USING GENAI-DERIVED MAJORITARIAN PATTERNS IN LEGAL ADJUDICATION

Generative AI (GenAI) systems today are primarily optimized to meet user expectations—emphasizing fluency, helpfulness, and alignment with prevailing norms (Lambert, 2022; Marcus, 2024). Yet this same design orientation obscures another crucial potential: GenAI's capacity to function as an epistemic instrument, capable of illuminating social regularities and offering new forms of empirical input for legal adjudication (Arbel & David A. Hoffman, 2024). This neglected potential exemplifies an *undone* area of computer science—an opportunity foregone by research priorities that privilege user experience and safety over interpretive and analytical exploration.

This paper argues that GenAI systems, despite their well-documented biases, can yield insight into the shared linguistic, cultural, and behavioral norms that underpin many judicial standards (Feldman, 2022; Bail, 2024). When approached with methodological rigor and procedural safeguards, GenAI-derived *majoritarian signals* can help courts interpret open-textured legal concepts that already defer to majoritarian practice—such as *reasonable care, ordinary meaning,* and *generic use* (Arbel, 2025; Snell, 2024; Hacohen & Elkin-Koren, 2024).

Across legal doctrine, many judicial standards rely—explicitly or implicitly—on the ordinary understandings of the majority. In tort law, courts assess "reasonable care" through prevailing social practices (Gibson, 1941); in contract law, they fill interpretive gaps based on terms most parties would likely have agreed upon (Barnett, 1992); in trademark and copyright law, determinations of genericity or originality depend on widespread usage and perception (Goldman, 2024). Yet the empirical foundations for identifying such majoritarian benchmarks have long been unstable. Surveys, dictionaries, and expert testimony—the traditional tools for gauging common understanding—are costly, inconsistent, and methodologically opaque (Lee & Mouritsen, 2018).

Foundation models offer a novel empirical alternative. Trained on massive corpora of human communication, they encode statistical regularities that approximate social consensus in language use (Haviv, 2023; Zhang, 2023; Chen, 2025). These regularities—what computational social scientists call *majoritarian signals*—capture how societies collectively produce meaning (Grossmann, 2023). When carefully interpreted, such signals can serve as epistemic mirrors of the linguistic and cultural expectations that legal standards seek to reflect (Gopnik, 2024). GenAI's outputs, in this sense, can help courts approximate ordinary meaning or customary practice without displacing normative judgment. Rather than dictating outcomes, these models can illuminate the linguistic and behavioral terrain within which judicial reasoning already operates.

Incorporating GenAI-derived signals into adjudication, however, raises substantial methodological challenges (Grimmelmann, Sobel, & Stein, 2025; Choi, 2025). Extracting reliable signals requires confronting model opacity, training-data bias, post-training alignment, and prompt sensitivity. Closed-source models like GPT offer limited transparency, while open-source alternatives may trade off performance for verifiability. Even small variations in architecture, fine-tuning, or prompt phrasing can produce divergent outputs—complicating any claim that a signal reflects genuine majoritarian understanding. Moreover, the data underlying foundation models may reproduce historical

inequalities rather than contemporary norms, risking the reinforcement of exclusionary practices (Shur-Ofry, 2025).

Yet these epistemic challenges are not unique to AI; they echo the fragilities of expert evidence, surveys, and dictionaries already used in court (Weatherall, 2017; Lee & Mouritsen, 2018). What distinguishes law is its procedural infrastructure for vetting, contesting, and contextualizing evidence through adversarial testing and evidentiary rules (Roth, 2017). Cross-examination can expose modeling assumptions; expert testimony can frame methodological limitations; and burden-shifting mechanisms can structure how parties use or rebut GenAI-derived signals. In this way, courts can transform AI's technical uncertainties into sites of reasoned evaluation—turning epistemic risk into institutional learning.

The paper's contribution is both conceptual and institutional. Conceptually, it reconceives foundation-model bias as a form of distributed cultural knowledge that, under the right conditions, can inform legal interpretation. Institutionally, it proposes that procedural safeguards—admissibility standards, expert scrutiny, and adversarial rebuttal—provide a pragmatic pathway for integrating GenAI into adjudication without undermining fairness or legitimacy. The result is a reframed vision of AI and law: one that treats GenAI not as an oracle of truth or a source of distortion, but as an empirical witness to the linguistic and normative patterns through which societies understand themselves.

References:

1. Adi Haviv et al., *understanding transformer memorization recall through idioms*, *in* ARXIV (2023), https://arxiv.org/abs/2210.03588.
2. Alison Gopnik, *on Imitation and Innovation in AI,* CITRIS AND THE BANATAO INSTITUTE. (2024).
3. Andrea Roth, *Machine Testimony*, 126 YALE L. J. 1972 (2017).
4. CA. Bail, *Can Generative AI Improve Social Science?* PROC NATL ACAD SCI USA 2024 May 21;121(21): e2314021121. doi: 10.1073/pnas.2314021121. Epub 2024 may 9. PMID: 38722813; PMCID: PMC11127003.
5. Eric Goldman, *The GOOGLE Mark Isn't Generic—Elliott v. Google,* TECHNOLOGY & MARKETING LAW BLOG (Forbes Cross-Post) (Nov. 10, 2024).
6. Gary Marcus, *Has Google gone too woke? Why even the biggest models still struggle with guardrails*, MARCUS ON AI (Feb. 21, 2024), https://garymarcus.substack.com/p/has-google-gone-too-woke-why-even (including screenshots of dialogue with ChatGPT).
7. Igor Grossmann, et al., *AI and the Transformation of Social Science Research,* SCIENCE 380: 1108 (2023).
8. James Gibson, *Doctrinal Feedback and (Un)Reasonable Care*, 94 VIRGINIA LAW REVIEW (1641).
9. James Grimmelmann, Benjamin L.W. Sobel, & David Stein, *Generative Misinterpretation,* 63.1 HARV. J. ON REG (forthcoming, 2025)
10. James Snell v. United Specialty Insurance Company,D.C. Docket No. 1:21-cv-00229-CG-M (05/28/2024).
11. Jonathan H. Choi, *Large Language Models Are Unreliable Judges* at 32 (work in progress).
12. Jiahao Chen, et al., *Rethinking the Bias of Foundation Model under Long-tailed Distribution* (2025) https://arxiv.org/abs/2501.15955.
13. Kimberlee Weatherall, *The Consumer as The Empirical Measure of Trademark Law,* 80(1) THE MODERN LAW REVIEW 57 (2017).
14. Michal Shur-Ofry, *Multiplicity as an AI Governance Principle* (forthcoming at the INDIANA LAW JOURNAL, 2025).
15. Philip Feldman, et al., *Polling Latent Opinions: A Method for Computational Sociolinguistics Using Transformer lLanguage Models.* ARXIV (2022).
16. Nathan Lambert et al., *Illustrating Reinforcement Learning from Human Feedback (RLHF)*, HUGGINFACE, (Dec 9, 2022), https://huggingface.co/blog/rlhf.
17. Randy E. Barnett, *The Sound of Silence: Default Rules and Contractual Consent*, 78 VA. L. REV. 821, 858–59 (1992)
18. Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 YALE L.J. 788, 801, 810–11 (2018).
19. Uri Y. Hacohen & Niva Elkin Koren, *Copyright Regenerated: Harnessing GenAI to Measure Originality and Copyright Scope* 37 HARV. J.L. & TECH. 556 (2024).
20. Yifan Zhang, et al., *Deep Long-Tailed Learning: A Survey*, 45 IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE 10795 – 10816 (2023).
21. Yonathan A. Arbel, *The Silicon Reasonable Person: Can AI Predict How Ordinary People Judge Reasonableness?*, ARXIV:2508.02766 [cs.CY] (2025).
22. Yonathan Arbel & David A. Hoffman, *Generative Interpretation*, 99 NEW YORK UNIVERSITY LAW REVIEW 451 (2024).