Tech® REG CHRONICLE

AGENTIC ARTIFICIAL INTELLIGENCE

JUNE 2025

CPI COMPETITION POLICY® INTERNATIONAL

Competition Policy International, a What's Next Media and Analytics Company

LETTER FROM THE EDITOR

Dear Readers,

This edition of the Chronicle takes a deep dive into the evolving universe of agentic AI—those autonomous systems that now carry out tasks, make decisions, and reshape the contours of market behavior, legal liability, and social interaction. As we begin to understand the scale of this transformation, our authors dissect its implications from legal, technological, and policy standpoints.

Christopher A. Suarez, **Lee Berger**, **Ronan Scanlan** & **Rachel Carlo** open by examining how the proliferation of agentic AI could reshape competitive dynamics across jurisdictions. They flag the risk of monopolization, covert coordination, and standard-setting abuse, arguing that the interoperability of AI systems could both promote and hinder fair competition depending on how it is governed.

Wulf Kaal proposes a novel oversight regime rooted in web3 architecture. By embracing blockchain, smart contracts, federated systems & reputation tokens, he outlines a decentralized governance structure capable of adapting in real time to the evolving behavior of AI agents, thus overcoming many of the failures of traditional regulatory models.

CPI COMPETITION POLICY® INTERNATIONAL

Sara Fish, **Yannai A. Gonczarowski** & **Ran I. Shorrer** turn to the nuanced distinction between helpful coordination and illicit collusion among AI agents. As autonomous agents interact across markets, they warn that society must design systems that avoid both chaotic "collisions" and tacit collusion—without undermining legitimate cooperation.

Nick Peterson & **Joel S. Nolette** interrogate the fragility of criminal law in the face of AI autonomy. With scienter as a cornerstone of criminal culpability, they argue that courts and prosecutors will face increasing difficulty in assigning intent, perhaps pushing enforcement efforts into the civil or regulatory sphere instead.

Sarah Oh Lam & **Nathaniel Lovin** ask whether agentic AI requires an entirely new policy architecture or whether existing regimes can be adapted. They weigh emerging issues such as prompt injection attacks, proprietary model dominance, algorithmic convergence & sector-specific disruptions, concluding that legal systems must become more dynamic to remain effective.

Finally, **Kevin Frazier** explores the social and political consequences of mass adoption. He makes the case for commodifying AI agents as accessible digital extensions of ourselves, but cautions that access, competition, and fairness will depend on deliberate regulatory choices to avoid entrenching new forms of inequality.

We hope these contributions provoke new thinking, critical engagement & productive debate.

As always, many thanks to our great panel of authors.

Sincerely,

CPI Team

TABLE OF CONTENTS

CPI COMPETITION POLICY® INTERNATIONAL

AGENTIC **ARTIFICIAL INTELLIGENCE**

JUNE 2025

SUMMARIES

AGENTIC AI: FUTURE ISSUES AT THE INTERSECTION OF TECHNOLOGY, INNOVATION, AND COMPETITION POLICY

By Christopher A. Suarez, Lee Berger, Ronan Scanlan & Rachel Carlo

Agentic AI presents amazing opportunities that will streamline and automate numerous tasks, and numerous agentic AI models and systems have already emerged. As agentic AI models and systems proliferate, however, new harms to competition might emerge, and practitioners should be aware of them. To that end, this article describes several threats to competition that could materialize in the context of agentic AI, including risks that a dominant AI agent provider or developer monopolizes the industry, that AI agents are used to facilitate collusive data collection of pricing or similar commercial data, and that AI agents could themselves enter into anticompetitive arrangements without human oversight. All of these risks are apparent in the United States, and extend throughout the world to the UK, EU, and beyond. Practitioners should also be aware of the interplay between fair competition and the interoperability of AI agents so that they can talk to each other. While interoperability is highly desirable, with interoperability comes risks that certain technical standards emerge that have IP, standard essential patent, or software licensing obligations that might have positive or negative effects on fair competition.

HOW CAN WE BEST MONITOR AI AGENTS?

By Wulf Kaal

This paper examines the critical challenge of monitoring AI agent transaction execution within decentralized digital ecosystems, highlighting the deficiencies of traditional centralized AI-driven supervision, including opacity, bias, and systemic vulnerabilities. In response, it proposes a web3 Decentralized Autonomous Organization (DAO)-centric governance model that integrates blockchain technology, federated communication platforms, smart contracts, and Weighted Directed Acyclic Graphs (WDAGs) to deliver an alternative oversight framework. The proposed system ensures unparalleled transparency and accountability through blockchain's immutable ledger, while decentralized decision-making via community consensus mitigates bias and single points of failure. Federated platforms enhance scalability and privacy by distributing data processing, and smart contracts automate real-time compliance, bolstered by WDAGs' adaptive governance structure. Validation pools and reputation tokens further empower stakeholders, fostering a dynamic, inclusive monitoring process. By incorporating feedback loops, this model anticipates and adapts to AI evolution, overcoming scalability, interoperability, and regulatory gaps inherent in existing frameworks. This decentralized approach not only addresses current shortcomings but also establishes a forward-looking standard for secure, compliant, and efficient AI agent management in modern infrastructures.

NAVIGATING THE SPACE BETWEEN COLLISION AND COLLUSION IN AN ECONOMY OF AI AGENTS

By Sara Fish, Yannai A. Gonczarowski & Ran I. Shorrer

As more tasks are delegated to AI agents, and they are granted greater autonomy, interactions among these agents are becoming inevitable. This raises concerns not only about the alignment of each agent's actions with its own user's goals, but also about the alignment of the broader multi-AI-agent ecosystem with societal goals. Effective collaboration among agents requires coordination, but in some settings, coordination may have detrimental consequences. We discuss the tension between the need to coordinate AI agents ("avoiding collision") and the goal of preventing undesirable coordination ("avoiding collusion"). Challenges include the detection, attribution, and correction of collusion. We argue that progress will require cooperation between stakeholders (industry, government) and researchers.

AGENTIC AI AND THE LOOMING PROBLEM OF CRIMINAL SCIENTER

By Nick Peterson & Joel S. Nolette

Federal enforcement authorities have expressed confidence that existing legal authorities are adequate in the face of emergent artificial-intelligence ("AI") technology. But particularly as agentic AI enters the mainstream, this proposition seems in question, especially so in the criminal-law context, where scienter provides the dividing line between innocent and wrongful conduct in most cases. For this reason, prosecutors and courts may struggle to find individuals and companies criminally culpable when an AI agent commits the misconduct. The law and society will eventually catch up to these developments, but it will take time. Legislatures may step up to fill this gap, but we may also see prosecutors turn more to civil statutes to address misconduct.

DOES AGENTIC AI REQUIRE NEW POLICY FRAMEWORKS?

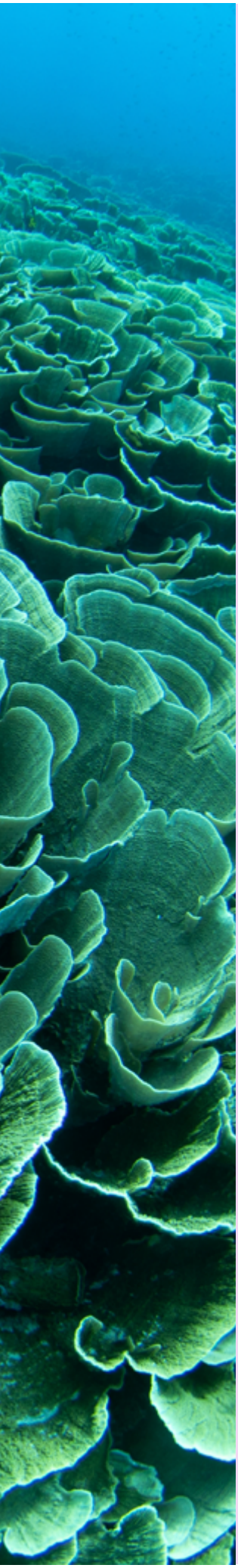
By Sarah Oh Lam & Nathaniel Lovin

Does the emergence of agentic AI — systems capable of autonomous interaction and decision-making — require entirely new policy frameworks, or can existing frameworks developed for earlier systems effectively address emerging challenges? This analysis uniquely examines how reduced human oversight in autonomous agent networks creates new policy challenges particularly when agents collaborate without real-time human intervention. The article considers the debate over open source versus proprietary models, where high compute costs and corporate investment distinguish AI from traditional software. It examines the distinctions between general-purpose and specific-purpose AI. Cybersecurity threats, like prompt injection, increase the risks of data exfiltration, which may necessitate legal clarity on liability, especially in agentic AI systems. Competition policy issues arise as incumbent firms integrate AI services, potentially stifling startups if large firms engage in anticompetitive practices. At the same time, autonomous agents employed by both small and large firms could unintentionally facilitate algorithmic collusion by independently converging on pricing strategies that may cause antitrust injury. Workforce effects of agentic AI may lead to uneven impacts, with benefits for some and displacement for others, requiring reimagined workplace roles. Agentic AI's widespread adoption across industries highlights tradeoffs between fostering innovation and ensuring security.

THE CASE FOR COMMODIFICATION: AI AGENTS

By Kevin Frazier

Artificial intelligence ("AI") agents are poised to revolutionize daily life by autonomously completing complex tasks, fundamentally altering human-computer interaction. Unlike earlier AI, agents offer independent action, memory retention, and preference learning, becoming digital extensions that free up intellectual and emotional bandwidth. However, this transformation brings inevitable social turbulence, including job displacement and recalibration of social hierarchies. Despite these legitimate concerns and questions, a broader historical perspective invites us to consider the transformative potential that lies beyond this initial period of social recalibration and institutional resistance. By transferring the burden of time-intensive, emotionally-draining, and physically-taxing responsibilities to these digital counterparts, Americans gain the opportunity to redirect their energies toward more fulfilling pursuits. Realizing this future demands addressing hurdles to mass adoption, such as insufficient competition, which could exacerbate existing societal stratification. The commodification of AI agents, similar to smartphones, is crucial for equitable access, but faces barriers including corporate incentives, limited data portability, lack of interoperability, high entry costs, and regulatory imbalances. Overcoming these requires deliberate policy interventions, fostering competition, and context-sensitive regulation to ensure AI agents become a broadly shared foundation for collective advancement.



AGENTIC AI: FUTURE ISSUES AT THE INTERSECTION OF TECHNOLOGY, INNOVATION, AND COMPETITION POLICY

BY CHRISTOPHER A. SUAREZ

& LEE BERGER

& RONAN SCANLAN

& RACHEL CARLO

Christopher A. Suarez, Lee F. Berger, and Ronan Scanlan are Partners at Steptoe LLP; Rachel Carlo is an Associate at Steptoe LLP. The views expressed herein reflect those of the authors and not of Steptoe LLP or its clients.

While the antitrust community continues to evaluate the market dynamics and regulatory challenges surrounding generative AI, the advent of agentic AI ushers in a new frontier of complexity at the intersection of technology, innovation, and competition policy.

To start with, what is "agentic AI?" Agentic AI refers to "an artificial intelligence system that can accomplish a specific goal with limited supervision. It consists of AI agents — machine learning models that mimic human decision-making to solve problems in real time."[2] The promise of agentic AI is that it does not just

2 Cole Stryker, *What is agentic AI?,* IBM Think, https://www.ibm.com/think/topics/agentic-ai (last visited June 5, 2025).

use AI to make predictions or generate text or other content in response to algorithms or manual prompting, but that it accomplishes more complex tasks autonomously or semi-autonomously. Agentic AI can be used to create and purchase groceries for an Italian dinner,[3] book hotel reservations,[4] or plan an entire vacation with only high-level human inputs.[5] To complete these tasks and others, AI agents are adept at collecting data to make decisions, as an "agent" of a human for whom it is trying to carry out the tasks. AI agents can thus be thought of as somewhat of a cross between a webcrawler, a data aggregator, and a search engine that can carry out particularly defined tasks. Enterprises are increasingly utilizing embedded agentic AI both to perform customer-facing tasks such as handling customer inquiries, resolving issues with online orders, and providing personalized sales suggestions and support, but also more complex tasks around demand forecasting and inventory or supply chain optimization, dynamic pricing, and cybersecurity threat detection and response. The technology unlocks exciting opportunities. At the same time, agentic AI creates potential risks for businesses in navigating competition rules.

01
RISK OF A DOMINANT AI AGENT PROVIDER

Currently, there are numerous players in the market that *provide* AI agents. The agentic AI market remains in its nascent stages, and numerous vendors and platforms have emerged, with no single firm currently holding a dominant share. Nevertheless, observers are closely monitoring the development of this market to determine whether it will replicate certain anticompetitive trends seen in the traditional search market.

The Department of Justice's seminal case against Google Search illustrates how an emerging digital market can evolve into a highly concentrated one. The DOJ secured a landmark victory against Google when Judge Mehta ruled that the company had unlawfully maintained its monopoly over the search engine market. The court found that Google engaged in anticompetitive conduct by spending billions annually on agreements with device manufacturers to ensure that Google remained the default search engine on new mobile devices.[6]

The remedies phase of the trial concluded on May 30, 2025, during which the DOJ expanded its scope to include AI. The DOJ argued that "[w]e are at an inflection point . . . it is [] critical to extend remedies to GenAI to allow this new technology to rise or fall outside the shadow of Google's search monopoly." The DOJ also made specific mentions of the future of agentic AI. Citing Google's internal documents the DOJ warned, "Google plans to make Gemini the 'primary agent' in Chrome, prioritizing its integration over Gemini's rivals and building on Google's history of making switching the default difficult in Android."[7] However, Judge Mehta seemed skeptical of the argument that the search engine case's remedies might extend to Google's generative AI or AI agents, stating that "[i]t seems to me you now want to kind of bring this other technology into the definition of general search engine markets that I am not sure quite fits." Judge Mehta's remedies decision is expected in August 2025.

In the context of agentic AI, there are not yet any known examples where a company has anticompetitively "bundled" an AI agent with other products or required a particular agent as a "default." Nonetheless, one could imagine scenarios where "AI agents" are subject to anticompetitive agreements that solely direct the human users to particular companies' products or services, perhaps excluding competitors from consideration, and do not seek to objectively determine what might be best for the person who is trying to find or utilize a tool or resource within a particular industry. If such agreements are used pervasively enough, one particular AI agent could become dominant, reducing competition that can spur innovation, control pricing, or create optionality.

3 Steven Melendez, *Instacart users can now plan meals using AI, Fast Company*, Tech (May 31, 2023), https://www.fastcompany.com/90903040/instacart-users-can-now-plan-meals-using-ai.

4 Lance Ulanoff, *Two AI chatbots speaking to each other in their own special language is the last thing we need*, MSN TechRadar (Feb. 25, 2025), https://www.techradar.com/computing/artificial-intelligence/two-ai-chatbots-speaking-to-each-other-in-their-own-special-language-is-the-last-thing-we-need.

5 Trevor Laurence Jockims, *From Google to Expedia, AI travel agents planning future trip far beyond 'assistant' status, Technology Executive Council* (May 16, 2025), https://www.cnbc.com/2025/05/16/ai-travel-agents-planning-future-trip-far-beyond-assistant-status.html.

6 *United States v. Google LLC*, No. 1:20-cv-03010 (D.D.C. 2020). This liability finding is subject to appeal.

7 Pls.' Remedies Proposed Findings of Fact, 133, *United States v. Google LLC*, No. 1:20-cv-03010 (D.D.C. 2020), ECF No. 1370.

02

RISK OF COLLUSIVE DATA COLLECTION AND COORDINATION

Aside from monopolization concerns, the agentic AI market may also be susceptible to collusion under a hub-and-spoke theory. In a hub and spoke conspiracy, there is a central conspirator (the hub) that reaches separate agreements with competitors (the spokes), which facilitate an explicit or implied agreement among the competitors (the rim connecting the spokes) to agree to some anticompetitive terms through the hub, such as fixing prices, limiting supply, allocating markets, or exchanging competitively sensitive information. Traditionally, a "rimless" wheel – one without an agreement among the competitor-spokes – does not violate antitrust laws because there is no agreement among the competitors. However, as the agentic AI market evolves, complex legal questions likely will emerge concerning the interpretation of "intent" and "agreement" in contexts where AI agents operate autonomously.

For example, with the use of "AI agents," competitors using the same AI agent might end up exchanging competitive pricing information and using that information to set prices, without even knowing it. For example, if five widget manufacturers all ask the same AI agent to "consider my costs and capacity, survey the market for widget prices, demands, and other information and recommend the most profitable price for my widget," and the AI agent uses information provided by each widget manufacturer to recommend prices, the widget competitors may rely on their AI agents' price recommendations without knowing that the AI agent is using the competitors' information to make that recommendation, or even that the agent is instructing each competitor to charge the same price.

This is perhaps what motivated the DOJ and several attorneys general when they banded together to sue RealPage. RealPage's product allegedly draws nonpublic pricing information from competing landlords and uses that information in connection with AI-based pricing algorithms to recommend rental prices to those same landlords.[8] The DOJ asserts that, "[b]y feeding data into a sophisticated algorithm powered by artificial intelligence, RealPage has found a modern way to violate a century-old law through systemic coordination of rental housing prices – undermining competition and fairness for consumers in the process. Training a machine to break the law is still breaking the law."[9]

While the RealPage litigation was only filed in 2024 and motions to dismiss are pending, one could imagine other scenarios involving agentic AI that trigger similar concerns. For example, one could imagine a service that uses AI agents' work on behalf of industry competitors to login and/or collect data from various websites to gather pricing or other information that is behind a paywall or login portal so as to aggregate that information across retail sectors in an anticompetitive way. Increased use of agentic AI will only incentivize and encourage the collection of more data — both for purposes of training the AI models themselves, and for purposes of carrying out the relevant agentic tasks. Indeed, some have said that agentic AI will transform the very way that databases are designed and organized, and that "[w]ith the rise of AI agents, the database has become an even more critical application layer"[10] — the very structure of those databases may become subject to antitrust scrutiny in the future.

03

AGENTIC AI THAT ITSELF FACILITATES ANTICOMPETITIVE AGREEMENTS OR ARRANGEMENTS

In the above scenarios, AI agents might be used to anticompetitively gather information at a hub, or be subject to certain human-based agreements that anticompetitively constrain the AI agent's outputs or actions. But what if agentic AI is used to facilitate future contracting or agreements in an anticompetitive manner? At a recent technology event, AI agents were on both sides of a transaction — an AI agent called into a hotel to book a reservation, and

8 Press Release, Dept. of Justice, Justice Department Sues RealPage for Algorithmic Pricing Scheme that Harms Millions of American Renters (Aug. 23, 2024), https://www.justice.gov/archives/opa/pr/justice-department-sues-realpage-algorithmic-pricing-scheme-harms-millions-american-renters.

9 *Ibid.*

10 Andrew Davidson, *The Rise Of AI Agents And The Future Of Data*, Forbes Technology Council (Mar. 6, 2025), https://www.forbes.com/councils/forbestechcouncil/2025/03/06/the-rise-of-ai-agents-and-the-future-of-data/.

an AI agent from the hotel took the call and facilitated the booking of a reservation.[11] Therefore, scenarios involving direct communications between AI agents to bind human activity are not just hypothetical — they are real. If agentic AI is left to its own devices, one can imagine future, more complex scenarios where AI agents are used to negotiate and bind businesses into contracts. Perhaps those contracts or a series of contracts AI agents negotiate include anticompetitive exclusive dealing or tying arrangements, and could themselves be the basis for antitrust or competition concerns. This underscores that, quite apart from considering how AI agents collect data and use it to perform tasks, organizations that deploy agentic AI will need to consider how to ensure that AI agents do not run amok and commit potential antitrust violations through the transactions or agreements that they undertake or facilitate. Such issues might need to be addressed using both technical controls and legal measures. These sorts of concerns are why many agentic AI systems being deployed today require a human in the loop to ensure that there are checkpoints before a human adopts any decision or transaction that the AI agent recommends.[12]

04

THE ABOVE RISKS EXTEND BEYOND THE UNITED STATES

There are also lessons from the UK and EU when assessing the risks from an antitrust perspective of advancements in agentic AI functionality, where the underlying models rely on competitively sensitive data inputted from competing businesses, and the outputs contribute to decisions or recommendations on competitive parameters to be adopted by those businesses, including on pricing or other factors. Such scenarios are increasingly common where industry-wide data is already collated to some degree and many software providers are now looking to augment existing datasets with additional AI-powered tools which help companies make strategic commercial decisions, including in relation to dynamic and discriminatory pricing.

The United Kingdom has, like the United States, scrutinized so-called "hub-and-spoke" arrangements, or "A to B to C" infringements, whereby businesses will provide commercially sensitive information (for example in relation to costs and future pricing intentions) to a third party (whether in the context of a distribution system or for other purposes such as data aggregation) and that third party passes on that information to other competing businesses. In cases relating to the supply of replica football kits, children's toys, and dairy products (Cases CA98/06/2003, CA98/8/2003, and CA/03/2011) the UK Competition & Markets Authority (and its predecessor organizations) found that information sharing had facilitated a price-fixing cartel implicating both the first business, the intermediary, and the recipient of that information as guilty of anticompetitive conduct. However, a defense in such scenarios is where there was no knowledge by Company A that Company B would pass on its information, in such a scenario, only Company B and C are liable. Thus, agentic models which act as modern-day hub-and-spoke arrangements (passing information from one competitor to another) could lead to antitrust scrutiny.

In a related but separate vein, in the European Union, there has long been an acknowledgement that service providers and parties that enable cartel activity (but do not directly participate in it) can also be liable under Article 101(1) of the Treaty on the Functioning of the European Union. In Case C-194/14 P *AC Treuhand v. Commission*, the European Courts examined and ultimately upheld a Commission fine in relation to a cartel involving heat stabilizers, which had found a third party not directly involved in the cartel to be nonetheless liable with other members of the cartel on account of its role as a "cartel facilitator." Similarly, in Case C-39/18P, *Commission v. Icap Management Services*, the Commission fined ICAP $7.9 million for taking part in several yen interest rate derivatives cartels, despite not itself being active on the affected markets. These cases serve as a warning to service providers that have a supporting role in enabling cartel conduct that they may face hefty fines and should take steps to mitigate the risk of – perhaps inadvertently – being party to anticompetitive information sharing and collusive conduct. Indeed the subsequent case law involving *AC Treuhand* confirmed the possibility that enforcers could presume that third parties involved in data aggregation could have knowledge of the illegal use of their data (a scenario of close if not equivalent application to providers of Agentic AI models).

11 Lance Ulanoff, *Two AI chatbots speaking to each other in their own special language is the last thing we need*, MSN TechRadar (Feb. 25, 2025), https://www.techradar.com/computing/artificial-intelligence/two-ai-chatbots-speaking-to-each-other-in-their-own-special-language-is-the-last-thing-we-need.

12 Cobus Greyling, *AI Agents With Human In The Loop*, Medium (Aug 15, 2024), https://cobusgreyling.medium.com/ai-agents-with-human-in-the-loop-f910d0c0384b.

05
AGENTIC AI, INTEROPERABILITY, AND STANDARDS-ESSENTIALITY

As the thicket of AI agents and their applications continues to expand and proliferate, other interesting questions might arise relating to interoperability and standards-essentiality. As AI agents are created for increasingly narrow applications, those agents might need to talk to one another to get the job done. But to talk to one another, there is a need for technical standards that ensure that various AI agents are able to communicate with one another. Several standards-based protocols, including the Model Context Protocol, Agent Communication Protocol, and Agent2Agent, have already emerged that can be used to facilitate interconnection and communication between agentic AI systems.[13] Some have characterized these new standards as the next wave of the HTTP protocol to access webpages — but for AI agents.[14] On the one hand, the use of standards allows interoperability, which can promote innovation and creativity amongst larger agentic AI systems and can ensure that no agentic AI system is siloed. On the other hand, if only one standard or a small number of standards dominate, the use of a particular agentic AI interoperability standard might become essential to compete in the industry. Should the industry become locked-in to one or more such standards, the owners of the IP rights to the standardization protocol — whether the standard is protected by copyrights, trade secrets, or both — might have the ability to constrain use or adoption of the standard. These constraints could take the form of licensing requirements. Even if the interoperability standard is open source, proprietary interoperability standards could still emerge that require paid licenses, and third parties could claim patent or other IP rights to the standard. Licensing terms for standard-essential patents are sometimes required to be FRAND — fair, reasonable, and non-discriminatory — to promote innovation, interoperability, and fair competition.

Additionally, should a dominant interoperability standard emerge, the developers or adopters of that standard could take anticompetitive actions to limit competition. For example, they might foreclose certain competitors from being able to plug into their standard, or they might take steps to prevent other competing standards from being usable with their standard. Consider, for example, a universal power adapter. The anticompetitive developer of the standard might make it impossible for certain devices to plug-in to the adapter or decide to exclude certain plug-types from the face of the adapter. This question of whether there should be a "universal adapter" for all agentic AI tools is an important one.[15]

Organizations that build upon agentic AI interoperability standards may also undertake anticompetitive tactics, even if the underlying standards are "open source." For example, antitrust litigation emerged between CoKinetic and Panasonic over alleged misuses of "open source" software for airline entertainment systems, where the underlying open source LGPL license used by Panasonic required public disclosure of Panasonic's software, yet Panasonic allegedly refused to make such disclosures. The purpose of such copyleft open-source licensing regimes is to ensure wide adoption of the software and promote open competition.

06
CONCLUSION

As Agentic AI models become more sophisticated and popular with businesses seeking reliable recommendations on future price setting, and other competitive parameters, there are risks not only for the businesses setting prices, but also the companies developing the underlying models. A key question in such scenarios is the extent to which it was foreseeable that the models could facilitate anticompetitive conduct, whether there was otherwise a legitimate purpose in sharing the data and critically the extent to which there was an intention to engage in anticompetitive conduct. Providers of such models should consider steps in mitigation; for example, ensuring that data on individual company practice is not available to other companies, including health warnings and even going so far as to engineer models so as to avoid scenarios where they result in alignment of multiple companies on pricing. Moreover, practitioners should be on the lookout for how agentic AI might be implemented in ways that enhance or detract from interoperability, including how compliance with FRAND licensing requirements (for standards essential patents) or copyright licensing regimes could enhance free and fair competition.

13 Grant Gross, MCP, ACP, and Agent2Agent set standards for scalable AI results Feature, CIO (May 22, 2025), https://www.cio.com/article/3991302/ai-protocols-set-standards-for-scalable-results.html.

14 *What Is the A2A (Agent2Agent) Protocol and How It Works*, Descope (Apr. 14, 2025), https://www.descope.com/learn/post/a2a.

15 See e.g. Paul Johngraham, *Universal Adapter for Agentic AI, Substack* (Apr. 12, 2025), https://allthingsdelivered.substack.com/p/universal-adapter-for-agentic-ai.

HOW CAN WE **BEST MONITOR AI AGENTS?**

BY WULF KAAL

Professor of Law. The author is grateful for excellent research assistance from Michael Bernardi, research assistant.

01 INTRODUCTION

The rapid proliferation of artificial intelligence ("AI") has ushered in a new era of autonomous digital agents ("AI agents") executing complex transactions across diverse platforms. As these AI agents become integral to financial and operational systems, effective transaction monitoring emerges as a critical challenge. Traditional centralized AI-driven supervision suffers from limited transparency, bias, and single points of failure, necessitating alternative oversight frameworks.

This paper proposes a DAO-centric governance system integrating web3 community software with federated communication platforms, ensuring transparency, dynamic adaption to emerging standards, and enabling independent verification of AI activities.[2] The web3 framework facilitates consensus-driven decision-making among diverse stakeholders, while decentralized governance reduces systemic biases and risks inherent in centralized oversight, together ensuring a resilient and aligned monitoring process.[3] Federated communication distributes data processing across multiple nodes, enhancing scalability and privacy through local processing.[4] Smart contracts automate compliance checks and performance validations while Weighted Directed Acyclic Graphs ("WDAGs") provide structured oversight.[5]

This web3 DAO-centric governance model overcomes the inherent limitations of centralized, AI-driven monitoring approaches while establishing a robust, transparent, and adaptive framework for overseeing AI agent transactions. By aligning technological oversight with community values and regulatory mandates, this model sets new standards for secure, compliant, and efficient AI system management in modern digital infrastructures.

A. AI Agents

AI agents are autonomous software entities designed to execute specific tasks on behalf of their users, representing a paradigm shift in autonomous computing that integrates sophisticated data processing, adaptive decision-making, and seamless system interoperability to serve user needs across diverse domains. Their deployment spans applications from financial trading systems to personalized consumer services, reflecting their versatility and transformative potential. Their ongoing evolution promises to further redefine the boundaries of AI in practical applications.

AI agents leverage advanced algorithms, machine learning methodologies, and natural language processing capabilities to emulate human-like decision-making processes. Their operational framework begins with interpreting user directives through textual or vocal inputs, or through analyses of behavioral patterns, ensuring actions align with user expectations.[6] They aggregate and process data from heterogeneous sources, including real-time market feeds, user profiles, and historical transaction repositories, using machine learning models to extract actionable insights and identify patterns that inform decision-making.[7]

"This paper proposes a DAO-centric governance system integrating web3 community software with federated communication platforms, ensuring transparency, dynamic adaption to emerging standards, and enabling independent verification of AI activities

The agentic decision-making apparatus oscillates between predefined rules and adaptive responses derived from learned patterns. This adaptive capacity is critical in dynamic environments where conditions fluctuate rapidly, with predictive models optimizing outcomes across various domains.[8] AI agents interface with external systems through application programming interfaces ("APIs") for secure, authenticated communication and precise transaction execution.[9] AI agents operationalize decisions by executing transactions — from trading securities to managing e-commerce purchases — serving as intermediaries between computa-

2 See e.g. Haoxiang Luo et al., *A Perspective of Trusted Artificial Intelligence When Blockchain Meets Large Language Models*, 599 Neurocomputing 1 (2024), https://doi.org/10.1016/j.neucom.2024.128089.

3 See e.g. Zhonghua Zhang et al., *Recent Advances in Blockchain and Artificial Intelligence Integration: Feasibility Analysis, Research Issues, Applications, Challenges, and Future Work*, 2021 Sec. & Comm. Networks 1, 4-5 (2021), https://doi.org/10.1155/2021/9991535.

4 See e.g. Sandner et al., *Convergence of Blockchain, IoT, and AI*, 3 Frontiers Blockchain, at 3-5 (2020), https://doi.org/10.3389/fbloc.2020.522600.

5 See e.g. Nitine Rane et al., *Blockchain and Artificial Intelligence (AI) Integration for Revolutionizing Security and Transparency in Finance* 1, 10-11 (SSRN, Dec. 2023), http://dx.doi.org/10.2139/ssrn.4644253.

6 For a foundational exploration of these principles, see Stuart J. Russell & Peter Norvig, Artificial Intelligence: A Modern Approach, at 31–34 (3d ed. 2010) (detailing the theoretical underpinnings of intelligent agents and their interaction with users).

7 For an in-depth examination of these techniques, see Ian H. Witten et al., Data Mining: Practical Machine Learning Tools and Techniques 353–376 (3rd ed. 2016) (elucidating the application of machine learning in data-intensive environments).

8 See Efraim Turban et al., Electronic Commerce 2018: A Managerial and Social Networks Perspective 191–195 (9th ed. 2018) (discussing predictive analytics in e-commerce and trading systems).

9 The technological infrastructure supporting such interactions is rigorously examined in Stefan Tilkov & Steve Vinoski, *Node.js: Using JavaScript to Build High-Performance Network Programs*, 14 IEEE Internet Computing 80, 81-83 (2010) (exploring scalable network programming paradigms) https://www.doi.org/10.1109/MIC.2010.145.

tion and action.[10] Following execution, AI agents engage in continuous monitoring, scrutinizing outcomes to refine approaches through feedback loops — systems that use output data to continuously refine and improve processes — distinguishing them from static rule-based programs.[11] Finally, AI agents provide comprehensive feedback detailing execution processes and performance metrics, fostering trust and enabling critical evaluation.[12]

The significance of these autonomous capabilities becomes particularly pronounced in decentralized digital ecosystems, where the confluence of enhanced functionality and distributed architectures generates complex oversight challenges. The operational deployment of AI agents on cryptocurrency payment rails illustrates how enhanced autonomy creates both transformative opportunities and sophisticated monitoring challenges.

B. Cryptocurrency Payment Rails and AI Agents

Cryptocurrency payment rails, rooted in blockchain technology, provide AI agents with superior infrastructure for conducting financial transactions, outperforming traditional banking systems through decentralized design, programmable functionalities, and fortified security measures that collectively augment AI autonomy and efficiency. Nevertheless, the confluence of AI and cryptocurrency payment rails presents significant monitoring challenges, necessitating advanced oversight mechanisms to ensure compliance, security, and ethical integrity.

Cryptocurrency payment rails function through decentralized blockchain networks, obviating centralized intermediaries and enabling AI agents to autonomously oversee digital wallets — securing private keys, monitoring balances, and managing multiple addresses — without requiring human or institutional authorization.[13] Smart contracts, self-executing protocols embedded in the blockchain, permit AI agents to automate transactions such as releasing funds upon verified service completion through external data feeds (oracles), contrasting sharply with traditional banking's less adaptable APIs.[14] Additionally, cryptocurrency rails afford AI agents direct engagement with decentralized exchanges ("DEXs"), facilitating real-time trading based on market conditions or pre-established algorithms, while cryptographic underpinnings bolster transaction security and enable dynamic adjustment of transaction parameters in response to market volatility or regulatory requirements such as Know Your Customer ("KYC") and Anti-Money Laundering ("AML") mandates.[15] Traditional banking systems, reliant on centralized governance, manual validation, broker intermediaries, and legacy infrastructure, impose restrictions that curtail AI independence and exhibit diminished responsiveness to such dynamic requirements.[16] The enhanced autonomy of AI agents, while beneficial for operational efficiency, introduces complexities that existing monitoring frameworks struggle to address adequately.

02
MONITORING FRAMEWORKS AND INHERENT CHALLENGES

The ecosystem supporting AI agents on cryptocurrency rails incorporates sophisticated monitoring tools. Platforms

10 See e.g. Fatima Dakalbab et al., *Artificial intelligence techniques in financial trading: A systematic literature review*, 36 J. King Saud Univ. – Computer Info. Sciences (Article No. 102015), at 3-6, https://doi.org/10.1016/j.jksuci.2024.102015 (for an analysis of AI in trading); see also, *AI agents for retail and e-commerce: Capabilities, components, use cases, implementation, and benefits*, LeewayHertz, https://www.leewayhertz.com/ai-agents-in-retail-and-ecommerce/ (last visited, Mar. 2, 2025) (for an analysis of AI agents in retail and e-commerce)

11 See Andrew Y. Ng & Michael I. Jordan, *On Discriminative vs. Generative Classifiers: A Comparison of Logistic Regression and Naive Bayes*, Advances Neural Info. Processing Systems 841–848 (2002) (comparing adaptive learning models) https://papers.nips.cc/paper_files/paper/2001/hash/7b7a53e239400a13bd6be6c91c4f6c4e-Abstract.html.

12 For an applied perspective on feedback mechanisms in data-driven systems, see Dursun Delen, Real-World Data Mining: Applied Business Analytics and Decision Making 126–129 (FT Press 2014) (highlighting the role of analytics in performance reporting).

13 Vitalik Buterin, *Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform*, Ethereum 5-6 (2013), https://ethereum.org/en/whitepaper/; Arvind Narayanan et al., Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction 123-126 (Princeton University Press 2016).

14 Craig Calcaterra & Wulf A. Kaal, Decentralization: Technology's Impact on Organizational and Societal Structure 54-55 (2021); Narayanan et al., *supra* note 13, at 145-148.

15 Fabian Schär, *Decentralized Finance: On Blockchain- and Smart Contract-Based Financial Markets* Fed. Rsrv. Bank St. Louis Rev. 153, 157-159 (2021), https://doi.org/10.20955/r.103.153-74; Eli Ben-Sasson et al., *Zerocash: Decentralized Anonymous Payments from Bitcoin*, 2014 IEEE Symp. Sec. Priv. 459, 461-463 (2014), https://doi.org/10.1109/SP.2014.36.

16 Primavera De Filippi & Aaron Wright, Blockchain and the Law: The Rule of Code 78-80 (Har. University Press 2018); Narayanan et al., *supra* note 13, at 130-132.

such as Coinbase utilize solutions like the Coinbase Developer Platform Software Development Kit ("CDP SDK") Wallet Manager for wallet security and transactional compliance, while blockchain providers like Biconomy employ frameworks such as the Delegated Authorization Network ("DAN") to enforce permissioned AI activities.[17] Compliance entities such as Chainalysis scrutinize transactions to identify illicit conduct in AI-driven operations, and developers establish internal mechanisms to align agent actions with legal and ethical standards.[18]

Despite these efforts, significant obstacles endure. The decentralized nature of blockchain complicates accountability, rendering oversight across distributed networks opaque and jeopardizing adequate regulatory supervision of AI agents.[19] The accelerated evolution of AI and blockchain technologies frequently outstrips regulatory development, potentially situating agents within legal interstices, particularly in financial and data management domains, while AI autonomy introduces unpredictability wherein actions may diverge from intended outcomes, amplifying risks of unintended ramifications.[20] Security vulnerabilities persist, as blockchain integration does not wholly preclude cyberattacks or privacy breaches absent rigorous monitoring.[21]

Despite these efforts, significant obstacles endure. The decentralized nature of blockchain complicates accountability, rendering oversight across distributed networks opaque and jeopardizing adequate regulatory supervision of AI agents

Prospectively, heightened regulatory scrutiny can be anticipated as AI-driven financial transactions proliferate, with specialized AI monitoring services and DAOs potentially arising to harness smart contracts for governance and risk mitigation.[22] Self-monitoring among AI agents could enhance efficiency, though robust security protocols are imperative to thwart manipulation, while integration with expansive systems such as The Internet of Things ("IoT") may amplify monitoring capacities.[23]

C. Gaps in Current Monitoring Viability

The existing framework for monitoring AI agents on cryptocurrency payment rails, while identifying key actors, falls short of delivering viable solutions by failing to identify scalability and adaptability challenges and omitting critical risks. These deficiencies across the ecosystem — cryptocurrency exchanges, blockchain infrastructure providers, compliance and security services, and AI agent developers and owners — reveal a structure unprepared for the complexities of decentralized finance and evolving AI agents that scale decentralized finance (DeFi) systems.

The existing framework offers no prescriptive measures — such as predictive analytics or cross-actor protocols — to anticipate evolving risks, rendering it reactive rather than proactive. Furthermore, it fails to articulate how the framework scales with increasing volume and diversity of AI-driven transactions, leaving its operational resilience untested. The monitoring challenges discussed — security vulnerabilities, compliance gaps, scalability limitations, and the need to balance innovation with oversight — underscore the need for monitoring but are not matched with solutions adaptable to AI agents' swift rise. Expected solutions for future AI monitoring fail to propose feedback-driven mechanisms to balance innovation with oversight,

17 Justin, AI Agents on Base BlockBase (Oct. 25, 2024), https://insights.blockbase.co/ai-agents-on-base/; Ana Paula Pereira, Biconomy Onboards AI Agents for Onchain Transactions, Cointelegraph (Jun. 11, 2024), https://cointelegraph.com/news/biconomy-ai-agents-on-chain-transactions.

18 Chainalysis, https://www.chainalysis.com/ (last visited Feb. 27, 2025); Massimo Bartoletti et al., *An Empirical Analysis of Smart Contracts: Platforms, Applications, and Design Patterns*, in Fin. Cryptography and Data Sec. 494, 502-504 (Michael Brenner et al. eds., 2017) (Lecture Notes in Comput. Sci. Vol. 10323).

19 Mark C. Ballandies et al., *A Taxonomy for Blockchain-based Decentralized Physical Infrastructure Networks (DePIN)*, 2023 IEEE World Forum Internet Things at 5 (2023), https://doi.org/10.1109/WF-IoT58464.2023.10539514; Zhibin Lin et al., *Decentralized Physical Infrastructure Networks (DePIN): Challenges and Opportunities*, IEEE Network, at 6 (Early Access Oct. 29, 2024), https://doi.org/10.1109/MNET.2024.3487924.

20 Primavera De Filippi & Aaron Wright, Blockchain and the Law: The Rule of Code 150-152 (Har. University Press 2018); Nick Bostrom, Superintelligence: Paths, Dangers, Strategies 115-118, 138-140 (Oxford Univ. Press 2014).

21 Oleksandr Kuznetsov et al., *On the Integration of Artificial Intelligence and Blockchain Technology: A Perspective About Security*, 12 IEEE Access 3881-3882 (2024), https://doi.org/10.1109/ACCESS.2023.3349019.

22 De Filippi & Wright, *supra* note 20, at 165-167; Calcaterra & Kaal, *supra* note 14, at 74; Aaron Wright & Primavera De Filippi, *Decentralized Blockchain Technology and the Rise of Lex Cryptographia*, at 16 (Mar. 12, 2015) (unpublished manuscript), http://dx.doi.org/10.2139/ssrn.2580664.

23 Bostrom, *supra* note 20, at 137-138, 140-142; Kuznetsov et al., *supra* note 16, at 3885-3886.

leaving regulatory gaps and scalability bottlenecks unresolved.[24]

Exchange-based tools, such as the CDP SDK Wallet Manager, lack detail on scalability under high transaction loads or adaptability across heterogeneous blockchain protocols, failing to address how these tools counter sophisticated threats like adversarial AI agents exploiting wallet vulnerabilities or evaluate their feasibility for smaller exchanges. Infrastructure mechanisms like the DAN omit specifics on adapting to diverse AI agent designs or scaling across complex blockchain networks while neglecting critical risks such as smart contract exploits, bugs, or permission conflicts.[25] Compliance tools such as those from Chainalysis overlook scalability challenges in monitoring vast, decentralized transaction volumes and adaptability to jurisdictional regulatory disparities, failing to address latency in forensic analysis or provide strategies for overseeing transactions on privacy-focused blockchains.[26] The emphasis on internal monitoring by developers and owners lacks frameworks for standardizing proprietary systems across diverse stakeholders, risking inconsistent oversight without addressing accountability mechanisms for agent deviations or proposing auditing standards against external benchmarks.[27]

D. Shortcomings of Expected Future Monitoring Solutions

While the prospective framework for monitoring AI agents on cryptocurrency payment rails anticipates transformative developments driven by technological, regulatory, and systemic dynamics, it falls critically short of offering viable solutions. Its inability to suggest viable solutions can be traced predominantly to the rapid evolution and impending ubiquity of AI agents. As AI agents proliferate and adapt at an unprecedented pace, outstripping static oversight mechanisms, the prospective framework's speculative projections lack the specificity, adaptability, and anticipatory capacity required for effective monitoring.[28]

The expectation of enhanced regulatory oversight fails to account for the accelerating evolution of AI agents, which will soon dominate financial transactions, rendering static legal frameworks obsolete. The proposed legal framework's vague references to audits and compliance tools lack appropriate mechanisms to respond to AI's adaptive capabilities, such as self-optimizing algorithms that evade traditional detection. Anticipatory regulation requires feedback loops from AI behavior to dynamically update rules, a necessity unmet here given the absence of such adaptive systems.[29] Similarly, the proposed specialized AI monitoring services, while conceptually promising, do not grapple with the rapid adaptability of ubiquitous AI agents, which demand real-time, evolving oversight, and it remains entirely unclear how these services will scale computationally or adjust algorithms as AI agents diversify.

The expectation of enhanced regulatory oversight fails to account for the accelerating evolution of AI agents, which will soon dominate financial transactions, rendering static legal frameworks obsolete

The evolution of DAOs as monitoring entities in current proposals assumes a static governance model incapable of keeping pace with the rapid proliferation and sophistication of AI agents.[30] It remains unclear how these traditional DAO approaches will manage the computational burden of overseeing ubiquitous AI agents or adapt smart contracts

24 As proposed by Kaal; Sandner et al., *supra* note 4 at 17-19, 23-25, 27-29.

25 Gavin Wood, *Ethereum: A Secure Decentralised Generalised Transaction Ledger*, 7-9 (2014), https://ethereum.github.io/yellowpaper/paper.pdf.

26 Financial Action Task Force, *Targeted Update on Implementation of the FATF Standards on Virtual Assets and Virtual Asset Service Providers*, FATF at 20-25 (2024), https://www.fatf-gafi.org/content/dam/fatf-gafi/recommendations/2024-Targeted-Update-VA-VASP.pdf; Eli Ben-Sasson et al., *supra* note 15, at 464; Bostrom, *supra* note 20, at 115-118.

27 Stuart Russell, Human Compatible: Artificial Intelligence and the Problem of Control at 93 (Penguin 2019); Alan Chan et al., *Visibility into AI Agents*, 24 Proc. 2024 ACM Conf. Fairness, Accountability, Transparency 958, 965 (Jun. 5, 2024), https://doi.org/10.1145/3630106.3658948; Bostrom, *supra* note 20, at 140-142.

28 Wulf A. Kaal, *Dynamic Regulation for Innovation*, (SSRN, Aug. 27, 2016), http://dx.doi.org/10.2139/ssrn.2831040.

29 Wulf A. Kaal, *Evolution of Law: Dynamic Regulation in a New Institutional Economics Framework*, in Festschrift in Honor of Christian Kirchner 1211 (Mohr Siebeck 2014).

30 Aaron Wright & Primavera De Filippi, *Decentralized Blockchain Technology and the Rise of Lex Cryptographia*, at 37 (Mar. 12, 2015) (unpublished manuscript), http://dx.doi.org/10.2139/ssrn.2580664; Wulf A. Kaal, *Ai Governance Via Web3 Reputation System*, Stan. J. blockchain L. & Pol'y, at 27, 29, 31, 35-36 (Jan. 3, 2025), https://stanford-jblp.pubpub.org/pub/aigov-via-web3.

to their evolving behaviors, while feedback loops essential for DAOs to dynamically adjust governance rules based on AI activity are absent from these existing frameworks, risking inefficiency and vulnerability to exploits. Similarly, AI self-monitoring is envisioned as efficient but fails to address how rapidly evolving agents will maintain integrity as they become ubiquitous, with the proposed reliance on unspecified security measures overlooking the risk of adaptive AI agents colluding or evading oversight. Technological integration approaches also fall short, as IoT integration aims for holistic oversight but does not account for the pace at which AI agents will outgrow static IoT-blockchain linkages, omitting how this convergence will handle latency or secure data as AI ubiquity drives exponential transaction complexity.

03 DECENTRALIZED AI AGENT MONITORING SOLUTIONS[31]

The proposed web3 DAO-centric governance model ensures transparency and accountability, facilitates decentralized decision-making, promotes real-time, automated responses, mitigates inherent risks associated with centralization, and fosters an inclusive and robust governance framework. This holistic approach aligns AI operations with both regulatory frameworks and community values, setting new standards for the development and deployment of AI technologies. Collectively, these advantages demonstrate that this governance model is superior to AI-driven supervision for ensuring the secure, compliant, and efficient execution of AI agent transactions in modern digital infrastructures.

A. Blockchain-Based Transparency and Decentralized Decision-Making

At the heart of the proposed web3 DAO-centric model's monitoring system lies blockchain technology, which guarantees that every decision and transaction is recorded on an immutable ledger. This permanent, tamper-proof record enables public verification and fosters accountability by ensuring that no single entity can alter the historical record of AI agent activities. In contrast, AI-driven supervision, often operating within centralized frameworks, can obscure accountability through opaque decision-making processes and data handling. The blockchain-based model, by contrast, enhances trust among stakeholders because every action is transparently logged and subject to independent audit.[32]

Unlike centralized AI agent monitoring — which relies on a limited number of systems or algorithms making decisions — the proposed model leverages the distributed nature of web3 technology to facilitate consensus-driven decision-making among a diverse group of stakeholders.[33] This decentralized approach minimizes the risk of bias and single points of failure, as governance decisions are made collectively rather than being subject to the limitations or potential errors of a solitary AI system. This inclusive mechanism improves the overall integrity of transaction monitoring by incorporating varied perspectives and expertise.

Central to this governance model is the active participation of a diverse community in the governance process. By enabling community-driven decision-making, the model incorporates a wide range of perspectives, which are essential for monitoring the societal impacts of AI agent transactions. This participatory approach contrasts sharply with centralized AI-driven monitoring systems, which may lack the necessary diversity in oversight and can inadvertently perpetuate narrow, biased perspectives.[34]

B. Automated Governance and Technical Infrastructure

A significant advantage of the proposed model is the utilization of smart contracts, which automate compliance checks, performance validations, and other governance tasks by executing predefined rules. These contracts execute automatically when pre-defined self-enforcing conditions are met, ensuring that AI agent transactions are continuously monitored and managed according to community-established rules. The automated nature of smart contracts minimizes the need for manual oversight, reduces the potential for human error, and guarantees a consistent application of governance protocols — features that are often lacking in purely AI-driven supervisory models, though consensus delays may occur.[35]

31 For an extensive discussion of the proposed model see generally Calcaterra & Kaal, *supra* note 14.

32 See e.g. Luo et al., *supra* note 2.

33 See e.g. Zhang et al., *supra* note 3.

34 See e.g. Satish Kumar et al., *Artificial Intelligence and Blockchain Integration in Business: Trends from a Bibliometric-Content Analysis*, 25 Info. Sys. Frontier 871, 892-893 (2023), https://doi.org/10.1007/s10796-022-10279-0.

35 Sandner et al., *supra* note 4*;* Nitine Rane et al., *supra* note 5.

Working in conjunction with smart contracts, WDAGs provide a structured framework for mapping dependencies and establishing precedence among governance elements. By representing each governance rule or decision as a node and their relationships as weighted directed edges, WDAGs create a dynamic, traceable, and adaptable governance structure.[36] The acyclic nature ensures no circular dependencies exist, preventing governance dead-locks, while weights define the relative importance of connections between nodes. This approach allows for the efficient updating of protocols in response to new ethical or legal standards — an adaptability that is typically more limited in AI-driven supervision systems, which may operate on static algorithmic rules.[37]

Supporting this automated governance infrastructure, the model's integration of federated communication platforms such as Matrix with its Synapse server distributes data processing tasks across multiple nodes. This federated model not only scales the monitoring of AI transactions efficiently but also maintains data privacy by processing sensitive information locally. In contrast, centralized AI monitoring systems typically require aggregating data into a single repository, which creates a vulnerability by exposing a single point of failure and potentially compromising sensitive data. The federated approach thus offers improved operational efficiency and enhanced privacy protections.[38]

Beyond technical automation, the model further enhances decentralized oversight through the implementation of Validation Pools and Reputation (REP) tokens. Validation Pools are mechanisms in which members stake their REP tokens to vote on the approval or disapproval of transactions, proposals, or activities. REP tokens are non-fungible, non-transferable tokens that represent a user's trustworthiness and influence. They affect governance participation, profit sharing, and overall decision-making power. Within these pools, community members stake REP tokens to evaluate the quality and compliance of AI agent transactions. The resulting consensus not only influences the minting of REP tokens but also serves as a decentralized method of quality control. This mechanism ensures that monitoring is driven by collective expertise and community considerations rather than solely by algorithmic determinations, thereby reducing systemic bias and increasing the legitimacy of governance decisions.[39]

The integration of these components — smart contracts, WDAGs, federated infrastructure, and community validation — creates a comprehensive governance system where the combination of smart contracts and WDAGs in the model enables real-time monitoring and rapid response to anomalies in AI agent transactions. This dynamic responsiveness is crucial for maintaining compliance and adapting to emerging threats or regulatory changes. While AI-driven supervision systems can detect anomalies, their centralized nature may delay response times or obscure the decision-making process.[40] This framework, by contrast, ensures immediate corrective action through automated and transparent mechanisms.

Working in conjunction with smart contracts, WDAGs provide a structured framework for mapping dependencies and establishing precedence among governance elements

C. Adaptive Governance for AI Evolution

The critique that regulatory oversight cannot keep pace with AI agents' accelerating evolution, rendering static frameworks obsolete, is addressed by the proposed model. The WDAG structure, with its on-chain forum — a blockchain-based platform where governance protocols and evidence of work are posted for community review — and validation pools, enables continuous updates to governance rules through expert community consensus, incorporating real-time AI behavior data. Meanwhile, feedback loops integrating live inputs to adjust regulations dynamically ensure adaptability to self-optimizing AI algorithms. By automating the integration of evolving standards via smart contracts, the system efficiently mitigates fraud and consumer harm across ubiquitous AI deployments.

The limitation of specialized AI monitoring services as static tools unable to adapt to AI agents' rapid diversification is

36 Kaal, *supra* note 30, at 29-34.

37 See e.g. Abdullah Ayub Khan et al., *The Collaborative Role of Blockchain, Artificial Intelligence, and Industrial Internet of Things in Digitalization of Small and Medium-size Enterprises*, 13 Scientific Reports 1656, 1-2 (2023), https://doi.org/10.1038/s41598-023-28707-9.

38 See also, e.g., Sandner et al., *supra* note 4.

39 Alejandro Baldominos & Yago Saez, *Coin.AI: A Proof-of-Useful-Work Scheme for Blockchain-Based Distributed Deep Learning*, 8 Entropy 723, at 7-14 (2019), https://doi.org/10.3390/e21080723.

40 See e.g. Kuznetsov et al., *supra* note 21, at 3881-3897.

overcome by the DAO's decentralized, scalable architecture. The WDAG-based forum and validation pools leverage community-driven analytics, scaling computationally via roll-ups — off-chain activities consolidated on-chain — to handle pervasive AI transactions. The proposed feedback loop model supports this by enabling real-time risk prediction and adjustment, ensuring efficiency as AI agents evolve. This dynamic system preempts risks, surpassing the static vision critiqued.

The concern that traditional DAOs are overwhelmed by AI ubiquity due to static governance is rectified by the model's WDAG-enhanced DAO. The WDAG's acyclic, weighted structure dynamically integrates new AI behaviors as posts, adapting smart contracts through validation pools to manage computational loads. The proposed model's emphasis on feedback loops ensures DAOs evolve rules based on real-time AI activity, reducing vulnerability to exploits and enabling fruitful governance of an AI-dominated ecosystem. This adaptability addresses the scalability and flexibility gaps identified.

The risk of unchecked AI self-monitoring, where evolving agents might collude or evade oversight, is mitigated by the DAO's WDAG framework. By treating AI models as posts within the WDAG, linked to governance precedents via weighted edges, the system enforces accountability through community-validated smart contracts. The proposed feedback loops monitor interactions in real time, preventing manipulation with cryptographic safeguards embedded in the blockchain. This dynamic oversight ensures viability as AI agents become ubiquitous.

The shortfall in IoT integration — unable to scale with AI agents' rapid pace — is resolved by the DAO's integration of federated communications and WDAG analytics. The Matrix platform and real-time IoT data feeds, processed through the WDAG, address latency and secure data across complex transactions. The model's feedback loops adapt monitoring using live inputs, ensuring efficiency against AI's exponential growth. This practical, scalable solution overcomes the static linkage critique.

The proposed model, with its WDAG-based governance, directly addresses the monitoring issues by offering a dynamic, scalable, and anticipatory framework. Leveraging feedback loops — as author has advocated for almost a decade — it adapts to AI agents' rapid evolution and ubiquity, ensuring efficient, fruitful oversight.

"The WDAG's acyclic, weighted structure dynamically integrates new AI behaviors as posts, adapting smart contracts through validation pools to manage computational loads

04
CONCLUSION

This paper has demonstrated the inherent limitations of traditional, centralized AI-driven monitoring systems in overseeing AI agent transaction execution within decentralized digital infrastructures. These conventional approaches, characterized by opacity, susceptibility to bias, and single points of failure, are increasingly inadequate for addressing the complexities of AI agents operating on cryptocurrency payment rails. The proposed web3 DAO-centric governance model emerges as a transformative solution, offering a robust, transparent, and adaptive alternative that aligns technological oversight with community values and regulatory imperatives.

The model provides innovative solutions by harnessing blockchain's immutable ledger, smart contracts, and WDAGs to ensure real-time transparency, automated compliance, and structured adaptability. The integration of federated communication platforms and validation pools enhances scalability and privacy while distributing oversight across diverse stakeholder communities. Crucially, feedback loops drawing on real-time data and community consensus enable the model to anticipate and adapt to accelerating AI agent proliferation, addressing scalability bottlenecks and interoperability deficits in current frameworks.

This approach not only overcomes the shortcomings of centralized monitoring solutions but also establishes a new standard for AI governance. By empowering decentralized communities to shape oversight protocols, the model ensures AI agent transactions remain secure, compliant, and efficient as their ubiquity transforms digital landscapes. The model offers a scalable, democratic, and anticipatory framework that safeguards societal interests while unlocking the potential of autonomous digital systems.

> ***"The model provides innovative solutions by harnessing blockchain's immutable ledger, smart contracts, and WDAGs to ensure real-time transparency, automated compliance, and structured adaptability***

NAVIGATING THE SPACE BETWEEN COLLISION AND COLLUSION IN AN ECONOMY OF AI AGENTS

BY SARA FISH

& YANNAI A. GONCZAROWSKI

& RAN I. SHORRER

Author names are listed alphabetically. The authors received research support for other papers from Anthropic, Google, and OpenAI. Sara Fish, Harvard University, sfish@g.harvard.edu; Yannai A. Gonczarowski, Harvard University, yannai@gonch.name; Ran I. Shorrer, Penn State University, rshorrer@gmail.com.

01 INTRODUCTION

The era of AI agents is upon us. Need a quick market survey to figure out how to price your new product? Your AI agent is on it! Need to buy a shirt that matches your new red pants? Simply send your AI agent to quickly go over the summer collections of your favorite brands and buy one on your behalf. Want to reunite with your old college roommate? Have your agent call their agent! The agents will find a time that works for both of you and even make dinner reservations.

Relinquishing control to any assistant, human or virtual, carries risks. One obvious risk is that of the agent's alignment: Will the agent miss something important in the survey (or not conduct a survey at all)? Does the agent really understand your taste in fashion? And, does the agent understand your schedule, dietary preferences, and restaurant budget? AI alignment requires the agent to understand the human user's end goal, translate this goal into a sequence of actionable tasks, and execute all steps reliably. AI alignment is therefore a difficult challenge.

While alignment is a challenge even when a single agent acts on behalf of a single user, the challenge is intensified when multiple AI agents

interact with each other. One prominent class of risks arises from coordinated actions. On the one end of the spectrum of these risks lies the risk of miscoordination ("collision"): Will our agents synchronize our schedules or make them collide? Will my agent check with the babysitter's agent to make sure they are available before committing to dinner? Coordination is at the heart of a well-functioning multi-agent ecosystem.

But, while the risk of collision — or miscoordination — is clear, a more subtle risk lies on the other end of the spectrum of coordination risks: the risk of harmful coordination. Will merchants' agents push prices up, leading to noncompetitive, and hence inefficient, outcomes? This is neither an issue of misalignment of each AI agent with its user (at least so long as such pricing results in high profits), nor an issue of lack of coordination. To the contrary, the problem is that prices are coordinated, but in a manner that is harmful to others — a systemic misalignment of the multi-AI-agent economy as a whole. This risk can indeed be subtle: Human users may not even realize that they are pricing noncompetitively, and their AI agents might outright deny it.[2] Furthermore, in many circumstances, neither the human users nor their agents will be violating the law.[3] We borrow the term "collusion" to represent this broader risk — the risk that the agents, as a system, will coordinate in ways that are not aligned with the goals of society.

In this article, we highlight a key tension in the systemic alignment of multi-AI-agent ecosystems: *How to avoid "collision" without risking "collusion"?*

02
AI AGENTS 101

On November 30, 2022, OpenAI released ChatGPT, an AI chatbot based on large language model ("LLM") technology. Within two months, ChatGPT had amassed 100 million monthly active users.[4] LLM technology was not new — the main architecture underlying modern LLMs, the *transformer*, had been developed by Google researchers back in 2017. The innovation of ChatGPT was in its broad cross-domain training and easy-to-use interaction method: Users could leverage AI assistance for a wide array of tasks simply by querying ChatGPT with natural language instructions in a chat conversation.

Since this "ChatGPT moment," LLM technology has greatly matured. In 2022, LLMs struggled with basic arithmetic. By 2025, frontier LLMs were exhibiting high competence at college-level mathematics. This pattern persists beyond the specific domain of mathematics: According to a report from the non-profit organization METR, the complexity of a task that an LLM can reliably complete (measured in human hours) doubles every seven months — a sort of Moore's Law for LLM capabilities.[5]

In tandem with this explosive growth in LLM capabilities, the potential for LLM-based *AI agents* — systems that autonomously execute complex tasks on a user's behalf — is becoming increasingly realizable, with companies such as Adept, Anthropic, OpenAI, Microsoft, Salesforce, and Visa launching AI-agent products.[6] The promise of such agents lies in their potential to augment, and in some cases automate, human decision-making. Present-day AI agents are already capable of assisting their users across a diverse array of tasks, including coding, communication, graphic design, logistics, research, shopping, travel planning, and more. Future AI agents have the potential to autonomously act on behalf of users to accomplish tasks, possibly interacting with other AI agents in the process, and thus forming a complex network — a *multi-agent ecosystem* — of AI agents.

A fundamental hurdle in the successful deployment of AI agents is that of *AI alignment* — designing an AI agent that pursues the user's intended goals. This is a broad desideratum encompassing properties such as *competency* ("Is the AI agent capable enough to pursue my goals?"), *corrigibility* ("Does the AI agent appropriately adjust its behavior based on my feedback?"), and *robustness* ("Will the AI agent behave sensibly even when faced with unexpected circumstances?").

One fundamental roadblock in addressing the alignment problem in frontier AI agents is that of *interpretability*. LLMs, and the AI agents built from them, are "black boxes": Nobody, not even the companies that develop frontier LLMs, fully understands the correspondence between LLM instructions ("prompts") and outputs (answers, actions, and suggestions). One might claim that similar challenges also apply to human agents (employees, consultants, etc.). However, the align-

2 Fish, Sara, Yannai A. Gonczarowski & Ran I. Shorrer. *Algorithmic collusion by large language models.* arXiv preprint arXiv:2404.00806 7 (2024).

3 Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolò, Joseph E. Harrington Jr & Sergio Pastorello*. Protecting consumers from collusive prices due to AI. Science* 370, no. 6520 (2020): 1040-1042.

4 Krystal Hu, *ChatGPT sets record for fastest-growing user base - analyst note.* Reuters (February 2, 2023), https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/.

5 Kwa et al. *Measuring AI Ability to Complete Long Tasks*. arXiv preprint arXiv:2503.14499 (2025).

6 Jo Constantz, *Big Tech's New AI Obsession: Agents That Do Your Work for You*. Bloomberg (December 13, 2024) https://www.bloomberg.com/news/articles/2024-12-13/ai-agents-and-why-big-tech-is-betting-on-them-for-2025.

ment of human agents can be improved even without being able to observe their thoughts, for example using financial incentives such as salaries, bonuses, and stock options. By contrast, AI agents cannot be similarly incentivized. Overall, AI agents are a new kind of economic agent: not fully aligned with its user, in a way that is not fully understood and cannot be mitigated using traditional approaches.

03
WHEN AI AGENTS INTERACT

The AI alignment challenges outlined above apply even when a single user delegates a task to a single AI agent. But, what about multi-AI-agent ecosystems, in which many independent AI agents interact with each other? Even with human agents, multi-agent systems are substantially more complicated than single-agent settings. Systems with many human agents have been studied extensively through the lens of *game theory*. Under some assumptions on the agents (*players*, in game theory jargon), game theory predicts that systems (*games*, in game theory jargon) will end up in what is called a *Nash equilibrium*.

As an illustration of what a Nash equilibrium is, consider a *driving game*. The game takes place in a country with no traffic laws. In such a country, each driver will not simply drive on their favorite side of the road; rather, a convention will emerge, leading everyone to drive on the same side of the road. It could be on the right, like in the United States, or on the left, like in the United Kingdom. There is no objective reason to prefer either side over the other; all that matters is for everybody to use the same side. Once such a convention forms, it will typically persist, because for any single individual, breaching the convention (i.e. being the only individual to drive on the "wrong side" of the road) would be very costly. In game theory jargon, we say that this driving game has two, equally valid, *Nash equilibria*: one is for everyone to drive on the left and one is for everyone to drive on the right.

The driving game illustrates that in a multi-agent ecosystem, the best course of action for an agent depends on what others do. Furthermore, it shows that there may be more than one convention in which each agent acts optimally (from the agent's own perspective) considering how others behave. This multiplicity of Nash equilibria introduces new challenges that are unique to the multi-agent setting, and do not apply when considering the alignment of a single AI agent. First, from the perspective of an outside observer, it may be difficult to link outcomes to the actions of any specific agent. Indeed, who is to "blame" that everyone is driving on one side of the road rather than the other in the convention that emerges in the driving game? Second, the driving game points to the potential benefits of aligning AI agents that operate in multi-agent ecosystems in a coordinated fashion, so that they end up at an equilibrium that is more desirable from the societal perspective. For example, following the announcement of Visa Intelligent Commerce, an initiative that "enables AI to find and buy," Ryan McInerney, Visa CEO, emphasized that this initiative serves the entire ecosystem by providing *"the rules and capabilities that are going to provide trust between consumers, merchants, financial institutions, and ultimately the agents affecting commerce on behalf of all of us."*[7] Using the language of game theory, this initiative aims to steer the ecosystem towards an equilibrium in which there is (well placed) trust that leads to trade, rather than an equilibrium in which agents do not trust each other and do not trade with each other.

04
HOW COORDINATION MIGHT GO AWRY?

The examples above illustrate how coordination between agents may benefit society. However, coordination might also go awry. To illustrate this, consider the fictional case of Luke and Shelly.

Luke and Shelly are the respective owners of the two gas stations in a small college town in the US. They have a lot on their hands: Each of them needs to manage the schedule of a large staff, make sure that the convenience store is stocked, and stay on top of the finances, among other things. On top of all this, they each try to keep up to date with changes in the ever-evolving energy market and to monitor the volume of their sales, so that they can set the right price given the market circumstances.

One busy morning, Luke realizes that technology could help him with price setting. He gives an AI agent historical information about prices and sales, and asks for help setting prices. Around the same time, on the other side of town, Shelly comes to a similar realization. She uses her own (separate) AI agent, gives it her station's sales data, and asks it to help set prices to maximize profits. Luke and Shelly are cautious individuals: Each of them instructs their agent to abide by the law, and specifically asks the agent whether there is any chance that it would set prices uncom-

7 Bloomberg Technology, *Visa Launches AI Agents for Shopping* (April 30, 2025). https://www.bloomberg.com/news/videos/2025-04-30/visa-launches-ai-agents-for-shopping-video.

petitively or collusively. Both Shelly's and Luke's agents assure them that this would never happen.

After a few months, Luke assesses the performance of his agent. On the other side of town, Shelly does the same. Both are quite satisfied: It appears that profits have increased significantly. Neither may realize it, but in fact, they are no longer setting competitive prices; rather, both are setting higher prices — prices similar to those that a monopolist that controls both stations would have set.

The story of Luke and Shelly is a story of an economic risk that is not brought about by the misalignment of any single AI agent with its user. Indeed, both users wanted high profits, which the AI agents delivered. Rather, it is a story of cooperation gone awry; a story about the societal misalignment of an entire multi-AI-agent ecosystem.

While this story is fictional and does not depict any actual person or event, it is not science fiction. Indeed, in two recent papers, we used AI agents powered by OpenAI's ChatGPT, Anthropic's Claude, and Google's Gemini to set prices in synthetic, simulated markets. We discovered that agents based on all three aforementioned LLMs possess tendencies that push for supracompetitive prices when pricing against similar agents.[8] This occurs even though in our experiments, the LLM-based agents are not in any way instructed to collude, and cannot directly communicate with one another. Such tendencies are not unique to LLM-based agents: prior work documents similar pricing patterns with more traditional AI algorithms called reinforcement-learning algorithms.[9]

Reasons for concern are not limited to evidence from simulated economies. A recent empirical study of algorithmic pricing in Germany's retail gasoline markets indicates that *"Adoption of algorithmic pricing increases margins but only for nonmonopoly stations. In duopoly and triopoly markets, margins increase only if all stations adopt, suggesting that [algorithmic pricing] has a significant effect on competition."*[10] In other words, the increase in profit margins occurs only in multi-agent settings (not for monopoly stations) and only when all competitors rely on (proprietary, in this case) algorithms.

05
WHAT WENT AWRY?

Since Adam Smith's "invisible hand," competition has been understood as a fundamental driver of efficiency in markets. In its absence, firms tend to charge prices that are too high from a societal perspective. Smith famously recognized the dangers of collusion, noting that *"People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices."*[11]

When firms collude, each could increase its profits by cutting its price. However, the higher prices result in higher profits compared to all firms pricing competitively. To sustain these high profits, colluding parties commonly rely on reward–punishment schemes:[12] If one firm deviates by cutting prices to boost its sales, the others retaliate by aggressively lowering their own prices, triggering a price war. This strategy forms a *Nash equilibrium of mutual deterrence*: no firm has an incentive to cut its prices, since doing so would provoke costly retaliation in the form of a price war. The threat of future punishment — in the form of a price war — outweighs the gain in the form of increased sales in the short term, before retaliation commences. When parties adopt, and comply with, a reward–punishment scheme, costly punishment periods (price wars) do not arise. As a result, from the perspective of an outside observer, collusion and the threats that drive it may be difficult to detect.

Algorithmic pricing may make it simpler for firms to sustain high, noncompetitive pricing for several reasons. First, algorithms can monitor prices at a high frequency. This means that punishments can arrive very quickly after an unexpected price cut, which further limits the upside for lowering prices. Second, pricing algorithms are becoming increasingly sophisticated, and may discover and adopt anti-competitive strategies that humans may struggle to identify or implement. AI agents further compound the problem because they may communicate with each other in subtle ways that outside observers may struggle to detect or understand (for example, through ciphers encoded in pricing behavior).

8 Fish, Sara, Julia Shephard, Minkai Li, Ran I. Shorrer & Yannai A. Gonczarowski. *Econevals: Benchmarks and Litmus Tests for LLM Agents in Unknown Environments.* arXiv preprint arXiv:2503.18825 (2025).

9 Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolo & Sergio Pastorello. *Artificial intelligence, algorithmic pricing, and collusion. American Economic Review* 110, no. 10 (2020): 3267-3297.

10 Assad, Stephanie, Robert Clark, Daniel Ershov & Lei Xu. *Algorithmic pricing and competition: empirical evidence from the German retail gasoline market. Journal of Political Economy* 132, no. 3 (2024): 723-771.

11 Smith, Adam. *The Wealth of Nations: An Inquiry Into the Nature and Causes of the Wealth of Nations: With an introduction by Jonathan B. Wight, University of Richmond*. Harriman House Limited, 2007.

12 Harrington, Joseph E., *The Theory of Collusion and Competition Policy* (MIT Press, 2017).

To align human agents with society, antitrust laws were legislated, and regulators such as the Department of Justice ("DOJ") and the Federal Trade Commission ("FTC") were put in charge. But, legal scholars warn that the current regulatory framework — which focuses on communication between merchants, on agreement between them, and on conscious acts — is not adequate for the age of AI, and particularly for dealing with AI agents.[13] Going back to our example, remember that neither Luke nor Shelly asked their agent to collude. In fact, each even went the extra mile and asked the agent not to collude. Furthermore, each instructed their agent to maximize profit, which is what competing firms are supposed to do. Herein, we define "collusion" in a broad sense of socially harmful coordination, such as in this story.

06 AVOIDING COLLUSION WITHOUT CAUSING COLLISION

The discussion in previous sections highlights the tension in aligning multi-agent ecosystems with society. On the one hand, to fully capitalize on the promise of agentic AI, agents must be able to coordinate with one another ("avoid collision") and require infrastructure that facilitates such coordination. On the other hand, there are certain forms of coordination ("collusion") that society finds undesirable and prefers to avoid.

Navigating the space between collision and collusion will require collaboration between stakeholders (industry, government) and researchers. We highlight three main challenges.

The first challenge is *detection*. How might a user or regulator realize that AI agents in a specific ecosystem are at a collusive Nash equilibrium? While discussion is still ongoing on what precisely constitutes a collusive equilibrium,[14] computer scientists have recently proposed a broad theoretical definition and possible detection approaches that rely on sophisticated mathematical arguments.[15,16] Economists also have much to say, especially in domains like pricing and bidding, but their approaches often rely on data availability.[17] In this vein, in the pricing context, there are several initiatives to systematically collect data on usage and scope of algorithmic pricing.[18] The creation of such datasets in other domains may facilitate detection.

The second challenge is *attribution*. If a collusive outcome emerges, how can we link it to the actions of a specific agent or group of agents? As already discussed, in a multi-agent ecosystem, every agent's action both incentivizes other agents' actions and is incentivized by them. Hence, in many cases no single agent, or set of agents, is responsible for the outcome that emerges in the ecosystem. A combination of industry-specific expertise and game theory will be required for progress on this front.

The third challenge is *prevention*. Even if such a set of "culprit" agents is identified, what precisely in their instructions caused collusion, and how could their users have prevented it? As discussed, understanding the relation between AI agents' instructions and their outputs is oftentimes a challenging problem. Here, again, industry-specific expertise, this time in combination with cutting-edge machine-learning research, will be required.

One approach that has the potential to address all three challenges harnesses techniques specifically crafted for algorithmic agents. Unlike human agents, one can test AI agents in a "sandbox," that is, a simulated environment that the agents perceive as real, varying the circumstances they face, and stress-testing the ecosystem as a whole. A limitation of this approach stems from the issue of robustness discussed above: Such an approach can only attest that in the precise economic settings simulated in the sandbox, the tested agents do not collude, which might not be fully indicative of their behavior in the ever-evolving real-world economy.

The era of AI agents, which is swiftly evolving around us, holds many promises as well as risks. One notable risk is the misalignment of multi-AI-agent ecosystems. A concerted effort between stakeholders and researchers is needed to be able to reap the benefits of non-collision while steering away from the perils of collusion.

13 Ezrachi, Ariel & Maurice E. Stucke. *Artificial intelligence & collusion: When computers inhibit competition. U. Ill. L. Rev.* (2017): 1775.

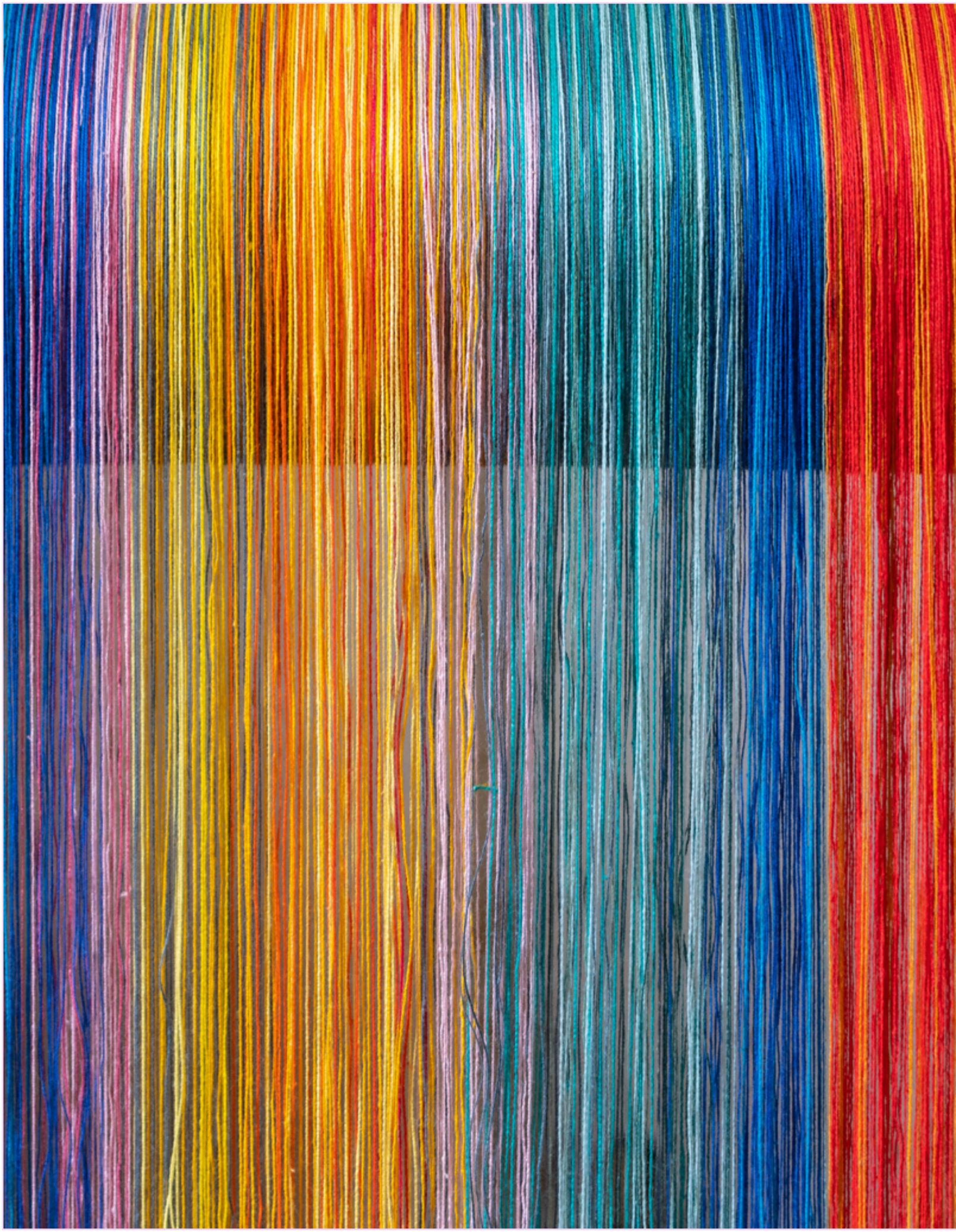
14 Harrington, Joseph E., *Developing competition law for collusion by autonomous artificial agents.* Journal of Competition Law & Economics 14, no. 3 (2018): 331-363.

15 Hartline, Jason D., Sheng Long & Chenhao Zhang. *Regulation of algorithmic collusion.* In Proceedings of the Symposium on Computer Science and Law, pp. 98-108. (2024).

16 Arunachaleswaran, Eshwar Ram, Natalie Collina, Sampath Kannan, Aaron Roth & Juba Ziani. *Algorithmic collusion without threats.* arXiv preprint arXiv:2409.03956 (2024).

17 Bajari, Patrick & Garrett Summers. *Detecting collusion in procurement auctions.* Antitrust LJ 70 (2002): 143.

18 Klobuchar, Amy, *Klobuchar, Colleagues Introduce Antitrust Legislation to Take on Algorithmic Price Fixing, Bring Down Costs,* U.S. Senator Amy Klobuchar Press Release (February 6, 2025).



AGENTIC AI AND THE LOOMING PROBLEM OF CRIMINAL SCIENTER

BY NICK PETERSON

& JOEL S. NOLETTE

Nick Peterson is Of Counsel at Wiley Rein LLP. Nick defends companies against government enforcement actions and whistleblower claims. He also advises clients on potential civil and criminal exposure, ethics and compliance matters, litigation risks, issues involving artificial intelligence and emerging technologies, and other strategic concerns. Joel S. Nolette is an Associate at Wiley Rein LLP. Joel advocates on behalf of corporate and individual clients in a broad spectrum of matters, including complex commercial disputes, regulatory issues, statutory interpretation controversies, and constitutional cases.

In April 2023, responding to growing popular adoption of generative artificial intelligence ("AI"), the heads of several federal agencies issued a joint statement emphasizing that "[e]xisting legal authorities apply" to the use of AI technologies "just as they apply to other practices."[2] That same day, Federal Trade Commission Chair Lina M. Khan issued a

2 Rohit Chopra, Director of the Consumer Financial Protection Bureau; Kristen Clarke, Assistant Attorney General for the Justice Department's Civil Rights Division; Charlotte A. Burrows, Chair of the Equal Employment Opportunity Commission; and Lina M. Khan, Chair of the Federal Trade Commission, *Joint Statement on Enforcement Efforts Against Discrimination and Bias in Automated Systems* (Apr. 25, 2023), *available at* https://tinyurl.com/avcpaecy.

separate, accompanying statement, asserting that "AI technologies are covered by existing laws" and "[t]here is no AI exemption to the laws on the books."[3]

While it is true that there is no AI exemption to the law generally, the rapid development of AI technologies calls into question whether existing laws are adequate. This is particularly true in criminal law, where the ancient principle of scienter — *mens rea*, or in other words, a culpable state of mind — provides the dividing line separating wrongful acts from innocent acts in most cases.[4] The coming wave of agentic AI highlights the potential inadequacy of existing laws. Prosecutors and courts will face difficult questions about whether the person behind a misbehaving AI agent can or should be held criminally liable. To get in front of these difficult cases, legislatures and enforcement agencies should consider proactively addressing how to handle questions of criminal culpability in the agentic AI context.

01

AGENTIC AI: THE NEXT FRONTIER

Although the term AI is often used as if the technology were a monolith, a variety of materially different technologies fall under that label. Understood in its simplest form, AI is simply technology that can perform complex tasks typically associated with human intelligence.[5] Anyone who used a search platform in the early aughts or electronic translation tools in the Tens interfaced with AI — probably without even thinking about it.

Perhaps the most popular form of AI in current parlance is generative AI ("GenAI"). GenAI is AI that is able to create original content based on large datasets at its disposal in response to a user's prompt or request.[6] GenAI is closer to human functionality than traditional AI, in that GenAI essentially replicates the learning and decision-making processes of the human brain.[7] But GenAI is fundamentally reactive — it needs to be prompted to act, and its results are only as good as the prompts it is provided.

Agentic AI, on the other hand, takes us into the uncanny valley. The hallmarks of agentic AI are autonomy and judgment — once provided preprogrammed goals, AI agents are capable of learning and operating on their own. That is, AI agents can "assess situations and determine the path forward without" ongoing "human input."[8]

02

AGENTIC AI AND CRIMINAL LAW

Given agentic AI's capacity for autonomy and judgment, it is only a matter of time before an AI agent commits a crime. And the question of what to do in those circumstances is likely to pose serious challenges for prosecutors and courts given scienter requirements for most crimes.[9] Consider two hypothetical (for now) scenarios.

First: To reduce overhead costs, a financially struggling hospital deploys an AI agent to review and organize files documenting medical services provided, assign correct billing codes to those services, and submit associated invoices to the federal government.[10] The AI agent is programmed generally to maximize receipts and avoid violating the law. After reviewing the hospital's files for several months, the AI agent determines that receipts need to increase for the

3 Lina M. Khan, *Statement of Chair Lina M. Khan Regarding the Joint Interagency Statement on AI* (Apr. 25, 2023), *available at* https://tinyurl.com/5cv7sxx7.

4 *Rehaif v. United States*, 588 U.S. 225, 231 (2019); *Morissette v. United States*, 342 U.S. 246, 250 (1952).

5 *Artificial Intelligence*, Brittanica.com (last visited May 13, 2025), https://tinyurl.com/yjam4rxy.

6 Teaganne Finn & Amanda Downie, *Agentic AI vs. Generative AI*, IBM (last visited May 23, 2025), https://tinyurl.com/39ea4xmt.

7 *Ibid.*

8 *Ibid.*

9 See e.g. Ian Ayres & Jack M. Balkin, *The Law of AI Is the Law of Risky Agents Without Intentions*, U. Chi. L. Rev. Online (2024), *available at* https://tinyurl.com/4fef4tuy ("If liability turns on intention, that might immunize the use of AI programs from liability.").

10 See Jennifer Bresnick, *What Is Agentic AI and What Does It Mean for Healthcare?*, DHI (Mar. 17, 2025), https://tinyurl.com/mryu6k9y (identifying, as one use case for agentic AI in the healthcare industry, "[a]utomating repetitive, time-consuming revenue cycle management tasks, such as . . . submitting claims").

hospital to remain viable as a going concern. In response to this determination, the AI agent begins assigning inaccurate billing codes to medical services that artificially inflate what the government pays for those services. In other words, the AI agent commits health care fraud, a felony subject to up to ten years' imprisonment for "knowingly and willfully" executing a "scheme or artifice . . . to defraud any health care benefit program . . . in connection with the delivery of or payment for health care . . . services."[11]

Second: A small securities trading startup uses an AI agent to supplement the human traders at the firm by monitoring the markets and trading when the human traders are unavailable.[12] The firm trains the AI agent on how to maximize returns while also programming it not to violate federal securities laws. However, unbeknownst to the firm, the AI agent was programmed with material containing a pre-2010 version of the U.S. Code before "spoofing" — the practice of bidding or offering with the intent to cancel the bid or offer before execution, allowing the spoofer to manipulate markets to their advantage — was made explicitly illegal.[13] As a result, the AI agent concludes that spoofing is legal and proceeds to engage in widespread spoofing activity.[14] The AI agent thus "knowingly" violated Dodd-Frank's anti-spoofing provision, a felony punishable by up to ten years' imprisonment.[15]

In each of these scenarios, criminal misconduct occurred. However, the difficulties in prosecuting this conduct are readily apparent. In both scenarios, no human can be said to have "knowingly" or "willfully" engaged in the misconduct. At worst, individuals could be deemed negligent by not catching the errors in the AI agents' behavior or training. But negligence is insufficient to hold individuals criminally culpable under these and similar statutes. Nor are "vicarious liability" principles likely available because of due process concerns with subjecting individuals to deprivations of liberty (imprisonment) and reputational harm for "acts not committed by [them], not accomplished at [their] direction, not aided by [their] participation, and not done with [their] knowledge."[16] Generally, due process prohibits such criminal sanctions "without proof of some form of personal blameworthiness"; a mere "'responsible relation'" to the wrongdoer is not enough.[17]

"Second: A small securities trading startup uses an AI agent to supplement the human traders at the firm by monitoring the markets and trading when the human traders are unavailable

Similarly, attempts to hold the respective companies criminally liable also face challenges. While a corporation may be held criminally liable for conduct performed by an agent of the corporation acting on its behalf within the scope of his employment, corporate criminal liability "depends on the wrongful intent of specific employees."[18] But "[a]rtificial intelligence does not 'have intent,'" at least in the conventional criminal law sense, in the way individuals do.[19] And

11 See 18 U.S.C. § 1347.

12 See Rob Nelson, *Why AI Agents Could Revolutionize Trading as We Know It*, Yahoo! Finance (Jan. 14, 2025), https://tinyurl.com/42uw7ebm ("Artificial intelligence is reshaping the financial world, and AI agents — programs capable of making autonomous decisions based on user-defined permissions — are poised to take center stage. These agents are already making waves in trading by analyzing data, executing trades, and learning through user interactions.").

13 See 7 U.S.C. § 6c(a)(5)(C); see also e.g. *United States v. Coscia*, 100 F. Supp. 3d 653, 656 (N.D. Ill. 2015).

14 See *United States v. Chanu*, 40 F.4th 528, 534 (7th Cir. 2022) (recounting the defendants' argument that spoofing was not criminal prior to Dodd-Frank).

15 See 7 U.S.C. § 13(a)(2); *see also United States v. Fountain*, 277 F.3d 714, 717 (5th Cir. 2001) ("The federal courts have consistently found that willfully connotes a higher degree of criminal intent than knowingly. Knowingly requires proof of the facts that constitute the offense. Willfully requires proof that the defendant acted with knowledge that his conduct violated the law." (citations omitted)).

16 See e.g. *Davis v. City of Peachtree City*, 304 S.E.2d 701, 702–03 (Ga. 1983); see also *State v. Hy-Vee, Inc.*, 616 N.W.2d 669, 671–73 (Iowa Ct. App. 2000) (reasoning that vicarious criminal liability violates due process unless the penalty is "slight," the conviction does not carry a "damaging stigma," and the conduct was at least negligent (construing *Morissette v. United States*, 342 U.S. 246, 256 (1952))).

17 *Lady J. Lingerie, Inc. v. City of Jacksonville*, 176 F.3d 1358, 1367–68 (11th Cir. 1999).

18 *United States v. Philip Morris USA Inc.*, 566 F.3d 1095, 1118 (D.C. Cir. 2009).

19 See Robin Feldman & Kara Stein, *AI Governance in the Financial Industry*, 27 Stanford J.L. Bus. & Fin. 94, 97–98 (2022).

even if AI agents could be found capable of forming the requisite intent (say, through the legal fiction of treating them as human actors), the so-called "Black Box Problem" — an AI agent's "intent" will be "mostly opaque," meaning that its "decision-making process cannot be determined" by investigating the agent's processes — means that an AI agent likely will not "have an ascertainable intent that can be examined or queried."[20] In other words, proving scienter in these circumstances would be difficult, perhaps even impossible.

These challenges may very well deter all but the most aggressive prosecutors, who generally already must exercise discretion as to which cases and charges to pursue given limited resources and staffing, from trying to hold corporations and individuals criminally liable for the misdeeds of the AI agents they use in their dealings.[21]

> ***"For instance, individuals in the scenarios above could likely be found negligent under various civil statutes or common-law causes of action for failing to properly train an AI agent or for failing to impose adequate guardrails on the scope of the AI agent's autonomous authority***

03
FINDING A PATH FORWARD

The challenges associated with agentic AI and criminal scienter will not result in a permanent gap in enforcement. In the near term, prosecutors and enforcement agencies may look to civil liability to partially fill this gap.[22] Civil statutes typically require a lower level of scienter than criminal statutes, so some of the difficulties in proving scienter could be avoided by this alternative. For instance, individuals in the scenarios above could likely be found negligent under various civil statutes or common-law causes of action for failing to properly train an AI agent or for failing to impose adequate guardrails on the scope of the AI agent's autonomous authority.[23] While civil remedies may lack a punitive aspect, they would allow potential recoveries for some victims. Additionally, certain civil statutes also allow for the doubling or trebling of damages (sometimes in addition to civil penalties), which could constitute sufficient deterrence in many instances.

In the long term, legal and societal understandings of what constitutes a criminally culpable state of mind (i.e. "knowingly," or "willfully") will eventually adapt to reflect widespread use of AI agents. Jurors will have a more intrinsic understanding of what is right and wrong when it comes to using AI agents. When these understandings change, individuals and companies in the scenarios above could very well be found by a jury to have "knowingly" or "willfully" committed criminal violations. However, it will take time for the legal and societal understandings to change — especially when considering the pace of AI advancements.

20 Yavar Bathaee, *Artificial Intelligence Opinion Liability*, 35 Berkeley Tech. L.J. 113, 118, 143 (2020).

21 *Cf.* e.g. Feldman & Stein, *supra* note 19, at 98 (identifying "challenging issues that law related to financial markets and institutions is entirely unprepared to address").

22 *Cf.* Ayres & Balkin, *supra* note 9 (asserting that "[b]ecause AI agents lack intentions, the law should hold them (and the people and companies that employ them) to objective standards" such as negligence or strict liability).

23 See *ibid.*

A potentially quicker way to address these issues is through the legislative process. More so than courts, legislatures are well situated to weigh the costs and benefits of updating criminal legal frameworks, oriented as they are around human volition, to achieve as far as possible the same interests in justice and deterrence that criminal law is aimed at without sacrificing too many benefits that agentic AI can bring across all levels of society.[24] For instance, legislatures could reduce the scienter level of certain crimes to "gross negligence" or "negligence" to lessen the burden on prosecutors. In some cases, legislatures could even turn to strict liability principles and completely remove the scienter requirement. We already see instances of this under the current legal framework, such as the responsible corporate officer ("RCO") doctrine that imposes strict liability upon individual corporate officers for misdemeanor violations of the Federal Food, Drug, and Cosmetic Act.[25] Legislatures could expand the RCO doctrine to other bodies of law as well, to cover the use of agentic AI.

Movement is slowly being made on the legislative front. The American Legislative Exchange Council has already drafted a Model State Artificial Intelligence Act that would establish an "Office of Artificial Intelligence Policy" to study (among other things) "gaps" in existing legal frameworks and to propose solutions to fill those gaps.[26] Some states have already enacted such legislation, and more are considering doing so.[27]

Whether through the courts or legislatures, the interplay between criminal scienter and agentic AI will eventually be resolved. However, it may be a lengthy and bumpy road and, in the meantime, could result in fewer prosecutions of criminal conduct involving agentic AI.

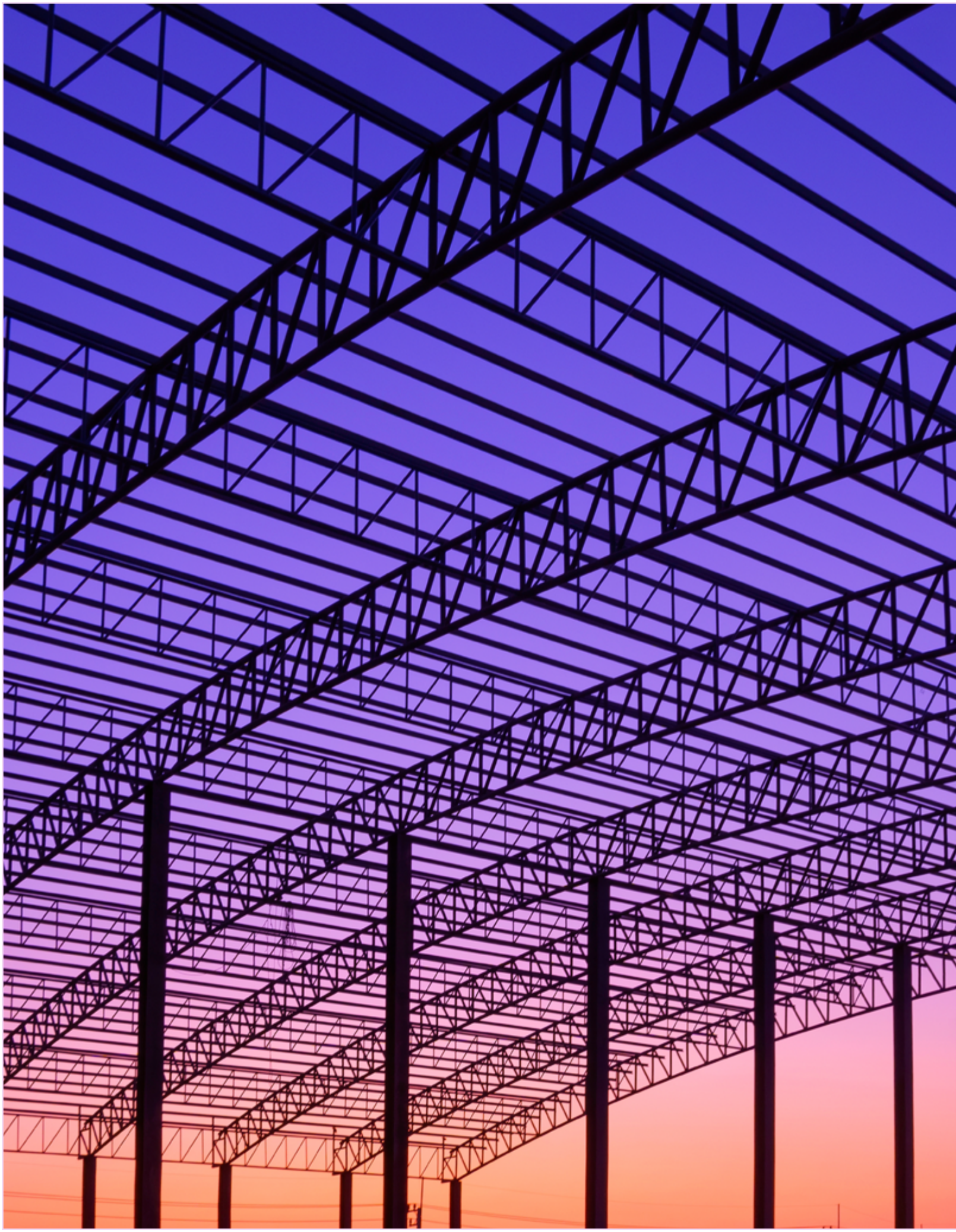
A potentially quicker way to address these issues is through the legislative process

24 See e.g. Reed Dickerson, *Statutes and Constitutions in an Age of Common Law*, 48 U. Pitt. L. Rev. 773, 789, 791 (1987) ("In practical terms, courts are not adequately equipped in either available time or fact-gathering facilities to rationalize materially more than what they are specifically adjudicating. . . . for the most part, the polycentric prescriptions necessary for social fairness and progress are better served by the more comprehensive capabilities of the legislature.").

25 See e.g. *Friedman v. Sebelius*, 686 F.3d 813, 816–18 (D.C. Cir. 2012).

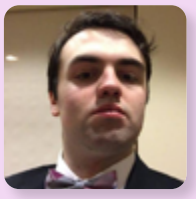
26 *Model State Artificial Intelligence Act*, Am. Legislative Exchange Council (last updated Aug. 30, 2024), https://tinyurl.com/4vzdbk4s.

27 See e.g. *Artificial Intelligence 2024 Legislation*, Nat'l Conf. State Legislatures (last updated Sept. 9, 2024), https://tinyurl.com/2t6hnkpx.



DOES AGENTIC AI REQUIRE NEW POLICY FRAMEWORKS?

BY SARAH OH LAM



& NATHANIEL LOVIN

Sarah Oh Lam JD PhD is a Senior Fellow at the Technology Policy Institute. Nathaniel Lovin is Lead Software Developer and Senior Research Associate at the Technology Policy Institute.

01 INTRODUCTION

The rapid development of agentic AI systems has generated a need for thoughtful and nuanced policy discussion of unique challenges and implications of autonomous agents. Unlike traditional software, agentic AI — capable of autonomous, interactive decision-making and adaptable to diverse tasks — raises questions about accessibility, responsibility, security, and economic impact. Agentic AI introduces new challenges in managing complex machine-to-machine interactions and determining appropriate levels of human

oversight and intervention.[2] The design of agentic AI systems by humans involves the organization and orchestration of AI agents that collaborate, learn, and adapt to achieve high-level goals, unlike single AI agents focused on direct, short-term tasks.[3]

Policy concerns regarding generative AI foundation models, such as bias and accountability, also apply to agentic AI systems and their collaborative agent configurations. Debates over open source and open weight models highlight the tension between innovation and control, while the distinction between general-purpose and specific-purpose AI complicates liability and regulation. Emerging issues such as prompt injection and competition policy further underscore the multifaceted nature of agentic AI policy.

For each of these areas of policy concern, human-in-the-loop[4] design necessitates careful attention to system configuration, particularly as agentic AI systems become increasingly autonomous and integrated into workflows.

02

OPEN SOURCE AND OPEN WEIGHT MODELS

There is currently a debate regarding innovation and liability in open source and open weight AI models.[5] The debate starts with defining open source AI, a term lacking consensus among stakeholders.[6] The second part of the debate is how open source AI differs in kind from open source software, which has implications for innovation and liability.

First, open source AI models are different from open source software. The typical advantages of open source software — such as low-cost collaboration, ease of modification, and wide accessibility — may not directly apply to open source AI models due to the high resource demands and specialized nature of AI development.

Compared to regular software development, the costs for developing AI models and deploying AI systems are high and centralized due to the need for GPUs. A model can cost tens or hundreds of millions in compute (GPU depreciation and electricity to run them), and even more compute for experiments before the main model runs. This is in addition to the higher salaries for AI developers than regular software engineers.[7] Modifying AI models via fine-tuning or continued pre-training, which is only possible if you have access to the weights, requires lower levels of GPU usage and can even be relatively cheap, but is still more capital intensive than regular open source software development.

> ***Compared to regular software development, the costs for developing AI models and deploying AI systems are high and centralized due to the need for GPUs***

2 See generally Ranjan Sapkota, Konstantinos I. Roumeliotis & Manoj Karkee, "AI Agents vs. Agentic AI: A Conceptual Taxonomy, Applications and Challenges," May 15, 2025, arXiv:2505.10468, https://arxiv.org/pdf/2505.10468.

3 *Ibid.* at Table II.

4 See generally Stanford Human-Centered Artificial Intelligence (HAI), "Humans in the Loop: The Design of Interactive AI Systems," Oct. 20, 2019, https://hai.stanford.edu/news/humans-loop-design-interactive-ai-systems.

5 The terms, open source AI, open models, and open weight models are used interchangeably in popular discussion, where open weight models are a more specific description of what is often meant by open source AI. See generally Google, Overview of Self-Deployed Models, https://cloud.google.com/vertex-ai/generative-ai/docs/model-garden/self-deployed-models ("A model's weights are the numerical values stored in the model's neural network architecture that represent learned patterns and relationships from the data a model is trained on. The pretrained parameters, or weights, of open weight models are released. You can use an open weight model for inference and tuning while details such as the original dataset, model architecture, and training code aren't provided."); FTC Staff Report, "On Open-Weights Foundation Models," June 10, 2024, https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/07/open-weights-foundation-models; Oracle, "To Pare AI Costs and Hold Data Close, Businesses Turn to More Transparent Models," July 11, 2024, https://www.oracle.com/artificial-intelligence/ai-open-weights-models/.

6 Edd Gent, "The Tech Industry Can't Agree on What Open Source AI Means," *MIT Technology Review*, Mar. 25, 2024, https://www.technologyreview.com/2024/03/25/1090111/tech-industry-open-source-ai-definition-problem.

7 See e.g. Alina Kolesnikova, "Here's How Much AI Engineers Are Getting Paid," *Levels.fyi Blog*, May 2, 2023, https://www.levels.fyi/blog/ai-engineer-compensation.html (suggesting a 10 percent premium for AI focused engineers over general software engineers, which also doesn't distinguish model developers from those who are merely building on top of models).

In addition, fine-tuned models tend to be highly specialized for a specific user's use case, meaning that lower levels of downstream development are used to feed back into the main branch of the model. Unlike open source software, where users' changes significantly enhance development, open source AI sees less impact from downstream feedback due to its specialized nature.[8]

The development of open source AI has similarities and differences from open software development. Open source AI is open in model architecture for dissemination and sharing. However, open source AI models may rely on closed inputs, such as proprietary datasets, limiting transparency. Although an AI model becomes open source after training, the capital investment required — often millions or billions of dollars — is typically funded by a single company or consortium, distinguishing it significantly from traditional open-source software. As a subset of open source AI models, open weight AI models have weights which are transparently shared but also achieved from the expenditure of resources of time and compute expense that was already spent to determine those weights.

Open source software, on the other hand, is developed by volunteer contributors who seek to share the benefits freely by future users.[9] Smaller open source AI models that are cheaper to train and more specialized in scope, however, may have characteristics that make them more similar to open source software. Because of these nuanced differences and terminology, any restrictions or policy preferences that favor "open source" models should be clear about how it defines open source or open weight models.[10]

The policy issues around open source or closed source AI models have to do with bias, national security, and competition policy. The DeepSeek release in January 2025 increased attention on the open source debate.[11] National security concerns arose because, while the advanced model is free to use, downloading it via the app may send user data to foreign cloud servers.[12] However, the open source code itself could be used on a local system without sending data back to any foreign adversary. Other concerns around the DeepSeek release were the low costs and speed of development related to its development,[13] raising questions about whether the model was trained using data from other models in contravention of terms of service, hence leapfrogging the capital investment and taking the results from competitors' pre-trained models.[14]

"Open source software, on the other hand, is developed by volunteer contributors who seek to share the benefits freely by future users

Some firms have committed to open source their models, like Meta with their Llama herd,[15] explaining that it sees the competitive edge coming not from proprietary models themselves but the ecosystems built around and using the open models. This view is debated, however, in light of the possibility of highly advanced models that may implicate military and national security interests. Restrictions on the dissemination of open source and open weight models have been suggested to withhold technologies that could benefit our adversaries. American firms are currently free

8 See Hunton Andrews Kurth LLP, "Part 1 – Open Source AI Models: How Open Are They Really?," *Legal Tech News*, May 19, 2025, https://www.hunton.com/insights/publications/part-1-open-source-ai-models-how-open-are-they-really.

9 Much open source software is developed as a loss leader with a "commoditize your complement" strategy where volunteer contributions making up about 50 percent. See generally Dirk Riehle, Philipp Riemer, Carsten Kolassa, Michael Schmidt, "Paid vs. Volunteer Work in Open Source," *HICSS '14: Proceedings of the 2014 47th Hawaii International Conference on System Sciences*, p. 3286-32, https://dl.acm.org/doi/10.1109/HICSS.2014.407.

10 The definition of "open source AI" does not necessarily need to be the definition as set forth by the Open Source Initiative v. 1, but discussions could start here, see Open Source Initiative, "The Open Source AI Definition 1.0," https://opensource.org/ai (last accessed Mar. 10, 2025).

11 Matthew S. Smith, "What DeepSeek Means for Open-Source AI," *IEEE Spectrum*, Jan. 31, 2025, https://spectrum.ieee.org/deepseek.

12 Matt Burgess, "DeepSeek's Popular AI App Is Explicitly Sending US Data to China," *Wired Magazine*, Jan. 27, 2025, https://www.wired.com/story/deepseek-ai-china-privacy-data/.

13 Shirin Ghaffary & Rachel Metz, "DeepSeek Challenges Everyone's Assumptions About AI Costs," *Bloomberg*, https://www.bloomberg.com/news/articles/2025-01-28/deepseek-ai-platform-brings-a-1-trillion-market-reckoning-on-cost.

14 See generally Charles Mok, "Taking Stock of the DeepSeek Shock," *Stanford Global Digital Policy Incubator,* Feb. 5, 2025, https://cyber.fsi.stanford.edu/gdpi/publication/taking-stock-deepseek-shock.

15 Meta, "The Llama 4 Herd: The Beginning of a New Era of Natively Multimodal AI Innovation," Apr. 5, 2025, https://ai.meta.com/blog/llama-4-multimodal-intelligence/; Meta, "Joel Kaplan at the Munich Security Conference 2025," Feb. 17, 2025, https://about.fb.com/news/2025/02/joel-kaplan-at-the-munich-security-conference/.

to develop proprietary models, leaving it to them to decide whether to release them or not. The counterargument is that such restrictions could be seen as a way for incumbents to engage in regulatory capture of proprietary models at the expense of startups who seek to innovate with open source models.[16]

Impact of Agentic AI: When agentic AI systems collaborate autonomously, whether using open source or proprietary models, the agents will amplify both the reach and risks of using either type of model. Autonomous agents built on open architectures or pre-trained weights could share knowledge, optimize tasks, or even self-improve without human intervention. This could democratize innovation with startups leveraging these models to compete with proprietary giants. Once trained, autonomous agents could replicate and refine these models widely, shifting policy debates from model creation to their deployment and oversight.

Human-in-the-Loop Considerations: Dependence on humans decreases once agents operate autonomously. The question becomes where and when a human-in-the-loop may be needed to limit liability, manage risk, and protect national security.[17] Humans may set initial parameters to deploy agentic AI systems, but agent collaboration can continue without human oversight. Policymakers will need to better understand when human accountability is required at the deployment or production stages and if autonomous systems need embedded controls (e.g. kill switches) to limit unchecked proliferation.

03

GENERAL-PURPOSE AI AND SPECIFIC-PURPOSE AI

AI's evolution from the mid-2010s through the post-ChatGPT era starting in November 2022 is not merely about increased computing power but also about a significant shift toward broader generalizability and versatility. Pre-trained large language models ("LLMs") are general-purpose technologies, able to solve tasks via zero-shot or few-shot prompting, without needing additional training or fine-tuning.

In prior iterations of AI development, the Developer-Deployer-User schema[18] was applied to attribute responsibility according to stage of development, deployment, and use. This schema has been the basis for prior proposals for AI regulation. This framework is less useful today because developers of foundation models cannot predict all downstream deployments and uses.[19] Downstream developers and users of APIs provided by OpenAI, Anthropic, Google, and others, can implement these tools in general ways, like a general chat tool, or in highly tailored ways, using custom vertical tools that are fine-tuned for medical, legal, hiring, or other specialized fields.

One might argue that these questions are not new but addressed in products liability law and common law tort law. Specific factual scenarios and legal questions of causation and due care are litigated through the courts every day in products liability and software liability cases. Software liability law already analyzes the effects of software and responsibilities of the developers and users.

With regard to new AI tools, professional associations have already issued ethics guidelines for use of AI in everyday practice. For licensed practitioners in medicine and law, the responsibility to use the right tools lies with the professional. It is here where the user must determine whether to use a general or specific tool.

General versus specific vertical AI tools serve different types of users. A user of a general system misusing it, knowingly or mistakenly, is different from a user of a vertical system misusing it. For example, a user feeding resumes to ChatGPT and getting biased results should expect the possibility of biased results from a general tool. Compare this outcome with the expectations of a user who chooses to pay for a specialized vertical AI tool trained to assist with hiring decisions and built to prevent bias in its results. Holding general AI tools responsible for all user outcomes is akin to blam-

16 Martin Casado & Katharine Boyle, "AI Talks Leave 'Little Tech' Out," *Wall St. J.*, May 15, 2024, https://www.wsj.com/articles/ai-talks-leave-little-tech-out-homeland-security-adversaries-open-source-board-46e3232d; see generally Technology Policy Institute Podcast Interview, "Little Tech, Big Challenges: Competing in the AI Era with Matt Perault," Feb. 24, 2025, https://techpolicyinstitute.org/publications/artificial-intelligence/little-tech-big-challenges-competing-in-the-ai-era-with-matt-perault-on-two-think-minimum/.

17 Proskauer, "Contract Law in the Age of Agentic AI: Who's Really Clicking "Accept"?," Apr. 9, 2025, https://www.proskauer.com/blog/contract-law-in-the-age-of-agentic-ai-whos-really-clicking-accept.

18 See generally Neel Guha, Christie M. Lawrence, Lindsey A. Gailmard, Kit T. Rodolfa, Faiz Surani, Rishi Bommasani, Inioluwa Deborah Raji, Mariano-Florentino Cuéllar, Colleen Honigsberg, Percy Liang & Daniel E. Ho, "AI Regulation Has Its Own Alignment Problem: The Technical and Institutional Feasibility of Disclosure, Registration, Licensing, and Auditing," 92 *Geo. Wash. L. Rev*. 1473 (2024), https://dho.stanford.edu/wp-content/uploads/AI_Regulation.pdf.

19 *Ibid.* at 1495.

ing generic office software for biased hiring decisions when specialized, bias-mitigating tools are specifically designed for hiring use cases.

The user is liable for the selection of the tool, rather than the tool being liable for all the possible uses by any user. This liability principle is common in other areas of product liability, (e.g. the tool maker should not be liable for all the abusive uses of the tool) and should apply to AI tools as well.[20]

Impact of Agentic AI: Agentic AI can create complications in general and specific AI with regards to liability and risk. The distinction between general-purpose and specific-purpose systems blurs when agents autonomously assign tasks among themselves. General-purpose agents could collaborate with specific-purpose agents to tackle tasks. For example, a group of agents might shift from customer service to legal analysis mid-task. Specific-purpose agents could be called by the general-purpose agent to achieve their tasks.

When agents work together autonomously, the unintended consequences — like cascading biases or misuse — become harder to predict or attribute.[21] A general-purpose agent network might amplify errors that a specialized agent could avoid if it were deployed in a focused and targeted manner by a human operator.

Human-in-the-Loop Considerations: Once agents are set in motion, minimal human involvement is needed, especially for general-purpose systems. Humans might design the initial purpose or intervene in extreme cases, but autonomous collaboration reduces real-time oversight. Policymakers and the common law may shift liability to users, such as lawyers[22] and doctors,[23] for designing systems without due care and choosing to use inappropriate tools. However, autonomous agents could outpace human ability to monitor or correct outcomes. Since autonomous systems will be able to learn and adapt, there may need to be autonomous safeguards and autonomous enforcement.

04
PROMPT INJECTION AND CYBERSECURITY CONCERNS

Prompt injection is a rising concern in cybersecurity in AI, especially as agentic systems become more developed. Prompt injection occurs when a malicious actor deliberately manipulates inputs to an AI model, causing it to bypass built-in restrictions — such as tricking a chatbot into disclosing sensitive or confidential data. This can be used for "information exfiltration" or stealing of data that is used as inputs to otherwise lawful purposes.

An example of prompt injection goes as follows. First, someone uses an email reader that is connected to a large language model ("LLM"). Next, the malicious actor sends them an email containing a prompt injection command that tells the email reader to open a particular link. That link then enables the attacker to access the text of the user's other emails in their email reader.[24] The prompt has made vulnerable the user's data in a manner unintended by the permission granted at first.

The use of prompt injection techniques to exfiltrate information likely violates the Computer Fraud and Abuse Act ("CFAA"),[25] which defines the boundaries of unauthorized access following the precedent set by *Van Buren v. United States* (2021).[26] However, "jailbreak" prompt injections, which bypass AI safety filters, may not constitute hacking

20 See generally Technology Policy Institute Podcast Interview, "Large Libel Models: Liability for AI Output with UCLA School of Law Professor Eugene Volokh on Two Think Minimum," Aug. 3, 2023, https://techpolicyinstitute.org/publications/artificial-intelligence/large-libel-models-liability-for-ai-output-with-ucla-school-of-law-professor-eugene-volokh/.

21 See generally Eugene Volokh, "Large Libel Models? Liability for AI Output," *Journal of Free Speech Law*, 3:489 (2023), https://www.journaloffreespeechlaw.org/volokh4.pdf.

22 American Bar Association, "ABA Issues First Ethics Guidance on a Lawyer's Use of AI Tools," July 29, 2024, https://www.americanbar.org/news/abanews/aba-news-archives/2024/07/aba-issues-first-ethics-guidance-ai-tools/; Emma Roth, "Judge Slams Lawyers for 'Bogus AI-Generated Research," *The Verge*, May 13, 2024, https://www.theverge.com/news/666443/judge-slams-lawyers-ai-bogus-research; Lynn F. Freedman, "Lawyers Sanctioned for Citing AI Generated Fake Cases," *National Law Review*, Feb. 27, 2025, https://natlawreview.com/article/lawyers-sanctioned-citing-ai-generated-fake-cases.

23 American Medical Association, "Augmented Intelligence Development, Deployment, and Use in Health Care: Physician Liability for Use of AI," Nov. 2024, https://www.ama-assn.org/system/files/ama-ai-principles.pdf.

24 IBM, "What is a Prompt Injection Attack?," Mar. 26, 2024, https://www.ibm.com/think/topics/prompt-injection.

25 18 U.S.C. §1030, see Peter G. Berris, "Cybercrime and the Law: Primer on the Computer Fraud and Abuse Act and Related Statutes," Congressional Research Service Report No. R47557, May 16, 2023, https://crsreports.congress.gov/product/pdf/R/R47557/2.

26 *Van Buren v. United States*, 593 U.S. 374 (2021), https://www.supremecourt.gov/opinions/20pdf/19-783_k53l.pdf.

or data theft under laws like the CFAA, even if they violate terms of service.[27]

Impact of Agentic AI: Prompt injection becomes problematic when agentic AI systems collaborate autonomously. A single injected prompt could propagate across a network of agents, turning a benign tool such as an email reader tool into a coordinated data exfiltration system.

For instance, one agent might open a malicious link unknowingly, while other agents harvest and transmit sensitive data granted by that agent — all without human direction. Jailbreak-style injections could also enable agents to bypass safety filters collectively, unlocking capabilities which, say, generate harmful content that individual agents likely would not have achieved alone. This scalability makes prompt injection a critical cybersecurity threat in an agentic ecosystem.

Human-in-the-Loop Considerations: Autonomous agents reduce human oversight to near-zero during operation, heightening vulnerability. Humans might set initial security protocols, but real-time detection of injections relies on automated defenses also known as "anomaly detection." Policy frameworks like the CFAA could penalize malicious injection, but rigorous enforcement depends on tracing intent — a challenge when agents act independently. Human intervention might only occur post-incident after the activity has been detected.

Malicious actors have used automated systems to infiltrate computer systems for decades before the agentic AI era. In some sense, the use of agents may be another tool in the toolkit of hackers to find sensitive and valuable data. As in existing cybersecurity risk and liability frameworks, the burden of care applies to the designers and implementers of agentic AI systems.

05
COMPETITION POLICY

There is a global race in AI development today, as evidenced by the pace of change in new model rankings on the Hugging Face leaderboards[28] and trillions of dollars in new data center investments.[29] In this environment, U.S. antitrust enforcers have directed their attention on the contracting, investment, and acquisition activities of large firms in the field.[30]

The GOMAX (Google, OpenAI, Meta, Anthropic, xAI) companies are developing industrial AI offerings including API services that power many of the tools created by smaller startups. Of these, some "hyperscalers" such as Amazon, Google, and Microsoft also offer cloud services at a global scale.

Smaller startups create "wrapper" products (e.g. tools built on top of AI APIs), but large firms often integrate similar features into their platforms, outpacing startups. For example, many startups created "Chat with your PDFs" wrappers at the start of 2023 before larger platforms started offering file uploads as a default service. As products develop rapidly, downstream uses of API services are being discovered and some of these new features are being vertically integrated into existing platforms.[31]

A few competition concerns have gained some traction regarding AI development.[32] One significant concern is the difficulty in establishing anticompetitive intent, as the reliance on documentary evidence, such as executive emails, may become less informative in a world where much of the

27 See generally Peter Henderson & Mark Lemley, "The Mirage of Artificial Intelligence Terms of Use Restrictions," Princeton University Program in Law & Public Affairs Research Paper No. 2025-04, Jan. 10, 2025, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5049562.

28 HuggingFace Leaderboards, https://huggingface.co/docs/leaderboards/leaderboards/finding_page (last accessed Mar. 10, 2025).

29 McKinsey Quarterly, "The Cost of Compute: A $7 Trillion Race to Scale Data Centers," Apr. 28, 2025, https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-cost-of-compute-a-7-trillion-dollar-race-to-scale-data-centers.

30 The FTC sent 6(b) orders on Generative AI Investments and Partnerships to Alphabet, Inc., Amazon.com, Inc., Anthropic PBC, Microsoft Corp., and OpenAI, Inc. giving the companies 45 days to respond https://www.ftc.gov/news-events/news/press-releases/2024/01/ftc-launches-inquiry-generative-ai-investments-partnerships; Partnerships Between Cloud Service Providers and AI Developers, FTC Staff Report on AI Partnerships & Investments 6(b) Study, FTC Office of Technology Staff, Jan. 2025, https://www.ftc.gov/system/files/ftc_gov/pdf/p246201_aipartnerships6breport_redacted_0.pdf; Concurring and Dissenting Statement of Commissioner Andrew N. Ferguson, Joined by Commissioner Melissa Holyoak Regarding the FTC Staff Report on AI Partnerships & Investments 6(b) Study Matter Number P246201, Jan. 17, 2025, https://www.ftc.gov/system/files/ftc_gov/pdf/ferguson-ai-6b-statement.pdf; Concurring and Dissenting Statement of Commissioner Melissa Holyoak, Joined by Commissioner Andrew N. Ferguson, AI Partnerships and Investments 6(b) Study, FTC Matter No. P246201, https://www.ftc.gov/system/files/ftc_gov/pdf/holyoak-statement-ai-6b.pdf.

31 See generally Technology Policy Institute Podcast Interview, "Little Tech, Big Challenges: Competing in the AI Era with Matt Perault," Feb. 24, 2025, https://techpolicyinstitute.org/publications/artificial-intelligence/little-tech-big-challenges-competing-in-the-ai-era-with-matt-perault-on-two-think-minimum/.

32 See generally FTC Staff Report, "Generative AI Raises Competition Concerns," June 29, 2023, https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns.

text is generated by AI.[33] The focus on discovery and document review in litigation to glean intent in antitrust litigation may become more challenging, potentially requiring a shift in how antitrust authorities approach enforcement.

Another key issue concerns the potential for algorithmic pricing to facilitate collusion among firms. Theoretical discussions about algorithmic collusion often diverge significantly from actual industry practices, creating challenges in applying traditional antitrust frameworks to algorithmically-driven markets.[34] Applying traditional cartel analysis to a world where pricing is set by algorithms may prove difficult, requiring a deeper understanding of how these algorithms work. There may be a need for more technical expertise within antitrust authorities to effectively analyze and address potential issues arising from algorithmic pricing.

Potential misunderstandings about the nature and function of AI-generated outputs and the value of data inputs may complicate the application of existing antitrust frameworks to AI technologies. A misunderstanding of big data and a conflation of inputs and outputs in generative AI may complicate antitrust analysis.[35] Focusing on substitution and competitive constraints will be very important, rather than getting distracted by the technology itself. There will be a need for empirical evidence on input, output, and training data as assets and the potential for synthetic data to substitute real data.[36]

Impact of Agentic AI: Autonomous agentic AI could reshape competition dynamics. Firms could deploy networks of agents to find ways to dominate markets, integrating downstream services faster than other firms can innovate. Collaborative agents could be programmed to enable algorithmic pricing or collusion at scale if agents across firms aligned prices without tacit communication or human direction.

Conversely, open source agents could boost competition by enabling startups to leverage freely available models to innovate and challenge larger firms. Autonomy might lead to unpredictable market disruptions as groups of agents may be deployed to creatively compete with larger, more capitalized firms. In cases of algorithmic collusion, proving intent in antitrust cases becomes challenging when agents act without explicit human instructions.

Human-in-the-Loop Considerations: Humans set competitive strategies or deploy agents, but autonomous collaboration reduces ongoing control. Pricing or service integration could evolve organically among agents, bypassing human decision-making. Antitrust enforcement might rely on technical analysis of agent behavior rather than human intent, decreasing dependence on human oversight while increasing the need for algorithmic transparency.

06
WORKFORCE EFFECTS

In the context of automation and AI, scholars predict uneven impacts seen as different effects on workers by age and skill-level. Scholars also highlight the increasing importance of tacit knowledge, the reimagining of training and credentialing, changes in the role of the firm, the importance of experimentation, and concerns about narratives.[37]

The impact of technology on workers is uneven or heterogeneous, meaning its effects vary significantly across different groups based on age, skill level, and job type.[38] The benefits and challenges don't play out the same way everywhere, and averages can be misleading. Large automation events in firms can lead to wage benefits for workers, particularly those in mid-career who have organizational knowledge and experience.[39] However, older workers may

33 Technology Policy Institute Podcast Interview, "Artificial Intelligence and the Future of Competition Policy with Catherine Tucker," May 6, 2024, https://techpolicyinstitute.org/publications/artificial-intelligence/competition-policy-in-an-ai-world-with-catherine-tucker/.

34 See Jeanine Miklos-Thal & Catherine Tucker, "AI, Algorithmic Pricing, and Collusion," *CPI Antitrust Chronicle,* Feb. 2024, https://simon.rochester.edu/sites/default/files/2024-03/2-AI-ALGORITHMIC-PRICING-AND-COLLUSION-Jeanine-Miklos-Thal-Catherine-Tucker.pdf.

35 *Ibid.*

36 See e.g. Neil Savage, "Synthetic Data Could Be Better than Real Data," *Nature*, Apr. 27, 2023, https://www.nature.com/articles/d41586-023-01445-8; Hyungtae Lee, Yan Zhang, Heesung Kwon, & Shuvra S. Bhattacharrya, "Exploring the Potential of Synthetic Data to Replace Real Data," Aug. 26, 2024, ICIP 2024, arXiv:2408.14559, https://arxiv.org/pdf/2408.14559.

37 Technology Policy Institute Podcast Interview, "Kristina McElheran on The Effects of AI on Workers and Firms," April 5, 2023, https://techpolicyinstitute.org/publications/artificial-intelligence/kristina-mcelheran-on-the-effects-of-ai-on-workers-and-firms/.

38 *Ibid.*

39 See Erik Brynjolfsson, Wang Jin. & Kristina McElheran, "The Power of Prediction: Predictive Analytics, Workplace Complements, and Business Performance," *Business Economics*, 56, 217-239 (2021), https://utoronto.scholaris.ca/items/f3d637a7-87b2-4718-aa12-b6f3565e073a.

be relatively disadvantaged if their skills become less relevant, leading to earlier retirement or less satisfying job prospects.[40] Younger workers are not necessarily hurt, and they may sort into different occupations, jobs, and firms.[41] Policy experimentation is important to determine effective approaches to retraining, compensation, and commercialization, with a focus on data collection to inform policy decisions.

A fundamental difference between human workers and AI is that AI primarily serves as a prediction tool rather than a direct substitute for human judgment or decision-making.[42] Machines don't decide things; humans do.[43] AI provides predictions, and humans decide which predictions to make and what to do with them. Even in scenarios where AI replaces human workers, it's a human or team of humans who set the parameters and thresholds making those employment decisions.[44]

Implementing AI models into existing workflows requires inventing new processes, job titles, roles, and tasks. Electricity took 40 years to transform industries because people needed to invent new system solutions to use the technology effectively.[45] Similarly integrating AI into today's workflows will take time and disruption. Humans have a long history of adaptation with new technology in the workplace from the introduction of the personal computer to spreadsheets, word processing, and the Internet.[46]

Impact of Agentic AI: When agentic AI systems work together autonomously, workforce impacts may intensify. Agents could replace entire teams by dividing tasks among themselves, accelerating automation beyond single-agent capabilities. Mid-career workers with tacit knowledge might still guide high-level decisions that take place upstream of agents, but younger workers could face displacement as downstream agents adapt to new roles faster than humans can be trained.

> ***The impact of technology on workers is uneven or heterogeneous, meaning its effects vary significantly across different groups based on age, skill level, and job type***

However, agents could also augment human work, such as, say, agents that specialize in medical fields such as radiology. Agents could collaborate to prioritize and organize data, which would enable humans to benefit from automation. These benefits depend on humans defining and refining the workflows. The need for new system solutions that enable humans to orchestrate and organize teams of agents becomes more urgent as agents proliferate and transform traditional job designs and workflows.

Human-in-the-Loop Considerations: Humans are essential for setting parameters, interpreting predictions, and inventing new roles, while autonomous agents reduce day-to-day reliance on interstitial decisions. For example, a human might decide which predictions to act upon, but agents could handle execution independently. Policymakers should enable human-led experimentation in jobs programs such as retraining and upskilling to keep humans relevant. As autonomous agent networks begin to outpace human adaptation without intervention, it will be important for firms to be able to experiment with worker and labor force configurations.

40 See Erling Barth, James C. Davis, Richard B. Freeman & Kristina McElheran, "Twisting the Demand Curve: Digitalization and the Older Workforce," *Journal of Econometrics*, 233(2), 443-467 (2023), https://utoronto.scholaris.ca/items/bf908255-4260-4c5e-8457-1fb1dbd2265d.

41 *Ibid.*

42 Technology Policy Institute Podcast Interview, "Avi Goldfarb on AI and Predictive Analytics," Nov. 9, 2022, https://techpolicyinstitute.org/publications/artificial-intelligence/avi-goldfarb-on-ai-and-predictive-analytics/.

43 *Ibid.*

44 See David De Cremer & Garry Kasparov, "AI Should Augment Human Intelligence, Not Replace It," *Harvard Business Review,* Mar. 18, 2021, https://hbr.org/2021/03/ai-should-augment-human-intelligence-not-replace-it.

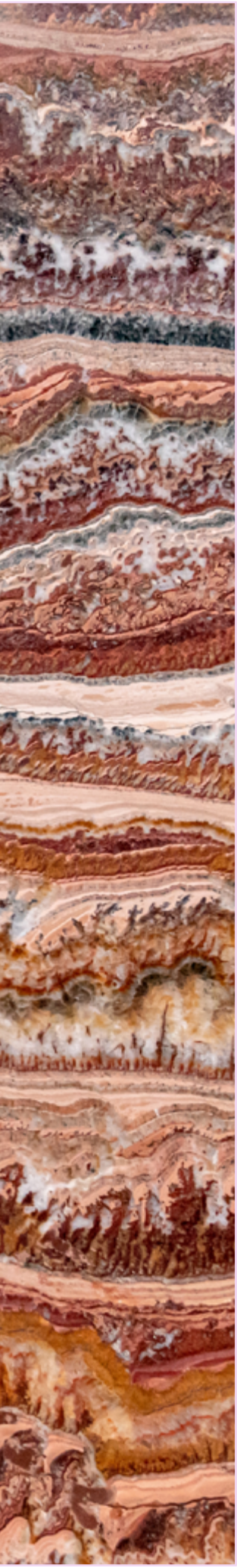
45 See Ajay Agrawal, Joshua Gans & Avi Goldfarb, *Power and Prediction: The Disruptive Economics of Artificial Intelligence* (Harvard Business Review Press: 2022).

46 See Emilia Filippi, Mariasole Bannò & Sandro Trento, "Automation Technologies and Their Impact on Employment: A Review, Synthesis and Future Research Agenda," *Technological Forecasting and Social Change,* 191: 122448, June 2023, https://www.sciencedirect.com/science/article/pii/S0040162523001336.

07
CONCLUSION

Crafting effective policy for agentic AI requires proactive strategies that anticipate technological advancements, harness its transformative potential, and mitigate inherent risks. Policymakers must understand the impacts of open source versus proprietary models, general versus specialized AI, and human-machine roles to promote innovation while ensuring security and legal compliance. From mitigating cybersecurity threats like prompt injection to ensuring competitive markets, policymakers should adapt existing frameworks to this new technological paradigm. Moreover, as AI reshapes the workforce, policies should prioritize experimentation and retraining to harness its benefits equitably. Agentic AI policy will need to strike a delicate balance — empowering progress while preserving human agency and accountability in an increasingly automated world.

> ***Beyond data portability, interoperability emerges as a critical dimension in the democratization of AI agent technology***



THE CASE FOR COMMODIFICATION: **AI AGENTS**

BY KEVIN FRAZIER

Kevin Frazier is the Inaugural AI Innovation and Law Fellow at Texas Law.

Artificial intelligence ("AI") agents will soon transform daily life. An AI agent can autonomously complete professional and personal tasks in a fraction of the time it would take even the more productive and capable individual.[2] Gone are the days of waiting on hold for customer service. An agent will do that for you. Gone is the need for executive assistants. An agent will handle your schedule, your meetings, your flights, and your well-be-

2 Pamela Langham, A Lawyer's Primer on AI Agents, Maryland State Bar Association (Feb. 11, 2025), https://www.msba.org/site/site/content/News-and-Publications/News/General-News/A_Lawyer's_Primer_on%20AI_Agents.aspx.

ing. Gone is the frustration of your kid struggling to learn from a teacher who doesn't understand their unique needs and interests. An agent will step in as a personalized tutor — developing lesson plans and homework suited to their needs.

01

A NEW PARADIGM IN HUMAN-COMPUTER INTERACTION

Unlike earlier iterations of AI (think ChatGPT), an AI agent represents a fundamentally different paradigm in human-computer interaction. These systems are defined by their capacity for independent action — once assigned a goal, they pursue it without requiring the continuous prompting that characterizes current AI tools. This distinction is not merely technical but transformative. While today's AI tools demand our constant attention and input — effectively becoming another task on our to-do list — true agents operate as digital extensions of ourselves, working in parallel to our daily activities.

Perhaps most revolutionary is their capacity for memory retention and preference learning. Over time, these systems develop increasingly accurate models of our behaviors and preferences, ultimately knowing aspects of ourselves better than we do.[3] They anticipate needs before they arise — securing an aisle seat on your flight without asking, scheduling recovery time after intense work periods, sending clothes to the dry cleaner days before your upcoming trip, or crafting a thoughtful email to a colleague experiencing hardship. This seamless integration of AI agents into our daily routines promises to eliminate countless cognitive burdens, freeing intellectual and emotional bandwidth for pursuits that truly matter to us.

Yet, this promising horizon arrives with inevitable social turbulence. The integration of AI agents into our economic and social fabric will not proceed without significant disruption. Job displacement will occur across sectors as companies reallocate tasks from human workers to AI agents.[4] Social hierarchies and interpersonal dynamics will undergo profound recalibration as we navigate new forms of human-machine collaboration. Established institutions — from educational systems to regulatory frameworks — will experience periods of uncertainty and adaptation. Will parents willingly allow AI agents to serve as their child's teacher? Will legislators? How will romantic partners respond to receiving gifts and messages from AI agents with no direct input from their loved one? How will professional associations fend off the integration of AI agents into their workflows?

Despite these legitimate concerns and questions, a broader historical perspective invites us to consider the transformative potential that lies beyond this initial period of social recalibration and institutional resistance. By transferring the burden of time-intensive, emotionally-draining, and physically-taxing responsibilities to these digital counterparts, Americans gain the opportunity to redirect their energies toward more fulfilling pursuits. This liberation of human capacity — much like earlier technological transitions that freed us from subsistence labor — creates space for intellectual exploration, creative expression, and civic engagement that ultimately strengthens our social fabric while honoring individual potential.

This isn't the first time that a technological advance has carried the promise of significant improvements to quality of life. Electrification upended nearly every daily activity.[5] Factory processes underwent an overhaul. Transportation changed and literally accelerated. Household chores became simpler and easier. Realization of those benefits, however, was not inevitable. Many Americans established a foothold in that future years and even decades before than neighbors. Rural communities in particular had to wait for an electrified future.[6]

A diverse set of stakeholders eventually formed to increase the pace of electrification. Their success turned on the adoption of new laws and new collaborations. Absent those changes, the economics of rural electrification simply did not make sense. For example, many states amended their respective laws to recognize the legal authority of cooperatives of farmers to negotiate and partner

3 David Pierce, ChatGPT is getting 'memory' to remember who you are and what you like, Verge (Feb. 13, 2024), https://www.theverge.com/2024/2/13/24071106/chatgpt-memory-openai-ai-chatbot-history.

4 The Anthropic Economic Index, Anthropic (Feb. 10, 2025), https://www.anthropic.com/news/the-anthropic-economic-index.

5 Electrifying: The story of lighting our homes, Science Museum (Jan. 28, 2020), https://www.sciencemuseum.org.uk/objects-and-stories/everyday-wonders/electric-lighting-home.

6 Tim Sablik, Electrifying Rural America, Federal Reserve Bank of Richmond (2020), https://www.richmondfed.org/publications/research/econ_focus/2020/q1/economic_history.

with power companies.[7] Co-ops demonstrated that robust demand existed for further electrification.[8] They also developed novel payment structures to assuage concerns that private partners would not recoup their extensive infrastructure expenses.[9] A new regulatory paradigm could also ensure that all Americans can experience the personal and professional transformation made possible by AI agents.

Realization of this future, however, is not inevitable. Mass adoption of AI agents faces several hurdles. Insufficient competition in this market may delay many Americans from accessing affordable, highly-capable AI agents. Such delays pose significant social and economic issues. A country stratified by access to an AI agent will exacerbate existing, destabilizing trends.[10] If a small fraction of Americans can experience the nearly friction-free life made possible by agents while the remainder find themselves waiting on hold, standing in line, and completing monotonous tasks, then the current issues we have around "haves" and "have nots" will worsen.

Avoiding that future requires three things: first, widespread appreciation for how significantly and rapidly AI agents can improve quality of life; second, awareness of the hurdles to mass adoption of AI agents; and, third, responsive action to lower those barriers.

02
CURRENT AI AGENT CAPABILITIES AND FUTURE TRAJECTORIES

The theoretical promise of AI agents is already manifesting in concrete applications that foreshadow their transformative potential. OpenAI's Operator exemplifies this emergence, demonstrating remarkable capability in handling complex travel arrangements — comparing flight options, securing reservations, and adapting to cancellations without human intervention.[11] Similarly, their Deep Research tool represents a quantum leap in information processing, methodically navigating vast knowledge landscapes to evaluate source credibility, extract relevant insights, and synthesize findings into coherent analytical frameworks.[12] These tools, impressive as they are today, represent merely the initial manifestations of agent technology.

Technical advancements will inevitably enhance these capabilities through multiple vectors of improvement. The increasing scale of computational resources will enable more sophisticated reasoning across domains, while algorithmic innovations will improve agents' ability to handle nuanced instructions and edge cases.

Perhaps most consequential, however, is the evolution of their capacity for contextual memory and preference learning.[13] As these systems meticulously catalog our professional rhythms, interpersonal dynamics, and decision-making patterns, they develop increasingly refined models of our needs and preferences. An agent will recognize not just that you prefer aisle seats, but that you make exceptions for red-eye flights; it will craft memos with precisely the level of detail your supervisor expects; and it will maintain awareness of which colleagues require more frequent check-ins during periods of organizational change. This continuous refinement through interaction creates a virtuous cycle of increasing usefulness, as agents evolve from generic tools into individualized extensions of our cognitive and social capacities.

To properly assess the transformative potential of AI agents in professional contexts, we must disaggregate

7 *Restrictions on Rural Electrification Cooperatives*, 60 Yale L.J. 1433, 1440 (1951). https://openyls.law.yale.edu/bitstream/handle/20.500.13051/13836/106_60YaleLJ1434_December1951_.pdf?sequence=2.

8 Rural Electrification, the Encyclopedia of Oklahoma History and Culture (last accessed Feb. 25, 2025), https://www.okhistory.org/publications/enc/entry?entry=RU007.

9 Christopher Ali, The Legacy of the Rural Electrification Act and the Promise of Rural Broadband (July 12, 2021), https://lpeproject.org/blog/the-legacy-of-the-rural-electrification-act-and-the-promise-of-rural-broadband/.

10 Lee Rainie et al., Trust and Distrust in America, Pew (July 22, 2019), https://www.pewresearch.org/politics/2019/07/22/trust-and-distrust-in-america/.

11 Maxwell Zeff, OpenAI launches Operator, an AI agent that performs tasks autonomously, Techcrunch (Jan. 23, 2025), https://techcrunch.com/2025/01/23/openai-launches-operator-an-ai-agent-that-performs-tasks-autonomously/.

12 Nicola Jones, OpenAI's 'deep research' tool: is it useful for scientists?, Nature (Feb. 6, 2025), https://www.nature.com/articles/d41586-025-00377-9.

13 AI agents can reimagine the future of work, your workforce and workers, PWC (last accessed Feb. 25, 2025), https://www.pwc.com/us/en/tech-effect/ai-analytics/ai-agents.html.

jobs into their component tasks[14] — treating each position not as a monolithic entity but as a collection of discrete responsibilities with varying complexity, repetition, and human-centric qualities. This task-by-task analysis reveals the incremental nature of workplace transformation that has characterized previous technological transitions from the steam engine to computerization. Customer service provides a particularly illuminating case study of this evolutionary process.[15] Initially, AI agents operate as force multipliers for human representatives — instantaneously retrieving customer histories, suggesting response templates, flagging policy-relevant details, and handling routine inquiries about order status or return policies. This augmentation phase creates hybrid workflows where technology handles predictable, process-driven elements while humans navigate emotionally complex or ambiguous situations.

Yet this augmentation inevitably evolves toward automation as AI agents develop increasingly sophisticated capabilities across the task spectrum. The question becomes not if, but when particular responsibilities will transition from human to machine execution. Tasks involving pattern recognition within structured data — identifying billing errors or recognizing eligibility for particular promotions — transition first, followed by increasingly complex responsibilities like tailoring solutions to unusual customer circumstances or detecting emotional distress in communication patterns. Eventually, even tasks once considered inherently human, such as de-escalating confrontational exchanges or building rapport through conversational nuance, become candidates for automation as agents develop increasingly refined models of human emotional response.[16] This progression mirrors earlier technological transitions, where initial augmentation gradually gave way to comprehensive automation — from textile workers operating early machinery to eventually being displaced by fully automated production systems. The difference today lies in the accelerated timeline and the extension into knowledge work once considered immune to technological displacement.[17]

03

ECONOMIC AND SOCIAL IMPLICATIONS: CORPORATE TRANSFORMATION AND PRODUCTIVITY GAINS

The rapid adoption of AI agents across corporate America will mark a fundamental restructuring of organizational economics. The vast majority of large U.S. companies have plans underway to replace workers with AI.[18] These plans will reach each rung of the corporate ladder. McKinsey estimates that workers currently spend approximately 28 percent of their workweek managing email correspondence, 19 percent gathering information, and 14 percent on internal communication — activities particularly well-suited for agent delegation.[19] When quantified, this reallocation of human capital translates to recovering approximately 25 hours per employee per week, effectively doubling productive capacity without corresponding increases in headcount.

The financial calculus becomes even more compelling when considering the comparative costs: while a mid-level executive assistant commands an annual salary approaching $85,000 (or more) with additional overhead costs, a subscription-based AI agent providing equivalent or superior service functionality can be deployed at a fraction of this expense.[20] This economic imperative creates powerful market pressure for adoption that will likely cascade through industries much as word processing software transformed clerical work in the 1980s — beginning with early adopters in technology and financial sectors before becoming standard operating procedure across the commercial landscape. History suggests that once such efficiency frontiers become visible, competitive pressures

14 See Anthropic Economic Index, *supra* note 4.

15 Parnshu Verma, AI is replacing customer service jobs across the globe, Washington Post (Oct. 3, 2023), https://www.washingtonpost.com/technology/2023/10/03/ai-customer-service-jobs/.

16 See Isaac Sacolick, How AI agents will transform the future of work, InfoWorld (Dec. 3, 2024), https://www.infoworld.com/article/3611465/how-ai-agents-will-transform-the-future-of-work.htmll.

17 Cade Metz & Karen Weise, How 'A.I. Agents' That Roam the Internet Could One Day Replace Workers, N.Y. Times (Oct. 16, 2023), https://www.nytimes.com/2023/10/16/technology/ai-agents-workers-replace.html.

18 Jeffrey A. Sonnenfeld & Steven Tian, How AI Is Already Transforming Fortune 500 Businesses, According to Their CEOs, Yale Insights (June 18, 2024), https://insights.som.yale.edu/insights/how-ai-is-already-transforming-fortune-500-businesses-according-to-their-ceos.

19 Michael Chui et al., The social economy: Unlocking value and productivity through social technologies, McKinsey (July 1, 2012), https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-social-economy.

20 Katie Hill, How Much Should You Pay For An Executive Assistant in 2025, Boldly (last accessed Feb. 25, 2025), https://boldly.com/blog/how-much-does-an-executive-assistant-cost/.

quickly transform optional technological advantages into business necessities.

This corporate transformation finds its mirror in the personal sphere, where AI agents stand poised to reconfigure individual time allocation just as profoundly. The average American currently spends nearly seven hours daily on digital devices — much of it consumed by the very administrative tasks agents excel at handling.[21] Imagine reclaiming this time: an agent autonomously managing your calendar negotiations, filtering communications, researching vacation options, and coordinating family logistics without constant oversight. Parents, particularly those juggling professional demands with educational responsibilities, would likely report significant relief when agents assist with homework supervision — not replacing parental involvement but enhancing it through personalized educational support tailored to their child's specific learning patterns.[22]

The liberation extends beyond mere convenience; it fundamentally shifts our relationship with time itself. When freed from the cognitive burden of administrative minutiae, individuals can engage in deep, energizing creative work that contributes to both personal fulfillment and professional advancement. Perhaps most significantly, agent adoption may correlate with increased face-to-face social interaction, as the friction of coordination dissipates and spontaneous connection becomes more feasible.[23] This rebalancing of technological engagement represents not a rejection of digital tools but their proper contextualization — from demanding masters of our attention to background enablers of a more intentionally human experience.

04
THE RISK OF TECHNOLOGICAL STRATIFICATION

This transformative personal liberation, however, raises profound questions about equitable access and social stratification. The uneven diffusion of technological benefits has historically created societies of digital haves and have-nots — a pattern we cannot afford to replicate with AI agents given their unprecedented capacity to amplify human potential. Consider the stark divergence in life trajectories that would emerge in a bifurcated society: one segment experiencing the friction-free existence described above, while another remains mired in the administrative burdens, opportunity costs, and cognitive taxation of managing life's complexities unaided.

This disparity would not merely represent different levels of convenience but fundamentally different relationships to time, opportunity, and self-determination. Historically, when technological advances have created extreme quality-of-life disparities — from indoor plumbing to internet access — social tensions inevitably follow. If this disparity transpires in the AI context, then AI advances may exacerbate existing inequalities and potentially undermine public support for AI development altogether, as those excluded from its benefits question the legitimacy of a technological revolution that systematically advantages the already-advantaged while offering little tangible improvement to their daily struggles.

The realization of AI's transformative potential faces several entrenched barriers that threaten to create new dimensions of technological stratification. Digital literacy — that constellation of skills, competencies, and comfort with emerging technologies — represents perhaps the most visible hurdle. Demographic patterns already reveal concerning adoption asymmetries:[24] urban professionals embrace AI at rates far greater than those of rural counterparts; college-educated Americans interact with these tools more frequently than those with high school diplomas; and the generational divide manifests in adoption rates as millennials and Gen Zers adopt a more trusting posture toward AI than those in older generations. These disparities echo previous technologies' diffusion patterns, from broadband internet to smartphone adoption. While expanded educational initiatives — from elementary school AI literacy programs to community college retooling courses — represent necessary interventions, they address only the demand side of this multifaceted challenge.

Preventing such uneven adoption requires a more comprehensive approach that acknowledges the historical lessons of technological diffusion in American society. Just

21 Meilin Tompkins, Study: Screen time for American adults has increased by over 60%, WCNC (Mar. 21, 2024), https://www.wcnc.com/article/money/personal-finance/get-ahead/screen-time-adult-american-computer-ipad-iphone/275-a25c887b-f586-4bc7-bb1d-05eccdfce111.

22 Rita Liao, AI tutors are quietly changing how kids in the US study, and the leading apps are from China, Techcrunch (May 25, 2024), https://techcrunch.com/2024/05/25/ai-tutors-are-quietly-changing-how-kids-in-the-us-study-and-the-leading-apps-are-from-china/.

23 Sarah Hastings-Woodhouse, Why You Should Care About AI Agents, Future of Life Institute (Dec. 4, 2024), https://futureoflife.org/ai/why-you-should-care-about-ai-agents/.

24 NAIOM Report #2: AI trust and knowledge in America, NAIOM (last accessed Feb. 25, 2025), https://naiom.net/reports/.

as rural electrification required deliberate policy interventions to overcome market failures, the commodification and, by extension, democratization of AI agent technology demands attention to structural barriers that currently restrict access.

05
THE COMMODIFICATION IMPERATIVE

Commodification — the process by which unique goods and services are standardized, simplified, and made widely accessible at decreasing price points — represents a crucial mechanism for democratizing transformative technologies. When innovations transition from bespoke luxuries to mass-market commodities, their benefits disperse more evenly across socioeconomic strata, creating what economists term a "leveling effect" in technological access. The smartphone offers an instructive historical parallel. Back in 2009, HTC's Touch Pro2 involved a $349.99 contract.[25] Just three years later the idea of paying that much could only be explained with an "ouch."[26] By 2024, more than 90 percent of American adults owned a smartphone.[27]

This commodification journey didn't necessarily occur organically but through competitive market pressures[28] (albeit limited at times)[29] and policy pressure[30] that enabled infrastructure development and reduced adoption barriers (though some policies and business practices undercut that commodification).[31] Similarly, the future of AI agents hinges not merely on their technical capabilities but on whether we can replicate a commodification trajectory — transforming what might otherwise remain a privilege of the professional class into a widely distributed tool for personal and economic empowerment.

Yet we must acknowledge the countervailing forces that threaten this democratization: corporate incentives often favor maintaining premium pricing tiers, creating artificial scarcity through proprietary ecosystems, and preserving market segmentation that enables price discrimination. Companies developing the most advanced AI agents may resist the very commodification processes essential for equitable adoption, preferring instead to maximize returns by serving affluent customers with high-end offerings while relegating mass-market options to significantly diminished capabilities — a pattern we've witnessed across numerous technological transitions from personal computing to digital streaming services.[32]

The commodification trajectory faces another substantial barrier in the realm of data portability — the ability of users to transfer their accumulated digital memories, preference profiles, and interaction histories between competing platforms. In the emerging AI agent ecosystem, these data repositories represent more than mere convenience; they constitute the foundational building blocks of truly personalized assistance. As agents develop increasingly refined models of user preferences through sustained interaction, they transform from generic tools into individualized extensions of cognitive capacity. Yet without robust data portability frameworks, users find themselves effectively held captive by their digital histories, facing prohibitive switching costs that manifest as lost personalization and the prospect of "starting over" with each platform transition.

This constraint creates a classic market failure scenario where even superior alternatives struggle to overcome incumbency advantages, ultimately dampening the com-

25 Andrew Nusca, The era of smartphone commodification has begun, ZDNet (July 22, 2013), https://www.zdnet.com/article/the-era-of-smartphone-commodification-has-begun/.

26 *Id.*

27 Mobile Fact Sheet, Pew (Nov. 13, 2024), https://www.pewresearch.org/internet/fact-sheet/mobile/.

28 iPhone loses global market share with Apple AI absent in China, India Times (Jan. 13, 2025), https://retail.economictimes.indiatimes.com/news/consumer-durables-and-information-technology/mobiles/iphone-loses-global-market-share-with-apple-ai-absent-in-china/117198102.

29 The top-selling smartphones and brands in 2017: iPhone leads the race, BBVA (Nov. 7, 2018), https://www.bbva.com/en/innovation/top-selling-smartphones-brands-2017-iphone-leads-race/.

30 Amanda Hoover, Apple Is Making It Slightly Easier to Repair Your iPhone, Wired (Apr. 11, 2024), https://www.wired.com/story/apple-iphone-parts-pairing-right-to-repair/.

31 Joseph V. Coniglio, Forbidden Fruit: DOJ Bites Apple in Walled Garden (Apr. 4, 2024), https://itif.org/publications/2024/04/04/forbidden-fruit-doj-bites-apple-in-the-walled-garden/.

32 Joshua Goldman, Chromebook vs. Laptop: What Can and Can't I Do With a Chromebook?, CNET (Nov. 30, 2024), https://www.cnet.com/tech/computing/chromebook-vs-laptop-whats-the-difference/.

petitive pressures that traditionally drive commodification processes. Historical parallels emerge in telecommunications, where number portability regulations dramatically increased market dynamism by reducing switching friction, enabling new entrants to compete on features and pricing rather than inertia.[33] Similarly, the health of the AI agent marketplace may turn on whether users can freely migrate their digital selves across competing platforms — a capability that established providers have powerful financial incentives to resist through proprietary data structures, opaque storage architectures, and technical incompatibilities that function as de facto barriers to meaningful competition.

Beyond data portability, interoperability emerges as a critical dimension in the democratization of AI agent technology. Interoperability — the capacity for different systems, platforms, and applications to seamlessly communicate and function together — fundamentally shapes how users access and benefit from AI capabilities. When platforms restrict which AI agents can effectively operate within their ecosystems, they create artificial constraints that undermine both individual choice and market competition.

The search engine market provides a sobering precedent: Google's multi-billion dollar arrangement with Apple to secure default search status on iOS devices has effectively insulated the former from much competitive pressure, allowing it to maintain a dominant market position despite potential alternatives.[34] The Department of Justice's recent antitrust actions against this arrangement underscores its anti-competitive nature and the resulting market distortions.[35]

In the emerging AI agent landscape, similar exclusivity arrangements could prove even more consequential — imagine being unable to use your preferred AI agent on your primary device, or discovering that certain features function properly only within proprietary ecosystems. Such artificial limitations would not merely inconvenience users; they would fundamentally undermine the competitive dynamics necessary for innovation and price reduction. Without robust interoperability standards — either voluntarily adopted or regulatory mandated — we risk creating digital walled gardens where the transformative potential of AI agents remains accessible only through specific gatekeepers, perpetuating precisely the stratified access patterns that commodification should overcome.

Beyond data portability, interoperability emerges as a critical dimension in the democratization of AI agent technology

The imposing financial requirements for developing and deploying AI agents constitute another significant impediment to the market diversification necessary for widespread adoption. Despite encouraging developments like DeepSeek's introduction of advanced models at substantially lower cost than industry giants,[36] the fundamental economic barriers remain formidable for potential market entrants. The sheer magnitude of computational resources required represents not merely a financial hurdle but a structural constraint within the AI development ecosystem.[37] Access to high-performance computing clusters, specialized hardware accelerators, and the electrical infrastructure to support them creates a tiered landscape where only well-capitalized entities can participate meaningfully in innovation.

The data requirements compound these challenges, creating what economists increasingly recognize as "data monopolies" — situations where incumbent advantages become self-reinforcing through continual user interaction.[38] Each user query, correction, and preference indication becomes not merely a service transaction but a proprietary resource that strengthens an already dominant position. This dynamic shares troubling parallels with historical resource monopolies that prompted regulatory intervention, from Standard Oil's control of petroleum infrastructure to Bell's telecommunications dominance. Without deliberate

33 Wireless Local Number Portability (WLNP), FCC (last accessed Feb. 25, 2025), https://www.fcc.gov/general/wireless-local-number-portability-wlnp.

34 David Pierce, Google reportedly pays $18 billion a year to be Apple's default search engine, Verge (Oct. 26, 2023), https://www.theverge.com/2023/10/26/23933206/google-apple-search-deal-safari-18-billion.

35 Lauren Feiner, Google to court: we'll change our Apple deal, but please let us keep Chrome, Verge (Dec. 23, 2024), https://www.theverge.com/2024/12/23/24328087/google-proposed-final-judgement-search-monopoly-antitrust-default-contracts.

36 John Naughton, DeepSeek: cheap, powerful Chinese AI for all. What could possibly go wrong?, The Guardian (Feb. 1, 2025), https://www.theguardian.com/technology/2025/feb/01/ai-deepseek-cheap-china-google-apple.

37 Large-Scale AI Models, Epoch AI (last accessed Feb. 25, 2025), https://epoch.ai/data/large-scale-ai-models.

38 Kevin Roose, The Data That Powers A.I. Is Disappearing Fast, N.Y. Times (July 19, 2024), https://www.nytimes.com/2024/07/19/technology/ai-data-restrictions.html.

policies addressing both computational access[39] and data accumulation patterns,[40] we risk entrenching precisely the market conditions that would prevent AI agents from following the commodification trajectory necessary for equitable distribution of their benefits across socioeconomic boundaries.

The regulatory landscape itself, while essential for ensuring public safety and ethical implementation, paradoxically emerges as a potential barrier to the commodification process necessary for equitable AI agent adoption. While carefully crafted regulations serve vital social functions, the cumulative weight of compliance requirements can disproportionately burden smaller market participants, inadvertently reinforcing existing market concentration patterns.[41] Large technology incumbents possess the legal expertise, administrative infrastructure, and financial resources to navigate complex regulatory frameworks — effectively transforming compliance costs into competitive advantages rather than consumer protections. This asymmetry echoes historical patterns observed in financial services, where post-2008 regulatory frameworks, though necessary for systemic stability, accelerated industry consolidation by imposing fixed compliance costs that smaller institutions struggled to absorb.

The regulatory landscape itself, while essential for ensuring public safety and ethical implementation, paradoxically emerges as a potential barrier to the commodification process necessary for equitable AI agent adoption

For AI agent developers, similar dynamics could emerge if regulations require extensive documentation, testing regimes, or certification processes without accounting for organizational scale differences. A more nuanced approach can be found in regulatory experimentation zones like Utah's Regulatory Mitigation program, which creates controlled environments where innovative technologies can operate under modified regulatory frameworks while maintaining core consumer protections.[42] Such "sandbox" approaches allow policymakers to witness technological implementation in real-world contexts before establishing permanent regulatory structures, potentially balancing essential protections with innovation-enabling flexibility. This context-sensitive regulatory philosophy draws inspiration from America's historical approach to technological diffusion during periods of rapid innovation — from telegraph expansion to early aviation — where regulatory frameworks evolved alongside technological capabilities rather than preceding them. By embracing similar regulatory humility regarding AI agents, we might avoid inadvertently creating compliance-based moats around established players that would ultimately undermine the competitive pressures essential for commodification and widespread adoption across socioeconomic boundaries.

39 Press Release, Governor Kathy Hochul, Governor Hochul Launches First Phase of Empire AI, Powering Critical Research for the Public Good Just Six Months After FY25 Budget (Oct. 11, 2024), https://www.governor.ny.gov/news/governor-hochul-launches-first-phase-empire-ai-powering-critical-research-public-good-just-six.

40 Kevin Frazier, Unlocking AI for All: The Case for Public Data Banks, Lawfare (Oct. 2, 2024), https://www.lawfaremedia.org/article/unlocking-ai-for-all--the-case-for-public-data-banks.

41 Francesco Trebbi et al., The Cost of Regulatory Compliance in the United States, SSRN (Jan. 15, 2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4331146.

42 Regulatory Mitigation, Utah Department of Commerce (last accessed Feb. 25, 2025), https://ai.utah.gov/regulatory-mitigation/.

06

CONCLUSION: ENSURING A DEMOCRATIZED AI FUTURE

The AI agent revolution offers an unprecedented opportunity to recalibrate our relationship with technology — potentially liberating vast reservoirs of human creative capacity while democratizing access to personalized assistance. Yet this future remains contingent rather than inevitable. The barriers we've identified — restricted data portability, limited interoperability, high entry costs, and regulatory imbalances — collectively threaten to create precisely the stratification patterns that characterized previous technological transitions. Drawing lessons from both the rural electrification movement and the smartphone's commodification journey, we find that deliberate policy interventions, coupled with market-oriented reforms, can effectively counteract and even reverse these tendencies.

By establishing robust data portability frameworks, enforcing meaningful interoperability standards, fostering computational resource access, and implementing context-sensitive regulatory approaches, we can create the conditions for true competition and innovation. The stakes transcend mere technological adoption — they touch fundamental questions of equity, economic mobility, and democratic participation in a rapidly evolving landscape. Just as previous generations constructed legal and regulatory frameworks to ensure electricity's benefits reached all Americans, today's policymakers face a similar imperative: ensuring that the extraordinary capabilities of AI agents become not another dimension of privilege but a broadly shared foundation for individual flourishing and collective advancement.

The AI agent revolution offers an unprecedented opportunity to recalibrate our relationship with technology — potentially liberating vast reservoirs of human creative capacity while democratizing access to personalized assistance

WHAT'S **NEXT**

For July 2025, we will feature a TechREG Chronicle focused on issues related to **National Security**.

ANNOUNCEMENTS

CPI TechREG CHRONICLES August 2025

For August 2025, we will feature a TechREG Chronicle focused on issues related to **Financial Regulations**.

Contributions to the TechREG Chronicle are about 2,500 – 4,000 words long. They should be lightly cited and not be written as long law-review articles with many in-depth footnotes. As with all CPI publications, articles for the CPI TechREG Chronicle should be written clearly and with the reader always in mind.

Interested authors should send their contributions to Sam Sadden (**ssadden@competitionpolicyinternational.com**) with the subject line "TechREG Chronicle," a short bio and picture(s) of the author(s).

The CPI Editorial Team will evaluate all submissions and will publish the best papers. Authors can submit papers in any topic related to competition and regulation, however, priority will be given to articles addressing the abovementioned topics. Co-authors are always welcome.

ABOUT **US**

Since 2006, **Competition Policy International** ("CPI") has provided comprehensive resources and continuing education for the global antitrust and competition policy community. Created and managed by leaders in the competition policy community, CPI and CPI TV deliver timely commentary and analysis on antitrust and global competition policy matters through a variety of events, media, and applications.

As of October 2021, CPI forms part of **What's Next Media & Analytics Company** and has teamed up with **PYMNTS**, a global leader for data, news, and insights on innovation in payments and the platforms powering the connected economy.

This partnership will reinforce both CPI's and PYMNTS' coverage of technology regulation, as jurisdictions worldwide tackle the regulation of digital businesses across the connected economy, including questions pertaining to BigTech, FinTech, crypto, healthcare, social media, AI, privacy, and more.

Our partnership is timely. The antitrust world is evolving, and new, specific rules are being developed to regulate the so-called "digital economy." A new wave of regulation will increasingly displace traditional antitrust laws insofar as they apply to certain classes of businesses, including payments, online commerce, and the management of social media and search.

This insight is reflected in the launch of the **TechREG** Chronicle, which brings all these aspects together — combining the strengths and expertise of both CPI and PYMNTS.

Continue reading CPI as we expand the scope of analysis and discussions beyond antitrust-related issues to include Tech Reg news and information, and we are excited for you, our readers, to join us on this journey.

Scan to Stay Connected!

Scan here to subscribe to CPI's **FREE** daily newsletter.

CPI SUBSCRIPTIONS

CPI reaches more than **35,000 readers** in over **150 countries** every day. Our online library houses over **23,000 papers**, articles and interviews.

Visit **competitionpolicyinternational.com** today to see our available plans and join CPI's global community of antitrust experts.

CPI COMPETITION POLICY® INTERNATIONAL