

The Monoculture (Homogenization) Trap in the Practice of Law

Why “AI First” Misses the Point in High-Stakes Deal Work

FlatironLaw.AI

Flatiron Law Group LLP

Executive Summary

New-model law firms are branding themselves “AI First.” The thesis: AI can learn to automate human thinking on legal issues and practices by mining legal data (i.e. negotiated agreements, opinions, laws and regulations) to predict outcomes, gradually phasing the lawyer out of the picture. At Flatiron, we are not on board with that thesis. To be clear, we believe that AI is a force multiplier – enabling talent to do more, better, faster and with less labor and less overhead.

But in high value, high stakes deal work (M&A, private equity, project finance, commercial real estate, venture capital) you need senior talent to drive the AI to deliver the requisite quality.

The challenges of relying on AI alone are well documented: hallucinations, bias, ethics. But one major issue gets overlooked — what some have called **the “monoculture trap”** or more simply – the **“homogenization of the practice”**. Because models draw from the same (or substantially the same) source data — agents left to analyze clauses or negotiate terms without human wisdom, judgment and direction will always converge on whatever consensus outcome the data produces. Everybody gets the same homogenized outcome – or what can be called “breeding mediocrity.”

Commonality of outcome may be just fine for commodity work at the low end of the legal spectrum – repeatable commercial contracts like MSAs, SaaS Terms, and NDAs. You do not need to break the bank to marginally improve your standard NDA. But it does not work for high value, high risk transactions, regulatory analysis or litigation, where talent, judgement and expertise are required.

Players in high-stakes dealmaking are not going to settle for the lowest common denominator, homogenized outcome that everyone else gets. Clients want creative, superior outcomes that maximize the value of their negotiating position — whether strong or weak. And they’re willing to pay for results. Wachtell Lipton is paid for value, not by the hour, and clients gladly pay it.

Those kinds of outcomes require senior talent. AI can help. AI can accelerate decision-making by surfacing optionality and analysis. But at the end of the day, talent drives the creative process.

This article evaluates the risks of the homogenization - monoculture trap, concentrated in seven overlapping areas: (i) substantive convergence of legal analysis (“race to mediocrity”); (ii) the subtle danger that the model is right in the abstract but wrong for the deal; (iii) standardization of drafting and the slow death of bespoke risk allocation; (iv) the risk of abdicating quality control to the models (the “fluency trap”); (v) the risk of cognitive de-skilling, particularly of junior lawyers; (vi) systemic fragility, in which a single model error propagates simultaneously across firms; and (vii) governance and ethical exposure under existing professional-conduct rules.

The recommended posture is not retreat from LLM use but deliberate management by subject matter experts, or what we call “Talent First”. Talent First keeps experienced human judgment — including the contrarian instinct to challenge a clean-sounding answer — at the center of the workflow, and it holds senior talent responsible for quality control rather than delegating that role to the model. It means measuring the model's output against the realities of the deal — what the leverage actually supports, what advances the client's position, and what merely reads as “rational” — and refusing to let cost pressure flatten the firm's analysis to the model's median.

What Monoculture and Homogenization Means for Lawyers

The term “algorithmic monoculture” comes from computer science — “homogenization” is just the plainer word for the same phenomenon —, but the agricultural analogy is the intuitive one: a field planted with a single cultivar can be highly productive but is catastrophically vulnerable to a single pathogen. Research has shown mathematically that even without an external shock, the shared use of one algorithm can degrade aggregate decision quality across a system — even when that algorithm is more accurate than any alternative in isolation (Kleinberg & Raghavan, <https://www.pnas.org/doi/10.1073/pnas.2018340118>).

For lawyers, three findings matter most.

First, LLMs systematically narrow their output relative to the diversity of their training data. researchers call this *generative monoculture* and trace it to the alignment process itself <https://arxiv.org/abs/2407.02209> — meaning you cannot prompt your way out of it. Adjusting temperature, rewording the question, asking for “creative” answers: none of it materially widens the distribution of model output.

Second, even firms using ostensibly different products do not escape monoculture. Research shows <https://arxiv.org/abs/2211.13972> that decision-makers who share components — above all, training data — reliably produce more homogenized outcomes; with foundation models the degree depends on how the model is adapted, but the shared-data effect is robust. Buying a different LLM or wrapper does not necessarily buy you a different brain.

Third, the legal-AI vendor market is already concentrated. Westlaw AI-Assisted Research, Lexis+ AI, Harvey, Legora, Perplexity, Co-Counsel, and the rest increasingly route through a small handful of frontier models. The convergence is not a future risk. It is the present default.

Why Law Is Especially Exposed

Several features of our work make legal practice more vulnerable to monoculture than other professional services.

Our output is text. Almost everything a lawyer produces is language — memoranda, briefs, contracts, opinions, regulatory submissions. LLMs are precisely positioned to do this work, so the share of legal output passing through them will outstrip what is seen in other professions.

The training data converges on the same canon. Frontier models are trained substantially on the same publicly available case law, statutes, treatises, and well-known briefs. Stanford’s RegLab documented <https://arxiv.org/abs/2401.01301> that LLMs return canonical, well-known cases at much higher rates than less-cited authorities. Every model points to the same cases.

Alignment pushes toward the median. Reinforcement Learning from Human Feedback (“RLHF”) and similar techniques are designed to make models produce confident, conventional, well-hedged output. That is precisely the opposite of what wins hard cases or breakthrough deal terms. A novel reading of a statute, an aggressive but defensible position, a creative distinction — by training-data construction, these are the outputs the model is least likely to generate.

Adversarial use produces convergent results. When opposing counsel both use the same model to brainstorm positions, summarize precedent, or red-line a contract, the output is structurally biased toward the same canonical analysis. The adversarial process — and the negotiation process — depends on cognitive diversity. A model that produces median analysis for both sides flattens the dynamic entirely.

Seven Risks in High-Stakes Practice for Law Firms

1. Race to mediocrity.

In an adversarial system, the value of legal advice is partly a function of how it differs from what opposing counsel produces. When both sides converge on the same canonical analysis, what looks like consensus is the loss of one side's edge. For deal work, the practical translation is stark: the difference between getting market terms and getting the deal your leverage entitles you to. Worse, the lawyers with the better-but-less-common argument may now find that their reasoning sounds “unusual” or “non-market” against the LLM-generated baseline — and be tempted to abandon it.

2. The "rational" answer that breaks the deal.

The subtler danger is not that the model is wrong, but that it is right in the abstract and wrong for the deal. Models optimize for the "rational" position — the one the data says is the best outcome for the client — and will advance it regardless of where the leverage actually sits. But one of the most important things a deal lawyer does is to calibrate reality for the client: what is realistically achievable, what is not, and where to spend finite negotiation capital given the parties' leverage. Models do not do this. Absent careful prompting by a subject-matter expert, they do not weigh leverage, goodwill, or timing risk — and left unfiltered, their "rational" output can do more damage than good, stressing deals and burning goodwill that a seasoned lawyer would have protected.

We have seen this failure mode more than once. Clients have used AI to generate comments late in a negotiation — comments that, in an ideal world, would have advocated for better terms, but in the real world would have stressed, and in one case nearly broken, the deal. This is where Talent First earned its keep: Triaging model output and reconciling it with reality. The intervention of subject-matter experts kept a rational-on-paper position from producing an irrational outcome.

3. The bespoke becomes generic.

Sophisticated transactional practice depends on negotiated, idiosyncratic provisions — unusual carve-outs, novel indemnities, calibrated reps and warranties, creative earn-out structures. Those are exactly the provisions least likely to be generated by an LLM trained on the model contract. And over time, LLM-drafted contracts become the next generation's training data, and the model contract narrows further. There is no equilibrium other than convergence. The pessimistic case is not that bespoke drafting disappears — it is that “bespoke” itself gets outsourced to the model, which produces a generic version of bespoke.

4. The fluency trap: Abdicating quality control.

The fluency of the model can be a trap: output that reads well invites the assumption that it is well-reasoned, tempting firms to relax the layer of senior review that firms historically deployed to ensure quality control. That is a dangerous concession: the output that reads so well may be flawed for all of the reason discussed, and its polish is exactly what lets those flaws slip past review.

5. Junior lawyer de-skilling.

Junior lawyers historically built judgment by doing the cognitive work that LLMs now perform: reading every cited case, identifying the doctrinal architecture, drafting the first analytical pass. What AI removes from junior lawyers is not the typing but the thinking — the formative practice that develops legal judgment. Today’s senior lawyers can use LLMs effectively because they were trained without them and can recognize when an output is wrong or shallow. The lawyers being trained today may not develop that recognition capacity. The senior partners of 2035 are the juniors who are skipping that work now. For another very interesting caution on the use of AI, see “[Why Our Firm Still Prohibits Generative AI for Legal Research and Written Advocacy](#)” | [Carlton Fields](#)

The remedy for this training gap is to create new virtual environments that replicate the real live deals that juniors historically trained on. Simulation is well suited to this. Simulators can give juniors repeated, consequential reps at the judgment call itself — interacting with virtual clients, opposing counsel, and other stakeholders, before the stakes are real. It restores the one thing an AI-assisted workflows tend to strip out: practice at exercising judgment under realistic pressure.

Simulation also addresses the monoculture problem directly. Scenarios can be built around off-market facts, incomplete information, and deliberately aggressive or unconventional counterparties, forcing the lawyer to hold and defend the better-but-less-common position rather than accept the consensus answer an LLM would hand to both sides. The goal is not to rehearse the median outcome but to build, early, the contrarian instinct this article argues the profession cannot afford to lose. We have built, together with other collaborators, our own agentic negotiation-training platform, called Deal Mentor, focused first on M&A. <https://dealmentor.ai/>. Deal Mentor supports training at Flatiron and is available for licensing by third parties.

6. Cascade failures.

The Stanford RegLab/HAI study “Hallucination-Free?” <https://onlinelibrary.wiley.com/doi/full/10.1111/jels.12413> tested the leading commercial legal-research products and found hallucination rates that, while lower than general-purpose models,

remained material — roughly one-third of queries for Westlaw’s AI-Assisted Research, and roughly one in six for Lexis+ AI. The sanctions cases are the consequence. *Mata v. Avianca* (S.D.N.Y. 2023) is the most-cited example. The consequences have only escalated since. In *Couvrette v. Wisnovsky* (D. Or. 2026), a federal magistrate judge imposed roughly \$110,000 — believed to be the largest AI-hallucination sanction in U.S. history — and dismissed the case with prejudice after counsel filed briefs riddled with fabricated citations and quotations. And in *Whiting v. City of Athens* (6th Cir. 2026), a federal appeals court fined two lawyers \$15,000 each and anchored its holding to a deliberately tool-agnostic principle: no filing may contain a citation, however generated, that a lawyer has not personally read and verified. The trend is not anecdotal. The tracker maintained by researcher Damien Charlotin now catalogues more than 1,300 filings worldwide tainted by AI-fabricated authority, with sanctions climbing from \$5,000 in *Mata* to six figures in under three years (<https://www.damiencharlotin.com/hallucinations/>). When both sides rely on the same model, no one is positioned to catch the other’s error: error rates do not aggregate linearly; they cluster.

7. Ethics and governance exposure.

[ABA Formal Opinion 512](#) (July 2024) applies six existing duties to generative AI use: competence, confidentiality, communication, candor to the tribunal, supervisory responsibility, and reasonable fees. The competence duty is the one monoculture most directly implicates. If “competent representation” comes to require LLM use — because everyone uses them and clients expect their efficiencies — lawyers and firms become structurally exposed to the failure modes of those LLMs. Opinion 512 also warns that several lawyers using the same self-learning tool may inadvertently disclose client information in others’ outputs: shared tools mean shared failure modes for confidentiality. And the supervisory duties under Rules 5.1 and 5.3 apply to AI “assistance” in ways most firms are not yet operationally meeting.

Since Opinion 512, the ground has shifted from guidance toward enforcement. More than two dozen state bars have now issued their own AI guidance — Texas (Opinion 705), Florida (Opinion 24-1), New York, North Carolina, and Pennsylvania among them — most tracking Opinion 512 but several going further on disclosure and supervision. Most consequentially for the kind of work described here, in August 2025 the California Supreme Court [directed the State Bar](#) to move AI duties out of advisory “practical guidance” and into the enforceable Rules of Professional Conduct — and, separately, to study the distinct obligations raised by agentic AI, systems that act without step-by-step human prompting. Courts, meanwhile, increasingly require affirmative disclosure of AI use through standing orders, and discipline is no longer confined to sanctions motions: bar authorities in several states have begun reprimanding and suspending lawyers directly for

unverified AI filings. The direction of travel is unmistakable — the duty to verify is becoming both non-delegable and independently enforceable, and “the model produced it” is not a defense.

The Non-Homogenized Approach

Four disciplines to follow.

Build devil’s-advocate discipline into the workflow.

Lawyers should not ask LLMs for “the best argument.” They should ask for “the strongest argument against the obvious answer.” Disciplined adversarial prompting alone will not solve generative monoculture — nothing at the prompt layer will — but it surfaces more diverse output than default prompting, and it trains lawyers to keep looking for the contrarian read.

Capture your distinctive analytical patterns.

Your firms’ deal teams undoubtedly have characteristic approaches — particular ways of structuring indemnities, distinctive theories of damages, non-obvious doctrinal frames, signature negotiating moves. Those are precisely what monoculture erodes. Capture and operationalize them, so training lawyers and your systems on your edge — not the median of the model.

Tier work by monoculture sensitivity.

Not every matter is created equal. High-volume commodity work (form NDAs, routine compliance, standard sub-doc review) tolerates convergence and benefits from speed. But high value, high stakes deal work (like M&A) requires LLM use with mandatory human review by subject matter experts. At the top of the legal work product food chain, LLM output should be a starting hypothesis to be triaged against the realities of the deal, argued against, and not relied on.

Preserve the cognitive load that builds junior judgment.

Junior lawyers are the future of your firm. A firm that stops developing them is quietly deciding not to have a next generation, and will slowly evaporate. While juniors may no longer be the profit centers that they once were, in the age of AI they remain critical assets to train and prepare. Their development is not overhead; it is an investment in the firm's future. That is why firms should resist the temptation to replace junior research and drafting entirely with LLM output. Some training tasks remain genuinely cognitive: first-pass analysis without AI, followed by AI comparison, followed by senior review. The efficiency cost *is* the training investment.

A Winnable Race

The race to mediocrity is not a metaphor. It is a structural prediction. When a small number of frontier models, trained on overlapping data and aligned through similar processes, become the de facto first draft of legal analysis, the profession's output cannot help but converge toward the central tendency of those models. That central tendency is, by training-data construction, the modal — not the best — legal answer. Median analysis is precisely what mediocrity means.

The deepest risk lies one level deeper still. The profession's diversity of thought is not preserved in the LLMs. It is preserved in the lawyers themselves: in habits of close reading, the discipline of writing without help, the willingness to advance an unconventional argument. If those habits atrophy, the loss is not recoverable by switching tools.

LLMs are great scriveners but not necessarily great thinkers.

The right response is not abstention. LLMs are astounding accelerators of productivity that competitive forces will push every firm to adopt. Those efficiency gains open real strategic opportunities: a firm can redeploy the freed-up capacity to take on more work, or pass the savings through as sharper pricing to win market share. Sitting out is not a neutral choice — it is ceding that ground to the firms that move. That said, the LLMs are flawed enough that no responsible firm can outsource judgment to them. Talent always needs to be first. The firms that win the next decade will be the ones that recognize the dominant failure mode of legal AI is not hallucination, which is easily solved — it is conformity — and that deliberately structure their practice, their training, and their governance to resist that conformity precisely where it matters most.

Further Reading

- Wu, Black & Chandrasekaran, *Generative Monoculture in Large Language Models* (2024) — <https://arxiv.org/abs/2407.02209>
- Bommasani, Creel, Kumar, Jurafsky & Liang, *Picking on the Same Person: Does Algorithmic Monoculture lead to Outcome Homogenization?* (2022) — <https://arxiv.org/abs/2211.13972>
- Kleinberg & Raghavan, *Algorithmic Monoculture and Social Welfare*, 118 PNAS e2018340118 (2021) — <https://www.pnas.org/doi/10.1073/pnas.2018340118>
- Creel & Hellman, *The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision-Making Systems*, 52 Canadian J. Phil. 26 (2022) — https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3786377
- Socol de la Osa & Remolina, *Artificial Intelligence at the Bench: Legal and Ethical Challenges of Informing—or Misinforming—Judicial Decision-Making Through Generative AI*, 8 Data & Policy e59 (2024) — <https://www.cambridge.org/core/journals/data-and-policy/article/artificial->

[intelligence-at-the-bench-legal-and-ethical-challenges-of-informing-or-misinforming-judicial-decisionmaking-through-generative-ai/D1989AC5C81FB67A5FABB552D3831E46](https://arxiv.org/abs/2407.13070)

- Ganesh, Daldaban, Cofone & Farnadi, *The Cost of Arbitrariness for Individuals: Examining the Legal and Technical Challenges of Model Multiplicity* (2024) — <https://arxiv.org/abs/2407.13070>
- Corbo, *LLMs in Interpreting Legal Documents* (Elsevier, 2026) — <https://arxiv.org/abs/2512.09830>
- Dahl, Magesh, Suzgun & Ho, *Large Legal Fictions: Profiling Legal Hallucinations in LLMs* (Stanford RegLab, 2024) — <https://arxiv.org/abs/2401.01301>
- Magesh et al., *Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools*, 22 J. Empirical Legal Stud. 216 (2025) — <https://onlinelibrary.wiley.com/doi/full/10.1111/jels.12413>
- *Couvrette v. Wisnovsky*, No. 1:21-cv-00157-CL (D. Or. 2026) — <https://www.wealthmanagement.com/estate-planning/estate-dispute-sees-potentially-largest-sanctions-imposed-for-ai-related-misconduct>
- *Whiting v. City of Athens*, No. 25-5425 (6th Cir. Mar. 13, 2026) — <https://www.opn.ca6.uscourts.gov/opinions.pdf/26a0080p-06.pdf>
- D. Charlotin, *AI Hallucination Cases Database* (2026) — <https://www.damiencharlotin.com/hallucinations/>
- ABA Formal Opinion 512, *Generative Artificial Intelligence Tools* (July 29, 2024) — https://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/ethics-opinions/aba-formal-opinion-512.pdf
- California ethics rulemaking on AI and agentic AI (State Bar of California, 2025–26) — <https://www.esquiresolutions.com/california-ethics-panel-turns-up-the-heat-on-artificial-intelligence/>
- Hedden & Raghavan, *Algorithmic Monoculture and Its Critics* (2026) — <https://arxiv.org/abs/2604.06047>
- Bommasani, Bana, Creel, Jurafsky & Liang, *Algorithmic Monocultures in Hiring* (2026) — <https://arxiv.org/abs/2605.27371>
- Carlton Fields, *Why Our Firm Still Prohibits Generative AI for Legal Research and Written Advocacy* (2026) — <https://www.carltonfields.com/insights/publications/2026/why-our-firm-still-prohibits-generative-ai-for-legal-research-and-written-advocacy>
- DealMentor — AI negotiation training for the legal profession (Deal Mentor LLC) — <https://www.dealmentor.ai>