hereinafter Findings of Fact].

<sup>48</sup> See WAVESCIENCE, <https://www.wavesciencescorp.com/solution> [<https://perma.cc/8ZB5-JXES>] (last visited Dec. 20, 2025) (explaining GLIMPSE is powered by machine learning).

materially change a voice to sound faster or more robotic.<sup>49</sup> After all, imagine if a human declarant said to police, “actually, I was there and what you don’t see in this grainy video is that the sign actually says X, the license plate was Y, and the guy actually had a shiny knife in his hand that you could easily see in the dark.” That would clearly be an additional layer of hearsay by a human declarant that would need to be confronted and scrutinized if admitted.<sup>50</sup> In key respects, the enhancement in *Puloka* and the resulting artifacts, if the court’s order is any indication, seem more like an additional (machine) declarant.

One could imagine other AI-“enhancement” systems at the margins that should be treated as adding a layer of independent analysis. For example, proprietary automatic transcription services such as Google Voice can be “extremely inaccurate” under certain conditions, such as when a speaker has a heavy accent.<sup>51</sup> Likewise, police reports generated by AI programs like Axon’s Draft One might be thought of as merely “clarifying” or “translating” body-worn camera audio into the template of a police report.<sup>52</sup> But such programs create their own narratives and might in the process add hallucinated details or be trained to smooth over events that appear unusual.<sup>53</sup>

In turn, AI-“enhancement” algorithms delivering the equivalent of independent analysis should be subject to *Daubert/Frye* when these algorithms are the method by which an expert renders an opinion. In such instances, the expert is not wielding an AI tool to change basic known inputs or conditions to view the evidence in a new light. Rather, the decisions about whether and how to change the evidence are being made by the algorithm itself, through a proprietary process that even the expert might not be able to understand, much less replicate.<sup>54</sup> To the extent we subject analytical AI (such as the LLM

---

<sup>49</sup> See, e.g., SCI. WORKING GRP. ON DIGIT. EVIDENCE, SWGDE BEST PRACTICES FOR THE ENHANCEMENT OF DIGITAL AUDIO 4 (Sep. 17, 2020), <https://www.swgde.org/documents/published-complete-listing/20-a-001-swgde-best-practices-for-the-enhancement-of-digital-audio/> [<https://perma.cc/QCD2-TSQZ>] (noting need to “apply the required processing without creating unwanted artifacts”); *id.* at 6 (noting possibility audio enhancement could impact speaker recognition results); Yan Ru Pei, Ritik Shrivastava & FNU Siddarth, *Raw Speech Enhancement with Deep State Space Model*, ARXIV (Sep. 5, 2024), <https://arxiv.org/html/2409.03377v1> [<https://perma.cc/432Q-US5G>] (noting “unnatural artifacts . . . are common with raw audio processing systems” and may include “‘humming’ noise” and “robotic tone of speech”).

<sup>50</sup> See FED. R. EVID. 801, 802.

<sup>51</sup> See Cornell, *supra* note 42, at 283, 283 n.185 (discussing evidentiary issues raised by inaccuracy of Google Voice and other automatic transcription services).

<sup>52</sup> See Andrew Guthrie Ferguson, *Generative Suspicion and Risks of AI-Assisted Police Reports*, 120 NW. U. L. REV. 301, 316-17 (2025) (explaining “Axon introduced a new data-driven policing service, Draft One—an AI-assistant that drafts police reports using data from police-worn body cameras”).

<sup>53</sup> See *id.* at 320-21.

<sup>54</sup> I am not sure I agree with Professor Fan’s characterization that the AI-enhancement in *Puloka* “was not the basis of expert opinion” and was instead “submitted to support Puloka’s



determining mental illness) to *Daubert/Frye*, there does not appear to be a principled way to distinguish those systems from AI-“enhancement” algorithms that engage in independent analysis.

The stakes for defendants of allowing AI-enhancement algorithms engaged in independent analysis to avoid reliability scrutiny might well be high in a given case. For example, while the Topaz Labs algorithm in *Puloka* was introduced by the defense,<sup>55</sup> what if the government had introduced the video and the program had created what appeared to be a shiny weapon in Puloka’s hand that was not there in the original video, which Puloka insisted he did not have? If the government’s expert could not meaningfully explain how and why that shiny object appeared, other than to say the proprietary algorithm “enhances” or “clarifies” the video, the defense would presumably want to access the tool themselves to see how it reacts to various inputs, to scrutinize the validation studies, and so forth. In other words, they would want a *Daubert* or *Frye* hearing.

## II. SUBJECTING SOME AI-“ENHANCED” EVIDENCE TO *DAUBERT/FRYE* SCRUTINY WILL NOT EXACERBATE THE “INEQUALITY OF ARMS”

As noted above, Professor Fan rightly critiques several proposals around the country that would impose highly burdensome requirements involving expert affidavits on proponents seeking to introduce nearly any image or video.<sup>56</sup> Indigent criminal defendants, among others, obviously lack the resources to meet such a high bar, while the state, with its spate of experts on payroll, would have an easier time of it.<sup>57</sup> Fan also correctly points out that the state wins a disproportionate number of *Daubert/Frye* hearings, both when it challenges defense evidence and when it seeks to put on its own expert testimony.<sup>58</sup>

Moreover, to the extent Professor Fan’s concern about imposing *Daubert* scrutiny reflects a more general skepticism of strict admissibility rules as compared to discovery and impeachment rules, I completely share that skepticism. Given how often defendants lose admissibility hearings and given that defendants themselves sometimes hope to put on evidence, it may be that a robust right to have *access* to the government’s evidence, to *discover* key evidence before trial that is material to the preparation of the defense, and to *impeach* the government’s evidence at trial is equally as important, if not more important, than robust admissibility requirements. As Rebecca Wexler has recently argued, the best way for the criminal system to regulate AI output in a

---

side of the story.” Fan, *supra* note 1, at 2182. The latter is simply what makes it *relevant*; all expert and lay testimony must be relevant. See FED. R. EVID. 402. The fact that an AI-enhanced result supports Puloka’s defense does not make it any less in need of *Daubert* or *Frye* scrutiny.

<sup>55</sup> See Findings of Fact, *supra* note 47, ¶ 3.

<sup>56</sup> See Fan, *supra* note 1, at 2191-92.

<sup>57</sup> See Eva S. Nilsen, *The Criminal Defense Lawyer’s Reliance on Bias and Prejudice*, 8 GEO. J. LEGAL ETHICS 1, 19-20 (1994).

<sup>58</sup> See Fan, *supra* note 1, 2185.

fair and rational way is to insist upon the creation of *metrics* by which the jury can accurately assess the probative value of the evidence, not merely *thresholds* such evidence must meet to be admitted.<sup>59</sup> Of course, we have a long way to go in ensuring that defendants have the tools to meaningfully learn the limits of and to impeach AI-generated opinions.<sup>60</sup>

Nonetheless, although I sympathize with Professor Fan's concern that imposing *Daubert/Frye* scrutiny in cases like *Puloka* and *Magbanua* will impose a higher burden on defendants,<sup>61</sup> I also think such scrutiny is critical to protecting the rights of those same defendants against being convicted based on information conveyed by undertested, insufficiently accessible proprietary algorithms.

First, realistically, it is typically the state that is putting on evidence, period.<sup>62</sup> In my public defender days, one of the first training sessions we had was on "evidence blocking"—the idea that the less evidence comes in, the better it is for the defense. After all, the government has the burden of proof. While defendants sometimes seek to put on evidence of their own, their ability to do so is typically secondary to their ability to block the government's own evidence. That is why California's Proposition 8, which in 1982 created a state constitutional right for both parties in a criminal case to present relevant evidence, which eliminated several longstanding statutory rules of exclusion in the process, was seen as a victory for the state, and the victims' rights movement, rather than the defense.<sup>63</sup>

In turn, in the age of AI, as Professor Fan herself has ably demonstrated through her own scholarship, the government is increasingly and routinely relying on computer-generated results in its investigations and prosecutions.<sup>64</sup> While the occasional Mr. Puloka will want to wield output of an algorithm's independent analysis to try to buttress a self-defense claim, the much more common scenario is the Ms. Magbanua defending against government-proffered

---

<sup>59</sup> See Rebecca Wexler, *Metrics Not Thresholds: How to Regulate AI and Other Expert Methods for Use in Criminal Investigations and Proceedings*, 8 J.L. & INNOVATION (forthcoming Jan. 2026).

<sup>60</sup> See Andrea Roth, *What Machines Can Teach Us About "Confrontation,"* 60 DUQ. L. REV. 210, 211 (2022) (arguing right of confrontation, including confrontation of machine "witnesses," should be deemed right to meaningful impeachment).

<sup>61</sup> See Fan, *supra* note 1, at 2191-93.

<sup>62</sup> The state carries the burden of proving every element of an offense beyond a reasonable doubt. See *In re Winship*, 397 U.S. 358, 364 (1970).

<sup>63</sup> See Diana Friedland, Note, *27 Years of "Truth-in-Evidence": The Expectations and Consequences of Proposition 8's Most Controversial Provision*, 14 BERKELEY J. CRIM. L. 1, 34-35 (2009) (explaining Proposition 8's rollback of many of California's exclusionary rules expanded evidence prosecutors could use both for impeachment and in their cases-in-chief).

<sup>64</sup> See Fan, *supra* note 1, at 2153-54 (recounting conviction of Katherine Magbanua using AI-enhanced audio); Mary D. Fan, *Suspecting with Data*, 109 MINN. L. REV. 2253, 2259-60 (2025) (discussing rise of keyword warrants collected from Google searches, geofence searches of cellphones, and facial recognition technology).

AI output. Consider this hypothetical. The government wins a disproportionate amount (say, 80%) of *Daubert* or *Frye* hearings, and 90% of *Daubert/Frye* hearings involve government-proffered experts. Assuming those statistics, the government is still overall disadvantaged by increased scrutiny. For every one hundred *Daubert/Frye* hearings:

- Ninety will involve a government expert, and eighteen will be excluded.
- Ten will involve a defense expert, and eight will be excluded.
- Ten more acquittals will occur than without such scrutiny (assuming exclusion or admission of evidence is dispositive in the case).

Thus, even if the defense loses a *disproportionate* number of *Daubert/Frye* battles, it does not follow that getting rid of *Daubert/Frye* for AI-enhancement tools engaging in “independent analysis” would be a victory for defendants. On the contrary, I would expect most defenders to say that increased *Daubert/Frye* scrutiny of proprietary AI-enhancement tools offering independent analysis is critically important.

Moreover, Professor Fan’s argument may prove too much, in the sense that most major rules of exclusion could be criticized for impeding evidence offered by the defense in the face of the government’s seeming advantage in admitting evidence. After all, the rules against admission of hearsay (FRE 802), character-for-propensity (FRE 404 and 608), and proof of liability insurance to show fault (FRE 411) all potentially impede the presentation of defense evidence, and the state surely wins a disproportionate number of admissibility battles involving these rules as well. But unless such rules directly jeopardize the defendant’s right to testify<sup>65</sup> or due process right to present a defense,<sup>66</sup> courts uphold these exclusionary rules as legitimate, whether or not they lead to an “inequality of arms.” And again, given that the state presents vastly more evidence in general than the defense, it is not obvious that eliminating such rules of exclusion would be on balance better for defendants. Of course, some rules seem to unfairly target evidence offered by criminal defendants in particular.<sup>67</sup> But *Daubert* does not seem to be one of those.

---

<sup>65</sup> See, e.g., *Rock v. Arkansas*, 483 U.S. 44, 55-57 (1987) (deeming Arkansas’s blanket ban on hypnotically refreshed testimony to violate Rock’s right to testify).

<sup>66</sup> See, e.g., *Chambers v. Mississippi*, 410 U.S. 284, 295-97 (1973) (deeming Mississippi’s “voucher” rule irrational and archaic and violative of Chambers’s due process right to present defense).

<sup>67</sup> For example, the polygraph’s status as an evidentiary pariah is arguably attributable to the fact that it is typically offered by criminal defendants. See Adam B. Shniderman, *You Can’t Handle the Truth: Lies, Damn Lies, and the Exclusion of Polygraph Evidence*, 22 ALB. L.J. SCI. & TECH. 433, 437-39 (2012); see also FED. R. EVID. 804(b)(3) (2009) (repealed 2010) (noting before 2010 amendments, rule subjecting statements against interest to special corroboration requirement only where they “expose the declarant to criminal liability and [are] offered to exculpate the accused”).

Second, in any given jurisdiction, the precedent established for a tool will apply to both sides. In Florida, a defendant seeking to offer GLIMPSE results to buttress a self-defense claim presumably will have an easier time of it after *Magbanua*, unless there is evidence that the defense expert unreliably applied the tool. Conversely, the prosecution seeking to offer output from Topaz Labs' AI-enhancement tool in Washington State will have a harder time after *Puloka*. Of course, if the state offers the tool to the defense and makes key disclosures about how the tool works and the assumptions it makes, perhaps the state will have a better chance of having the evidence be admitted. But I am not sure that is the wrong result.

Third, and relatedly, *Puloka* and *Magbanua* are not the most obvious examples of an inequality of arms. That is, there are alternative reasons the defense evidence in *Puloka* was excluded while the evidence in *Magbanua* was admitted. First, Washington is a *Frye* state, in which a cutting-edge but reliable AI program might be more likely to be excluded than in a *Daubert* state like Florida.<sup>68</sup> Indeed, Fan cites a recent federal case, *United States v. Harris*,<sup>69</sup> in which a defense expert's AI-enhancement of a video was admitted over objection.<sup>70</sup> Second, as Fan herself notes, the attorneys in the Florida case failed to object to the evidence, whereas the prosecution in *Puloka* moved to exclude the Topaz Labs enhancement on *Frye* grounds.<sup>71</sup> Finally, as Fan notes, the defense in the Florida case did not even receive notice of the evidence until shortly before trial.<sup>72</sup> Of course, there is an "inequality of arms" problem here, but it looks to be more along the lines of overworked defense attorneys failing to make the objections they need to make, the impotence of ineffective assistance of counsel doctrine, and insufficient means of enforcing discovery obligations. Or perhaps the problem could be solved simply by adopting *Daubert* and allowing the defense and prosecution both to introduce cutting-edge evidence it can prove to be reliable. But the solution to the inequality of arms problem here is not necessarily to eliminate any reliability requirement for analytical AI-enhanced evidence.

Finally, getting rid of *Daubert/Frye* scrutiny for AI-enhanced evidence engaging in independent analysis will not eliminate the need for an expensive "battle of the experts." I would expect, instead, that the battle will simply move

---

<sup>68</sup> Fan does cite to another Washington case, *State v. Hayden*, 950 P.2d 1024 (Wash. Ct. App. 1998), in which the state won a *Frye* challenge to an expert's use of a digital forensic tool to enhance handprints for comparison purposes. See Fan, *supra* note 1, at 2176-77. But *Hayden* was decided in 1998, before the 2009 National Academy of Sciences report on forensic evidence, the 2016 President's Council of Advisors on Science and Technology ("PCAST") Report, and other reports on the limits of subjective pattern identification methods. Moreover, that case appeared to be well litigated by the proponent and the defendant appeared to have offered nothing in terms of rebuttal. See *Hayden*, 950 P.2d at 1027-28.

<sup>69</sup> No. 20-CR-00021, 2022 WL 1642498, at \*7 (D. Colo. May 24, 2022).

<sup>70</sup> See Fan, *supra* note 1, at 2167-68.

<sup>71</sup> See *id.* at 2152-54.

<sup>72</sup> See *id.* at 2154-55.

to the trial stage, where the prosecution will be allowed to present such evidence, and the defense would still have to counter it. Thus, the defense in such cases would still be at a disadvantage if the state can afford a fancy expert and the defense cannot. What *Daubert/Frye* offers is the chance to argue against such evidence getting to the jury at all. Furthermore, because hearsay is admissible in *Daubert* hearings, at least in federal court,<sup>73</sup> winning a *Daubert* motion with expert affidavits is cheaper than having to pay for the expert to testify at trial. Moreover, even where the state wins a *Daubert* or *Frye* hearing, the defense typically learns through that hearing at least more about the method's limits, potential error rates, and extent of validation. That is critical discovery that can then be used at trial to impeach the AI's output.

#### CONCLUSION

Ultimately, notwithstanding where Professor Fan and I disagree, our desired solutions look quite similar. Although she argues against *Daubert/Frye* scrutiny of AI-enhanced evidence, Fan appears to agree that such evidence should be subject to more than merely authentication requirements.<sup>74</sup> She suggests that judges implement standing orders to require the proponent of such evidence to disclose the anticipated admission of the evidence ninety days before trial, and to disclose "how it operates, and what steps were taken to enhance the evidence with that technology."<sup>75</sup> Earlier, she even suggests that the defense needs to know the program's "track record."<sup>76</sup>

For what it is worth, although Professor Fan touts these requirements as lenient and minimal, they are well beyond what Federal Rule of Criminal Procedure 16 requires. These standing order proposals would not explicitly require expert testimony, but it is hard to imagine a criminal defendant being able to put a lay witness on, or merely a Rule 902(13) certificate, to explain how an AI tool operates, how it was used in this case, and its track record. Indeed, even *Daubert* and *Frye* do not explicitly require evidence of a tool's "track record" for admission, however much defense attorneys and scholars have tried to insist on it.<sup>77</sup>

If judges follow Professor Fan's lead and require key performance metrics and meaningful explanation of how these enhancement tools work, that would be a major improvement over the status quo, even for the most complex analytical AI. Whatever quibbles I have with Fan's article with respect to

---

<sup>73</sup> See Ross Andrews Oliver, *Testimonial Hearsay as the Basis for Expert Opinion: The Intersection of the Confrontation Clause and Federal Rule of Evidence 703 After Crawford v. Washington*, 55 HASTINGS L.J. 1539, 1539-40 (2004).

<sup>74</sup> See Fan, *supra* note 1, Section III.B.

<sup>75</sup> *Id.* at 2202.

<sup>76</sup> See *id.* at 2199-200.

<sup>77</sup> See Wexler, *supra* note 59 (arguing for requirements of performance "metrics" before admitting AI opinions, so jury knows how much to credit opinion).

*Daubert* and *Frye* and the line between clarifying and analyzing data, we agree on the need to ensure adversarial fairness in criminal trials in an age of AI.