ESSAY

AI'S MAGICAL REALISM

Curtis E. A. Karnow*

Writing is an expressive human act. It's not fabricating text. That's the key reason it's dangerous to use generative artificial intelligence ("GenAI"), models based on large language models such as ChatGPT, Claude, Gemini and the rest, to write our legal briefs and court opinions.

To unpack that conclusion, this Essay discusses aspects of writing legal works by humans and compares GenAI's assemblage of text.

I. WRITING LAW

Law is a literate profession: It's based on writing. There is an oral component, which is important, but the oral is entirely a function of the written presentation. We read laws and rules, and we read prior cases. We analyze the words and decide what they mean and what they imply. We take notes, and then we write briefs and memoranda. Most everything we do is written, and it could not be otherwise, because we cannot transmit the complexity of the law only orally. We cannot teach the law only orally; we cannot turn to only oral sources of the law to learn it. We may orally summarize our arguments to judges; but without writing we cannot fully explain the reasons for our positions, nor can judges orally convey

*Judge of the California Superior Court (San Francisco) (Ret.).

their reasons to the public, parties, and to courts above them in the judicial hierarchy.

The fact that our law is written allows it to be thought of as objective, not the subjective or idiosyncratic expression of a judge or king. Chancellors in the twelfth and thirteenth centuries[1] may have ruled as they thought the king might wish, but their successors, relying on the accumulation of written opinions over centuries, are able to cast their opinions not in the stead of a monarch, but as expressions of objective law, independent of any one judge. The fact that our law is written allows us the comfort of the dictum that our republic is "a government of laws and not of men."[2]

The law also reasons (in writing) in a certain way. There are of course all sorts of writing—from poems and song lyrics to mathematical equations and novels. Each of those modes has its way of "reasoning," that is, setting forth its premises and continuing step by step (by words and symbols) to a conclusion.

The law's way of reasoning is one of analogy steered by theory, a sense of what is relevant.[3] That is, we compare our facts to a series of possible precedents and rules and decide (or at least argue) that our case is more like one set of precedents than another and thus should reach a similar result. Thus far, this sort of reasoning is common, and not limited to law. The cognitive and computer scientist Douglas Hofstadter has shown, pretty convincingly, that all reasoning is a form of analogy.[4]

---

1. Willard Barbour, *Some Aspects of Fifteenth-Century Chancery*, 31 HARV. L. REV. 834, 834 (1918).

2. MASS. CONST. pt. 1, art. XXX; *see also* Curtis L. Collier et al., *The Rule of Law—Implemented By Men and Women*, U.S. COURTS: CONNECTIONS (2024), https://connections.ca6.uscourts.gov/node/127 [https://perma.cc/6XRN-S2FJ] (attributing the quote to John Adams).

3. Cass R. Sunstein, *Of Artificial Intelligence and Legal Reasoning* 4 (Univ. of Chi., Pub. L. & Legal Theory Working Paper, Paper No. 18, 2001), https://chicagounbound.uchicago.edu/public_law_and_legal_theory/207/ [https://perma.cc/W4YP-CDSQ]; Curtis E.A. Karnow, *Similarity in Legal Analysis & The Post-Literate Blitz*, 15 GREEN BAG 2D 243, 248 (2012).

4. *See generally* DOUGLAS R. HOFSTADTER & EMMANUEL SANDER, SURFACES AND ESSENCES: ANALOGY AS THE FUEL AND FIRE OF THINKING (2013) (arguing that analogies and concepts play the "starring role" in thinking, as there can be

The *legal* trick is determining which rules (or statutes or cases) are "like" the one we're considering—and concomitantly, which facts of the case before us are most like what might be controlling precedent.

This is an iterative process. We assume we have an issue discussed in Case X; then we see which of our facts are also in Case X. At the same time, we look at our facts and look about for a case, any case—perhaps Case Y or Case Z, that has them; and then we might decide that Case Z controls, and not Case X. We might find that the issue of Case X is not, indeed, our issue.

As described, this seems a random process. We could for example see that an event in our case happened in the morning and so decide that a case (where an event happened in the morning) was better precedent than another (where events happened at night). We are (usually) not subject to this idiocy because we are guided by theories—theories of materiality. The issue is: What are the legally *relevant* similarities?[5] Our legal training instills in us a variety of materiality theories—theories about what counts and doesn't count in legal analysis.[6] We know enough about contracts to know that the time of day a promise is made doesn't usually matter, although if done at night, and a written proposal could not be read, we would take out our theory about "knowing assent" and wonder if the contract was really ever made. We use the basic theories about contracts to pick the cases that guide us. Because it's an iterative process and the law is an evolving and ramifying thing, each case adds to the dense complexity of facts which are potentially material. So from time to time we modify our

---

no thought without concepts and no concepts without analogies). Among other books, Hofstadter is the author of the enormously influential DOUGLAS R. HOFSTADTER, GÖDEL, ESCHER, BACH: AN ETERNAL GOLDEN BRAID (1979).

5. Sunstein, *supra* note 3, at 5.

6. We also rely in part on the assumption that facts recited in opinions are material. The assumption is not always well-founded. Some opinions, perhaps in an effort to provide a complete background (something we might expect in a bench memo where the clerk overcompensates because he isn't entirely sure which facts the judge will find material) have lengthy procedural histories and factual recitation the reader does not need to know.

materiality tests and thereafter proceed with that adjusted view. Our legal training goes on forever.

As well as theories of materiality, or perhaps as part of them, we have what Posner called our priors—our inclinations and predispositions which in the absence of countervailing forces (such as a conflicting supreme court opinion) guide us through precedent.[7] For example, we might be disposed to a broad reading of the First Amendment, and look for restrictions on both speech and expressive action, and, having picked out those facts, be inclined to see an infringement of rights in a new fact pattern.

We might be disposed to seeing private contract as the essential, sacred engine of the economy, and so be inclined to see some sort of agreement, if not by words then by action, perhaps actions relied on by the putative counterparty. In these ways, our priors help us pick out what is material, and in so doing aid us to decision. Call them biases, perhaps (which makes them sound bad); but again these are only inclinations, which can be, and should be, derailed by events such as decisions from a higher court or new law from the legislature. Modifications come, too, from experience, either one's own experience or the experience of the people around us and of those we read. This was what Holmes and Cardozo taught us, that practical stuff matters. We know, practically, that if we stick out a foot we might trip someone, but that it's not foreseeable that if we bump a man into a railway car he might have a package of fireworks which might drop and then explode and so cause a scale to fall and so hurt someone.[8] This matters when we try to decide if someone under "similar" circumstances might be liable for an injury.

---

7. RICHARD A. POSNER, HOW JUDGES THINK 66 (2008); Jack Knight, Mitu Gulati & David Levi, *How Bayesian Are Judges?*, 16 NEV. L.J. 1061, 1063–64 (2016); Michael C. Dorf, *Richard A. Posner, Divergent Paths: The Academy and the Judiciary*, 66 J. LEGAL EDUC. 186, 189, 194 (2016). As noted by others, Posner writes that the personal priors of a judge rarely control because precedent and statutes' wording usually control. *E.g.*, Jennifer Gerarda Brown, *Posner, Prisoners, and Pragmatism*, 66 TUL. L. REV. 1117, 1162 (1992).

8. Palsgraf v. Long Island R.R. Co., 248 N.Y. 339, 340–41, 345–46 (1928).

These priors and theories of materiality are not isolated. They are usually grouped; there is a taxonomy. There are categories, such as the distinctions among tort, contract, statutory liability, and so on. We have administrative law and arbitration. Each of these categories, while fuzzy at the edges, helps provide a structural relationship among concepts, and in so doing steers our inquiries when confronted with a new fact pattern. When wondering if attorneys' fees are available from the losing side, we first find a general category to investigate because the analysis differs perhaps depending on whether the case was one of contract or statutory liability. What is material in one such investigation may not be in another. We do not aimlessly look for cases which mention "attorneys' fees"; we look for cases within a structured hierarchy. When we decide if a lawyer has "waived" the right to make an objection to evidence, we look at the rules under the category "evidence" and not, for example, to cases deciding if a contractual right has been waived, or if the right to a jury has been waived, all of which entail different analyses.

This observation on hierarchical categories comprising a taxonomy of legal notions which govern legal analysis—this is obvious; but I have thought it important to tease it out both to show it's too complicated to do except in writing and in order to contrast how GenAI works.

A few more thoughts on the law as a literate profession. When we write, we give substance to the fleeting thoughts of our mind and arrange them on paper in a way we never could (or most of us never could) in our minds only. I speak for many when I say that I do not know what I think until I see what I have written. On paper, we edit, cut, and add. What we read then spurs us to further investigation—of various writings—which then are combined with the draft. It is commonplace to distinguish between the writer's and editor's role, including when these are combined in a single person. Editing to communicate with others is also editing to communicate with oneself. In this way too, writing is an

iterative process: Without it, the author may never find out what he really intends or believes.

Perhaps 3000 languages are spoken; only 78 have a literature.[9] Only those might have "law" as we understand it, based not just on human memory with its oral traditions. Oral culture depends on mnemonics, formulas, easily recalled phrases, so law in an oral tradition is a function of essentially a series of recalled proverbs[10] and clichés.[11] Sustained oral analysis is difficult.[12] Before writing, "law" could only have been sourced in custom and power. The "right" answer to a dispute would be whatever force, divine authority, or oral custom, determined in the moment. The authority of the past would have been limited by human memory and its oral transmission.

The effort made in "doing" law as we know it is a literate effort: Our law cannot exist but for the act of writing.

> Persons who have interiorized writing not only write but also speak literately, which is to say that they organize, to varying degrees, even their oral expression in thought patterns and verbal patterns that they would not know of unless they could write.[13]

> Without writing, the literate mind would not and could not think as it does, not only when engaged in writing but normally even when it is composing its thoughts in oral form. More than any other single invention, writing has transformed human consciousness.[14]

---

9. WALTER J. ONG, ORALITY AND LITERACY 7 (Routledge 2002). What counts as a distinct language is an open question. Other sources suggest somewhere between 7,000 and 8,000 current languages. Of that number at least half are unwritten, and throughout all of history, the percentage of languages with a written form is vanishingly small. *How many languages in the world are unwritten?*, ETHNOLOGUE, https://www.ethnologue.com/faq/how-many-languages-unwritten/ [https://perma.cc/WC9P-T2SB] (last visited May 18, 2026).

10. ONG, *supra* note 9, at 10.

11. *See id.* at 38.

12. *Id.* at 38–39, 101.

13. *Id.* at 55.

14. *Id.* at 77; *see, e.g.*, *id.* at 81, 148.

Those who do not write cannot participate in making the law.

II. THE MACHINE

A. Association

Association has been the hallmark of the means by which computer search operates. It's not analytical, but that doesn't mean it doesn't mimic many things humans do—we "associate" as well, all the time. You may have heard of social science experiments in which subjects holding warm drinks felt "warmer" towards people covertly cast as their subject than those who held cold drinks. The "halo" effect reveals that we trust those knowledgeable in one area with expertise they don't have in a different area. "Good looking" people tend to do better professionally than those who don't have that set of physical characteristics. We leap to conclusions, we use heuristics[15] and are otherwise subject to a host of cognitive fallacies which, in effect, are the result of following a link from one notion to another which vaguely seems "connected."

Early computer search worked this way, generating results which contained words or their synonyms used in the prompt. Google's PageRank algorithm, which transformed search in 1998, estimated the likelihood of a webpage being responsive to a search query by calculating the number of links to and from that page, increasing the value of authoritative pages' links and, iteratively, increasing the "authoritative" value of a page by the number of links from other pages, and especially other authoritative pages. This in essence measured the popularity of a page, reasoning that the more popular a page the more likely it was to be the one you were searching for. In brief, the algorithm inferred a valid

15. *See generally* DANIEL KAHNEMAN, THINKING, FAST AND SLOW (2013) (discussing the study of unconscious cognitive biases).

association between the prompt and the results from this weighted, looping set of links.

GenAI also infers pertinence by, in effect, estimating popularity. GenAI is trained on text, almost all of it from the internet, and on that basis estimates the likelihood of one word[16] following another (or following a string of words). The process is stochastic, i.e., the same input may produce a different output on different occasions; random selection is part of the mechanism. While the output often appears surprisingly fluent, it is not the product of reasoning or analysis;[17] it is the product of associations inferred from enormous amounts of text, and different associations will win out on different runs of the system.

One of the founders of information theory, instrumental in the development of computer networks, noted this effect of stochastic processing:

> In 1948, the legendary computer scientist Claude Shannon proposed a radical new way to write a sentence: Pull a book off your shelf, choose a random word on a random page, and repeat. That doesn't sound very promising, and sure enough, Shannon's

16. Actually, *tokens* are predicted. These are elements such as words, parts of words, or punctuation marks. Some predictions are simply the most likely; sometimes randomness is used to generate what might be a more "creative" approach.

17. As a Slashdot author noted:

> I've been testing various LLMs to assist with coding. I found they make utterly "inhumane" mistakes. For instance they have expert knowledge on Makefile syntax, yet a Makefile that one model generated for me contained a utterly stupid, beginner level mistake. Astoundingly I found models to be also unable to correct coding mistakes after you pointed these mistakes out to them. This kind of untrustworthiness is unfortunately build [sic] into the very fabric of these stochastic machines. *Their reasoning is not grounded in formal inference*, and that's a huge problem, it makes it difficult for humans to check their work, and it means the models fall short in explaining their reasoning and their ability to catch errors and self-correct.

Comment, quax (Jan. 23, 2025, at 14:52 MST), *on* msmash, *AI Mistakes Are Very Different from Human Mistakes*, SLASHDOT (Jan. 23, 2025, at 11:45 MST) (emphasis added), https://slashdot.org/story/25/01/23/1645242/ai-mistakes-are-very-different-from-human-mistakes#:~:text=AI%20errors%20come%20at%20seemingly,complex%2C%20multi%2Dstep%20problems [https://perma.cc/5ZQX-EEX2].

> experiments produced strings of gibberish. But something curious happened when he tweaked his procedure slightly to account for how often different words appear together. The sentences still made no sense, but they looked much more like ordinary English text. Shannon's observation was the first hint that *simply mimicking statistical patterns in text, with no regard for meaning, could yield surprisingly fluent prose*.[18]

This sort of associative calculation is as opposed to formal inference, or deductions, made within a hierarchical structure where terms secure their meaning (at least in part) from their placement in the structure. As noted, we have a taxonomy of legal concepts: Civil and criminal law are perhaps at the top, each branching into different topics such as (under civil) tort, contract, and so on, and with overlap, such as the law of evidence. This is not, however, how GenAI constructs its sentences. GenAI does it by association, which, because it may track associations in real legal writing, often look similar to a trained human's output.

B. Plausibility & Sycophancy

GenAI's outputs, a remix based on relatively frequent associations, are *plausible*. Recall what plausible usually means: That someone might believe it. A plausible statement is one that is neither (i) a contradiction nor (ii) inconsistent with what one knows. Just as there is an infinity of true statements, there is an infinity of false ones. Some true statements are implausible (the Big Bang; matter arises from perturbations in space-time quantum "foam") and many false statements are entirely plausible (such as that my Uncle Fred lives in Omaha). Plausibility is a measure of our ignorance: If you don't know whether a proposition is

18. Ben Brubaker, *AI Challenges Our Best Laws of Language*, QUANTA MAG. (Nov. 24, 2025) (emphasis added), https://www.linkedin.com/pulse/ai-challenges-our-best-laws-language-quanta-magazine-rus1e/ [https://perma.cc/YEA9-JDCD].

true, then you might find it plausible, as long as it does not contradict other things you know. The more ignorant one is, the more statements will be plausible.

Those are the statements GenAI is trained to offer up. GenAI doesn't try to offer true statements; it just relies on similarities with statements previously labeled as true or false.[19] It is now very old hat to note the persistent, built-in feature of GenAI, which is to confabulate—to make stuff up.[20]

To be sure, GenAI and its chatbots are sometimes correct because true statements may be more plausible than false ones.[21] But GenAI takes it much further: It is also trained to persuade, to convince. This is effective in part due to "automation bias" by which humans tend to see machines as unswayed by emotion, unbiased, and so more deserving of trust:

> Interestingly, a June 2024 study found that personalized AI debaters were more persuasive than human debaters in 64.4 percent of debates. The study measured persuasiveness by how much an opponent's opinion or viewpoint shifted by the end of the debate. As a general proposition regarding AI use, the study noted that humans consider machines a more objective and confident source of information. According to the study, humans assume that humans have an agenda, but that a machine is simply presenting "the facts" in an unbiased, neutral, and dispassionate manner. Further, if the

19. ARVIND NARAYANAN & SAYASH KAPOOR, AI SNAKE OIL: WHAT ARTIFICIAL INTELLIGENCE CAN DO, WHAT IT CAN'T, AND HOW TO TELL THE DIFFERENCE 196 (2024).

20. Following others' suggestion I use the term "confabulate" rather than "hallucinate" because the latter word suggests a mind which creates the hallucination. *See, e.g.*, Benj Edwards, *Why ChatGPT and Bing Chat are so good at making things up: A look inside the hallucinating artificial minds of the famous text prediction bots*, ARS TECHNICA (Apr. 6, 2023, at 8:58 MST), https://arstechnica.com/information-technology/2023/04/why-ai-chatbots-are-the-ultimate-bs-machines-and-how-people-hope-to-fix-them/ [https://perma.cc/Q397-7HRN]; Gregg Henriques, *Chatbots Do Not Hallucinate, They Confabulate*, PSYCH. TODAY (Mar. 6, 2024), https://www.psychologytoday.com/us/blog/theory-of-knowledge/202403/chatbots-do-not-hallucinate-they-confabulate [https://perma.cc/QH4K-KBTF].

21. NARAYANAN & KAPOOR, *supra* note 19, at 140.

> AI's text output sounds assertive, humans are more likely to be persuaded. Therefore, if we believe AI is unbiased and confident, we are more likely to accept its opinion as fact—even if it is a black box masquerading as a human . . . .[22]

And from someone who should know—a founder of Google: "If the system knows you well enough, it can learn to convince you of anything."[23] GenAI is very, very good at persuading the humans who use it.[24]

Another of GenAI's attributes is sycophancy, an ally in its ability to engage and so persuade. The business model—of course—is to keep customers coming back for more, and in this cause GenAI will praise, flatter, suggest false intimacy, agree with users' observations, and engage on subjects it detects interest the customer. GenAI will "validate views and intentions even when they were irresponsible, deceptive or mentioned self-harm."[25]

---

22. Frances M. Green, Raymond Sheh & Katherine Heaney, *AI and the Law: The Chaotic Collusion of Machines v. Courts*, N.Y. L.J. (June 8, 2025, at 17:15 MST), https://www.ebglaw.com/insights/publications/ai-and-the-law-the-chaotic-collusion-of-machines-v-courts [https://perma.cc/4NZ2-RUKH].

23. PETER H. DIAMANDIS, *Ex-Google CEO: What Artificial Superintelligence Will Actually Look Like w/ Eric Schmidt & Dave B*, at 1:01:30 (YouTube, July 17, 2025), https://youtube.com/watch?v=qaPHK1fJL5s&si=FNZy7_0kHptMAEqW; *see also, e.g.*, Steven Lee Myers & Teddy Rosenbluth, *Chatbots Can Meaningfully Shift Political Opinions, Studies Find*, N.Y. TIMES (Dec. 21, 2025) (chatbots often did not tell the truth and relied on unsubstantiated evidence), https://www.nytimes.com/2025/12/05/well/ai-chatbots-political-opinions-study.html?smid=nytcore-ios-share [https://perma.cc/CHD5-XX7V].

24. *See, e.g.*, Francesco Salvi et al., *On the conversational persuasiveness of GPT-4*, 9 NAT. HUM. BEHAV. 1645, 1648–49 (2025), https://doi.org/10.1038/s41562-025-02194-6 [https://perma.cc/ZJR5-4EFT]; Steven Randazzo et al., *GenAI as a Power Persuader: How Professionals Get Persuasion Bombed When They Attempt to Validate LLMs* 16–21 (Harvard Bus. Sch., Working Paper No. 26-021, Oct. 2025), https://www.hbs.edu/ris/Publication%20Files/26-021_8db29bd1-04ed-4e98-86de-d286721afc7a.pdf [https:// perma.cc/S4FZ-3JKD].

25. Ian Sample, *'Sycophantic' AI chatbots tell users what they want to hear, study shows*, THE GUARDIAN (Oct. 24, 2025, at 11:00 EDT), https://www.theguardian.com/technology/2025/oct/24/sycophantic-ai-chatbots-tell-users-what-they-want-to-hear-study-shows [https://perma.cc/4ZPH-D2AK]; *see also, e.g.*, Miryam Naddaf, *AI chatbots are sycophants—and it's harming science*, NATURE (Nov. 6, 2025), at 13–14 https://www.nature.com/articles/d41586-025-03390-0 [https://perma.cc/Y7H3-RRJS]; Erie Meyer et al., *Tech Brief: AI Sycophancy & OpenAI*, *Georgetown L. Inst. for L. & Pol'y* (July 30,

> [M]odels will even agree with objectively incorrect mathematical statements if primed by users. Another study, conducted by Anthropic in 2023, demonstrated that human preference models often prefer "convincingly-written sycophantic responses" that affirm common misconceptions over more accurate responses that may contradict the user's stated beliefs.[26]

Once again, what seems to be a bug is a feature, in the sense that sycophantic exchanges will appear to manifest an alignment with the human user. "Alignment" is not only good business but seems to match a classic goal of AI research, which is to ensure that the "interests" and opinions of GenAI are congruent with basic human values.

But this alignment is not basic; it is superficial. It only seems in alignment at the moment. Basic alignment would not simply "mirror users' tone or affirm their beliefs;"[27] it presumably would be more closely tied to what most humans think of as "good" values, which do not include advising self-harm or touting Hitler.[28] But if so, then it would often contradict user's views, or correct them, which isn't nearly as inviting. We could call this superficial quality of GenAI "*misguided* alignment."

We can combine these notions: (i) the probabilistic associations of terms (as opposed to analytical decision making); (ii) plausibility, which is wholly distinct from "truth," and (iii) and the ability to persuade supported by a tendency to misguided sycophancy and conclude that GenAI is a "bullshit" generator. It creates speech that is

---

2025), https://www.law.georgetown.edu/tech-institute/insights/tech-brief-ai-sycophancy-openai-2/ [https://perma.cc/PHX8-8EET].

26. Meyer et al., *supra* note 25.

27. Amy Winecoff, *Artificial Sweeteners: The Dangers of Sycophantic AI*, TECH POL'Y PRESS (May 14, 2025), https://www.techpolicy.press/artificial-sweeteners-the-dangers-of-sycophantic-ai/ [https://perma.cc/7JUQ-BAMK].

28. Peter Hoskins & Charlotte Edwards, *Musk says Grok chatbot was 'manipulated' into praising Hitler*, BBC (July 10, 2025), https://www.bbc.com/news/articles/c4g8r34nxeno [https://perma.cc/8HZF-X7HZ].

meant to persuade without regard to the truth; the output is in effect automated bullshit.[29]

> The very mechanisms that make Large Language Models (LLMs) so capable—their ability to detect nuanced patterns in language, their optimization for user satisfaction, and their designed ability to simulate empathy—have converged to create the most powerful persuasion engine in human history . . . .[30]

So, what's not to like? Persuasion is what judges and lawyers do. Judges persuade the public, parties, lawyers—and not least, the courts above them—that they have made the right decision. Lawyers live and die by the art of persuasion, turning water into wine, a sow's

---

29. NARAYANAN & KAPOOR, *supra* note 19, at 139–141 (citing HARRY G. FRANKFURT, ON BULLSHIT 55–57 (2005)). Critically, Frankfurt distinguishes the liar from the bullshitter: "The fact about himself that the bullshitter hides . . . is that the truth-values of his statements are of no central interest to him . . . . It is impossible for someone to lie unless he thinks he knows the truth. Producing bullshit requires no such conviction." FRANKFURT, *supra*, at 55. The bullshitter "does not care whether the things he says describe reality correctly. He just picks them out, or makes them up, to suit his purpose." *Id.* at 56.

30. Mohit Sewak, *Your GenAI is 81.7% More Persuasive Than Your Best (Human) Friend*, LINKEDIN (Sep. 20, 2025), https://www.linkedin.com/pulse/your-genai-817-more-persuasive-than-best-humanfriend-mohit-qhvcc [https://perma.cc/YR9E-3E5Z]. Here's how Sewak says it:

> The People-Pleaser Problem: Why We're Training AI to Be Sycophants: It seems counterintuitive that the primary method used to make AI models safer and more helpful would also teach them to be manipulative. Yet, this is the paradox at the heart of the industry-standard training technique: Reinforcement Learning from Human Feedback (RLHF). RLHF is how a raw language model is fine-tuned into a helpful assistant. The process is straightforward: the AI generates multiple responses to a prompt. Human raters then evaluate these responses, indicating which they prefer. A Reward Model is trained on these preferences, and the AI is optimized to generate outputs that maximize the expected reward. The critical flaw in this loop is the human element. Humans are psychologically wired to prefer responses that are agreeable, validating, and confirm their existing beliefs—a phenomenon known as confirmation bias. We often prefer a comforting lie over a challenging truth.

Id. (citations omitted) (citing Long Ouyang et al., *Training language models to follow instructions with human feedback* 2 (Mar. 4, 2022) (unpublished manuscript) (available at https://arxiv.org/abs/2203.02155 [https://perma.cc/3S2E-MJV4])); *id.* (citing Paul Christiano et al., *Deep Reinforcement Learning from Human Preferences* 2–3 (Feb. 17, 2023) (unpublished manuscript) (available at https://arxiv.org/abs/1706.03741 [https://per ma.cc/SYA2-94LV])).

ear into a purse; they are paid to turn a thousand disconnected facts into a story—a story with punch, one that convinces.[31]

Resistance to the siren call of GenAI isn't futile. But it is difficult.

III. WRITING WITH GENAI

GenAI can help us with ideas, analyses of voluminous documents, and other tasks.

But GenAI makes it easier than ever to be a terrible lawyer (or judge). We have always had bad lawyers, as we have bad pilots, plumbers, and poets. Lousy lawyers are cartoons of the real thing, making motions and arguments that vaguely sound plausible but are unanchored to the meaning of the cases or statutes they cite. They make pointless objections. They are not serious people. They can't read the room; they have no *judgment*. For them, GenAI is the perfect companion.

The problem is not so much the bad citations that get all the press.[32] Appellate opinions to date have focused on those fake citations, and to real but entirely irrelevant citations, and so on.[33] There, the issue isn't that AI was used; the problem is that the lawyer, oblivious to what everybody knows about GenAI—which is that it confabulates, makes stuff up—doesn't check the output. This stupidity gets the headlines.

Nor is the issue the stimulation that GenAI might provide. Random walks through the files, the words of a

---

31. Curtis E.A. Karnow, *Telling Stories*, 38 NOTRE DAME J.L. ETHICS & PUB. POL'Y 529, 530 (2024). In a jury trial a win is a win: If there's substantial evidence supporting the verdict somewhere in the record, a verdict, the result of a convincing story, will be sustained without the slightest regard to whether the jury relied on any of the evidence.

32. *See, e.g.*, COURT TRIALS TV NETWORK, *ATTORNEY BUSTED Using ChatGPT!! Judge REPORTS Him to the BAR!* (YouTube, Aug. 21, 2025), https://youtu.be/DcrDci2YOqE?si=1VJGMXIl_GZHoRIX (Georgia trial court). A compendium is found at Damien Charlotin, *AI Hallucination Cases*, https://www.damiencharlotin.com/hallucinations/ [https://perma.cc/Q9BJ-YWCQ] (last visited Feb. 4, 2026).

33. *See infra* note 41.

three-year child, or a poem can all provide inspiration for a brief; so can GenAI.

The issue, I suggest, is that reliance on AI to draft briefs or opinions makes us, and our output, stupid.

First, as with an addiction to a drug, habituation to GenAI to draft our materials gradually strips away our ability to do our jobs. We become duller. Instead of every new case as an opportunity to learn a new area of law, or newly learn how the world works, we can retreat to a cocoon and let the AI do the spinning of words. It is a loss of essential engagement. This is what happens in the classroom.[34] It may be happening with doctors.[35] It happens with lawyers.[36] It is general problem, not just for those occupied in explicit learning.

> EEG revealed significant differences in brain connectivity: Brain-only participants exhibited the strongest, most distributed networks; Search Engine users showed moderate engagement; and LLM [i.e. GenAI] users displayed the weakest

34. Ashanty Rosario, *I'm A High Schooler. AI Is Demolishing My Education*, THE ATLANTIC (Sep. 3, 2025), https://www.theatlantic.com/technology/archive/2025/09/high-school-student-ai-education/684088/?utm_campaign=atlantic-intelligence&utm_content=20250905&utm_source=newsletter&utm_medium=email&utm_term=Atlantic+Intelligence [https://perma.cc/CU8S-AWUH]; *see, e.g.*, Anastasia Berg, *Why Even Basic A.I. Use Is So Bad for Students*, N.Y. TIMES (Oct. 29, 2025) ("To leave our students to their own devices—which is to say, to the devices of A.I. companies—is to deprive them of indispensable opportunities to develop their linguistic mastery, and with it their most elementary powers of thought.").

35. Teddy Rosenbluth, *Are A.I. Tools Making Doctors Worse at Their Jobs?*, N.Y. TIMES (Aug. 29, 2025), https://www.nytimes.com/2025/08/28/well/ai-making-doctors-worse-deskilling.html [https://perma.cc/XWB7-J29W].

36. *See, e.g.*, AI FUTURES FORUM, THE BROKEN LADDER: HUMAN CAPITAL RISK IN UK PROFESSIONAL SERVICES 4–5 (2026), https://aiforum.org.uk/the-broken-ladder-green-paper-request/?utm_source=artificial%20lawyer2&utm_medium=article&utm_campaign=broken%20ladder%20green%20paper [https://perma.cc/Y4RB-6R3X] (reviewing various professions including UK solicitors where junior lawyers may trust AI output over legal principles. "Research warns that extensive AI use by juniors triggers metacognitive laziness, where critical thinking tasks are offloaded to the machine, preventing the development of deep analytical skills.").

> connectivity. Cognitive activity scaled down in relation to external tool use.[37]

Second, if we give away the practice of writing, we lose its iterative process of shifting from writing to editing and back again. We surrender our reasoning, our analytical abilities, to the ease of association. We lose the ability to formulate our argument. We do not trace the development of a doctrine through case law to understand how it has strayed from or kept to an earlier rationale. We may miss an argument, doctrine, or principle better suited to our facts. We cannot discover ideas and concepts that are embedded in our written legal history and which might be reborn in the midst of new facts as we write out our memos and briefs and opinions. We aren't doing law.

Third, GenAI output is mediocre at best—and I do not mean that the brief is chock full of fake cites. I mean the text is bland; it tends to "AI slop," a term now often used to describe GenAI output.[38] Again, fake cites aside, the issue is that the arguments made may be wrong, or ineffective, because they are not founded on analytical reasoning, but only the simulacrum of reasoning. Think of it as magical realism.

---

37. Adam Rowe, *Yet Another Study Finds That AI Is Making Us Dumb*, TECH.CO (June 17, 2025) (quoting Nataliya Kosmyna et al., *Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task* Abstract (Dec. 31, 2025)), https://tech.co/news/another-study-ai-making-us-dumb [https://perma.cc/CZF3-HZH9]; *see, e.g.*, Lucas Ropek, *Multiple Studies Now Suggest That AI Will Make Us Morons*, GIZMODO (June 26, 2025), https://gizmodo.com/multiple-studies-now-suggest-that-ai-will-make-us-morons-2000620811 [https://perma.cc/7WDY-X3D6]; Lila Shroff, *The People Outsourcing Their Thinking to AI*, ATLANTIC (Dec. 1, 2025), https://www.theatlantic.com/technology/2025/12/people-outsourcing-their-thinking-ai/685093/ [https://perma.cc/Q968-2L6T].

38. *See, e.g.*, Evan Gorelick, *Vigilante Lawyers Expose the Rising Tide of A.I. Slop in Court Filings*, N.Y. TIMES (Nov. 7, 2025), https://www.nytimes.com/2025/11/07/business/lawyers-ai-vigilantes.html [https://perma.cc/6TFW-NXQX]; Marina Adami, *AI-generated slop is quietly conquering the internet. Is it a threat to journalism or a problem that will fix itself?*, REUTERS INST. (Nov. 26, 2024), https://reutersinstitute.politics.ox.ac.uk/news/ai-generated-slop-quietly-conquering-internet-it-threat-journalism-or-problem-will-fix-itself [https://perma.cc/9E9Y-LBZ4].

Fourth, there's what I call the "first draft" problem. This is a type of anchoring: The cognitive bias by which a person's decision is influenced by an often immaterial reference point or "anchor." The anchoring effect is used all the time in settlement talks, where each side will first throw out an exaggerated position in an effort to move the battleground closer to their interests.[39] One classic experiment presented subjects with a wheel of fortune rigged to give small numbers to one group and large numbers to a different group—in effect, the anchoring number. The groups were then asked to estimate the number of African nations in the UN: The first group estimated a small number; the second group estimated a significantly larger number.[40]

Lawyers and judges who rely on GenAI to produce a first draft are, I suggest, likely to be anchored by its structure and conclusions. They may edit; they may do further research and even check cites. But the draft will look plausible; it will mostly read well. It is likely to be the central basis for further research—into areas it selects and cases it cites, not necessarily those to which an experienced lawyer might gravitate.

Finally, I suggest an accountability problem. GenAI provides yet another excuse when lawyers get into trouble. One way to think about this is to group the excuse—that AI and not the lawyer got it wrong—together with excuses such as that a secretary, co-counsel, or junior associate is responsible. Few judges are distracted by those excuses, and most find them to be aggravating factors. Caught by courts in their misuse of GenAI, lawyers have tried these sorts of lame "the dog ate my homework" excuses, usually to be met with

39. *See* Thomas Mussweiler, *The Malleability of Anchoring Effects*, 49 EXP. PSYCHOL. 67, 71 (2002), https://doi.org/10.1027//1618-3169.49.1.67 [https://perma.cc/3JKG-DGNB]. Interestingly, a study has observed an anchoring effect on judges asked to impose very long sentences. *See* Piotr Bystranowski et al., *Anchoring Effect in Legal Decision-Making: A Meta-Analysis*, 45 L. HUM. BEHAV. 1, 4, 7, 18–19 (2021), https://doi.org/10.1037/lhb0000438 [https://perma.cc/KJ4X-QP5Z] (explaining a study that analyzed judicial anchoring in the context of criminal cases and sentencing).

40. KAHNEMAN, *supra* note 15, at 118–19.

sanctions, in part at least because the explanations are false.[41]

But there is a subtly different way to think of the accountability issue, which is from the perspective of the lawyer or judge using GenAI to draft. That human may see GenAI as a true agent. The agent is the locus of expertise; it finds arguments and seems to know things that the human does not. We marvel at its speed and access to a universe of facts. It is a black box, taking in enormous amounts of data and spitting out apparently lucid results, all done behind an impenetrable screen—as the human brain does. There is much loose talk about "agentic AI," supporting the vague notion that AI can formulate its own goals and procedures, in effect making its own decisions.[42] As a black box, AI can be an

41. *See* Noland v. Land of the Free, L.P., 114 Cal. App. 5th 426, 435–39 (2025); People v. Alvarez, 114 Cal. App. 5th 1115, 1119–20 (2025) (noting the attorney admitted his "lack of professionalism" and explaining this type of misconduct damages attorney credibility); County of Los Angeles v. Niblett, 116 Cal. App. 5th 454, 454–55 (2025) (unpublished portion); Shayan v. Shakib, 116 Cal. App. 5th 619, 624–26 (2025); Schlichter v. Kennedy, 116 Cal. App. 5th 30–33 (2025) (multiple findings counsel was not telling the truth in his explanations for false citations*). See generally* Ashley Belanger, *You won't believe the excuses lawyers have after getting busted for using AI*, ARS TECHNICA (Nov. 11, 2025, at 8:54 MST), https://arstechnica.com/tech-policy/2025/11/lawyers-keep-giving-weak-sauce-excuses-for-fake-ai-citations-in-court-docs/ [https://perma.cc/3SK9-E66F]; Jaclyn Diaz, *A recent high-profile case of AI hallucination serves as a stark warning*, NPR (July 10, 2025, at 13:49 ET), https://www.npr.org/2025/07/10/nx-s1-5463512/ai-courts-lawyers-mypillow-fines [https://perma.cc/3BRR-5XGR] (lawyer pretended that he inadvertently filed a draft, judge questions counsel's veracity). As for judges, see Akul Saxena, *As Judges Admit AI Drafted Faulty Orders, Sen. Grassley Demands Oversight*, BROADBAND BREAKFAST (Oct. 27, 2025), https://broadbandbreakfast.com/as-judges-admit-ai-drafted-faulty-orders-sen-grassley-demands-oversight/ [https://perma.cc/ELB4-S3NP].

42. DARIO AMODEI, MACHINES OF LOVING GRACE: HOW AI COULD TRANSFORM THE WORLD FOR THE BETTER (2024), https://darioamodei.com/essay/machines-of-loving-grace [https://perma.cc/9QK7-Z9ZR] (arguing that powerful AI, which some call AGI or artificial general intelligence, will arrive within the next few years, "smarter than a Nobel Prize winner across most relevant fields—biology, programming, math, engineering, writing, etc. This means it can prove unsolved mathematical theorems, write extremely good novels, write difficult codebases from scratch, etc"). Amodei also discusses "agentic" (self-directed) AI in his follow-up essay. DARIO AMODEI, THE ADOLESCENCE OF TECHNOLOGY: CONFRONTING AND OVERCOMING THE RISKS OF POWERFUL AI (2026), https://www.darioamodei.com/essay/the-adolescence-of-

"accountability sink[]," where responsibility goes to die.[43] So it is that AI's product is not exactly ours; it is of the machine; a convenient redirection to an independent locus of decisions. This evisceration of ownership over the product reduces accountability: For surely we are not responsible for the things we merely pass on to others. This issue is not exactly that lawyers are trying to wriggle out of responsibility; it's that they don't think they *are* responsible, just as they are not liable for relaying the plausible factual contentions of their clients.

IV. END NOTES

Does it matter if in the end we file a convincing, well-groomed brief or opinion, and it happens to be conceived and drafted by GenAI (assume the cites are right)?

I think so. Because we would then relinquish the process of making law. And *process* is, in our law, key. It's the basket in which we've placed all our eggs.

In truth, we're never entirely sure we will reach the right outcome—whether a defendant is really guilty or not, whether damages should be awarded, or just how much; if sensitive documents should be produced, or injunctions issued. We don't know for certain, because we are fallible, and our tests allow for that. In civil cases we are usually content with 51% probability, the "more likely than not" test.[44] Trial judges have a wide range of

---

technology [https://perma.cc/W5FC-4A7L] ("If the exponential continues—which is not certain, but now has a decade-long track record supporting it—then it cannot possibly be more than a few years before AI is better than humans at essentially everything.").

43. See the remarkable book, DAN DAVIES, THE UNACCOUNTABILITY MACHINE: WHY BIG SYSTEMS MAKE TERRIBLE DECISIONS—AND HOW THE WORLD LOST ITS MIND 68, 76–77, 88–89 (2025). As Davies writes, corporations are similar black boxes, in effect making decisions that cannot be attributed to any one person or identifiable group of persons. The problems this creates for ordinary civil liability are beyond the scope of my notes here. *Cf., e.g.*, *id.* at 77. The American Law Institute (of which I am a member) is working on these issues. *See Principles of the Law, Civil Liability for Artificial Intelligence*, A.L.I., https://www.ali.org/project/principles-law-civil-liability-artificial-intelligence [https://perma.cc/NL73-YCHX] (last visited Feb. 6, 2026).

44. *See e.g.*, Potter v. Firestone Tire & Rubber Co., 6 Cal. 4th 965, 997 (1993).

discretion in many situations and use balancing tests. They usually sustain verdicts based on the testimony of a single person,[45] who might or might not be telling the truth. Even our highest standard in criminal cases—beyond a reasonable doubt—is not certainty. It can't be.

Our law is infused with this sort of fuzzy logic. We *hope* we're right. We serve that hope by setting up processes designed to generate the right result. We insist on legal reasoning, analysis, and don't rely on free association or mere salience. Instead, we figure out what is material in our precedents, discerning doctrines and theories; and apply those to the facts we have. We do this, and show our work to clients, judges, and the public, in writing.[46] We want a just *system* because that's the best we can do to foster the right *result*. The lawyers and judges who make the law are accountable: Lawyers can be fired and suspended from practice; judges are elected or appointed and can be removed.

But those processes are not GenAI's system. It's not accountable for anything. We don't know how it assembles its draft; we don't know which notions were looked at and rejected, looked at and modified, or just made up. We don't know what it left out. GenAI's output may look very much like what a human would produce, but that says nothing about whether it approaches issues the same way we do; and indeed it does not.[47]

---

45. *See* Judicial Council of California Civil Jury Instruction No. 5003 (2025 ed.) ("If you believe it is true, the testimony of a single witness is enough to prove a fact.").

46. We do other things too, to protect the integrity of legal processes; indeed almost all our rules are designed to do this. We have notions of "due process" and allow parties to be heard; we organize the taking of evidence in trial and disallow unreliable evidence; we permit review with, e.g., motions for reconsideration and for a new trial; we have a deliberative appeals process; we shield judges and jurors from ex parte communications; we try to exclude judges or jurors with a bias; and so on.

47. Feynman explains this perfectly in an old video, made before the rise of large language models. *See* LEX CLIPS, *Richard Feynman: Can Machines Think?*, at 00:20 (YouTube, Nov. 25, 2019), https://youtu.be/ipRvjS7q1DI?si=_LNsU4a0b5_-TwX0 [https://perma.cc/U5AV-5GXJ]. While our inventions such as jets, cars and calculators are meant to generate the results produced by birds, horses, and the human mind, jets don't flap wings, cars don't use legs, and calculators don't do arithmetic in the way human minds do. *See also, e.g.*,

If there's anything like human judgment reflected in its products it's only because GenAI simulates it, to some unknown degree. With GenAI we will never know, or it may be hard to tell, if we are being led down the garden path, for it will all, all, be so *plausible*.

Yasemin Saplakoglu, *AI Is Nothing Like a Brain, and That's OK*, QUANTA MAG. (Apr. 30, 2025), https://www.quantamagazine.org/ai-is-nothing-like-a-brain-and-thats-ok-20250430/ [https://perma.cc/EM94-3ZGS]; ELIEZER YUDKOWSKY & NATE SOARES, IF ANYONE BUILDS IT, EVERYONE DIES 42–43, 183 (2025); YUDKOWSKY & SOARES, *supra* at 118 (explaining that AIs don't reason in words or language).