Agentic Workflows in the Practice of Law—AI Agents as Ethics Counsel

CATHERINE GAGE O'GRADY* AND CASEY SIMS O'GRADY**

ABSTRACT

Generative AI is reshaping legal practice as law firms invest in AI technology and prepare for a future where AI agents operate alongside human lawyers. While such a future raises numerous ethical concerns, it also opens new opportunities for ethical practice. This article explores the potential for AI agents to improve ethical decision-making within legal practices. We start by defining key concepts such as AI agents and agentic workflows. We then provide a brief overview of the role of ethics counsel in law firms and discuss critical behavioral challenges in the human practice of law. Finally, we introduce a new model of AI agents acting as dedicated ethics counsel. These agents should be specialized, accountable, and systematic to provide comprehensive ethical guidance in a firm's legal workflows. These ethical agents have the potential to mitigate the human biases in traditional legal practice and offer an efficient and scalable approach to ensure ethical compliance in an increasingly AI-driven legal landscape. To demonstrate this model, we end with real-world examples of what initial ethical agent contributions could look like in practice.

TABLE OF CONTENTS

* Professor of Law, James E. Rogers College of Law, University of Arizona.

** Founder and CEO of Blueshoe, a legal technology company (Y Combinator X25). We thank Joan Casciari, Mark Casciari, Scott Killingsworth, Cas Laskowski, Megan Nichols, Andrew M. Perlman, Stanley S. Reynolds, Keith Swisher, and Kai Yee Wan.

INTRODUCTION: GENERATIVE AI AND THE PRACTICE OF LAW

Even in its infancy, generative AI is impacting the practice of law. Although traditionally reluctant to adopt new technologies, law firms are already making significant investments in internal and third-party approaches to generative AI.[1] While the future of AI development in law's workplaces is far from clear, we believe legal practice will start to move toward agentic systems that combine human lawyers with AI agents. This shift will bring undeniable changes to the

1. *See* Lisa Shuchman, *Gen AI Adoption Is Taking Off. Law Firms Are Finally Ready*, LAW.COM INT'L (June 2, 2024), www.law.com/2024/06/02/gen-ai-adoption-is-taking-off-law-firms-are-finally-ready/ [https://perma.cc/G5XT-W62T] (noting that large law firms are increasingly hiring experts in generative AI to lead their technological transformation efforts); *CoCounsel: The Legal AI Assistant and Tool Essential for Legal Teams*, THOMSON REUTERS LEGAL (June 13, 2024), https://legal.thomsonreuters.com/blog/legal-ai-tools-essential-for-attorneys/[https://perma.cc/X7AG-5BV6] (explaining that law firms are increasingly adopting AI tools to automate routine tasks, enhance legal research, and improve overall efficiency, making AI an essential part of modern legal practice).

practice of law, and the possibilities have raised a variety of ethical concerns.[2] Even before the advent of ChatGPT, scholars worried about how to manage potential public risks posed by 'autonomous machine' tools.[3] Recent scholarship, however, has argued that the current use of AI tools is generally consistent with lawyers' ethical obligations and future AI tools may increase efficiency so substantially that competent attorneys will have to use them.[4]

This article makes a unique contribution by exploring how *specialized*, *accountable*, and *systematic* AI agents operating as ethics counsel can overcome human behavioral challenges and improve ethical decision making in the practice of law. Part II of this article starts by defining the key concepts needed to describe what agentic workflows are and how they may be used by law firms. Part III provides a brief overview of American law's current ethical regime, including the behavioral challenges lawyers face in ethical practice and the role of ethics counsel in a law firm. Finally, Part IV examines how agentic workflows might impact ethical practice and advocates for specialized ethical agents that can simplify accountability practices and provide systematic support to perform the work of ethics counsel.

Before we describe agentic workflows and the role of AI ethics counsel, the testimonial below will provide some context. It is from an attorney recounting a mistake made in the practice of law:

> In the process of filing for asylum, I was supposed to send some documents to an agency to get the biometric process started. I forgot to do it. By the time I realized my mistake, it was much too late to get it done in time for the asylum hearing.[5]

Lawyers, like all humans, occasionally make mistakes—all sorts of lawyering mistakes are possible from minor slip-ups to serious ethical mistakes. The above

---

2. *See* Benjamin Alarie, Anthony Niblett & Albert H. Yoon, *How Artificial Intelligence Will Affect the Practice of Law,* 68 U. TORONTO L.J. 106 (discussing many impacts of generative AI on the practice of law).

3. *See, e.g.*, Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J. L. TECH. 353, 356 (2016); Harry Surden, *Artificial Intelligence and Law: An Overview*, 35 GA. ST. U. L. REV. 1305, 1320 (2019).

4. Andrew M. Perlman, *The Legal Ethics of Generative AI,* SUFFOLK U. L. REV. (forthcoming 2025) (manuscript at 2) (concluding that "the duty of competence may eventually *require* lawyers' use of generative AI."); *see also* Keith Swisher, *The Right to (Human) Counsel: Real Responsibility for Artificial Intelligence,* 74 S.C. L. REV. 823, 825-27 (2023) (assuming, for the sake of argument, that fully AI lawyers will exist in 100 years and arguing for new approaches to ethical rules for the next-gen practice of law to address AI, not humans, as primary legal counsel). Recently, in its first and (so far) only ethics opinion addressing lawyers' use of AI, the American Bar Association noted: "As [Generative AI] tools continue to develop and become more widely available, it is conceivable that lawyers will eventually have to use them to competently complete certain tasks for clients." ABA Comm. on Ethics & Prof'l Resp., Formal Op. 512 at 5 (2024), https://www.americanbar.org/content/dam/aba/administrative/professionalresponsibility/ethics-opinions/aba-formal-opinion-512.pdf [https://archive.ph/GnUE2].

5. Catherine G. O'Grady, *A Behavioral Approach to Lawyer Mistake and Apology,* 51 NEW ENGL. L. REV. 7, 13 (2017) (quoting an immigration defense attorney with eleven years practice experience who was describing a mistake made in the practice of law).

testimonial suggests a certain type of mistake—perhaps a busy, experienced lawyer, who was overwhelmed, forgot to start an important process in sufficient time for a scheduled hearing. A similar mistake could have been made by an inexperienced lawyer who did not realize they needed to send documents to an agency well in advance of the hearing. This type of mistake, which may result in a missed deadline, is a commonly understood malpractice risk.[6] An experienced immigration attorney working on the case might have reminded the lawyer to submit documents in advance. If the firm had a full-time ethics counsel in place, a cautious lawyer *might* have discussed the case with ethics counsel; however, unless the ethics counsel is also an experienced immigration lawyer, they would not likely inform the lawyer to start the biometric process several weeks or months in advance. At best, ethics counsel would ensure that the firm has docketing systems in place and that the lawyer uses the docketing systems to prevent this kind of error.

Now imagine a law firm operating with an AI agent acting as ethics counsel.[7] The AI ethics counsel agent can look over every step of the lawyering process. Thus, in reviewing the case activities, the AI agent would note that the lawyer filed for asylum on behalf of a client and immediately send a notice to the lawyer reminding them that they need to submit documents by a certain date to begin the biometric process in time for the hearing. The AI agent could further provide the lawyer with relevant case law and other authorities to review. With the AI agent's ethical reminder, the mistake would likely be averted without the lawyer ever needing to ask for guidance.

From this rather straightforward example of a well-known malpractice risk—a missed deadline—we can imagine an AI ethics agent operating to manage more complicated ethical risks. For example, an AI ethics counsel reviewing the asylum application might note a potential conflict with one of the firm's new clients which was not in place when opening the asylum matter and was not caught in any updated conflict check. In Part IV of this article, we discuss several examples of AI ethics counsel operating to manage both malpractice risks and ethical risks.[8]

---

6. *See id*. at 12-14 (describing a variety of attorney mistakes).

7. In this paper, we use the term "AI ethics counsel" to refer to an AI agent acting as ethics counsel or assisting a firm's ethics counsel. This is in contrast to a different concept, not widely used, which refers to an AI ethics counsel as someone charged with the responsibility of ensuring the ethical use of AI at a firm or organization. *See Establishing & Maximising an AI Ethics Counsel Within Your Organisation,* AI DATA & ANALYTICS NETWORK, 3rd Annual Responsible AI Summit (Sept. 2024), https://www.aidataanalytics.network/events-responsible-ai-summit/downloads/establishing-maximising-an-ai-ethics-counsel-within-your-organisation [https://archive.ph/6UNKc].

8. *See infra* notes 85–92 and accompanying text (describing various ways an AI ethical agent might operate within a law firm to answer questions, prevent mistakes, and manage both malpractice and ethical risks).

I. AGENTIC WORKFLOWS

This section first presents our working definitions of AI and AI agents—both of which aim to simplify often broadly defined concepts. Our definition of AI focuses exclusively on generative technologies, namely LLMs. An AI agent is merely any tool that uses these technologies. With these building blocks in place, we elaborate on what an agentic workflow looks like and how such workflows might be integrated in a legal practice to work alongside human lawyers.

A. DEFINITION OF AI GENERALLY

We define the umbrella term Artificial Intelligence (AI) in a specific way: AI describes technology that employs a foundation model to process, understand, and generate unstructured data, including texts, images, and videos.[9] This is often labeled generative AI, but for simplicity we simply refer to it as AI.[10] For text-driven AI, which will be the primary focus in this paper, the foundation models employed are large language models (LLMs)—well-known examples include GPT, Claude, LLaMA, and Gemini.[11] Any software, tool, or system that calls upon one of these models would be described as an AI software, tool, or system under our definition.

This definition precisely captures the technology we are interested in without dragging us into the broader etymological contexts of "artificial intelligence" prior to the advent of LLMs. For example, we are not defining AI in terms of its ability to mirror "human intelligence," and therefore, avoid the longstanding philosophical contentions and science-fiction visions associated with AI.[12] Furthermore, our definition is more focused than the broader computer science use of "AI," which is used as a catchall for an array of machine learning, deep learning, and other algorithmic techniques.[13] While some of these techniques

---

9. Cole Stryker & Mark Scapicchio, *What is Generative AI?*, IBM TOPICS (Mar. 22, 2024), https://www.ibm.com/topics/generative-ai [https://perma.cc/WF52-73CY] (advancing a similar definition of generative AI in terms of its ability to "create original content—such as text, images, video, audio or software code").

10. *See id.*

11. OPENAI, CHATGPT, GPT-4 (2024), https://www.openai.com/chatgpt [https://archive.ph/4jKKl]; ANTHROPIC, CLAUDE 3.5 SONNET (2024), https://www.anthropic.com/claude [https://perma.cc/ZC7G-NFCK]; META, LLAMA (LARGE LANGUAGE MODEL META AI), VERSION 1 (2023), https://ai.meta.com/llama; GOOGLE, GEMINI, VERSION 1 (2023), https://ai.google.com/gemini.

12. *See* STUART J. RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH §1.1, 19-23 (4th ed. 2020) (discussing various definitions of AI, including those based on fidelity to human performance and those emphasizing rationality—broadly defined as doing the "right thing"—and noting disagreements over whether intelligence should be characterized by internal thought processes or external behavior). Science fiction is teeming with imaginative notions of how such technology might play out; *see, e.g.*, ARTHUR C. CLARKE, 2001: A SPACE ODYSSEY (1968); PHILIP K. DICK, DO ANDROIDS DREAM OF ELECTRIC SHEEP? (1968); WILLIAM GIBSON, NEUROMANCER (1984). Philosophers have long debated the possibilities of machines' ability to simulate human cognition and decision-making, but such debates are unnecessary here; Alan M. Turing, *Computing Machinery and Intelligence*, 59 MIND 433, 433 (1950); HUBERT L. DREYFUS, WHAT COMPUTERS CAN'T DO: A CRITIQUE OF ARTIFICIAL REASON (1972).

13. *See, e.g.*, *Artificial Intelligence*, DEPT. OF COMPUT. SCI. AND ENG'G, OHIO ST. U., https://cse.osu.edu/

helped build current LLMs and other foundation models, we focus on LLMs and models independently to best examine their contribution and impact on human work. Of course, there could be future advances in AI technology that reorient the focus away from LLMs, but contemplating such possibilities is beyond the scope of this paper.

B. DEFINITION OF AI AGENTS

We define AI agents as discrete tools that employ AI technology (e.g. LLMs) to complete tasks. This definition differs from those that emphasize the *autonomy* or *independence* of agents.[14] In a recent article entitled *Governing AI Agents*, Noam Kolt describes AI agents as "not mere tools, but *actors*" that do not simply "produce synthetic content" but "can independently accomplish complex goals on behalf of humans."[15] Certainly, what makes agents interesting is their ability to operate in the dynamic agentic workflows we discuss in the next section. But the distinction between "tool" and "autonomous or semi-autonomous actor" seems unimportant and potentially confusing.[16] The difference between a "tool" that can produce synthetic content, like an email, and an "autonomous" agent can be little more than the ability to hit send.[17] While automation is often important in agentic workflows, it need not characterize each agent individually. Some agents may operate like actors, while others are mere tools within the workflow.

Recognizing the difficulty of isolating pure agents, other definitions focus on systems, not agents. An OpenAI whitepaper, for example, sidesteps the notion of an individual agent altogether and instead defines *agenticness* as "the degree to which a system can adaptably achieve complex goals in complex environments with limited direct supervision."[18] Under this definition, there is "no clear line along which to draw a binary distinction between 'agents' and current AI systems."[19] Similarly, a Fairness, Accountability, Transparency, and Ethics (FATE) research paper identifies "[four]

---

research/artificial-intelligence [https://perma.cc/S79S-RJY9] (last visited Nov. 21, 2024) (defining AI as "a broad multidisciplinary area drawing from computer science, neuroscience and cognitive science, linguistics, statistics, applied mathematics and many other areas of research"); DAVID L. POOLE & ALAN K. MACKWORTH, ARTIFICIAL INTELLIGENCE: FOUNDATIONS OF COMPUTATIONAL AGENTS (2d ed. 2017); IAN GOODFELLOW, YOSHUA BENGIO & AARON COURVILLE, DEEP LEARNING (2016); CHRISTOPHER M. BISHOP, PATTERN RECOGNITION AND MACHINE LEARNING (2006).

14. *See* Noam Kolt, *Governing AI Agents* 12 (Apr. 2, 2024) (working paper), https://ssrn.com/abstract=4772956 [https://archive.ph/zqjbe] (stating that AI agents can "independently accomplish complex goals").

15. *See id.*

16. It is also unclear what autonomy or independence means for an AI agent, given that AI is not capable of choosing or executing its own goals in ways not tied back directly to human design or input.

17. For instance, Chat-GPT, perhaps the most well-known AI agent, is frequently used by people as a tool; for example, for research, coding, or to present questions and receive answers. At the same time, Chat-GPT performs a number of actions autonomously to complete its task as a tool; for example, determining the right approach to an answer based on the query, retrieving information from the internet when needed, and even understanding the sentiment of the user based on past interactions.

18. *See* Yonadav Shavit et al., *Practices for Governing Agentic AI Systems* 4 (Dec. 12, 2023) (working paper), https://cdn.openai.com/papers/practices-for-governing-agentic-ai-systems.pdf [https://perma.cc/ZG3V-3ADD].

19. *Id.*

key characteristics which, particularly in combination, tend to increase the agency of a given algorithmic system: under-specification, directness of impact, goal-directedness, and long-term planning."[20] Perhaps in purely technical sense, these systemic approaches are correct—it is often difficult to isolate a truly pure agent from the systems that constitute it and those in which it operates. Yet, we find value in the term "agent" as a discrete entity within a broader system, even if that agent can also be considered a system itself.[21] Although there is some arbitrariness in what is the "system" and what are the "agents," we find the designations meaningful in describing and understanding how AI workflows operate.[22]

The important thing about agents under our conception is that they direct and use LLMs (or foundation models) to complete particular tasks. Regardless of their degree of autonomy or how much human input they need to function, AI agents apply LLMs within a particular context to complete specific work.[23] Functionally, AI agents differentiate themselves in numerous ways to adjust and calibrate for these specific needs.[24] As we unpack agentic workflows, we see the importance of this specialization as agents rely upon others for specific tasks.

C. DEFINITION OF AN AGENTIC WORKFLOW

An agentic workflow describes how multiple AI agents work *independently* and *iteratively* to manage complex tasks. As noted in our discussion of Open AI's definition of "agenticness," these workflows can have lower or higher degrees of independence, complexity, and flexibility.[25] For us, the important thing is that a workflow employs multiple agents to complete tasks as part of a process. Technologists increasingly see the benefits of agentic workflows for multi-part, collaborative tasks as they provide greater degrees of accuracy, flexibility, and

---

20. Alan Chan et al., *Harms from Increasingly Agentic Algorithmic Systems*, 2023 PROC. OF THE 2023 ACM CONF. ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 651.

21. For instance, GPT-4 can be considered both an AI system, as it employs numerous and iterative processes, and an AI agent, in so far as it is used in a discrete way within a large system, for instance within an AI workflow that makes an application programming interface call to GPT-4.

22. As mentioned before, an AI tool can be a system of its own and also an agent within a larger system. The way it's viewed depends on the "level" of our analysis. If we want to understand how ChatGPT works, we might talk about it as a system and describe the many agents that operate within it. If we want to understand the use of ChatGPT in a larger workflow, we can think of it as an agent.

23. *See* Chhavi Chawla et al., *Agentic AI: The Building Blocks of Sophisticated AI Business Applications*, 3 J. OF AI, ROBOTICS & WORKPLACE AUTOMATION 1, 1–15 (2024), https://doi.org/10.69554/XEHZ1946 [https://perma.cc/PQG8-YDJQ] (defining agent as language models organized into workflows to perform specific tasks).

24. These include, but are not limited to the following: (1) selecting the foundation model (e.g., LLM), (2) prompt engineering to provide the AI agent with specific instructions to guide outputs, (3) retrieval augmented generation ("RAG") to incorporate contextual data or additional knowledge from sources outside the model's training, (4) fine-tuning to further train the LLM in order to structure precise outputs, and (5) integration APIs that allow the AI agent to interact with other systems, execute external actions, and integrate seamlessly with other software.

25. *See* Shavit et al., *supra* note 18, at 4 (breaking down agenticness into components of goal complexity, environmental complexity, adaptability, and independent execution).

thoroughness.[26] This is not too surprising. Indeed, a forthcoming study suggests that humans perceive an agentic workflow model as more trustworthy than other systems.[27] Like humans, AI agents benefit from working with independence, specialization, and iteration.

While singular AI agents can draw upon extensive knowledge to generate impressive outputs without much input or iteration, complex processes are still best mastered by breaking them into discrete tasks and workflows for specialized agents.[28] These workflows provide structure for agents to prevent missteps and create feedback loops that allow agents to review, refine, and improve their work. Thus, an agentic approach breaks down a task, such as writing an essay, into logical parts and has agents intervening across the workflow.[29] For instance, the initial agent might outline an essay. A second agent would gather and summarize relevant research. A third agent would draft the essay, and additional agents would give feedback and revise. This workflow would likely center around AI agents leveraging LLMs, but it could also incorporate human or non-AI agents as well.

D. THE AGENTIC LAW FIRM

Expanding the agentic concept further, we can start to imagine what an agentic workflow at a law firm might look like—an array of AI agents working on discrete tasks and sharing information within a broader legal workflow. The level of human supervision or involvement could vary. Early workflows might incorporate only a handful of AI agents; indeed, law firms are already experimenting with agent adoption.[30] As technology improves and develops, more of the workflow might be accomplished by agents.

An agentic legal workflow, for example, might begin with *client engagement agents*. These agents would engage in email and online dialogues with prospective or current clients to determine the nature of the client's case and the basic

---

26. *See* Tula Masterman et al., *The Landscape of Emerging AI Agent Architectures for Reasoning, Planning, and Tool Calling: A Survey,* ARXIV (Apr. 17, 2024), https://arxiv.org/abs/2404.11584 [https://perma.cc/KH3A-DCG5] (contending that multi-agent architectures tend to thrive more when collaboration and multiple distinct execution paths are required).

27. *See* Bart S. Vanneste & Phanish Puranam, *Artificial Intelligence, Trust, and Perceptions of Agency*, ACAD. OF MGMT. REV. (forthcoming).

28. *See* Masterman et al., *supra* note 26 ("Multi-agent architectures are generally well-suited for tasks where feedback from multiple personas is beneficial in accomplishing the task.").

29. *See* Chawla et al., *supra* note 23 (describing how agentic approach mimics human workflows by breaking down problems into structured tasks for agents that work reflectively to self-correct and iteratively to incorporate feedback).

30. *See* LEXISNEXIS, 2020 ALM-LEXISNEXIS LEGAL ANALYTICS STUDY: BRINGING LEGAL ANALYTICS INTO FOCUS 22 (2020) (stating that 92% of those surveyed planned to increase their adoption of AI for legal analytics in the next year); *see also* LEX MACHINA & LEXISNEXIS, 2024 IMPACT OF LEGAL ANALYTICS SURVEY 4 (2024) (suggesting that 80% of legal professionals reported that clients required or expected them to use legal analytics, and over two-thirds are excited about generative AI).

factual details surrounding the matter.[31] Next, *summary of facts agents* would determine whether there is enough information to assess a new case's prospects. If these agents conclude that the firm lacks sufficient information, it will return the matter to the *client engagement agents* to gather more information. If the case moves on, *legal analysis agents* would use the summarized facts to outline the potential legal claims and note the elements necessary to prove the claims. These agents would work iteratively with *relevant law agents* that complete legal research on relevant case law, statutes, and specific legal issues. This work would be used by fine-tuned *generation agents* to create well-structured pleadings or briefs. Additional *auditing agents* could review the work in a variety of ways to ensure completeness of the claim, correct legal reasoning, and to address potential counterarguments. Finally, as described in Section IV below, an *ethics agent* could be inserted throughout this workflow to provide ethical guidance when necessary.

Of course, human attorneys would still be instrumental in working alongside and maintaining these agentic workflows. In the early days of AI adoption, attorney work will largely proceed unchanged with agents only serving as co-pilots to aid human work. Even in the long-run, agentic workflows are too intricate and complex to imagine them operating without significant human management and oversight. The same way robotic assembly lines require human design, oversight, and fine-touch work, an AI-driven system will require people to design workflows, audit agents, oversee sensitive work (e.g. client communication) and review final outputs (e.g. pleadings and briefs). In some of these areas, human review might diminish as AI agent work improves, but human oversight and review will remain necessary to ensure quality and accountability.[32]

II. HUMAN LAWYERING—BEHAVIORAL CHALLENGES AND ETHICS COUNSEL IN A LAW FIRM

A. INDIVIDUAL ATTORNEY RESPONSIBILITY AND BEHAVIORAL MOTIVATIONS

Under the current regime, all attorneys are individually responsible for ethical compliance and all lawyers are required to have ethical training and an

31. Recently, the Florida State Bar issued an ethics opinion discussing the permissibility of using a chatbot for communicating with prospective and current clients. *See* Fla. State Bar Ethics, Op. 24-1 (2024), https://www.floridabar.org/etopinions/opinion-24-1/ [https://perma.cc/7XHF-LQYB] ("Generative AI chatbots that communicate with clients or third parties must comply with restrictions on lawyer advertising and must include a disclaimer indicating that the chatbot is an AI program and not a lawyer or employee of the law firm."); *see also* ABA Comm. on Ethics and Prof'l Responsibility, *supra* note 4 (the American Bar Association's first ethics opinion on lawyers' use of AI).

32. *See* Michael D. Murray, *Artificial Intelligence and the Practice of Law Part 1: Lawyers Must be Professional and Responsible Supervisors of AI* (June 14, 2023) (working paper), https://ssrn.com/abstract=4478588 [https://archive.ph/UZHSI] (contending that AI systems require responsible supervision and should not be asked to exercise discretionary judgment).

understanding of the ethics rules.[33] Before admission to the bar, a lawyer must pass the Multistate Professional Responsibility Exam (MPRE), a two-hour multiple-choice examination based solely on legal ethics. In addition, most states require that all active barred lawyers take a certain number of continuing legal education classes each year, which focus on ethics and fulfill mandatory ethics credits.[34] Even brand-new lawyers are assumed competent; they must comply with the rules of ethics, and it is no defense to claim that a "supervising attorney told me to do it."[35] This regime is based on individual responsibility and autonomy. Practically speaking, the individual responsibility model is necessary to ensure compliance and enforcement of the rules of ethics. It is not possible, for example, for a law firm's human ethics expert to look over the shoulder of every human lawyer in the firm at every juncture of the lawyering process. Individual responsibility is a necessity.

Lawyers, however, are not perfect ethical actors. Individual choices are often influenced by the power of underlying and unconscious implicit biases, heuristics, individual motivations, or situational dynamics. For any number of reasons, a lawyer may fail to be properly guided by a relevant rule of legal ethics. Perhaps the lawyer did not see an ethical issue or know about a relevant rule, or perhaps the ethics rule simply was not top of mind at the moment of the lawyering decision. The growing field of behavioral legal ethics looks beyond the rules of ethics to understand how various decision-making biases, situational pressures, and personal motivators might explain lawyers' decision-making.[36]

All lawyers confront behavioral challenges in their work, but the impacts of such pressures differ in law firms' hierarchical practice of law. New and mid-level associates, for example, are particularly susceptible to situational pressures

---

33. In the United States, lawyers' ethical conduct is overseen by states, and all states have adopted rules of professional responsibility, which are typically modeled on the American Bar Association's Model Rules of Professional Conduct. Generally, the body of law governing lawyering includes a jurisdiction's rules of professional responsibility, comments to the rules, selected statutes, the Restatement on Lawyering, court opinions, and ethics opinions issued by state bar associations. The ethics rules and principles found in these sources guide attorneys, judges, and law firms in the current ethical practice of law.

34. In Arizona, for example, each lawyer must acquire a total of 15 hours of mandatory continuing legal education each year, "three of which must be professional responsibility (ethics)." Ariz. Sup. Ct., *MCLE Regulations*, STATE BAR OF ARIZ. (June 27, 2022), https://azbar.org/media/k1jljdut/mcle-regulations-effective-june-27-2022.pdf [https://perma.cc/Y7FR-K5JE].

35. *See* Connick v. Thompson, 563 U.S. 51, 64-65 (2011) (noting that the presumption of competence applies "even to young and inexperienced lawyers in their first jury trial and even when the case is complex," *citing* United States v. Cronic, 466 U. S. 648, 658, 664 (1984)); *see also* Andrew M. Perlman, *The Silliest Rule of Professional Conduct: Model Rule 5.2(b)*, 19 PRO. LAW. 14, 14 (Sept. 6, 2009) (arguing that Model Rule 5.2 (b) provides no protection for new lawyers following a supervisor's unethical instructions).

36. *See infra* Part IV(A) for an analysis of several key behavioral influences. For a fuller exploration of behavioral legal ethics, see generally CATHERINE G. O'GRADY & TIGRAN W. ELDRED, BEYOND THE RULES – BEHAVIORAL LEGAL ETHICS AND PROFESSIONAL RESPONSIBILITY (2021); Jennifer Robbennolt & Jean Sternlight, *Behavioral Legal Ethics*, 45 ARIZ. ST. L.J. 1107 (2013) (surveying a wide range of ways that biased reasoning can influence lawyers' ethical decision-making).

and the desire to earn recognition and promotion.[37] Senior lawyers, on the other hand, may be unconsciously influenced by pressures to bring in new business and please established clients, which can lead to "ethical fading"—the framing of issues as business decisions that do not focus adequately on ethical issues.[38] Moreover, senior lawyers may not know recently adopted ethical rules as well as junior lawyers who recently studied the current ethical rules for their law school's professional responsibility course or have recently taken the MPRE. Overall, behavioral dynamics operate at all junctures of the lawyering process and may serve to explain or influence unintentional unethical conduct or mistakes.[39]

B. LAW FIRM ETHICS COUNSEL AND ETHICAL INFRASTRUCTURE

> "We have an ethics counsel at our firm. The ethics counsel is a lawyer who advises attorneys in the firm on how to proceed and who is called up to teach mistake prevention."[40]

Law firms are increasingly using one or more "ethics advisors" or "ethics counsel" to work with firm lawyers and firm systems.[41] A firm's ethics counsel can play a vital role in ensuring the firm's ethical practice of law. Ethics counsel can help the firm avoid unethical conflicts of interest, manage existing and potential claims against the firm, promote compliance with firm procedures, enhance lawyers' ethical decision-making, and reduce attorney mistakes. For many firms, two or more attorneys, typically partners, serve as part-time ethics counsel; other firms, especially large law firms, might devote a full-time senior partner to work as ethics counsel.[42] Some firms also establish an ethics committee, with ethics

---

37. *See* Catherine G. O'Grady, *Behavioral Legal Ethics, Decision Making, and the New Attorney's Unique Professional Perspective*, 15 NEV. L.J. 671, 680–89 (2015) (contending that a new attorney is "particularly vulnerable to many of the social dynamics of today's legal workplaces" including the culture of their work groups, obedience to senior lawyers, cognitive overload, and financial pressures).

38. *See id.* at 689-92 (contending that senior lawyers are more likely to be dependent on business schemas making them susceptible to ethical fading and fast (System 1) thinking).

39. *See* O'GRADY & ELDRED, *supra* note 36; Robbennolt & Sternlight, *supra* note 36.

40. O'Grady, *supra* note 5, at 38 (quoting an attorney at a general commercial litigation firm with fifty lawyers).

41. We use the term "ethics counsel" to refer to this person, but there are many titles used for this in-house compliance specialist including firm counsel, general counsel, ethics partner, ethics advisor, or attorney to the firm. *See* Elizabeth Chambliss & David B. Wilkins, *The Emerging Role of Ethics Advisors, General Counsel, and Other Compliance Specialists in Large Law Firms*, 44 ARIZ. L. REV. 559, 565 (2002) (noting a long list of titles all meant to designate what a lawyer in the firm might recognize as "our ethics [person]").

42. *See* Kimberly Kirkland, *Ethical Infrastructures and De Facto Ethical Norms at Work in Large US Law Firms: The Role of Ethics Counsel*, 11 LEGAL ETHICS 181, 187-88 nn.46-47 (2009) (citing Altman Weil's 2008 Flash Survey on Law Firm General Counsel, which "indicates that of the 86 Am Law 200 law firms responding to the survey, 62% had part-time general counsel" and "82% reported that their general counsel were partners in the firm"); *2008 Results of Confidential 'Flash' Survey on Law Firm General Counsel*, ALTMAN WEIL, INC. (Apr. 2008), https://www.altmanweil.com/wp-content/uploads/2022/05/Law-Firm-General-Counsel-2008-An-Altman-Weil-Flash-Survey.pdf [https://perma.cc/4N9H-WYK4]; *see also* Chambliss & Wilkins, *supra* note 41, at 585 (noting that "[a]ll but one of the specialists in our sample is a partner or in a position of comparable seniority . . .").

counsel reporting to or chairing the committee.[43] As discussed in more detail below, given limited resources, small and medium-sized firms are less likely than larger firms to designate an ethics counsel, and less likely to compensate or credit ethics counsel for their ethics work if they have designated ethics counsel.[44]

A law firm's ethics counsel may perform a number of functions; however, studies indicate that identifying and managing conflicts of interest are the primary responsibilities of a law firm's in-house ethics specialist.[45] In addition to handling conflicts, ethics counsel may guide and empower firm attorneys in ethical decision-making by confidentially listening to and answering lawyers' ethical questions and by helping to establish the firm's ethical infrastructure, which will be designed to overcome counter-ethical drives and remind lawyers of ethical obligations. The topics and issues ethics counsel may encounter (other than conflicts) cover a wide range, including: "attorney-client privilege and work product; advertising and solicitation; communication with represented parties; lateral hiring and departure; fees, billing and trust accounts; mandatory and permissive withdrawal; and the duty to report misconduct by other lawyers."[46]

In general, a firm's ethics counsel encounters ethical questions and issues in two broad ways—through individual attorney consultation, as the first point of contact at the firm for lawyers facing an ethical question or dilemma, and through systemic planning and monitoring, in other words, helping to create the firm's ethical culture, disseminate ethical norms, and establish the firm's ethical infrastructure. Under the "individual consultation" umbrella, ethics counsel can confidentially advise firm attorneys who seek their input on ethical concerns. Indeed, one of ethics counsel's greatest challenges is getting lawyers to reach out to them when they have ethics questions, issues, concerns, or even fears that they may have made a mistake.[47] Under the "systemic" umbrella, ethics counsel might procure in-house ethics training for all lawyers; provide specialized associate training; engage in regular educational emails and newsletters to firm attorneys;

---

43. *See* Chambliss & Wilkins, *supra* note 41, at 568 (noting that some of the ethics counsel surveyed were "chairs" of committees but did the bulk of the work while others co-existed with committees that rarely met).

44. *See id.* at 570, 576 (concluding that firm size is likely to be "an important determinant of firms' investment in in-house compliance specialists" and noting that the larger firms in their sample "tend to have more ethical infrastructure than the smaller firms" and are "be more likely to compensate in-house specialists").

45. *See id.* at 566 (noting that previous studies suggest that conflicts of interest issues are not only the primary substantive issue for in-house specialists, but that many firms "may invest relatively little in other types of ethical infrastructure."); Kirkland, *supra* note 42, at 184 (noting findings that "the decreasing loyalty of clients to firms and the increasing lateral mobility of large law firm partners has made conflicts of interest the primary focus of ethics counsel"). Increasingly, firms are employing full-time "conflicts check" attorneys to perform conflicts of interest functions, and they may report to the firm's ethics counsel.

46. Chambliss & Wilkins, *supra* note 41, at 566–67.

47. *See* O'Grady, *supra* note 5, at 29–31(noting the significant hurdles lawyers face in acknowledging that they made a mistake); Chambliss & Wilkins, *supra* note 41, at 587 (noting that ethics counsel in their study report that the "thing that keeps them 'up at night' is how to get people, especially partners, to raise rather than ignore ethical questions" and that some participants expressed a fear that partners "do not even recognize ethical questions").

develop standardized firm forms, such as conflicts waivers; survey and adopt programs and products to support the firm's ethical infrastructure (such as conflicts systems, billing programs, and docketing systems); review firm advertisements, marketing materials, and websites for compliance with the rules of ethics; monitor trust accounts; and respond to motions to disqualify firm attorneys.[48]

Across all responsibilities, ethics counsel will likely encounter tensions between establishing "professional independence norms" at the firm, through ethical decision-making input, and making or assisting with "business decisions" for the firm to remain competitive in the legal marketplace by taking on new clients and new business.[49] As an expert in ethics, ethics counsel will presumably recognize ethical issues and be knowledgeable about the body of laws that govern the resolution of such issues. Moreover, ethics counsel typically works outside the case's working group or team; thus, they can provide a fresh eye on an ethical question and dilute the pressures to conform to the working group. But, like all human decision-makers, ethics counsel may be unknowingly impacted by behavioral challenges and pragmatic pressures. If ethics counsel is a senior partner in the firm, which is typical, they may be unconsciously influenced by business schemas and the desire to ensure the firm is successful and competitive in the market by bringing in new business and winning cases.[50] Reliance on ingrained business schemas and a "business-decision" focus can lead to unintentional "ethical fading" and motivated reasoning.[51] Thus, ethics counsel can be extremely helpful in making a positive impact on ethical decision-making, but ethics counsel will nonetheless face their own unique behavioral dynamics.

Overall, assigning and employing a firm lawyer, typically a partner, to be ethics counsel in a law firm is an expensive solution. In general, full-time ethics counsel do not generate new clients or bill hours, and part-time ethics counsel likely generate fewer clients and billable hours than other firm lawyers due to the time demands of the ethics counsel position.[52] The role is understandably considered a "burden" if it is assigned to lawyers who must simultaneously keep up with the firm's new business generation and billing responsibilities.[53] Thus, having a dedicated ethics

48. *See* Chambliss & Wilkins, *supra* note 41, at 575.

49. *See* Kirkland, *supra* note 42 at 196, 199 (identifying these two tensions).

50. *See id.* (concluding through interviews of ethics counsel that a majority of large firm ethics counsel concentrate on business interests like remaining competitive and taking on new business instead of managing professional independence norms).

51. *See supra* Part III (A).

52. *See* Kirkland, *supra* note 42, at 182 n.5 (noting that full time ethics counsel do not practice law for clients of the firm); Chambliss & Wilkins, *supra* note 41, at 572–73 (noting that most part-time ethics counsel in their study were not be compensated directly for their in-house compliance work, while some part-time specialists were compensated directly, and full-time specialists maintained no outside practice and were compensated only for in-house compliance work).

53. Chambliss & Wilkins, *supra* note 41, at 573–74 (noting that uncompensated part-time ethics counsel "tend to characterize their in-house work as a 'burden'" and generally focus primarily on conflicts issues while those who are compensated directly "tend to play a much broader and more proactive role in their firms"). The time constraints are unique to this role in that lawyers with an ethical dilemma feel a sense of urgency to consult

counsel in place may not be a viable solution for many firms—especially small and medium-sized firms. That is unfortunate. The presence of ethics counsel at a firm provides a number of advantages and sends a clear signal to all firm lawyers that the firm's senior management values ethical decision-making.[54] In particular, a full-time ethics specialist is likely to be viewed as more approachable, especially by associates who may be reluctant to raise ethics questions with a partner who is directly evaluating their work.[55] And importantly, the presence of ethics counsel has been shown to decrease costs associated with malpractice claims, suggesting a positive impact on ethical decision-making.[56]

III. A NEW VISION: AI AGENTS AS ETHICS COUNSEL

An agentic law firm with specialized AI ethics agents presents an entirely new model in the ethical practice of law. AI agentic systems will have the capacity to provide immediate, affordable advantages to law firms and other legal workplaces looking to ensure ethical decision-making and support individual lawyers. In this section, we explore the role of *specialized AI ethics agents*. We contend that such agents will be less prone to some behavioral counter-ethical challenges, and that independent ethical agents provide a focused technical platform for ethical design, training, testing, updating, and supervising. When deployed properly, AI ethics agents can proactively alert other AI agents and human lawyers to ethical issues arising from their work and review actions across the firm to ensure ethical compliance. We conclude with some examples to illustrate the instant ethical feedback and guidance such agents might contribute at a scale unthinkable in the current regime.

A. REDUCING THE IMPACT OF DECISION-MAKING BIASES AND BEHAVIORAL INFLUENCES

AI does not replicate the human thinking process; rather, it is artificial, simulation-based thought processing.[57] Thus, in theory, it seems likely that some key decision-making biases will not impact AI to the same extent as humans. Unfortunately, however, if AI is relying on a bias-laden prompt or a limited data set that incorporates bias

---

with someone and want answers to their ethical questions immediately. *See id.* at 586.

54. *See, e.g.*, Elizabeth Chambliss & David B. Wilkins, *A New Framework for Law Firm Discipline,* 16 GEO. J. LEGAL ETHICS 335, 336 (2003) (proposing that the Model Rules require all firms to designate one law firm partner as the compliance specialist to avoid Model Rule of Professional Responsibility 5.1(a) concerns).

55. *See* Chambliss & Wilkins, *supra* note 41, at 580 (noting comments from participants that firm lawyers are more comfortable asking questions of a full-time specialist).

56. *See* Anthony E. Davis, *The Emergence of Law Firm General Counsel and the Challenges Ahead*, 20 PRO. LAW. 1, 1 (2010) (empirical study showing, over a five-year period, that law firms with general counsel spend "$1 million less on defense costs and indemnity payments" associated with malpractice claims).

57. Artificial intelligence has been described as the ability of "computers to imitate cognitive human functions." Clara Piloto, *Artificial Intelligence vs Machine Learning: What's the difference?*, MIT PRO. EDUC. (2024), https://professionalprograms.mit.edu/blog/technology/machine-learning-vs-artificial-intelligence [https://perma.cc/2RXU-QBWJ].

within it, the result will likely reflect that baked-in bias.[58] As Professor Keith Swisher recently noted, "[e]ven though AI counsel in theory will be unbiased, in practice AI could learn and repeat bias from humans," if humans "feed bias into AI counsel."[59] Yet, as AI's source material data set expands, and as AI is subjected to conscientious, periodic testing for bias, unwanted biases will be diluted and perhaps eliminated.[60] Moreover, AI can be trained to flag certain issues, including ethical issues that raise biases or conflicts in values, to ensure human supervision and review.[61]

Even accepting these present challenges, certain implicit reasoning errors and situational challenges apply uniquely to humans; thus, they do not seem likely to be automatically replicated in or applicable to AI.[62] To illustrate this point, we have selected a few common behavioral challenges to decision-making and asked whether an AI decision-maker might be better positioned to avoid the behavioral implications.[63]

---

58. *See* Agnieszka McPeak, *Disruptive Technology & The Ethical Lawyer*, 50 U. TOL. L. REV. 457, 467 n.85 (2019) (noting that "historical data may contain built in biases" thus reliance on the data would perpetuate the bias and discussing examples of AI bias including one where Amazon's AI recruiting tool analyzed ten years' worth of resumes, the vast majority of which were from men, and concluded that women were less qualified in a particular field than men); *see also* IBM Data and AI Team, *Shedding Light on AI Bias with Real-World Examples*, IBM (October 16, 2023), https://www.ibm.com/blog/shedding-light-on-ai-bias-with-real-world-examples [https://perma.cc/N85Z-WERT] (noting that "[i]ndependent research at Carnegie Mellon University in Pittsburgh revealed that Google's online advertising system displayed high-paying positions to males more often than to women" suggesting gender bias was built into the AI algorithms); Amy B. Cyphert, *A Human Being Wrote This Law Review Article: GPT-3 and the Practice of Law*, 55 U.C. DAVIS L. REV. 401, 442-43 (2021) (analyzing AI against Model Rule of Professional Conduct 8.4(g), which provides that it is misconduct for a lawyer to engage in conduct that is discriminatory on the basis of race, sex, gender identity, and socioeconomic status); Julia Klar, *Research - Artificial Intelligence Can Fuel Racial Bias in Health Care, But Can Mitigate It Too*, THE JOURNALIST'S RES. (2022), https://jheor.org/post/1590-research-artificial-intelligence-can-fuel-racial-bias-in-health-care-but-can-mitigate-it-too [https://perma.cc/9LSF-QLQ5] (noting that "[w]hen AI is trained by data that lack diversity, then it is more likely to mimic the same racial bias that healthcare professionals can themselves exhibit").

59. Swisher, *supra* note 4, at 860 (concluding that "[i]n sum, AI counsel in theory could finally be the truly unbiased lawyer, but humans will need to ensure that we do not feed bias into AI counsel").

60. *See* Emilio Ferrara, *Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies*, 6 SCI (Dec. 26, 2023) at 5-6, https://doi.org/10.3390/sci6010003 [https://perma.cc/P3CP-X2QH].

61. McPeak, *supra* note 58, at 467 (noting that judicial AI tools have the potential to "unearth the extra-legal (and perhaps improper) factors that judges might be using" to make decisions).

62. *See* Swisher, *supra* note 4, at 829-30 (noting that AI could potentially avoid human flaws in decision-making such as those grounded in bounded rationality and other cognitive biases).

63. Of course, whether AI effectively reduces human decision-making bias is an empirical question that is ripe for consideration and needs more detailed exploration. *See* Andrew K. Woods, *Robophobia*, 93 U. COLO. L. REV. 51, 59 n.29 (2022) (noting one example of an empirical project demonstrating that human doctors' implicit biases impacted patient diagnosis and resulted in mistakes, in contrast to AI's diagnosis (citing Emma Pierson et al., *An Algorithmic Approach to Reducing Unexplained Pain Disparities in Underserved Populations*, 27 NATURE MED. 136 (2021))); *see also* Sklar, *supra* note 58 (collecting research on AI in the medical profession and suggesting that when properly trained, AI can help mitigate racial bias).

1. MOTIVATED REASONING AND SELF-INTEREST

AI agents likely do not pursue "self-interest," which "sets them apart from traditional human agents for whom self-interest is the primary cause of agency problems."[64] Motivated reasoning and self-interest are indeed key underlying and often unconscious influences on our decision-making processes. They describe the human tendency to seek out and interpret information in a manner that is consistent with our personal wishes, wants, and desires.

Consider, for example, one early and well-known study (appreciated by football fans) that evaluates the influence of partisanship on viewers of a 1951 football game between Dartmouth and Princeton. After watching the exact same game, students from each school were asked to rate the number and severity of infractions. The results were as follows:

> Dartmouth students recorded approximately the same number of infractions between the two teams and did not view their team as responsible for rough play. In contrast, Princeton students attributed twice as many infractions to the Dartmouth team than to the Princeton team and found Dartmouth's fouls to be more flagrant.[65]

The research demonstrated that partisan motivation shaped how the students sought out and interpreted information. And the influences were unconscious—students believed they were objective.[66] Today, numerous studies on motivated reasoning have shown "the many tricks people use to reach the conclusions they want to reach."[67]

Motivated reasoning and self-interest can impact lawyers in various professional settings and at any number of points along the lawyering process. For example, such dynamics can impact lawyering decisions regarding whether information sought in discovery is relevant or privileged; whether a non-waivable conflict of interest exists; and whether information is exculpatory such that prosecutors must turn it over to the defense. Prosecutors, defense attorneys, in-house counsel, negotiators, and lawyers of all experience levels have been shown to be susceptible to the unconscious impacts of motivated reasoning and self-interest.[68] Moreover, a variety of situational pressures frequently operate in law's workplaces, and lawyer decision-making can be influenced by context. In a hierarchical law firm, for example, associates often receive work assignments from senior lawyers, rather than from

---

64. *See* Kolt, *supra* note 14, at 26 (noting that "there is only limited evidence that current AI agents pursue 'self-interest'").

65. O'GRADY & ELDRED, *supra* note 36, at 14.

66. *Id.* (noting also the similar tendency of "biased assimilation," which describes our tendency to be skeptical of data we do not want to believe).

67. JONATHAN HAIDT, THE RIGHTEOUS MIND: WHY GOOD PEOPLE ARE DIVIDED BY POLITICS AND RELIGION 98 (2012).

68. *See generally* O'GRADY & ELDRED, *supra* note 36 (analyzing motivated reasoning and self-interest implications for all these lawyers in a variety of contexts).

clients, and they are evaluated by senior lawyers for retention, raises, and promotion. Thus, in the context of a workplace, new lawyers will be inclined to respond to the human need to project their jobs, establish an excellent reputation, build their salaries, and secure their advancement within the firm. They will also be inclined to avoid personal embarrassment or shame, which makes admitting mistakes, when they occur, difficult.[69]

In thinking about how AI would respond in these and similar situations, we begin with the premise that AI should not have a "stake in the game" or a partisan self-interest in the outcome (unless such an interest was baked in by humans). By default, AI does not value financial resources, pursue self-interests, or care about reputational consequences.[70] With regard to the football game, for example, a well-instructed AI agent could simply count formal infractions and assess the physicality of the game without partisanship, thus acting without motivated reasoning or self-interest. AI would be working with the data—for example, observing the football game—and AI's decision-making posture would carry few, if any, self-serving underpinnings.

Of course, for some assessments, the exercise of discretion may be important in the decision-making process, and discretionary assessment is where human involvement is likely to be important.[71] AI can assist in this process by flagging discretionary areas for human input. Thus, while AI may not be perfectly trained and would likely need human input in some areas, the human tendencies that can lead to motivated reasoning are not present in AI. Ultimately, AI assists decision-making by ensuring that a decision is based on data, not on suboptimal partisan and situational forces, and AI provides a check on the human impulses that might lead to decisions grounded in motivated reasoning or self-interest.

2. CONFIRMATION BIAS

Confirmation bias is a ubiquitous reasoning error that describes a "powerful human tendency which causes us to seek out and interpret information in a manner that is consistent with our pre-existing beliefs" or conclusions.[72] Without being consciously aware of it, we all tend to start a project with a belief and then actively look for evidence that confirms our belief as true, rather than looking for or valuing evidence of its falsity.[73] In the law firm workplace, for example, imagine an associate working on a legal research project under the influence of

---

69. *See* O'Grady, *supra* note 5, at 29-31.

70. Kolt, *supra* note 14, at 26.

71. Current AI is thought to be limited at navigating certain discretionary decisions. *See* W. Bradley Wendel, *The Promise and Limitations of Artificial Intelligence in the Practice of Law*, 72 OKLA. L. REV. 21, 40-41 (2019) (arguing that AI would have "considerable difficulty replicating human decision-making" that requires discretion and the resolution of a number of considerations including black-letter law, interpretation norms, analogical reasoning with multiple sources of legal authority, and policy arguments).

72. *See* O'GRADY & ELDRED, *supra* note 36, at 12.

73. *Id.*

confirmation bias. Without realizing it, the associate will focus research efforts on finding sources that confirm the conclusions already reached and fail to look for or value cases that cut the other way. To make matters worse, the bias will build on itself—as confirming research mounts, each confirming case or authority serves to cement the associate's initial decision. Ultimately, the associate runs out of time or energy and the research project ends with disconfirming cases left unexplored or undervalued.

If given an objective prompt, a properly trained AI agent should not be operating under any sort of presumed narrative or pre-existing belief. For example, a prompt to "conduct legal research on whether for-profit corporations have First Amendment rights to free speech and free association" would provide results from an AI agent on all sides of the question that are free of confirmation bias. Of course, a prompt may be intentionally designed to confirm a desired result, such as "provide a list of cases that support a for-profit corporation's First Amendment rights to free speech and free association." Such a results-oriented prompt is distinguishable from confirmation bias because the prompt is conscious and intentional. It is an explicit request for just one side in a universe of research as opposed to the under-the-radar workings of an unconscious bias operating naively to seek out and value only certain findings.

Moreover, research on confirmation bias has found that a conscious, deliberate effort to consider alternative views and disconfirming evidence does help to counter the effects of confirmation bias.[74] Although it is difficult for humans to consciously fight against confirmation bias,[75] AI can help. An AI agent can be trained at the outset to look for disconfirming evidence as a specific subset of all its work. Thus, with increasingly expansive data sets and conscientiously designed prompts, AI could help guard against this ubiquitous decision-making bias.[76]

---

74. *See* JENNIFER K. ROBBENNOLT & JEAN R. STERNLIGHT, PSYCHOLOGY FOR LAWYERS: UNDERSTANDING THE HUMAN FACTORS IN NEGOTIATION, LITIGATION, AND DECISION MAKING, 35-36 (2012) (citing Charles G. Lord et al., *Considering the Opposite: A Corrective Strategy for Social Judgment*, 47 J. PERSONALITY & SOC. PSYCH. 1231 (1984); Laura J. Kray & Adam D. Galinsky, *The Debiasing Effect of Counterfactual Mind-Sets: Increasing the Search for Disconfirmatory Information in Group Decisions*, 91 ORG. BEHAV. HUM. DECISION PROCESSES 69, 76 (2003)).

75. *See* Masterman et al., *supra* note 26 (discussing how multi-agent architectures and specialized agents enable intelligent division of labor, dynamic team construction, and effective information sharing, leading to greater accuracy and efficiency in task completion).

76. It will be important to ensure that AI ethics counsel is permitted to render objective advice rather than being used by lawyers primarily to confirm whatever decision is best for the firm or its clients. Thus, programming AI to produce counter authority and counter arguments automatically can help achieve that goal. Another approach might be for State Bars to make their presumably objective AI ethics counsel available to lawyers at low or no cost, perhaps incentivizing lawyers to take advantage of the offer by protecting a lawyer from discipline for action taken on the advice of the State Bar's AI ethics counsel. We are grateful to Professor Keith Swisher for this intriguing idea.

B. SPECIALIZED ETHICS AGENTS IN LEGAL WORKFLOWS

To use AI agents at all, they must be competently trained on ethical rules and understand their proper application within workflows. This is no easy task.[77] The ability of AI technology to provide these capabilities is an open question, but there are reasons to be optimistic.[78] Assuming the technology is capable, we believe that *specialized ethics agents* are the best way to provide competent and comprehensive ethical guidance within AI workflows. *Specialized ethics agents* are independent AI agents designed to recognize potential ethics violations and provide appropriate guidance or correction.[79] They can intervene across a workflow to review the work of other agents, flag potential ethical issues, and ensure that all outputs and activities comply with ethical standards. Properly designed, trained, and deployed agents can provide significant benefits to ethical oversight across a firm's activities with a speed and scale unattainable in non-AI workflows.

We do not take a view on exactly *how* specialized ethical agents should come to prevalence. Law firms may adopt specialized ethics agents independently given their compliance and cost benefits. Alternatively, regulatory intervention may end up requiring such an approach, especially in circumstances where human mistakes are costly, less transparent, and where accountability is more difficult to assign. In either case, management and oversight of agentic workflows will present challenges.[80] Specialized agents simplify these challenges and make accountability in agentic workflows easier than alternative approaches that would build ethical accountability directly into *all* agents. As we discuss below, this approach would create added complexities for agent design, accountability and oversight. A model embracing specialized AI agents is a more natural fit as it leverages the benefits of agent specialization and agentic iteration to better ensure ethical compliance.[81]

---

77. Sayash Kapoor et al., *AI Agents That Matter*, PRINCETON UNIV., (July 2, 2024), https://agents.cs.princeton.edu/[https://perma.cc/JC39-M7C4] (elaborating the importance and challenges of benchmarking to improve AI agent performance).

78. Erik Brynjolfsson et al., *Generative AI at Work*, (Nat'l Bureau of Econ. Rsch., Working Paper No. 31161, 2023), https://doi.org/10.3386/w31161 [https://perma.cc/2DN5-WLV2] (noting productivity success of generative for enterprise users currently).

79. This approach contrasts with the *non-specialized approach* that tries to embed ethical knowledge and capabilities into *all agents* that need it.

80. *See generally* Alan Chan et al., *Visibility into AI Agents*, *in* FACCT '24: PROC. OF THE 2024 ACM CONF. ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 958 (2024). The authors discuss measures to increase the visibility of AI agents in systems with minimal direct supervision, including agent identifiers, real-time monitoring, and activity logging.

81. Aditi Singh et al., *Enhancing AI Systems with Agentic Workflows Patterns in Large Language Model*, *in* 2024 IEEE WORLD AI IOT CONG. (AIIoT) (2024), https://www.techrxiv.org/users/796234/articles/1236723-enhancing-ai-systems-with-agentic-workflows-patterns-in-large-language-model?commit=fe0eb80fe3202e80c143c36fc3f8a0b857066eaa [https://archive.ph/Dna05] (describing how agentic workflows improve the application of LLM technology through iteration).

1. ADVANTAGES OF SPECIALIZED ETHICS AGENTS

Like other AI agents, ethical agents benefit from specialization that allows them to operate in dynamic multi-agent workflows.[82] They can be designed and trained precisely, tested extensively and updated easily—all of which improves their performance and makes ethical compliance more transparent.[83] Designers can build focused agents that incorporate ethics-specific knowledge, such as state bar ethics opinions.[84] Further, they can test and train agents specifically on ethics cases to ensure accurate capture and correction of ethical violations.[85] If updates are needed, only a handful of agents require adjustment—not the workflow's entire roster of agents. Each step in creating, testing, and managing agents for ethical guidance is more straightforward and transparent with a specialized approach.

Conversely, trying to embed ethical guidance across non-specialized agents convolutes the process considerably. Certain design options are significantly more difficult, if not entirely off the table, in non-specialized agents.[86] Furthermore, non-ethics agents have to optimize for other tasks, and ethical considerations can be just one of many demands. While these non-ethics agents could be given access to databases with ethics rules, regulations and case studies, they would likely need further refinement to ensure appropriate application of these standards.[87] Testing and training would require curating examples that only apply to the specific set of tasks of the agent in question. If an agent needs to be adjusted to improve performance, one must consider the impact on performance of the agent's core tasks. Instead of a handful of updates exclusively to ethics agents, numerous agents would require complicated and bespoke adjustments.

While AI workflows offer a host of performance benefits, they can quickly become complex and require careful design from system architects.[88] Specialized agents are a simplifying mechanism needed to understand and manage how ethical guidance operates in these workflows. In addition to their advantages for building and developing ethical AI, specialized agents offer several benefits for oversight of and accountability in agentic systems.

---

82. Masterman et al., *supra* note 26, at 6–7 (discussing how multi-agent architectures and specialized agents enable intelligent division of labor, dynamic team construction, and effective information sharing, leading to greater accuracy and efficiency in task completion).

83. *See id.* at 3.

84. *See id.* at 3–4 (describing creation of effective agents using reasoning/planning and effective tool-calling).

85. *See id.* at 10–11 (describing importance of testing and benchmarking agents working beyond trained knowledgebase).

86. Fine-tuning, for instance, would be difficult for an agent with a model primarily designed for a non-ethical task. Prompt engineering might also get complicated or have diminished effectiveness in non-specialized agents.

87. *See* XIAO YANG ET AL., CRAG: COMPREHENSIVE RAG BENCHMARK 2 (2024), https://arxiv.org/abs/2406.04744 [https://perma.cc/RG3U-698T] (noting the poor benchmark performance of standard RAG and the need for more elaborate solutions for knowledge addition).

88. Singh et al., *supra* note 81 (describing how agentic workflows improve performance but require careful design measures—namely reflection, planning, multi-agent collaboration, and tool utilization).

2. ACCOUNTABILITY AND OVERSIGHT OF AI SYSTEMS

Regardless of how AI systems are managed or supervised, specialized agents will ease the burden of overseers by providing a locus of control where standards and best practices can be attached. Regulation of AI is already top of mind and debates about the best approach are well underway.[89] While we do not take an opinion on what the regulations should be or how they should be enforced, we believe the regulatory benefits of specialized agents are numerous. With appropriate adjustments to the ethical infrastructure supporting agentic workflows, law firms can use specialized agents to maintain a high level of accountability and provide more comprehensive ethical guidance than ever before.

As standards for accountability and the use of AI develop, enterprise leaders will need ways to monitor and enforce these standards. Specialized agents provide a fairly transparent way to ensure best practices are being followed.[90] Instead of having to monitor every agent in a workflow, overseers need only examine the specialized ethics agents and understand how they are being employed in broader workflows. Thus, accountability can be largely focused on two distinct sources—(1) the creators of the ethics agents and (2) the designers of the workflows or use cases of those agents. This greatly simplifies effective enforcement. Instead of policing every agent in a workflow and every human user of these agents, oversight can be applied to specific agents and specific individuals to ensure proper AI use in practice.

This approach is a bit different than our current ethical regime—the individual attorney accountability model discussed in Part III(a). An individual accountability approach has been necessary to ensure ethical legal practice, and we expect it to continue to remain necessary in the near term. Significant advancements in AI technology, however, could make it a problematic approach in workflows dominated by AI agents. As agents take on more tasks within a legal workflow and require less human oversight, it becomes unreasonable and impractical to expect individual lawyers to maintain accountability for the work of an AI agent.[91] In these workflows, we can rework accountability toward ethics agents and systems designers. Fortunately, the speed and scale capabilities of AI help facilitate comprehensive accountability in agentic workflows and bring new possibilities to how firms ensure ethical practice.

3. SYSTEMATIC BENEFITS OF AI IN ETHICAL PRACTICE

Specialized AI ethics agents can provide immediate efficiency gains to legal workflows and scale the ethical infrastructure of law firms in ways human ethics

89. *See* Aurelia Tamò-Larrieux et al., *Regulating for Trust: Can Law Establish Trust in Artificial Intelligence?*, 18 REGUL. & GOVERNANCE 780, 780–81 (2024).

90. *See* Chan et al., *supra* note 20, at 961 (noting that agent identifiers and activity logs attached to specific agents can improve visibility into that agent's effect on a system).

91. *See, e.g.*, Swisher, *supra* note 4, at 845.

counsels simply cannot.[92] Unlike human actors, AI agents can provide nearly instant feedback at a virtually unlimited scale. As noted above, human-centric systems require something like individual responsibility to help safeguard ethical practice.[93] A human ethics counsel cannot review every action or advise every lawyer across the firm to ensure ethical behavior, but a properly deployed AI agent can. For the first time, the agentic model allows us to imagine a professional practice of law where rapid ethical guidance is available at all key junctures of the legal process.

Leveraging this speed and scale, law firms can use AI ethics counsel for applications that go well beyond that of traditional law firm ethics counsel. With the ability to offer ethics support through AI agents, smaller firms and sole practitioners could bring an ethics expert into their practice, something they may find cost-prohibitive today.[94] In-house general counsel and compliance officers in corporations and entities, ethics advisors working with ethics hotlines in state bars, and agents working with malpractice insurance carriers could all find immediate uses for an AI ethics counsel. For the law firm ethics counsel discussed in Part III, the AI agents would have an immediate impact that would only expand as agentic workflows developed.

As elaborated in the examples in the next section, AI can provide all firm attorneys with instant consultation. This ensures lawyers get timely ethical guidance no matter how busy the human ethics counsel. As AI workflows develop, agents can provide unprompted ethical reminders precisely when such communications are most salient. Relevant ethical communication that is received close to a lawyer's decision to act can "profoundly influence behavior" and prevent unethical decisions or mistakes.[95] This single example of AI assistance could alleviate significant problems for ethics counsel who struggle to keep attorneys educated on ethical issues and try to impress on firm attorneys the importance of coming to them when facing ethical problems. It may also free up some valuable time for ethics counsel—some of which might be needed for systemic planning and monitoring.

While AI ethics agents will increase the scope and speed of ethical guidance, they will also radically rework the ethical infrastructure of a law firm. Instead of focusing on in-house training, educational emails, and internal best practices,

---

92. For a more in-depth analysis of the benefit of AI over human work, *see generally* Walid L. Bousmaha, *Exploring the Benefits of Artificial Intelligence (AI) in Developing Applications for Humans*, 4 UNITED INT'L J. FOR RSCH. & TECH. 122 (2023).

93. *See supra* notes 35–41 and accompanying text.

94. *See* Shavit et al., *supra* note 18, at 17 ("It is also possible that agentic AI systems can increase the agency and productivity of individual workers or small firms more than traditional AI systems have done, such as by increasing the availability of previously rare expertise.").

95. *See* Wallace J. Mlyniec, *Lawyering Practice: Uncovering Unconscious Influences Before Rather Than After Errors Occur*, 51 NEW ENG. L. REV. 81 (2017); Scott Killingsworth, *Modeling the Message: Communicating Compliance Through Organizational Values and Culture*, 25 GEO. J. LEGAL ETHICS 962 (2012).

ethics counsel might assume a more precise role to ensure adequate deployment and use of ethics agents across a firm's workflows. For instance, instead of educating a firm's marketing team on ethical rules related to advertising or manually reviewing marketing materials for compliance, an ethics counsel would ensure that the appropriate ethics agent is built into that part of the firm's workflow. This agent could automatically flag ethical issues to individuals working on marketing materials and perform ongoing instant reviews before any final works are disseminated. The upfront work to implement these agentic systems and the ethical agents to monitor them will be challenging. The long-term benefits, however, would be immense. A firm could have consistent and comprehensive ethical input, with minimal intervention, to limit the risks of unintentional or intentional unethical behavior.

C. AI AGENTS AS ETHICS COUNSEL—ILLUSTRATIVE EXAMPLES

To illustrate the potential impact of AI ethics counsel, imagine the following ethical questions raised in the practice of law. We expect that AI responses will vary depending on the task at hand. The first set of ethical issues are brought directly to ethics counsel through questions from attorneys. The second set of issues, however, demonstrate how AI ethics counsel can offer guidance on ethical considerations independently, spotting potential ethical issues through the circumstances of a case, without anyone needing to bring a question forward.

1. ETHICAL REMINDERS IN RESPONSE TO QUESTIONS

Some ethical questions are direct, and we can anticipate that AI counsel would provide a straightforward answer. For example, a new associate with a billing question could quickly seek input from AI ethics counsel as follows:

Question New Associate Posed to AI Ethics Counsel: I am about to travel by air on behalf of a client, XYZ Corporation, and I was told to count the travel time as billable hours for XYZ. I'd like to spend several hours on the airplane working on another client's billable matter. Can I bill both clients for the same hours worked?

AI Ethics Counsel Response: No You should not bill both clients for the same hours worked. That is called double billing, and it is unethical. Ethical Rule 1.5 requires that legal fees be "reasonable," and double billing is considered unreasonable. Instead, you should bill the other client only for time spent actually working on that client's matter and bill XYZ corporation for the rest of the travel time, if any. Or you could bill XYZ for all the travel time and either not work on the other client's matter or write off that time if you do. Please follow up with me if you have any questions at all. Have a good trip![96]

---

96. *See What Lawyers Need to Know About Double Billing*, ABA (Oct. 2, 2023), https://www.americanbar.org/groups/law_practice/resources/law-technology-today/2023/what-lawyers-need-to-know-about-double-billing [https//archive.ph/xE79r].

In this example, the ethics advice is received quickly, and the input is timely—it comes right before the new associate's travel. The communication prevented unethical behavior—the associate will not make the mistake of double billing clients.

In many other instances, answers to ethical questions require the use of discretion and professional judgment. In those instances, AI can provide immediate guidance and flag the areas that require a human lawyer's professional discretion. For example:

Question Lawyer Posed to AI Ethics Counsel: My client, the CEO for XYZ Corporation, just told me in confidence that one of their employees accidentally dumped toxic waste in the desert outside of Smithtown. I am concerned that this dumping polluted the water table that supplies drinking water to residents of Smithtown. I know this is confidential client information, but should I tell someone?

AI Ethics Counsel Response: Perhaps. [The State's] Rule of Professional Conduct 1.6(b)(1) permits you to disclose confidential information if you "reasonably believe" that disclosure is "necessary . . . to prevent reasonably certain death or substantial bodily harm."[97] In this jurisdiction, disclosure is not *required* under these circumstances; rather, the lawyer has *discretion* to make the disclosure. Thus, if you reasonably believe that disclosure is necessary to prevent reasonably certain death or substantial bodily harm, you may disclose the confidential information to the officials of Smithtown without violating your ethical obligations.

To assist you in making this assessment, I have attached data on the risks of harm caused by the sort of toxins that XYZ Corporation would produce. I have also attached current information about the water table that supplies drinking water to Smithtown. As you can see from the data, the water table is already slightly polluted. You will need to assess whether the additional toxins increase the risk to the point of causing "reasonably certain death or substantial bodily harm." For example, if you conclude that some vulnerable people, such as elderly citizens and infants, will likely get a short-term, uncomfortable stomach illness, that likely does not rise to the level of "substantial bodily harm" that is required to trigger the exception to confidentiality. I have consulted with one of our firm's Relevant Law agents, and I attached case law the agent provided to me from this jurisdiction that analyzes "substantial bodily harm."[98] Please note that the client should generally be consulted before making any disclosure. Let me know if you would like any additional information as you make this determination.[99]

---

97. *See, e.g.*, MODEL RULES OF PROF'L CONDUCT R.1.6 (b) (1) (2018).

98. *See supra* Section II (identifying possible agents operating in the workflow of an agentic law firm).

99. AI ethics counsel might also address or inform the lawyer to investigate further whether the dumping, even if accidental, or the client's failure to report it, is criminal in that jurisdiction.

2. INDEPENDENT AND TIMELY ETHICAL REMINDERS

In the two examples above, a lawyer sought answers to questions from AI Ethics Counsel. This is consistent with the human model and use of ethics counsel in firms—many ethics counsel make themselves available to the firm's attorneys when they have ethics questions (and, as noted above, attorneys with ethical dilemmas understandably hope for rapid responses to their questions). Unlike the human model, however, AI ethics counsel can operate independently to accomplish goals and will be in attendance at all junctures of the lawyering process.[100] Thus, unlike human ethics counsel, AI ethics counsel can interject at any time in the workflow to flag ethical issues in advance for the lawyers, perhaps by pinging lawyers on their laptops with notifications and ethical reminders. For example, in the toxic waste hypothetical, AI ethics counsel will be in attendance during the meeting with the CEO of XYZ Corporation. AI ethics counsel can immediately flag the ethical issues for the attorney's consideration by sending the attorney a message similar to this:

Message From AI Ethics Counsel to Lawyer:

Ethical Reminder. In a recent meeting, Mr. CEO said that XYZ Corporation accidentally dumped unpermitted toxic waste into the desert just outside of Smithtown. Ethical Rule 1.6(b)(1) . . . ."

In addition, and perhaps most importantly, in some circumstances, AI ethics counsel can timely flag ethical considerations *prior* to a lawyering event, thereby preventing mistakes and unethical behavior. For example:

Message From AI Ethics Counsel to Lawyer:

Ethical Reminder. I see from a calendar review that you have a negotiation coming up next week. Prior to your negotiation, you may wish to review [the State's] Rule of Professional Conduct 4.1, which sets a minimum standard for ethical conduct during negotiations. In addition to other information, Rule 4.1 provides that "puffing" about price or a bottom-line settlement number is permissible but lies about the subject matter of the negotiation are unethical. Please let me know if you have any questions or would like additional information.

Or

Message from AI Ethics Counsel Agent to Lawyer:

Ethical Reminder. I see from a calendar review that you have set up a meeting next week with George Smith, a former employee of ABC, Inc. ABC is the represented defendant in a lawsuit brought by our client, Pamela Jones. Please review Rule of Professional Conduct 4.2, the so-called "no contact" rule, and related case law to see if you may contact the former employee, Mr. Smith, directly or if you should schedule this meeting through ABC's attorney. Even though Smith is a former employee, he may still be covered under the no-contact rule. I have

---

100. *See supra* Section II; Kolt, *supra* note 14, at 4 (describing AI agents as "autopilots" and "actors" that can "independently take actions to accomplish complex goals on behalf of users").

provided the case law in this state that governs this situation. Please let me know if you have any questions or would like additional information.

As workflows incorporate more AI agents and become more agentic, these reminders may be given, not to humans, but to other agents. Working with the adaptive and iterative workflows described in Part II of this paper, ethical agents will provide inputs or feedback to other agents whose work might implicate or violate ethical guidelines. For example, if an agent were tasked with conducting the interview of George Smith above, it would be prompted in advance by an ethical agent to contact ABC's attorney if the lawsuit by client Pamela Jones required this.

CONCLUSION

The use of AI in law firms presents considerable challenges, but it also presents considerable benefits. This article contends that generative AI opens up new opportunities for the ethical practice of law and posits that *specialized AI ethics agents* offer considerable advantages over alternative approaches. A model integrating specialized AI agents as ethics counsel in a law practice's workflow is a uniquely efficient and scalable approach to improving ethical decision-making and compliance in American legal practice.