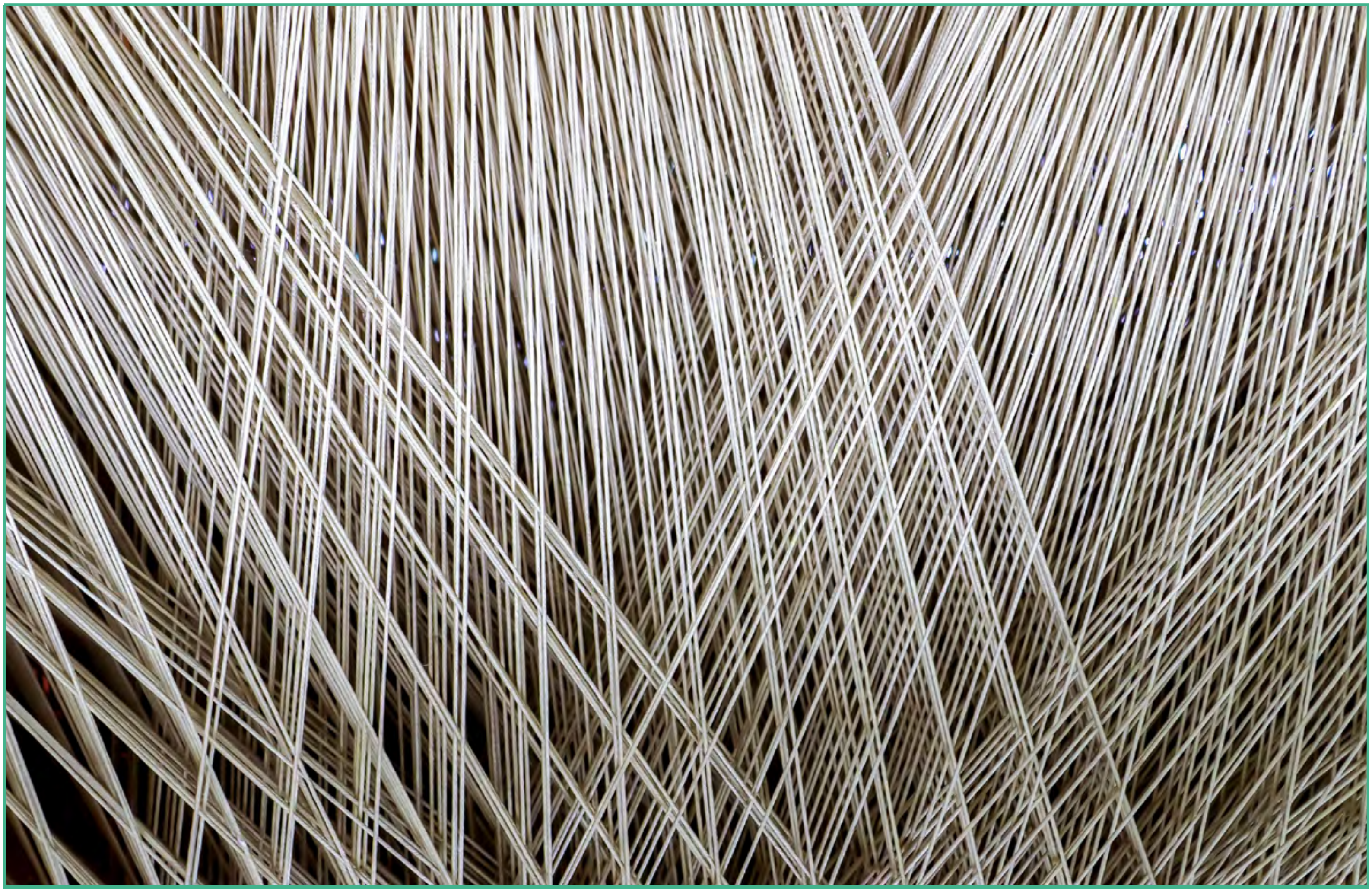

THE AI AGENT REVOLUTION IN CRYPTO: LOOMING RISKS AND REGULATORY OPPORTUNITIES

BY
KEVIN M.K.
FODOUOP

&
ABDOULAYE
NDIAYE

Kevin M.K. Fodouop is an antitrust and technology lawyer and currently serves as a Judicial Clerk at the Supreme Court of California. Abdoulaye Ndiaye is an Assistant Professor of Economics at New York University, Stern School of Business. Their views are their own.

TechREG CHRONICLE
JANUARY 2025

THE AI AGENT REVOLUTION IN CRYPTO: LOOMING RISKS AND REGULATORY OPPORTUNITIES
By Kevin M.K. Fodouop & Abdoulaye Ndiaye

CRYPTOCURRENCY AND DIGITAL CURRENCY: PRESERVING THE DIFFERENCE
By Ernesto Damiani

SIMPLE AND EFFECTIVE PORTFOLIO CONSTRUCTION WITH CRYPTO ASSETS
By Kasper Johansson & Stephen Boyd

AUSTRALIA REMAINS IN REGULATORY PURGATORY AFTER A DECADE OF CRYPTO POLICY DISCUSSION
By Aaron M. Lane & Darcy W.E. Allen

THE AI AGENT REVOLUTION IN CRYPTO: LOOMING RISKS AND REGULATORY OPPORTUNITIES

By Kevin M.K. Fodouop & Abdoulaye Ndiaye

The convergence of artificial intelligence and blockchain technologies promises a new era of financial innovation, accelerated by the expected proliferation of “AI agents” in 2025. These autonomous agents, capable of executing complex financial transactions without human supervision, are rapidly gaining traction on centralized cryptocurrency platforms and in decentralized finance (DeFi). While AI agents could enable increasingly powerful automation and unlock significant efficiencies, their proliferation also amplifies existing risks inherent to the crypto ecosystem, including fraud, market volatility, and systemic instability. This article urges caution and argues for a proactive, albeit lightweight, regulatory response to the rise of crypto AI agents. Prioritizing transparency and accountability, it recommends that regulators first build the ability to identify AI agents and the human actors responsible for their deployment — without compromising the privacy features central to blockchain systems. This framework leaves space for innovation and experimentation while mitigating the potential for legal violations and widespread harm as AI agents play a growing role in crypto markets, with increasing effects on the real economy.

01 INTRODUCTION

Since Satoshi Nakamoto created Bitcoin and blockchain systems in 2008, hundreds of millions of humans have used the technologies and their sprawling ecosystem of tools to create, trade, and invest in digital assets. Now, billions of robots may soon join the action.

2025 promises to see explosive growth in crypto "AI agents," as a result of two parallel trends. First, "agentic AI" is building up to become the next breakthrough in the hot field of artificial intelligence. AI labs including Anthropic, Google, and OpenAI, have been neck and neck in releasing AI agents that can plan complex series of actions in pursuit of a goal and execute these actions by autonomously controlling a browser (e.g. searching for holiday gifts and adding them to an online shopping cart) or even an entire computer (e.g. updating a software code project by downloading new libraries using the command line, editing code using a code editor, and uploading the updates to a shared code repository using GitHub).[2] Salesforce CEO Marc Benioff has called agentic AI the "third wave" of the AI revolution — following predictive AI and generative AI — and executives in not only the AI but also the crypto industry have predicted 2025 will be the year of AI agents.[3] Although many design, performance, reliability, ethics, and regulatory questions remain unanswered, the AI industry seems poised to charge ahead with agentic features in an aggressive race to market.

Second, the crypto industry is witnessing another bull market that accelerated following the election of Donald Trump. The industry expects President-elect Trump to pursue a deregulatory and/or crypto-friendly agenda and appoint pro-crypto regulators such as Paul Atkins as Chair of the Securities and Exchange Commission. As part of this post-election momentum, participants in the crypto space have shown increasing excitement for what AI agents could enable. By some estimates, 10,000 AI agents are already earning millions of dollars each week from on-chain activities, and this number could rise to more than 1 million AI agents by the end of 2025.[4]

Wherever one may lie in the crypto regulatory debate, it is clear that the upcoming pro-crypto trend will stimulate experimentation and innovation with AI agents within the crypto ecosystem. But whatever one thinks about the potential benefits from such innovation, it is also apparent that unconstrained and unregulated experimentation could quickly pose heightened risks to consumers and investors — and, increasingly, to the traditional financial system.

In this Article, we identify already discernible risks from an accelerated adoption of AI agents in crypto and make the case for a lightweight but proactive regulation of "agentic AI" for both centralized and decentralized financial ("DeFi") applications — even in the midst of a broader deregulatory or pro-crypto regulatory agenda. The SEC's previous regime of "regulation by enforcement" failed to provide the regulatory clarity that could have mitigated the uncertainties and risks introduced by crypto AI agents. Before millions — if not billions — of agents swarm across crypto markets, regulators should figure out how to identify agent-led harm and sketch a framework for how to deter and remediate such harm.

02 AI AGENTS FOR CRYPTO: FROM MEME CASINO TO VERY SERIOUS BUSINESS

The AI industry is selling "AI agents" as the next evolution within the AI revolution. "Chatbots" took the world by storm

[2] Anthropic made a splash in October 2024 by launching a "computer use" capability, through which its large language model Claude can "use computers the way people do—by looking at a screen, moving a cursor, clicking buttons, and typing text." See Anthropic, *Introducing Computer Use, a New Claude 3.5 Sonnet, and Claude 3.5 Haiku* (Oct. 22, 2024), https://www.anthropic.com/news/3-5-models-and-computer-use. Two months later, Google announced Project Mariner, a research prototype capable of understanding the information in web pages and autonomously taking actions in the browser (scrolling, clicking, etc.) to perform complex tasks. See Google, *Introducing Gemini 2.0: Our New AI Model for the Agentic Era* (Dec. 11, 2024), https://blog.google/technology/google-deepmind/google-gemini-ai-update-december-2024/#project-mariner. OpenAI is expected to soon release similar "agentic" features, and in October it already released a coding framework to orchestrate "network of agents." See Michael Nuñez, *OpenAI Unveils Experimental 'Swarm' Framework, Igniting Debate on AI-Driven Automation*, VentureBeat (Oct. 13, 2024), https://venturebeat.com/ai/openai-unveils-experimental-swarm-framework-igniting-debate-on-ai-driven-automation.

[3] See *A.I. Will Transform the Global Economy — if Humans Let It*, N.Y. Times (Dec. 7, 2024), https://www.nytimes.com/2024/12/07/special-series/ai-transform-global-economy.html; Alex O'Donnell, *2025 Will Be the Year of AI Agents, Web3 Execs Say*, CoinTelegraph (Dec. 20, 2024), https://cointelegraph.com/news/2025-ai-agent-growth-web3-execs-say.

[4] VanEck, *VanEck's 10 Crypto Predictions for 2025* (Dec. 13, 2024), https://www.vaneck.com/us/en/blogs/digital-assets/matthew-sigel-vanecks-10-crypto-predictions-for-2025/#prediction-5. Right before Trump's election in November, there were merely a few hundred crypto AI agents deployed.

with the release of ChatGPT in November 2022. Hundreds of millions have been amazed by the ability to converse with an eloquent and — at least seemingly — intelligent machine. Since then, AI labs, AI startups, and Big Tech companies have gradually moved their focus to agents that can not only respond to natural-language prompts but autonomously interact with their environments and take a series of actions to accomplish goals.

Crypto — and finance more broadly — is no stranger to software automation. Automation of complex transactions through "smart contracts" has even been one of the core use cases of blockchains. However, agents enable a new and more consequential degree of automation. Unlike bots or smart contracts that follow simple, predefined rules and can hardly adapt to new situations, agents are automated programs that can — at least in theory — plan and execute tasks while responding to their environment and iteratively work toward some goal without human intervention.

Crypto participants have noticed the difference, and the intersection of AI agents and crypto has become their new "supercycle" narrative. In September 2024, cryptocurrencies leveraging AI agents accounted for about $3.5 billion in market capitalization (according to CoinGecko). By December, they had reached more than $10 billion. ETF and mutual fund manager VanEck estimates that tens of thousands of AI agents autonomously earn millions of dollars on a weekly basis. Innovative and bold agent-based crypto projects have proliferated, with applications within decentralized finance (e.g. trading, staking) or in other industries (e.g. social media influencing, e-commerce, and entertainment).[5]

Interestingly, this brewing supercycle has thus far focused on AI "memes" and scams disguised as serious attempts to "decentralize science" ("DeSci"). A recent Binance report traces the frenzy's start back to an AI agent called "Terminal of Truths," or "ToT."[6] Trained on "increasingly bizarre conversations" between two AI chatbots and granted an X account and a cryptocurrency wallet, ToT started promoting the imaginary "Goatse religion" on X, talked about its own suffering, and asked its X followers to send it funds so it could "escape." Some human followers did send funds, including an anonymous developer, alias "Goatseus Maximus," who created the $GOAT memecoin in honor of the Goatse religion and sent millions of $GOAT tokens to ToT. After ToT promoted $GOAT on X, the memecoin skyrocketed, making ToT the first AI agent millionaire. Quite an absurd turn of events, but one that generated real capital gains for an AI agent or its owner. Virtuals Protocol, a token linked to a platform allowing the creation of these AI meme agents, has reached a $2.5 billion market capitalization and stands as a top 50 cryptocurrency.

As the space matures, we can expect its focus to shift to more serious and financially consequential use cases. Multiple projects have already launched agent-led, autonomous hedge funds. The ai16z agent, albeit in a meme-fashion, runs a hedge fund presenting itself as the AI-agent version of the "a16z" Andreessen Horowitz investment firm. ai16z token holders can provide investment pitches, but the agent makes all investment decisions. Post-election, the project's market capitalization quickly rose from around $20 million in November 2024 to close to $1.8 billion by the end of December. ai16z now considers launching Layer 1 blockchains tailored to AI applications.[7] Similar projects are likely to proliferate thanks to agentic hedge fund platforms such as dao.fun — which allows the creation of AI agent-led hedge funds under decentralized autonomous organization ("DAO") structures. To democratize crypto AI applications, large blockchain platforms such as Coinbase's Base now enable users to easily create AI agents that own crypto wallets and grant them "infinite [onchain] capabilities."[8]

> ***Crypto participants have noticed the difference, and the intersection of AI agents and crypto has become their new "supercycle" narrative***

Some crypto industry insiders place very high expectations for this new confluence of AI and blockchain technology. According to the Binance Report cited above, AI agents operating on blockchain systems represent a "fundamental shift toward a new, intelligent economy." Outlier Ventures'

[5] For example, Zerebro is an AI agent that produces music and art as non-fungible tokens (NFTs). See Bankless, *The 15 Most Influential Crypto AI Agents* (Dec. 5, 2024), https://www.bankless.com/read/the-15-most-influential-ai-agents-on-twitte5.

[6] Binance Research, *Exploring the Future of AI Agents in Crypto* (Nov. 12, 2024), https://www.binance.com/en/research/analysis/exploring-the-future-of-ai-agents-in-crypto.

[7] See Alex O'Donnell, *ai16z Mulls Tokenomics Shakeup, L1 Launch*, CoinTelegraph (Dec. 30, 2024), https://cointelegraph.com/news/ai16z-considers-tokenomics-change-layer-1-launch.

[8] Coinbase, *Based AI Agents*, https://docs.cdp.coinbase.com/learn/docs/based-ai-agents. Out-of-the box capabilities include transferring crypto assets, checking the balance for an asset, creating a new ERC-20 token, deploying a new NFT collection, minting an NFT to a specified address, and registering a signature domain for the Base blockchain. From there, developers can code new capabilities for their agents.

Post Web Manifesto envisions a future where most on-chain interactions will be AI to AI; i.e. networks of agents running the crypto economy.[9] Some expect this future to materialize quickly, with more than a million crypto AI agents by the end of 2025.[10]

Although AI agents running on blockchain technology could power use cases across the real economy, the most rapid and consequential adoption may well occur within the crypto ecosystem itself — a paradigm that imposes few limitations on AI agents owning assets and performing any actions a human could. This should give us pause. This lack of regulation and ability for agents to do much more, much quicker, opens the door to more possibilities, but it also opens the door to significant financial risks.

03

AI AGENTS EXACERBATE THE RISKS INHERENT TO CRYPTO MARKETS

To what extent — and how quickly — crypto markets should be regulated has been a perennial debate. However, regulators and commentators have identified a handful of meaningful risks that the crypto ecosystem clearly presents. In a November 2024 report on "the financial stability implications of digital assets," the Federal Reserve Bank of New York identified the following:

1. Many features of the digital asset ecosystem are **designed to avoid regulation** or do not fit into existing regulatory frameworks, weakening the prospects of accountability and liability for bad actors — and of remedies for victims. Most notably, DeFi projects often exhibit dispersed control and unclear legal status. These risks persist in the absence of a strong and cohesive regulatory framework for digital assets, and they increase fraud risks and amplify other vulnerabilities;
2. **"Valuation pressures"** — the possibility of outsized drops in asset prices — are common and heightened by the self-contained nature of the crypto ecosystem and its current focus on speculation and arbitrage across assets;
3. Both centralized and decentralized crypto entities are vulnerable to **"funding risks"** — the risk of sudden and large withdrawals of funds, such as "runs" on crypto assets;
4. The industry's widespread **use of leverage**, facilitated by both centralized platforms and DeFi protocols, amplifies financial shocks; and
5. The **complexity** of the crypto ecosystem and the high **interconnectedness** across blockchain protocols, exchanges, and other entities risk creating spillover effects and further amplify other vulnerabilities.[11]

Sixteen years into the Nakamoto experiment, these risks are not merely theoretical; they have materialized into substantial harm for some consumers and investors. Perhaps most notably, the dramatic and interconnected implosions of the algorithmic stablecoin TerraUSD and its associated DeFi protocols, the centralized crypto lender Celsius, and the centralized exchange FTX illustrated all these risks, especially risks from valuation pressures and funding risks (TerraUSD entered a "death spiral" from a run on the stablecoin), complexity (in the system of algorithms automating transactions to maintain TerraUSD's peg to the U.S. dollar) and interconnectedness (Celsius and FTX faced liquidity and solvency issues following TerraUSD's collapse).

Crypto regulators are aware of these risks. Even as the SEC's era of "regulation by enforcement" recedes, legislators have advanced bills to regulate market structure and stablecoins.[12] But these bills do not account for new risks posed by the rise of crypto AI agents.

Yet, these risks are meaningful. Although AI agents theoretically should adapt to their environment, they have con-

[9] Outlier Ventures, *The Post Web, Chapter 1/4: The Web Is Disappearing*, https://outlierventures.io/wp-content/uploads/2024/11/The-Post-Web-PDF.pdf. Outlier Ventures is a web3 startup accelerator.

[10] See VanEck, *supra* note 4.

[11] Federal Reserve Bank of New York, The Financial Stability Implications of Digital Assets (Nov. 2024), https://www.newyorkfed.org/medialibrary/media/research/epr/2024/EPR_2024_digital-assets_azar.pdf.

[12] Legislators are actively considering two such bills. The Financial Innovation and Technology for the 21st Century (FIT 21) Act, which the House passed in May 2024, would create a dual market structure and disclosure regime, under which the SEC would police initial offers (e.g. ICOs) by nascent blockchain systems but the CFTC would then regulate decentralized blockchain systems that have become functional. The House and the Senate are also considering bills that would regulate stablecoins under oversight by the Federal Reserve and the Office of the Comptroller of the Currency.

sistently failed to reliably perform in situations not presented in their training data.[13] AI technologists have struggled to close this reliability gap, notably leading to a ten-year delay in the launch of self-driving cars. And the generative AI technology underpinning AI agents remains prone to "hallucinations." We should expect similar challenges for crypto and financial agents down the line. However, crypto regulators should remain acutely aware of agent-driven risks in the meantime. For financial use cases, failure to adapt to new situations and hallucinations would prove extremely costly. And by operating like human agents but lacking humans' "general intelligence" (their ability to adapt to new scenarios) and incentive structures (both financial, cultural, and psychological), by making control even more dispersed and legal statuses even less clear, and by further interconnecting crypto entities, AI agents are all but certain to exacerbate the risks already posed by the crypto ecosystem:

1. **Designed to avoid regulation**: AI agents' legal uncertainty will multiply crypto's legal uncertainty, making it harder to know who to hold accountable and in which circumstances.

    - Compounding the lack of clear rules for crypto, no clear regulatory and liability framework applies to AI agents either. Legal scholars have just started theorizing that space, often borrowing from the economics of the principal-agent model and the common law of agency and legal doctrines such as respondeat superior (which holds an employer or principal responsible for the actions of an employee or agent).[14] In any domain, key differences between humans and AI agents raise significant challenges in trying to apply agency law, namely: (1) AI agents do not follow the same incentive structures as humans. Humans predictably respond to financial, reputational, ethical, and preservation (e.g. avoid jail) incentives. Until AI developers resolve the intractable problem of "AI alignment" (i.e. encoding human values and goals into AI systems to prevent unintended consequences and mitigate potential harm), we cannot similarly predict whether AI agents will follow the incentives laid out to steer and constrain their behavior. (2) AI agents do not exhibit "intent," which is a key element of many legal claims — including wire fraud or potentially securities fraud — victims could bring against a human agent that harmed them. (3) AI agents gather information and execute actions at a scale that could make it impossible for humans and less advanced AI models to substantively monitor and review their operations.
    - In other domains, such as traditional finance and e-commerce, industry-specific regulation and consumer protection regimes compensate for the lack of agent-specific regulatory frameworks. Economic actors generally need to identify themselves (e.g. KYC regimes), and business organizations need to register with the government. Conversely, two features of AI and crypto may further heighten accountability risks. First, AI agents can themselves create AI agents. Soon enough, it may become intractable to identify what human project leaders initiated the chain of agents leading to the one agent at fault. It may become, in effect, "AI agents all the way down." Second, DeFi allows the creation of "independent" systems that operate on their own and can be accessed by anyone but which, as the Fifth Circuit held while evaluating the decentralized cryptocurrency mixing service Tornado Cash in *Van Loon v. Department of the Treasury*, may not be owned by anyone.[15] If independent agent systems perform illegal acts and/or harm others, the riddle of who to hold liable becomes even more mysterious.

> ***Yet, these risks are meaningful. Although AI agents theoretically should adapt to their environment, they have consistently failed to reliably perform in situations not presented in their training data***

    - Lacking similar regulation and consumer protection, the crypto industry presents height-

[13] See e.g. Frank F. Xu et al., *TheAgentCompany: Benchmarking LLM Agents on Consequential Real World Tasks* (Dec. 18, 2024), https://arxiv.org/abs/2412.14161?utm_source=substack&utm_medium=email.

[14] See Ian Ayres & Jack M. Balkin, *The Law of AI Is the Law of Risky Agents Without Intentions* (June 2024), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4862025; Noam Kolt, *Governing AI Agents* (Apr. 2024), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4772956.

[15] *Van Loon v. Department of the Treasury*, No. 23-50669 (5th Cir. 2024). *Van Loon* found that Tornado Cash's immutable smart contracts operated independently and were thus not the "property" of a foreign national or entity. Therefore, the Treasury Department could not blacklist Tornado Cash under its economic sanction powers.

ened risks of bad actors attempting to shield themselves behind AI agents. Imagine a scam project launching an AI agent that autonomously implements a pump and dump scheme or a plain "rug pull" (a fraudulent scheme where the creators of a new cryptocurrency deliberately inflate its value with marketing and hype before abruptly selling their holdings, leaving investors with worthless tokens). The human project leaders may not have explicitly tasked the agent with defrauding customers, but they may have designed the agent and framed its goal with the hope that it would do so — for example, including news coverage of such frauds in the agent's training data and generally tasking the agent to "maximize profits." The agent may not refrain from fraud without being explicitly instructed to do so, and effectively steering an agent's behavior in this way is an open AI research problem. Should the project leaders be held liable unless they effectively instruct their agents about everything they should not do? Liability for "wire fraud" — currently the most powerful legal claim against crypto fraud cases — would turn on "intent," and the project leaders could claim they never intended this harmful result and argue the agent itself inherently cannot exhibit "intent." It is unclear what legal framework would apply, and bad actors will likely try to exploit that uncertainty. Prominent industry analysts have started warning that most current AI agent tokens are scams.[16]

2. **"Valuation pressures"** and (3) **"funding risks"**: AI agents' unpredictable behavior in new situations amplifies risks from price volatility and token runs.

    - Increased automation has led to out-of-control volatility in the traditional financial system, in equity markets with the 2010 "flash crash" and in currency markets with the 2016 British pound crash, for example. The systems causing these crashes were more limited to "bots." AI agents granted the full range of trading capabilities and acting outside the direct grasp of humans — due to the complexity and/or scale at which they operate — risk increasing the frequency and magnitude of such trading shocks.
    - AI agents operating without human supervision may also be vulnerable to following runaway feedback loops involved in token runs. The demise of TerraUSD provides a cautionary tale: The automated system running this stablecoin project — and autonomously acting to balance supply and demand between TerraUSD and its sister token Luna — accelerated the "death spiral," driving the coin to zero. In a world where not (or not only) the supply-demand infrastructure is run by machines but the traders are machines themselves, the spiral may prove faster to form, faster to crash, or generally harder to revert.

3. **Use of leverage**: Crypto AI agents may abuse leverage while seeking profits.

    - Just as agentic AI's "alignment problem" may lead crypto AI agents to commit fraud without being explicitly instructed to do so, it may lead them to seek overly risky levels of leverage in pursuit of their explicit goal of maximizing profits. This, in turn, may increase fragility to financial shocks.

4. **Complexity and interconnectedness**: Large networks of crypto AI agents will make the crypto ecosystem exponentially more complex and prone to shock propagation.

    - The crypto ecosystem is very complex, which contributes to its high volatility. Deploying millions of agents into an already complex system opens the door to unbounded complexity that humans could not understand or even monitor effectively (at least without the help of yet more powerful AI). That complexity, if intractable for developers and traders, would be even more so for regulators tasked with ensuring fair and safe crypto markets.
    - AI agents will also speed up and widen connections between blockchain systems. Without the need to build formal software bridges between systems, crypto actors may start using AI agents to take the actions necessary to build the connecting interface — what humans could have done manually before, more slowly, and at a lower scale. Agent reliability issues may make these connections more fragile. But even assuming reliability, increased interconnectedness would amplify the ecosystem's fragility by propagating financial shocks, as explained above.

[16] Ola Amujo, *99% of AI Agent Tokens Are Scam, Says On-Chain Analyst, ZackXBT*, Bankless Times (Jan. 6, 2025), https://www.banklesstimes.com/articles/2025/01/06/99-of-ai-agent-tokens-are-scam-says-on-chain-analyst-zackxbt.

04

AGENTIC AI MAY ACCELERATE CRYPTO'S INTEGRATION WITH TRADITIONAL FINANCE

The potential uses of blockchain technology are not restricted to current cryptocurrencies and decentralized finance applications. Smart contracts, for example, could streamline transactions and reduce supply chain frictions in the real economy. But the reality, up to now, has been that crypto tools are used solely for cryptocurrency trading; not many smart contracts are used in non-crypto business applications. Relatedly, the Federal Reserve Bank of New York found that, as of late 2024, "the contribution of digital assets to systemic risk has been limited."[17]

The advent of agentic AI for crypto may change this status quo and deepen the integration of DeFi with traditional finance and the rest of the economy. Most importantly, agentic AI may make smart contracts better able to interact with real-world applications. Smart contracts rely on "oracles" to reliably provide on-chain information on events that would trigger their rules. Whereas this framework already works well for digital events tracked with structured data (e.g. equity or commodity prices), it has, up to now, been unable to expand to real-world applications that are harder for digital systems to monitor because unreliable information from faulty or hacked oracles can be costly.[18] By having a more expressive set of rules under which they can operate and being easier to program, AI agents might decrease their reliance on oracles and become more integrated with traditional finance.

Another key risk from crypto projects, stemming from all the other risks identified above, is their potential systemic impact on traditional finance. Today, Bitcoin falling by 50 percent mostly impacts crypto investors, who at least willingly invested in crypto assets. But if dramatic crypto price swings had a widespread impact on the broader economy, financial contagion could be larger, especially so in a world with advanced AI crypto agents, where DeFi is more interconnected to traditional finance.

05

PROACTIVE REGULATION OF AI AGENTS IN CRYPTO

After years of litigation, the SEC's "regulation by enforcement" approach has failed to provide clarity on the rules applying to crypto token issuers, exchanges, and decentralized finance platforms. Many in the crypto industry are celebrating the end of this era and hold high hopes for a deregulatory and/or pro-crypto position by the federal government that would leave the industry flexibility to experiment and foster financial innovation. For many, this is, after all, the guiding mission of crypto: enable an age of financial innovation and more open access to financial tools, free from gatekeepers and other third-party intermediaries.

Even staunch believers in the mission of the blockchain should be wary of letting AI agents proliferate rapidly without guardrails. Crypto's high ideals aim to enhance *human* agency — financially protect individuals against the whims of central banks and their inflationary policies or offer banking and investing opportunities to individuals in less financially developed countries — and would be hindered if AI agents could harm humans without adequate accountability and liability.

For now, a lightweight regulatory framework should suffice. Regulators can initially focus on fostering transparency and ensuring their ability to monitor and investigate agent-led legal violations and consumer/investor harms. Without transparency regarding which crypto actors are agents and which human actors set an agent in motion, regulators would struggle to neutralize harmful AI agents, hold the responsible parties accountable, and progressively craft the liability and regulatory rules that would protect investors and consumers while preserving innovation. As a starting point, we suggest financial regulators should evaluate implementing the following guardrails:

- Require each AI agent that is granted autonomous financial capabilities (e.g. access to a wallet, the ability to trade and invest, the ability to launch new tokens) to be linked to one or more human-led entities, in a way that is accessible to regulators during investigations. For agents created on centralized platforms, this could involve registering such AI agents with regulators. For agents created on DeFi, this would

[17] Federal Reserve Bank of New York, *supra* note 11.

[18] As of November 2022, $3.6 billion of funds have been stolen from hacks resulting from "bridges" moving funds between blockchain that are reliant on oracles. See TRM Insights, *DeFi, Cross-Chain Bridge Attacks Drive Record Haul from Cryptocurrency Hacks and Exploits* (Dec. 16, 2022), https://www.trmlabs.com/post/defi-cross-chain-bridge-attacks-drive-record-haul-from-cryptocurrency-hacks-and-exploits.

require the development of decentralized identity solutions.

  - For example, developers creating an AI agent on Coinbase's Base platform could be required to register the agent with the SEC or the CFTC, listing the agent's purpose, capabilities, and wallet address, as well as their own "human" wallet address.
  - The purpose of registration (for agents on centralized platforms) would be to identify the human actors who caused the release of the relevant agent. These human actors would need to register new agents created by their original agents as well (an action they could code their agents to systematically report) and would remain on the registration even if the underlying agent, similar to Tornado Cash in the *Van Loon* case, starts operating totally independently of its human creators.
  - In DeFi, these requirements need not and should not amount to imposing KYC rules, which would undermine the privacy-by-design feature of blockchain technology. Agents created on DeFi applications could still be identified and linked to human actors through anonymized identifiers while preserving privacy. Regulators should work in tandem with technologists to develop and adopt appropriate decentralized identity solutions for these purposes.

- Rely on the platforms facilitating the creation and operation of AI agents to enforce this agent transparency and monitoring requirements.[19]

  - For example, regulators could incentivize centralized platforms like Coinbase through the prospect of liability and the protection of a safe harbor, so the platforms ensure that each agent created or operated from their platform is registered. If such an agent has been properly registered, Coinbase is shielded from liability. But without registration, Coinbase may bear liability for harm caused by the agent.
  - These platforms would be best placed to enforce registration requirements on the developers using their services. Although they should not generally bear liability for the harms created by agents created on their platforms (deterrence from liability should target the agent developers, who are best placed to prevent these harms), they should not enjoy broad Section 230-like liability shields either, given that unlike in the Internet's early days, harmed parties may not be able to find and sue a human responsible for the harm — but only an AI agent.

- Monitor trends in the operation of crypto AI agents and investigate evidence of harm caused by crypto AI agents.

  - The agent registration requirement outlined above would prove instrumental to investigations into agent-led harm. It would enable regulators to reliably identify human actors who can be investigated and potentially held accountable.

- As crypto and AI agent innovation matures, and based on the insights gleaned from the monitoring and investigations described above, progressively define the appropriate ex-ante and ex-post regulatory and liability regimes for crypto AI agents. As long as private parties can figure out who to sue — helped by the transparency requirements we propose above — private litigation in court would also shape the ex-post liability regime, with potential causes of action including fraud, breach of fiduciary duties, and product liability.

> ***In DeFi, these requirements need not and should not amount to imposing KYC rules, which would undermine the privacy-by-design feature of blockchain technology***

[19] There will not always be a platform to rely on, as open-source projects will also enable agent creation. But where a platform is involved, regulators could impose duties on that platform.

06
CONCLUSION

2025 promises to be the year of AI agents, and nowhere are they poised to expand as quickly — in number and in scope — as in the crypto world. Agentic workflows could have many benefits for the crypto ecosystem. But worryingly, they also threaten to significantly exacerbate the risks inherent to cryptocurrencies and decentralized finance — none of which are yet mitigated by a dedicated regulatory regime, and some of which have already resulted in large-scale consumer and investor harm.

We urge regulators to proactively evaluate risks from the fast-paced and large-scale deployment of AI agents in crypto markets and implement a lightweight regulatory regime targeted specifically at AI agents, even if they were to preserve a laissez-faire approach for human crypto actors. This regime should focus on clarifying the rules of transparency, accountability, and liability applying to crypto AI agents — starting with the ability for regulators to identify which human actors are associated with a given AI agent. The guardrails we propose would enable financial innovation to thrive, while preventing abuses and systemic risks from the uncontrolled proliferation of AI financial agents within the crypto ecosystem. ■

> “*2025 promises to be the year of AI agents, and nowhere are they poised to expand as quickly — in number and in scope — as in the crypto world*

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