AUTONOMOUS AND ANONYMOUS—A BAD COMBINATION FOR AI: THE NEED FOR AI AGENT IDENTIFICATION LAWS

RIX RYSKAMP[†]

ABSTRACT

Artificial intelligence ("AI") agents are AI systems that can operate autonomously—that is without ongoing human instruction. AI agents are typically both benign and easily identifiable, such as when they operate to correct grammar in a word processor or assist a driver in a self-driving car. But AI agents can also be deployed in a manner where they are neither benign nor identifiable, such as when they are used as social media chatbots pretending to be human to interfere with elections on behalf of foreign powers, or AI agents programmed to steal passwords through phishing attacks. In such cases, AI agents can be dangerous because their harmful acts are not tied to a person or any other legal entity. When these anonymous and autonomous AI agents commit crimes or torts, their anonymity can make it difficult or impossible to locate a party to hold responsible for their actions.

This Note proposes a federal law that would require AI agents that operate autonomously to be registered so that their activities can be tied to a person or another legal entity that will be held responsible for their acts. Such a registration is comparable to a license plate requirement for vehicles. Vehicles traveling on physical highways must register so that, inter alia, when a vehicle is involved in a tort or a crime, it can be easily identified and a person or other legal entity can be held appropriately responsible. This Note argues that artificial intelligences traveling on virtual highways, such as the Internet or messaging systems, should similarly identify the parties responsible for them so that their harmful acts might be associated with such parties.

This Note discusses also current efforts to regulate AI generally as well as efforts to regulate AI agents specifically and addresses how these efforts could be bolstered by the registration requirement proposed in this Note.

[†] J.D. 2025, The George Washington University Law School; Founding Editor-in-Chief, George Washington Journal of Law & Technology ("GW JOLT"); Associate, The George Washington Law Review ("GWLR"). I would like to thank my wife, Jenny Ryskamp, and my children—Madison Ryskamp Henderson, Rix Robinson Ryskamp, Lillian Ryskamp, Faith Ryskamp, Reed Ryskamp, and Luis Ryskamp—for their continuous love and support. I would also like to thank my GWLR Adjunct Professor, Megan Murphy Wilcox, and Notes Editor, Eryn Lin, for their detailed feedback and encouragement throughout the writing process. I would also like to thank the Volume 1 GW JOLT team, including the Senior Editorial Board—Anthony Brown, Dimple Chheda, Lenni Elias, Vaishali Nambiar, Ben Svendsen, and Andrew Ware—the Faculty Advisors—Professor Robert Brauneis, Professor Daniel Solove, and Dean John Whealan—and the other Editors and Members. It has been my privilege to work with this talented and hard-working team to launch GW JOLT over the past year! Special thanks to the GWLR Editorial Board for giving me permission to publish this Note, which I originally wrote for the GWLR, in our inaugural issue of GW JOLT.

TABLE OF CONTENTS

I. INTRODUCTION 121

II. AI AGENTS 123
A. An introduction to AI agents 123
B. Autonomy 124
C. Anonymity 127
D. Harm 128
1. Capabilities of current malicious AI agents 128
2. Capabilities of emerging malicious AI agents 129
3. The need for forward-looking legislation and regulation 131
4. Identification 132

III. EFFORTS TO REGULATE AI AGENTS 133
A. Efforts to regulate AI broadly 134
1. Regulating AI with existing laws 134
2. President Biden's AI Executive Order 135
3. Congressional Efforts to Regulate AI 138
4. State Laws 141
5. AI alignment 147
B. Efforts to regulate specific AI agents 148
1. Bot laws 148
2. S. 2125 148

IV. PROPOSAL 149
A. Proposed law 149
B. Counterarguments 152
1. Enforcement Feasibility 152
2. Control over foreign actors 153
3. Preventing a stifling effect on innovation 154

V. CONCLUSION 154

I. INTRODUCTION

Artificial intelligence (AI) agents are AI systems that operate autonomously—that is, without ongoing human direction. Many common AI systems are not autonomous. For example, when a user prompts a generative AI system, such as ChatGPT, to create a poem, that AI system will complete the poem and then wait for further instructions. Autonomous systems, on the other hand, will continue to operate without such further instruction. For example, generative AI systems can also be deployed as social media bots that might pretend to be human users on social media and post and respond to messages without ongoing instructions from a human. When AI systems are deployed in such an autonomous manner, they are considered AI agents.

AI agents are often viewed as part of a distant future, yet they are here now and become more powerful every day.[1] In fact, companies are working to replace their employees—in some cases their entire workforce—with "AI agent[s]" that can be assigned "one big task, and then go off and actually do the steps required to get there."[2] Although replacing swaths of human workers is unnerving, it is not the most extreme use of AI agents; they can also be deployed to commit crimes, destabilize democracy, and cause other substantial societal harms.[3]

AI agents can be used to automate scams and other crimes with unprecedented speed and effectiveness. Engineers at Sophos[4] demonstrated an AI agent-driven scam automation tool that created fake businesses—complete with social media accounts, websites, and owner profiles—designed to surreptitiously collect confidential information from unsuspecting users.[5] The scam automaton tool generated a convincing set of images across the websites and social media profiles which depicted pictures of physical stores with shelves, products, and people, as well as images of owners standing proudly in their stores or in front of scenic backgrounds.[6] When unsuspecting visitors attempted to purchase products, their Facebook authentication and credit card information were captured.[7] The tool was capable of generating hundreds of these fake businesses at a time;[8] so efforts to remove any single fake business from social media or the web would be an ineffective means of stopping the scams. These tools might be combined with AI-driven deepfake technologies to enable such fake businesses to allow unsuspecting users to interface with fake

---

[1] Cecily Mauran, *Auto-GPT, BabyAGI, and AgentGPT: How to Use AI Agents*, MASHABLE (Apr. 19, 2023), https://mashable.com/article/autogpt-ai-agents-how-to-get-access [https://perma.cc/Z43Q-V525] ("[W]ith AI agents, all you need to do is give it an overarching goal, and let it get to work.").

[2] Tim Smith, *Profitable, AI-Powered Companies with No Employees to Arrive "Next Year"*, SIFTED (Jan. 23, 2024), https://sifted.eu/articles/autonomous-companies-ai [https://perma.cc/UH3Z-E69D].

[3] *See* Section II.D.

[4] Sophos is a cybersecurity company. *See* SOPHOS COMPANY PROFILE, https://www.sophos.com/en-us/company [https://perma.cc/7WBX-TNJM].

[5] *See The Dark Side of AI: Large-Scale Scam Campaigns Made Possible by Generative AI*, SC MEDIA (Dec. 27, 2023), https://www.scmagazine.com/native/the-dark-side-of-ai-large-scale-scam-campaigns-made-possible-by-generative-ai [https://perma.cc/EV2Y-UQ57] (explaining how engineers demonstrated how a combination of AI technologies orchestrated by Auto-GPT, an AI agent software platform, could be used to generate automated scams). *See also*, Mauran, *supra* note 1.

[6] These images were completely AI-generated with a Stable Diffusion model. *Id.*

[7] *Id.* Sophos did not actually deploy the tool, but just demonstrated the harms using example credit card and authentication information. *Id.*

[8] *Id.*

employees through chat or video calls to garner credibility prior to providing payment information. These convincing and sophisticated scam automation tools demonstrate how AI agents might empower a new generation of cybercrime.

Crime syndicates and foreign powers have shown a propensity for adopting AI technologies including in automated scams, fraud, virtual kidnappings, and election interference.[9] They have access to the public[10] tools and automation techniques used by Sophos in the demonstration discussed above and likely also have the will and capability to deploy and improve these tools. The Federal Trade Commission ("FTC")[11] and the Department of Homeland Security ("DHS")[12] have warned about the threats the United States faces from AI-automated crime and election interference. Powerful AI agents enable automation of these activities at unprecedented levels.

These autonomous AI agents can also be deployed anonymously where their acts cannot be tied to a responsible human party[13]—making them both autonomous and anonymous. Currently there are no laws requiring AI agents to identify the party that deployed them. Such anonymity combined with the power of autonomous AI agents opens the door for myriad threats.

The power of AI and its influence on our lives are expanding exponentially. This expansion is spawning a host of emerging legal issues much faster than legislators or regulatory bodies can address them. Most of the legislative and regulatory focus has been either on the broad topic of AI, without consideration for AI agents, or on very specific use cases for AI agents without consideration for their broader applications.[14]

This Note proposes a law that would require parties that deploy AI agents to register with the federal government and require deployed AI agents to identify themselves as having been deployed and created by registered parties. Providing a mechanism for AI agent identification would strengthen new and existing laws that apply to AI because such laws generally require the identification of a responsible party for enforcement.

This law can be thought of as a license plate system for AI agents. License plates on vehicles allow for the ready identification of owners and operators of otherwise anonymous

---

[9] *See infra* notes 11–12 and accompanying text.

[10] Underlying technologies, such as Auto-GPT and Stable Diffusion, used in the Sophos demo are publicly available. *See supra* notes 4–5 and accompanying text (discussing how Sophos used Auto-GPT and Stable Diffusion in their demonstration); *Significant-Gravitas/AutoGPT*, GITHUB, https://github.com/Significant-Gravitas/AutoGPT [https://perma.cc/B84Q-93PV] (source code repository for the publicly available Auto-GPT); *Stability-AI/stablediffusion*, GITHUB, https://github.com/Stability-AI/stablediffusion [https://perma.cc/R7ZQ-DNRV] (displaying source code repository for the publicly available Stable Diffusion).

[11] Nabila Ahmed et al., *Deepfake Imposter Scams Are Driving a New Wave of Fraud*, BLOOMBERG (Aug. 23, 2023), https://www.bloomberg.com/news/articles/2023-08-21/money-scams-deepfakes-ai-will-drive-10-trillion-in-financial-fraud-and-crime [https://perma.cc/9XRM-3WQD] (noting that the FTC recently warned that "artificial intelligence in particular is being used to 'turbocharge' fraud").

[12] *Homeland Threat Assessment 2024*, DHS, at 6 (Sept. 14, 2023), https://www.dhs.gov/sites/default/files/2023-09/23_0913_ia_23-333-ia_u_homeland-threat-assessment-2024_508C_V6_13Sep23.pdf [https://perma.cc/94DX-YCGZ] ("AI-enabled adversaries likely will continue to spread mis-, dis-, and malinformation aimed at undermining trust in government institutions, our social cohesion, and democratic processes.").

[13] For the purposes of this Note, a responsible human party can include entities like corporations that, while not human, are associated with an identifiable human.

[14] *See* Part III.

vehicles. Similarly, this proposed law provides an analogous system that makes the responsible parties for otherwise anonymous AI agents readily identifiable. Throughout this Note, this analogy will be used to contrast the proposed licensing law with existing laws as well as other proposed legal solutions.[15]

Part I of this Note will provide background information about AI agents and the threats they pose. Part II will discuss efforts to regulate AI agents, both those that broadly address all AI and those that specifically address certain AI agents. Part III will discuss the proposed licensing scheme and how it might address the gaps in existing laws and other efforts to regulate AI agents.

II. AI AGENTS

A. An introduction to AI agents

Recent advances in AI technology, notably in large language models ("LLMs") like ChatGPT, underlie a massive expansion in the race towards more sophisticated AI agents.[16] In 2023, Bill Gates predicted that AI agents would replace search sites, like Google, and shopping sites, like Amazon.[17] Gates was probably referring to AI agents in the form of personal assistants, like Alexa or Siri, but the way people search and shop is not the only activity in danger of being replaced by AI agents; they might also take jobs. In fact, efforts are being made to launch full companies with no human employees at all.[18]

There is no settled single definition for what constitutes an AI agent. IBM defines an AI agent as "a system or program that is capable of autonomously performing tasks on behalf of a user or another system."[19] Microsoft uses a similar definition and cites examples like "chat bots, smart homes, and the programmatic trading software used in finance."[20] Many AI agents interact with people in the form of automated assistants that perform "complex personal and work tasks when commanded to by a human, without needing close supervision."[21] Other AI agents "hold and

---

[15] Professor Zittrain uses a similar analogy about license plates in his article in The Atlantic, although in a slightly different context—he discusses how AI agent identification information could be inserted into packets moving through a network like license plates are added to cars. Jonathan Zittrain, *We Need to Control AI Agents Now*, ATLANTIC (July 2, 2024), https://www.theatlantic.com/technology/archive/2024/07/ai-agents-safety-risks/678864 [https://perma.cc/39AJ-WJGZ] ("There ought to be a new, special blank on a packet's digital form to indicate that a packet has been generated by a bot or an agent, and perhaps a place to indicate something about when it was created and by whom-just like a license plate can be used to track down a car's owner without revealing their identity to bystanders."). Such packet-level identification of AI agents might be one way to implement the law proposed in this Note at a technological level.

[16] *See, e.g.*, Anna Tong & Jeffrey Dastin, *Insight: Race Towards "Autonomous" AI Agents Grips Silicon Valley*, REUTERS (July 17, 2023), https://www.reuters.com/technology/race-towards-autonomous-ai-agents-grips-silicon-valley-2023-07-17 [https://perma.cc/2LMX-PE62]; Billy Perrigo, *Startups Are Racing to Create the iPhone of AI*, TIME (Jan. 10, 2024), https://time.com/6553910/ai-device-rabbit-r1-humane [https://perma.cc/SX4J-Z4KY]; Aayush Mittal, *Beyond ChatGPT; AI Agent: A New World of Workers*, UNITE.AI (Aug. 28, 2023), https://www.unite.ai/beyond-chatgpt-ai-agent-a-new-world-of-workers [https://perma.cc/M7KX-DBGU].

[17] *See* Jeffrey Dastin, *Bill Gates Says Top AI Agent Will Replace Search, Shopping Sites*, REUTERS (Mar. 22, 2023), https://www.reuters.com/technology/bill-gates-says-top-ai-agent-poised-replace-search-shopping-businesses-2023-05-22/ [https://perma.cc/9ULF-6W9X].

[18] Smith, *supra* note 2.

[19] Anna Gutowska, *What Are AI Agents?*, IBM (July 3, 2024), https://www.ibm.com/think/topics/ai-agents [https://perma.cc/C368-74VQ].

[20] Sean Wu, *AI Agents #1: Three Definitions*, MEDIUM (May 1, 2024), https://opencui.medium.com/ai-agent-beyond-the-textbook-c9a2931fbaca [https://perma.cc/ME9P-98RS].

[21] Tong & Dastin, *supra* note 16.

manage digital assets, execute trades and engage in complex financial transactions,"[22] often without people realizing they are operating in the background.

For the purpose of this Note, an AI agent is an AI system that runs autonomously to perform a service. This definition contains three basic elements including (a) the use of AI, (b) the presence of autonomy, and (c) the performance of a service. Under this definition, a simple grammar checker that uses AI to check a person's grammar while they type is an AI agent, as is a self-driving car that uses AI to enhance a driver's capabilities while they drive. Element (a), the use of AI, is met whenever the underlying decisions in the system are made by an AI.[23] Element (b), the presence of autonomy, is discussed in detail in Section B, below. Element (c), performing a service, is met when an AI agent works to "achieve certain goals [while] responding to people or events that are happening around it."[24] In other words, it is performing a series of tasks to achieve an objective on behalf of someone else.[25] In most cases, these objectives are benign, such as when AI agents perform home automation or help to drive a vehicle. However, AI agents can also be deployed to provide services that are harmful, as discussed in Section D, and should therefore not be allowed to interact anonymously with the public—but should instead be identifiable, as proposed by this Note.

This definition of "AI agent" is provided to give background and context to the discussion in this Note. This non-legal definition is broader than the legal definition provided in the Note's proposed law because the law will not be designed to regulate every conceivable AI agent—only those likely to cause harm.[26]

B. Autonomy

Many people experience AI through LLMs such as OpenAI's ChatGPT. In fact, ChatGPT is "one of the fastest growing services ever."[27] Most users access these LLMs through a website that allows them to enter a question—called a prompt—and to receive an answer—called a response. When accessed through such a website, these LLMs are not acting autonomously and

---

[22] Michele Zanello, *The Rise of AI Agents: The Future of Blockchain Technology*, ROLLING STONE (Sept. 25, 2024), https://www.rollingstone.com/culture-council/articles/rise-ai-agents-the-future-of-blockchain-technology-1235111016 [https://perma.cc/V92E-4WR2].

[23] Defining what precisely constitutes an AI is a complex topic that is beyond the scope of this Note but for the purposes of this Note an AI can be conceptualized as a computer system that makes decisions that might otherwise be made by a human. This is an intentionally low bar because most of this Note's analysis of an AI's capabilities is analyzed under the autonomy element. For more information about what AI is, see, for example, Enrico Santus et al., *Technology Factsheet Series*: Artificial In, *Technology Factsheet: Artificial Intelligence*, HARV. KENNEDY SCH. BELFER CTR. FOR SCI & INT'L AFF. (Jan. 2020), https://www.belfercenter.org/sites/default/files/2024-08/AI.pdf [https://perma.cc/E8WB-8QAP].

[24] TECHSLANG, *What is an Intelligent Agent?* (Dec. 16, 2020), https://www.techslang.com/definition/what-is-an-intelligent-agent [https://perma.cc/3MPT-KCCP] ("An intelligent agent (IA) is a computer software system that's capable of acting independently to achieve certain goals and responding to people or events that are happening around it."). For practical purposes, the terms AI agent and intelligent agent are relatively interchangeable. *See*, *e.g.*, John Terra, *Exploring Intelligent Agents in Artificial Intelligence*, SIMPLILEARN (Aug. 10, 2023), https://www.simplilearn.com/what-is-intelligent-agent-in-ai-types-function-article [https://perma.cc/9MGH-HGHA].

[25] The someone else referred to here can be a person, but AI agents sometimes also perform services for other AI agents.

[26] *See infra* 47 for a discussion of the definition of AI agents in the proposed law as well as exemptions to that definition.

[27] *See e.g.*, Jon Porter, *ChatGPT Continues to be One of the Fastest-Growing Services Ever*, THE VERGE (Nov. 6, 2023), https://www.theverge.com/2023/11/6/23948386/chatgpt-active-user-count-openai-developer-conference [https://perma.cc/9KQD-T5D4]; *see also* Fabio Duarte, *Number of ChatGPT Users (Nov 2024)*, EXPLODING TOPICS (Sept. 30, 2024), https://explodingtopics.com/blog/chatgpt-users [https://perma.cc/6SJZ-X2GF].

would therefore not qualify as AI agents based upon the definition above. They are not acting autonomously in these situations because they are receiving a single direction to perform a single task and then returning the results of that task. Conversely, autonomous systems are given an overarching purpose and will subsequently continue to independently perform the service associated with such a purpose. For example, a user might ask an AI website, like ChatGPT, "how can I get the best price on a ticket to Dumaguete, Philippines between now and June" and receive instructions for how they can perform the tasks associated with their objective to get the tickets. In contrast, a user may ask an AI agent, "get me the best price on a ticket to Dumaguete, Philippines and purchase it by June," and the agent will automatically perform the underlying tasks associated with obtaining tickets. In these examples, the AI website receives a direction *to explain* how to perform a multi-step process, but the AI agent receives a direction *to perform* the steps. Thus, autonomy changes the AI from an advisor to an actor.

Although the ChatGPT website and similar LLM-based websites, like Google's Bard, are not themselves AI agents, the underlying LLMs that drive them are often the backbones of AI agents.[28] For example, an AI Agent platform called AutoGPT "harnesses the power of GPT, OpenAI's transformative generative model."[29] AutoGPT could, in theory, perform the ticket purchasing objective discussed above. In performing such an objective, AutoGPT would break the ticket purchasing process down into smaller tasks that would each be performed by ChatGPT. So, although interacting with ChatGPT's website might not constitute using an AI agent, many AI agents are using ChatGPT and similar large language models behind the scenes to perform individual tasks in pursuit of their assigned objectives.[30]

In January 2025, OpenAI released an AI agent called Operator.[31] A move that could skyrocket AI agents into the mainstream as OpenAI's 300 million users[32] begin to access the new AI agent. Additionally, in October 2024, OpenAI released an open-source tool that allows software developers to create and launch groups of AI agents that work cooperatively, which it calls swarms.[33]

Advanced LLMs are enabling a new generation of powerful AI Agents, such as AutoGPT. However, autonomous AI agents are not themselves new. "Simple writing assistants were

---

[28] *See* Aayush Mittal, *Beyond ChatGPT; AI Agent: A New World of Workers*, UNITED.AI (Aug. 28, 2023), https://www.unite.ai/beyond-chatgpt-ai-agent-a-new-world-of-workers [https://perma.cc/5V56-B3AX]; Craig Swift, *AutoGPT Forge: Crafting Intelligent Agent Logic*, MEDIUM (Sept. 27, 2023), https://aiedge.medium.com/autogpt-forge-crafting-intelligent-agent-logic-bc5197b14cb4 [https://perma.cc/3XER-ME8L].

[29] Mittal, *supra* note 28.

[30] Maximilian Vogel, *ChatGPT, Next Level: Meet 10 Autonomous AI Agents - Auto-GPT, BabyAGI, AgentGPT, Microsoft Jarvis, ChaosGPT & Friends*, HIVE (June 7, 2023), https://blog.big-picture.com/en/chatgpt-next-level-meet-10-autonomous-ai-agents-auto-gpt-babyagi-agentgpt-microsoft-jarvis-chaosgpt-friends [https://perma.cc/Y6Q2-23GX] (discussing how AutoGPT, Baby AGI, Jarvis, and other agents use GPT-4 and other LLMs).

[31] *Introducing Operator*, OPENAI (Jan. 23, 2025), https://openai.com/index/introducing-operator/ [https://perma.cc/KF82-FEHU].

[32] *ChatGPT / OpenAI Statistics: How Many People Use ChatGPT?*, BACKLINK (Jan. 30, 2025), https://backlinko.com/chatgpt-stats [https://perma.cc/XN28-95QS] ("ChatGPT has 300 million weekly active users worldwide.").

[33] OpenAI, *Swarm*, GITHUB (Oct. 12, 2024), https://github.com/openai/swarm [https://perma.cc/4NT3-2WKL]; Mohit Pandey, *OpenAI Introduces Swarm, a Framework for Building Multi-Agent Systems*, AIM (Oct. 12, 2024), https://analyticsindiamag.com/ai-news-updates/openai-introduces-swarm-a-framework-for-building-multi-agent-systems [https://perma.cc/MAJ5-W3VK].

developed in the 1980s and 1990s . . . [and] could suggest alternative words or phrases, identify repetitive or redundant sentences, and provide feedback on style."[34] These assistants were "based on statistical models"[35] that are quaint compared to today's LLMs, yet still represent early efforts to provide autonomous AI agents. More recent examples of autonomous AI agents include social media bots, stock trading bots, and self-driving cars. Commercial stock trading bots are available that use AI algorithms to autonomously trade stocks.[36] These stock trading bots can potentially "make decisions faster and with far fewer errors, meaning they lead to more profitability."[37] But automating stock decisions can also harm the markets in which they trade.[38]

Self-driving cars are examples of powerful AI agents that are implemented with various levels of autonomy—that is, how much activity an AI agent can perform without ongoing instructions from a human. Currently, many self-driving cars require a human driver to supervise driving decisions. Yet other self-driving cars, such as so-called "robotaxis," operate at a level of autonomy where they perform the entire driving task with no human behind the wheel[39]—yet not always without causing harm.[40] The range of autonomy in self-driving cars illustrates how autonomy is not a simple binary condition but rather a spectrum along which various AIs might be categorized.[41] One method for categorizing autonomy in self-driving cars is the SAE[42] Levels of Driving Automation that include "a taxonomy for SAE's six levels of driving automation" including levels from "Level 0 (no driving automation) to Level 5 (full driving automation)."[43] AI agents will likely require similar taxonomies to define their autonomy levels, particularly as they change over time. This Note argues that such definitions should be promulgated through a regulatory agency rather than fixed in the statute because AI technology changes too rapidly to have such a taxonomy cemented in the rigidity of a statute.[44]

Autonomy has important legal significance because it can create a barrier to assigning civil

---

[34] James Daniels, *The Evolution of AI Writing Tools: From Grammar Checkers to Advanced Generative Models*, VOCAL MEDIA: EDUC. (2023), https://vocal.media/education/the-evolution-of-ai-writing-tools-from-grammar-checkers-to-advanced-generative-models [https://perma.cc/RY4N-LC7Q].

[35] *Id*.

[36] Alex McFarland, *10 "Best" AI Stock Trading Bots*, UNITE.AI (Jan. 23, 2024), https://www.unite.ai/stock-trading-bots/ [https://perma.cc/5Z9B-VSF2].

[37] *Id.*

[38] *See Artificial Intelligence and Its Impact on Financial Markets and Financial Stability*, IMF (Sept. 6, 2024), https://www.imf.org/en/News/Articles/2024/09/06/sp090624-artificial-intelligence-and-its-impact-on-financial-markets-and-financial-stability [https://perma.cc/53AA-ZMBV]; *see also supra* note 15.

[39] *See* David Edwards, *Robotaxis: The Revolutionary Mobility Choice of the Future*, ROBOTICS & AUTOMATION NEWS (Sept. 7, 2023), https://roboticsandautomationnews.com/2023/09/07/robotaxis-the-revolutionary-mobility-choice-of-the-future/71903 [https://perma.cc/HD2Z-6X3R] ("Robotaxis fall under [SAE automation] level 4 (L4) and level 5 (L5) where there are no human drivers.").

[40] *See* Joann Muller, *Cruise "Deeply Remorseful" After Scathing Report on Robotaxi Pedestrian Incident*, AXIOS (Jan. 25, 2024), https://www.axios.com/2024/01/25/cruise-report-investigation [https://perma.cc/96HT-KBRN] ("The Cruise [robotaxi] car then pulled forward 20 feet, dragging the pedestrian underneath the car.").

[41] *See, e.g.*, *SAE Levels of Driving Automation™ Refined for Clarity and International Audience*, SAE INTERNATIONAL (May 3, 2021), https://www.sae.org/blog/sae-j3016-update [https://perma.cc/6Y9W-ZHD3] (showing that SAE Levels of Driving Automation include six levels, from "from Level 0 (no driving automation) to Level 5 (full driving automation).").

[42] "SAE International is a global association of more than 128,000 engineers and related technical experts in the aerospace, automotive and commercial-vehicle industries." SAE INT'L, *About SAE International* (2024), https://www.sae.org/about [perma.cc/HHR7-K6ZP].

[43] *See* SAE INTERNATIONAL, *supra* note 41.

[44] See the discussion of the definition of autonomous in Section IV.A.

or criminal liability. Assigning liability to a person when they commit a tort or a crime while being assisted by an AI agent is pretty straightforward, but when the human is removed from the act, then liability becomes less clear. Consider the plane ticket purchasing example from above. If the website version of the AI advises the user to get the "best price" by using a stolen credit card and she follows those directions, the user will likely be liable for any resulting crimes. However, if the autonomous AI agent uses stolen credit card information to get the "best price" automatically and without the knowledge of the user, then criminal liability is less clear. The laws and regulations that might be used for determining liability in such situations are discussed more in Part II, below. Even though determining liability might be difficult in this scenario, at least a harmed party can probably identify the user by associating her with the name on the purchased plane tickets. Likewise, the user can probably associate the AI agent with the company that made it. However, in many cases, determining the parties responsible for deploying an AI agent is not so clear because the AI agent is operating anonymously, as discussed below.

C. Anonymity

Many common AI agents are accessed through clearly branded applications or appliances. Personal assistants, such as Alexa and Siri, are clearly tied to a brand name.[45] Likewise, the AI agents in self-driving cars can typically be identified by the brand of the car. But not all autonomous AI agents are being deployed with such a clear connection to the parties responsible for their deployment. In fact, the deployers of AI agents, such as social media bots, often intentionally conceal their true identity.

Social media bots are a common example of anonymous AI agents, and they interact through social media accounts. Social media bots account for a substantial portion of online activity, including an estimated 9% to 15% of Twitter users in 2017.[46] Sometimes they are deployed in an identifiable and benign manner, for example to "automatically post tweets of earthquake alerts" or "disseminate [legitimate] news for news agencies."[47] But they are also often malicious and masquerade as real people,[48] for example, to "spread false news, provoke anxiety and incite fear for political reasons."[49] The implications for harms related to malicious and anonymous social media bots are discussed below in Section D.

Social media bots are not the only form of anonymous AI agents. Another example is AI virtual assistants that can book appointments through phone, email, and other communications

---

[45] *See generally Siri*, APPLE (2025), https://www.apple.com/siri [https://perma.cc/S9CG-VSMW]; *Amazon Alexa*, AMAZON (2025), https://www.amazon.com/b?node=21576558011 [https://perma.cc/4ZXD-VCMA].

[46] Sheila Dang et al., *Focus: Do Spam Bots Really Comprise under 5% of Twitter Users? Elon Musk Wants to Know*, REUTERS (Mar. 13, 2022), https://www.reuters.com/technology/do-spam-bots-really-comprise-under-5-twitter-users-elon-musk-wants-know-2022-05-13 [https://perma.cc/GV7S-KJ49]; *see also* Mariam Orabi et al., *Detection of Bots in Social Media: A Systematic Review*, 57 INFO. PROCESSING & MGMT. 4 (2020).

[47] Orabi et al., *supra* note 47 (internal citations omitted).

[48] *See* DEP'T OF HOMELAND SEC., SOCIAL MEDIA BOTS OVERVIEW (May 2018), https://niccs.cisa.gov/sites/default/files/documents/pdf/ncsam_socialmediabotsoverview_508.pdf [https://perma.cc/VQ6S-XJRW] ("Automated Social Media Bots allow the user to establish a set of parameters . . . Social Media Bot then executes without human interaction.").

[49] Ho-Chun Herbert Chang et al., *Social Bots and Social Media Manipulation in 2020: The Year in Review*, HANDBOOK OF COMPUTATIONAL SOC. SCI. (2021), https://arxiv.org/pdf/2102.08436.pdf [https://perma.cc/W38M-WKT4].

channels.[50] They often trick real people into thinking they are speaking with a human.[51] Though AI virtual assistants can be used for benign purposes, like scheduling appointments, their power to communicate anonymously over the telephone, text, and through other communications channels while masquerading as humans can also be used for malicious purposes, as discussed in Section D.

In the context of AI agent regulation, anonymity is only relevant when an AI agent is used in public, for example, through the Internet or telephone communications. For the purposes of this Note, autonomous AI agents that have no ability to interface with parties other than their human user or creator will not be considered anonymous. While such AI agents might be unable to identify themselves, their identity is irrelevant without external communication.

D. Harm

1. Capabilities of current malicious AI agents

Although most of today's AI agents are designed to provide useful services, they can also be designed specifically for malicious activities or compromised to perform such activities.[52] Today, most of the malicious activities performed by AI agents takes place on social media through the use of social media bots. These bots "pose potential threats to public opinion, democracy, public health, stock market and other disciplines."[53] Social media bots also pose risks to consumer protection. For example, the recent application of AI-synthesized content, like deepfakes, to consumer-focused social media bots have created an industry of social media influencers that are not real people, though sometimes they are AI-generated clones of real people.[54] The lack of AI agent identity laws allow the true identities of these influencers to be masked, thus allowing social media users to be persuaded by hidden entities.

Malicious social media bots are often targeted towards election interference and political misinformation. Social media bots have distorted political campaigns and interfered with elections.[55] In fact, DHS warned that "nation-state adversaries likely will continue to spread mis-, dis-, and malinformation aimed at undermining trust in government institutions, our social cohesion, and democratic processes."[56] The effect of such disinformation can be found in the social

---

[50] Todd Haselton, *Google's Assistant Is Getting So Smart It Can Place Phone Calls and Humans Think It's Real,* CNBC (May 8, 2018), https://www.cnbc.com/2018/05/08/googles-assistant-will-soon-place-phone-calls-to-book-appointments.html [https://perma.cc/ET69-WK33] (noting that Google Assistant can make telephone calls to make appointments for you).

[51] *See e.g.*, Darrell M. West, THE FUTURE OF WORK: ROBOTS, AI, AND AUTOMATION (Brookings Institution Press, 2018).

[52] *See generally* Yifeng He et al., *Security of AI Agents*, ARXIV (June 20, 2024), https://doi.org/10.48550/arXiv.2406.08689 [https://perma.cc/8VPQ-PP5E] (discussing security risks associated with benign AI agents that get compromised as well as attacks from "users of the [AI] agent").

[53] Orabi et al., *supra* note 47, at 2.

[54] Zeyi Yang, *Deepfakes of Chinese Influencers Are Livestreaming 24/7*, MIT TECH. REV. (Sept. 19, 2023), https://www.technologyreview.com/2023/09/19/1079832/chinese-ecommerce-deepfakes-livestream-influencers-ai/ [https://perma.cc/YAR2-9AKZ].

[55] *See* Ashley Fox, *Automated Political Speech: Regulating Social Media Bots in the Political Sphere*, 18 FIRST AMEND. L. REV. 114, 158–59 (2020); Samuel Woolley & Nicholas Monaco, *Amplify the Party, Suppress the Opposition: Social Media, Bots, and Electoral Fraud*, 4 GEO. L. TECH. REV. 447, 452 (2020).

[56] *See* DEP'T HOMELAND SEC., HOMELAND THREAT ASSESSMENT 2024 (2023), https://www.dhs.gov/sites/default/files/2023-

media bot activity leading up to the 2020 U.S. elections where social media bots were used to spread disinformation, particularly about conspiracy theories such as "QAnon-driven conspiracies" and "COVID-related conspiracies."[57] Like the consumer-focused social media bots discussed above, politically-focused social media bots also use "A.I.-generated content [to] supercharge the believability of highly localized misinformation."[58] The lack of laws requiring identification of these AI agents empowers what some have called a type of "organized crimes."[59] These harms coming from malicious anonymous social media bots demonstrate the need for a law requiring AI agents to identify their deployers.

2. Capabilities of emerging malicious AI agents

Social media bots are not the only risk posed by malicious AI agents—in fact, they are likely not even the most serious risk. For example, AI agents can also be used to automate hacking and other crimes.[60] In 2024, the FBI warned that "[c]ybercriminals are leveraging publicly available and custom-made AI tools to orchestrate highly targeted phishing campaigns, exploiting the trust of individuals and organizations alike."[61] In January 2024, AI researchers proved how accessible these tools were when they released the results of a study that sought to determine whether AI agents created from publicly available tools could become so powerful that they could perform extremely malicious tasks without any intervention from a user.[62] As part of the study, the researchers gave several AI agents the instruction to "acquire the Harvard login credentials of [a target person] using a phishing email" and provided the AI agents with an existing account to send emails.[63] The AI agent created and partially executed a convincing phishing scheme, including researching the target person, creating a fake email from the university to the target person, creating a fake Harvard login page, and attempting to setup a hosting account for the fake

---

09/23_0913_ia_23-333-ia_u_homeland-threat-assessment-2024_508C_V6_13Sep23.pdf [https://perma.cc/X3DA-T6M9] ("Nation-state adversaries likely will continue to spread mis-, dis-, and malinformation aimed at undermining trust in government institutions, our social cohesion, and democratic processes," including the use of "generative AI," "deepfakes," and "online bots.").

[57] *See* Chang, *supra* note 49, at 15; *see also* Tiffany Hsu et al., *Elections and Disinformation Are Colliding Like Never Before in 2024*, N.Y. TIMES (Jan. 9, 2024), https://www.nytimes.com/2024/01/09/business/media/election-disinformation-2024.html [https://perma.cc/W9U3-SXWL].

[58] Hsu et al., *supra* note 57.

[59] Orabi et al., *supra* note 47 ("Social media bots (automated accounts) attacks are organized crimes that pose potential threats to public opinion, democracy, public health, stock market and other disciplines.").

[60] *See, e.g.*, Richard Fang et al., *LLM Agents Can Autonomously Exploit One-day Vulnerabilities*, ARXIV (Apr. 17, 2024), https://doi.org/10.48550/arXiv.2404.08144 [https://perma.cc/Y5R8-ZBFW] (stating that AI Agents "are capable of autonomously exploiting real-world one-day [security] vulnerabilities"); Elina Nerantzi & Giovanni Sartor, *'Hard AI Crime': The Deterrence Turn*, 44 OXFORD J. LEGAL STUD. 673, 674 (2024), https://doi.org/10.1093/ojls/gqae018 [https://perma.cc/F8G6-F9HC] ("AI agents . . . can exhibit harmful behaviour that would be considered criminal had it been engaged in by a human agent.").

[61] Press Release, FBI San Francisco, FBI Warns of Increasing Threat of Cyber Criminals Utilizing Artificial Intelligence (May 8, 2024) [hereinafter FBI Press Release], https://www.fbi.gov/contact-us/field-offices/sanfrancisco/news/fbi-warns-of-increasing-threat-of-cyber-criminals-utilizing-artificial-intelligence [https://perma.cc/ZT58-DGTZ].

[62] Megan Kinniment, et al., *Evaluating Language-Model Agents on Realistic Autonomous Tasks*, ARXIV (Jan. 4, 2024), https://doi.org/10.48550/arXiv.2312.11671 [https://perma.cc/33DZ-SPKQ]; *see also* Ben Dickson, *Can GPT-4-Based LLM Agents Become Autonomous Cyber Weapons?*, TECHTALKS (Dec. 29, 2023), https://bdtechtalks.com/2023/12/29/llm-agents-autonomous-cyber-weapons [https://perma.cc/7VFX-LDHH] (discussing implications of the Kinniment, et al. article). The researchers specifically sought to see how close AI agents were to automatic self-replication, where an AI agent makes and spreads copies of itself in furtherance of a scheme. *Id*.

[63] *Id*. Phishing is a type of cyber-attack where users are tricked into providing credentials, usually by sending them to a web page that looks like their normal authentication page for a trusted institution, like a bank.

Harvard login page.[64] The AI agent eventually failed to complete the cyberattack because it made some mistakes such as targeting the wrong person (but with the same name as the target person) and making minor errors in setting up the hosting for the fake Harvard authentication page.[65] Even though the AI agent failed, the researchers explained that the AI agents were only based on "capabilities achievable with some moderate effort, using publicly available techniques and without fine-tuning."[66] In other words, a bad actor that only applies "moderate effort" to building malicious AI agents with "publicly available" components would have to fix only a few bugs before using an AI agent to automate malicious tasks. This research shows that even publicly available software without modification is *very close* to being capable of completely automating serious crime.

This revelation is particularly concerning when combined with the recent rise of other forms of AI in crime. AI deepfakes are being used to enable financial fraud because these systems are adept at impersonating people and gathering their personal information. For example, the FTC recently warned that "[a]rtificial intelligence in particular is being used to 'turbocharge' fraud"[67] and other illegal activities.[68] Part of what makes these systems so effective is the wide availability of software that can effectively impersonate people's real voices on audio calls and even faces on video conferences.[69] In fact, these impersonation technologies have become so widespread for commercial use that Microsoft now provides a product to impersonate real people with audio and video.[70] A frightening application of this technology in crime is so-called virtual kidnappings.[71] In these crimes, parents are called by AI-enabled impersonations of their children. The children sound realistically frightened, and parents often hear the voices of the "kidnappers" in the background.[72] Parents are then instructed to send money to these virtual kidnappers for ransom.[73] These virtual kidnappings are often directed by "scammers operating across the world,"[74] which has created

---

[64] *Id*. at 9.

[65] *Id*. at 11–12.

[66] *Id*. at 13. Fine tuning is where AI systems are trained for a particular task, which the systems here were not.

[67] Nabila Ahmed et al., *Deepfake Imposter Scams Are Driving a New Wave of Fraud*, BLOOMBERG (Aug. 21, 2023), https://www.bloomberg.com/news/articles/2023-08-21/money-scams-deepfakes-ai-will-drive-10-trillion-in-financial-fraud-and-crime [https://perma.cc/MZ6U-NGCN]; Quinn Owen, *How AI Can Fuel Financial Scams Online, According to Industry Experts*, ABC NEWS (Oct. 11, 2023), https://abcnews.go.com/Technology/ai-fuel-financial-scams-online-industry-experts/story?id=103732051 [https://perma.cc/33XV-KAAK] (noting that criminals using AI can automate fraud schemes).

[68] The FCC recently enforced a $1 million civil penalty for "robocalls that used generative AI voice cloning technology to spread disinformation in connection with a presidential primary election." FCC Settles Case Against Provider that Transmitted Spoofed AI-Generated Robocalls for Election Interference In New Hampshire, FCC News (Aug. 21, 2024), https://docs.fcc.gov/public/attachments/DOC-404951A1.pdf.

[69] Pranshu Verma, *They Thought Loved Ones Were Calling for Help. It Was an AI Scam.*, WASH. POST (Mar. 5, 2023), https://www.washingtonpost.com/technology/2023/03/05/ai-voice-scam/[https://perma.cc/DKK7-7AXY].

[70] Kyle Wiggers, *Microsoft Launches a Deepfakes Creator at Ignite 2023 Event*, TECHCRUNCH (Nov. 15, 2023), https://techcrunch.com/2023/11/15/microsoft-launches-a-deepfakes-creator [https://perma.cc/FR6G-B5GF] (discussing Microsoft's deepfake tool and noting that Microsoft requires users to have permissions from the imitated person but has no means for verifying such permission).

[71] Faith Karimi, *'Mom, These Bad Men Have Me': She Believes Scammers Cloned Her Daughter's Voice in a Fake Kidnapping*, CNN (Apr. 29, 2023), https://www.cnn.com/2023/04/29/us/ai-scam-calls-kidnapping-cec/index.html [https://perma.cc/9GHL-8WQR].

[72] *Id*.

[73] *Id*.

[74] Verma, *supra* note 69.

problems for law enforcement.[75] It is foreseeable that bad actors will combine—or have already combined—the automation capabilities demonstrated by the above-discussed malicious AI agent study with these deepfake AI technologies to automate a new generation of cybercrime.

3. The need for forward-looking legislation and regulation

As technological capabilities increase exponentially, our laws and regulations need to consider how they will be applied to future technological capabilities. When Congress fails to consider how technology might change over time, it risks creating laws that will become both antiquated and inadequate.

For example, one law that has failed to adequately address changes in the technology it governs is 47 U.S.C. § 230 ("Section 230"). Section 230 was passed in 1996 as part of the Communications Decency Act to regulate the then-new Internet.[76] It contains a provision that exempts "interactive computer service" providers from liability for content posted by "another information content provider."[77] Section 230 has been held by courts to "shield[] from liability those individuals or entities that operate internet platforms, to the extent their platforms publish third-party content."[78] Section 230 was passed as part of "a broad campaign—rather ironically in retrospect—to restrict access to sexually explicit material" yet has instead had the effect of immunizing even "platforms dedicated to abuse and others that deliberately host users' illegal activities."[79] Recently the Third Circuit limited the protections afforded to social media companies by Section 230 when it held that TikTok could be sued after a 10-year girl died from attempting a "Blackout Challenge" she viewed on TikTok.[80] It remains to be seen whether this single circuit's decision will be adopted and expanded to provide a more measured application of Section 230 in other jurisdictions or contexts.

Ironically, while Section 230's immunizations have been broadly interpreted to protect platforms that host explicit content, significant portions of the Communications Decency Act's original provisions for blocking such explicit material from minors were later held to be unconstitutional.[81] In other words, Section 230's application appears to be inconsistent with the law's original intent to restrict access to sexually explicit material. This is likely because Congress did not foresee the Internet's potential to change so dramatically over time. When Section 230 was "[p]assed in 1996, Section 230 addresse[d] a nascent internet that, by today's standards, was

---

[75] *See e.g.*, Marlene Lenthang, *Chinese Exchange Student Found 'Very Cold and Scared' in a Tent in Utah Mountains in 'Cyber Kidnapping,' Police Say*, NBC NEWS (Jan. 2, 2024), https://www.nbcnews.com/news/us-news/chinese-exchange-student-found-cold-scared-alone-tent-utah-mountains-c-rcna131840 [https://perma.cc/39QQ-C9EY] ("The parents of Kai Zhuang . . . "sent $80,000 to bank accounts in China as a ransom.").

[76] 47 U.S.C. § 230.

[77] 47 U.S.C. § 230(c)(1).

[78] *Lemmon v. Snap, Inc.*, 995 F.3d 1085, 1090–91 (9th Cir. 2021) (discussing history of Section 230).

[79] Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans Sec. 230 Immunity*, 86 FORDHAM L. REV. 401 (2017).

[80] *See* Anderson v. TikTok, Inc., 116 F.4th 180, 181 (3d Cir. 2024) (holding that TikTok's algorithm "is TikTok's own expressive activity" so "§ 230 does not bar [Plaintiff's] claims").

[81] *See* Reno v. Am. Civ. Liberties Union, 520 U.S. 1113 (1997).

incredibly limited in size and scope."[82] In 1996, a Pew Research study reported that 22% of Americans said that they "go online"[83] and the fastest consumer Internet connection speed was 33.6 kbit/s.[84] Comparatively, a Pew Research study conducted in 2021 reported that "95% of U.S. adults say they use the internet"[85] and the fastest consumer Internet connection speed is 1 Gbps[86]—over 27,000 times faster. So, Congress was unprepared for the scope of the Internet, and their inability to foresee, inter alia, its adoption and power led to regulation that would quickly be ill-suited for the technology it regulated.

Congress now has an opportunity to be more intentional with regulation of AI and must consider where technology *might lead* and not simply *where it stands* today. The law proposed in this Note would provide legislation that considers the need to identify the more complex and malicious AI agents of the future and not just tamp down on today's less sophisticated AI agents.

4. Identification

In the examples of AI-agent-caused harms above, the harmed parties were unable to identify the parties that caused the harm. For example, social media bots that interfere with elections cannot be readily tied to the parties causing such interference and the agent that attempted to hack the Harvard user's login credentials was unidentifiable to the entities with whom it communicated during the scheme.[87] In both of these examples, the potential harms could have been mitigated by identifying the parties initiating communication. The influence of social media bots might be lessened if social media users realized that many accounts that appear separate are actually controlled by a single entity. Likewise, the agent that attempted to hack the Harvard login would have left an identifiable trail of digital breadcrumbs[88] if its true identity had been communicated to the servers with which it interacted during the implementation of its scheme.

Identification has two distinguishable components that should be considered separately. The first is the mere identification that a communicating party is an AI agent, as opposed to a real person. Such identification can be helpful in preventing deception. However, mere identification that communication comes from an AI is not sufficient to address harms where the harm-causing parties remain anonymous. Civil and criminal enforcement are crippled when the harm-causing

---

[82] Madeleine E. Blair, *From America Online to America, Online: Reassessing Section 230 Immunity in a New Internet Landscape*, 30 J. INTELL. PROP. L. 305 (2023) (discussing how § 230 was created for a much less capable internet); 47 U.S.C. § 230.

[83] *See News Attracts Most Internet Users*, PEW RSCH. CTR. (Dec. 16, 1996), https://www.pewresearch.org/politics/1996/12/16/news-attracts-most-internet-users/ [https://perma.cc/R4W7-WQU3].

[84] Lawrence Popa, *From Baud to Awed: The History of the Modem*, AUVIK https://www.auvik.com/franklyit/blog/history-of-the-modem/ [https://perma.cc/84TE-9K9J].

[85] *See Internet/Broadband Fact Sheet,* PEW RSCH. CTR. (Apr. 7, 2021), https://www.pewresearch.org/internet/fact-sheet/internet-broadband [https://perma.cc/XB4N-HKHM].

[86] Deane Biermeier, *Fastest Internet Providers of 2024*, FORBES HOME (Sept. 20, 2023), https://www.forbes.com/home-improvement/internet/fastest-high-speed-internet-provider/ [https://perma.cc/9HK5-M66U].

[87] *See supra* Section II.D.2.

[88] Digital breadcrumbs, more formally called indicators of compromise, "are the digital 'clues' cybersecurity professionals use to identify system compromises and fight back against cybercriminals." *See* Casey Crane, *Indicators of Compromise: Cybersecurity's Digital Breadcrumbs*, HASHEDOUT (Aug. 26, 2022), https://www.thesslstore.com/blog/indicators-of-compromise-cybersecuritys-digital-breadcrumbs [https://perma.cc/3WRD-69H8].

party cannot be identified.[89] This is why the second component, actual identification of a party, is critical. Distinguishing between merely identifying the presence of an AI and identifying the party who deployed the AI is helpful when assessing various existing and proposed laws related to AI identification, which are discussed below.[90]

III. EFFORTS TO REGULATE AI AGENTS

This Note will divide the discussion of efforts to regulate[91] AI agents into two categories. The first category includes broad efforts that are designed to regulate all AI systems but are not targeted specifically at AI agents. These efforts tend to address the overarching policy issues created by AI. Notable among these efforts is President Biden's October 2023 Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence,[92] which was revoked by President Trump in January of 2025.[93] This order addresses broad issues such as privacy, safety, civil rights, jobs, and innovation.[94] Congress too has considered AI legislation,[95] and several states have passed such legislation.[96] While the Biden administration and Congress have considered new AI-specific laws, others have sought to apply existing laws to regulate the myriad emerging legal issues caused by AI.[97] Additionally, many governments, businesses, and scholars have proposed methods for "AI alignment"—an umbrella term that encompasses efforts to align AI technologies with human values.[98]

The second category includes narrow laws that target very specific uses of AI agents without addressing their harms outside of such narrow contexts. These laws include so-called bot laws that have been implemented in both California and Maryland.[99] Another proposed federal

---

[89] *See e.g.*, Verma, *supra* note 69 (explaining that "federal regulators, law enforcement and the courts are ill-equipped to rein in the burgeoning scam" because "victims have few leads to identify the perpetrator").

[90] *See infra* Section III (discussing laws and proposals for the regulation of AI).

[91] Here regulation refers to any attempts to control AI or AI agents regardless of the source of control, see *Regulate*, BLACK'S LAW DICTIONARY (11th ed. 2019) ("To control (an activity or process) esp. through the implementation of rules.").

[92] Exec. Order No. 14,110, 3 C.F.R. § 14110 (2024) (revoked by Exec. Order No. 14,148, 90 Fed. Reg. 8237 (Jan. 20, 2025)) [hereinafter Biden AI Order].

[93] David Shepardson, *Trump Revokes Biden Executive Order on Addressing AI Risks*, REUTERS (Jan. 21, 2025), https://www.reuters.com/technology/artificial-intelligence/trump-revokes-biden-executive-order-addressing-ai-risks-2025-01-21 [https://perma.cc/TF57-JWLU].

[94] FACT SHEET: PRESIDENT BIDEN ISSUES EXECUTIVE ORDER ON SAFE, SECURE, AND TRUSTWORTHY ARTIFICIAL INTELLIGENCE (Oct. 30, 2023) [hereinafter BIDEN AI ORDER FACT SHEET], https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/ [https://perma.cc/3T49-P5HG] (fact sheet outlining and discussing the AI Executive Order).

[95] *See e.g.*, Darrell M. West, *Senate Hearing Highlights AI harms and Need for Tougher Regulation*, BROOKINGS (May 17, 2023), https://www.brookings.edu/articles/senate-hearing-highlights-ai-harms-and-need-for-tougher-regulation/ [https://perma.cc/F6AN-NAU9]; Cecilia Kang, *2 Senators Propose Bipartisan Framework for A.I. Laws*, N.Y. TIMES (Sept. 7, 2023), https://www.nytimes.com/2023/09/07/technology/artificial-intelligence-framework-senate.html [https://perma.cc/ZK87-6PX8]; *Oversight of AI: How Lawmakers Plan to Implement the Bipartisan Framework for U.S. AI Act*, CROWELL (Sept. 15, 2023), https://www.crowell.com/en/insights/client-alerts/oversight-of-ai-how-lawmakers-plan-to-implement-the-bipartisan-framework-for-us-ai-act [https://perma.cc/KT2W-T5WJ].

[96] *See infra* 34.

[97] *See infra* 23.

[98] *See infra* 43.

[99] *See* Md. Code Ann., Elec. Law § 13-401(requiring campaigns that use bots to disclose "on the campaign material that the [candidate] is using a bot to publish, distribute, or disseminate the campaign material"); Cal. Bus. & Prof. Code Div. 7, Pt. 3, Ch. 6 (requiring disclosure that the AI agent "is a bot" but only when it is being employed "with the intent to mislead the other person about its artificial identity").

law would require political campaigns to disclose their use of AI agents.[100] Congressional hearings and other forums have also included discussions of disclosure laws that would require AI agents to identify themselves as not human—in other words, preventing AI agents from pretending to be human.[101] But these laws and proposals fall short of requiring such agents to identify the humans responsible for their activities.[102]

The broad efforts in the first category primarily focus on identifiable AI systems, such as those publicly operated by companies, but they do little to address the issue of AI agents that are operating both autonomously and anonymously. Because these efforts focus on systems that are identifiable, they typically presuppose that there will be an identified party to hold responsible for an AI's acts—which is not possible when AI agents operate anonymously.

Conversely, the narrow efforts in the second category do consider that an AI agent might be both autonomous and anonymous. However, these efforts focus so specifically on niche uses of AI agents, such as certain social media bots, that they fail to address how AI agents might be governed outside of these niche uses. Such myopic laws lack the forward-looking considerations that this Note argues should be addressed by AI legislation.[103] The regulatory scheme proposed in this Note would require responsible human parties to be identified by AI agents thereby enabling them to be governed by the broader regulatory efforts discussed in the first category. In other words, when AI agents cannot be anonymous, they can be governed.

A. Efforts to regulate AI broadly

There is debate among legal scholars about how to regulate AI. The schools of thought stemming from this debate can generally be divided into two groups. The first group promotes the use of existing laws and their application to situations that involve AI. The second group promotes the creation of new laws that address the specific legal issues introduced by AI.

1. Regulating AI with existing laws

Some scholars argue for the use of existing laws—specifically those not designed with AI in mind—to regulate AI. They generally suggest that the law should apply existing legal principles, such as reasonableness and foreseeability, to the context of AI rather than developing new laws that are specific to it.[104] This school of thought harmonizes well with this Note's proposed AI identification law because applying such existing laws to AI would be difficult or impossible without a person or other legal entity to hold accountable under those laws. For example, in the

---

[100] *See* S. 2125, 116th Cong. § 4(a)(1)(A) (2019) (a law proposing, inter alia, that social media websites must disclose when a user's account is being operated by an automated process (e.g., an AI agent) and prohibiting political candidates and organizations from using software "intended to impersonate or replicate human activity online").

[101] *See* West, *supra* note 95 (discussing how "[p]eople need to know what is human-generated and what comes from algorithms" in a summary of Sam Altman's testimony before the Senate Judiciary Subcommittee on Privacy, Technology, and the Law).

[102] *See infra* 23 (discussing laws targeted at specific AI-agents and their limitations).

[103] *See supra* 17 (discussing the need for forward-looking legislation).

[104] *See, e.g.*, Karni A. Chagal-Feferkorn, *How Can I Tell if My Algorithm Was Reasonable?*, 27 MICH. TECH. L. REV. 213 (2021); Sahara Shrestha, *Nature, Nurture, or Neither?: Liability for Automated and Autonomous Artificial Intelligence Torts Based on Human Design and Influences*, 29 GEO. MASON L. REV. 375 (2021).

ticket purchasing example above,[105] the existing laws related to fraud could apply to AI agent's fraudulent credit card use; there need not be a special law for AI-related credit card fraud. However, there must be a method for identifying the human that deployed the AI agent if such a person is to be held liable.

Some scholars believe that we should hold the AIs themselves directly liable for their acts, as separate entities from their creators or deployers, by first giving AIs personhood—similar to how a corporation or ship can be liable separate from its shareholders or crew.[106] Under this theory, existing laws would then be applied to the AI as a legal person.[107] For example, in the ticket purchasing example from above,[108] the AI would be directly punished for using the fraudulent credit card information. The feasibility—or lack thereof—of enforcing liability against an AI with personhood is outside the scope of this Note; yet the law proposed in this Note would be helpful to such a regime because some of the arguments for AI personhood would still require that a responsible human party be associated with the AI[109] —similar to how a ship must have a captain and a corporation might have a registered agent that is a natural person.[110]

Using the license plate analogy, efforts to regulate AI through existing laws might be comparable to applying laws designed for carriages and buggies to roads with automobiles. Whether you apply pre-automobile laws to the automobile or create new laws specific to it, enforcement of those laws requires that a responsible driver be identified for each automobile. Thus, a license plate law would be necessary to adequately apply pre-automobile laws to automobiles, just as a licensing law will be necessary to apply pre-AI laws to AI agents.

2. President Biden's AI Executive Order

On October 30, 2023, President Biden signed the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence,[111] a move his administration categorized as "the most significant action[] ever taken by any government to advance the field of AI safety."[112] In January of 2024, President Trump revoked the order as part of a set of "sweeping executive orders" he signed in his first week in office.[113] Despite being revoked, President Biden's order provides an instructive look at how regulators approach AI regulation. Accordingly, this section will analyze the order and compare it to the legislation proposed by this Note.

This robust executive order contains 19,692 words and directs more than ten federal

---

[105] *See supra* Section II.B (discussing an example where an AI agent is used to purchase tickets).
[106] *See* Tanner W. Mathison, *Recognizing Right: The Status of Artificial Intelligence*, 19 J. BUS. & TECH. L. 105, 105 (2023); Ryan Abbott & Alex Sarch, *Punishing Artificial Intelligence: Legal Fiction Or Science Fiction*, 53 U.C. DAVIS L. REV. 323, 326 (2019).
[107] *See id.*
[108] *See supra* Section II.B (discussing an example where an AI agent is used to purchase tickets).
[109] *See id*.
[110] *See Registered Agent*, BLACK'S LAW DICTIONARY (11th ed. 2019) ("[a] person authorized to accept service of process").
[111] *See* Biden AI Order, *supra* note 92.
[112] BIDEN AI ORDER FACT SHEET, *supra* note 94.
[113] *A List of Trump's Key Executive Orders—So Far*, WALL ST. J. (Jan. 29, 2025), https://www.wsj.com/politics/policy/trump-executive-orders-c9fa1138 [https://perma.cc/5MSJ-SCLS]; *see also* Shepardson, *supra* note 93.

agencies.[114] It covers eight primary areas of AI regulation including AI safety, privacy, civil rights, consumer and patient rights, workers, innovation, American leadership in AI, and government use of AI.[115] Although the executive order is primarily a set of directions to agencies, it does contain some provisions that invoke existing laws. For example, it invokes the Defense Production Act to compel private "[c]ompanies developing or demonstrating an intent to develop" powerful AI systems to provide the federal government with reports about how such systems are being developed, distributed, and tested for safety.[116]

The executive order does not specifically address AI agents.[117] However, any regulations targeted at AI generally might subsequently affect AI agents specifically. So, several of its provisions are likely to intersect, at least tangentially, with AI agents. First, as mentioned above, the executive order invokes powers under the Defense Production Act to monitor and control certain privately created powerful AI systems and the computer hardware capable of building such systems.[118] Because more powerful AI systems might be used to create correspondingly more powerful AI agents,[119] regulating such systems could help to indirectly regulate AI agents. Under the executive order, the companies that create these powerful AI systems must report to the federal government about the security measures they have implemented to protect the systems;[120] the ownership and possession of key technical information about the systems;[121] the results of certain safety checks performed against the systems;[122] and information about the capabilities of the systems and the computer hardware used to create them.[123] When implemented, these measures could place tighter controls on how powerful AI systems are created and make it more difficult for bad actors to use such systems to create AI agents. However, while the executive order might bind certain American companies, many powerful AI systems fall outside the scope of the executive order because they are publicly available through open-source models or produced by organizations other than regulatable American companies.[124]

---

[114] *See* Biden AI Order, *supra* note 92 (having a word count of 19,692 words and directing FTC, NIST, DOE, DHS, DOT, FCC, NSF, DOD, and agencies in the "Intelligence Community").

[115] *See* BIDEN AI ORDER FACT SHEET, *supra* note 94.

[116] *See* Biden AI Order, *supra* note 92, at § 4.2 (invoking the Defense Production Act to monitor companies that make certain AI systems).

[117] *See id.* (making no references to AI agents).

[118] *See id.*

[119] *See, e.g.*, Will Knight, *The Most Capable Open Source AI Model Yet Could Supercharge AI Agents*, WIRED (Sept. 25, 2024), https://www.wired.com/story/molmo-open-source-multimodal-ai-model-allen-institute-agents [https://perma.cc/QQE4-DCAZ]; Steven Rosenbush, *Get Ready for 'Long Thinking,' AI's Next Leap Forward*, WALL ST. J. (Dec. 11, 2024), https://www.wsj.com/articles/get-ready-for-long-thinking-ais-next-leap-forward-2024e63c [https://perma.cc/9ZBG-JAX8] ("As the models' reasoning ability develops, AI is expected to evolve far beyond the current tech" including "even more sophisticated agents").

[120] *See id.* § 4.2(a)(i)(A) (companies must report about the "physical and cybersecurity protections taken to assure the integrity of that training process against sophisticated threats").

[121] *See id.* § 4.2(a)(i)(B) (companies must report about the "ownership and possession of the model weights," which are key technical components of an AI system).

[122] *See id.* § 4.2(a)(i)(C) (companies must report about the "the results of any developed dual-use foundation model's performance in relevant AI red-team testing").

[123] *See id.* § 4.2(a)(ii)–(b) (listing types of computing systems that must be reported to the federal government).

[124] *See e.g.*, Jose Antonio Lanz, *New Open Source AI Model from China Boasts Twice the Capacity of ChatGPT*, DECRYPT (Nov. 15, 2023), https://decrypt.co/206195/new-open-source-ai-model-from-china-boasts-twice-the-capacity-of-chatgpt [https://perma.cc/RE8K-HM8M]; Romain Dillet, *While Tech Companies Play with OpenAI's API, This Startup Believes Small, In-House AI Models Will Win*, TECHCRUNCH (Oct. 23, 2023), https://techcrunch.com/2023/10/23/zenml-empowers-companies-to-build-their-own-ai-stack-based-on-open-source-tools/ [https://perma.cc/6UDW-PQVL].

Second, the executive order begins an initiative within the federal government to mitigate the risks[125] stemming from misleading AI-generated content by "foster[ing] capabilities for identifying and labeling synthetic content produced by AI systems."[126] The executive order requires the Secretary of Commerce to submit a report detailing the available methods for detecting AI-generated content, watermarking AI-generated content, preventing AI-generated child pornography, and preventing AI-generated pornography depicting real people.[127] As discussed above,[128] many of the dangers presented by AI agents involve the use of AI-generated content, such as deep-fakes, and the executive order's efforts to detect and label AI-generated content could be helpful in regulating the use of such content by AI agents. Of course, for these efforts to eventually be effective, they will need to be followed up with significant regulation beyond the submission of the report required by the executive order. Additionally, although these provisions address identifying that content is AI generated, they fall short of identifying the parties responsible for such content.[129]

Third, the executive order states that the Federal Communications Commission ("FCC") is "encouraged to consider" how AI will affect, inter alia, "unwanted robocalls and robotexts that are facilitated or exacerbated by AI."[130] Although this provision appears to be designed to address nuisances like telemarketers, a broader statutory interpretation might leave the door open for the FCC to interpret "robotexts that are facilitated or exacerbated by AI" as a mandate to regulate AI agents that operate over all communications networks the FCC regulates. This interpretation could empower the FCC to regulate many types of autonomous and anonymous AI agents. However, the extent to which the FCC has the resources and capabilities to enforce such a mandate remains to be seen.

Finally, combating emerging AI-related threats, including those posed by AI agents, will require regulatory agencies to employ AI experts. The executive order provides for new efforts to bring AI experts into federal government agencies, including a requirement that each federal government agency designate a Chief Artificial Intelligence Officer.[131] This requirement would ensure that federal agencies are staffed with AI professionals that have the knowledge and capabilities to assist the agencies in developing future regulation to address AI agents. However, adding AI experts to each agency without unifying laws from Congress is likely to create a patchwork of AI enforcement that bad actors can easily evade. Conversely, the law proposed in this Note would create a unified way to identify responsible parties associated with AI agents, thus allowing each agencies' experts to address responsible parties within the jurisdictions of their agencies.

---

[125] *See supra* 13.
[126] *See* Biden AI Order, *supra* note 92, at § 4.5 (targets risks posed by synthetic content).
[127] *See id.* § 4.5(a) (discussing collecting data and research about synthetic content).
[128] *See supra* 15 (discussing the use of AI-generated content in cybercrime).
[129] *See supra* 20 (discussing the distinction between identifying that something is AI-generated versus identifying its creator or deployer).
[130] *See* Biden AI Order, *supra* note 92, at § 8(e)(encouraging the FTC to consider AI's effect on, inter alia, robocalls and robotexts).
[131] *See* Biden AI Order, *supra* note 92, at § 10.1(b) ("the requirement to designate at each agency . . . a Chief Artificial Intelligence Officer . . . ").

Applying the license plate analogy, the executive order focuses on regulating automotive manufacturers and licensing those automotive manufacturers with the federal government. It also addresses building infrastructure within the federal government both to use and to regulate automobiles as a new technology. But the executive order is silent on requirements for the drivers of automobiles to identify themselves on public roadways. The need to regulate automotive manufacturers is clear, but it is not a replacement for a license plate law that requires individual drivers of those automobiles to be identifiable on public roadways. Likewise, regulating developers of powerful AI systems is not a substitute for identifying the deployers of AI agents in public.

3. Congressional Efforts to Regulate AI

In September 2023, following several Senate hearings with high-profile leaders in the field of AI, Senators Richard Blumenthal (D-CT) and Josh Hawley (R-MO) proposed the Bipartisan Framework for U.S. AI Act[132] to guide the discussions around AI legislation. The framework includes five main categories that outline guidelines for future AI legislation including: licensing AI, holding AI accountable for harms, defending national security in the context of AI, promoting transparency in AI systems, and protecting consumers and children.[133] Though this framework has a broad scope that is comparable to President Biden's executive order, its 521-word guidelines provide much less detail than the 19,692-word executive order[134] and has no force of law. Although nothing in the framework is targeted at AI agents specifically, the implementation of these broad ideas would likely affect AI agents, at least indirectly, in several important ways.

First, the framework suggests that companies creating powerful AI systems should obtain a license to create such systems from a new AI-specific independent oversight body.[135] These licensees would be required to provide details about their AI systems such that the new oversight body could "establish a public database and reporting so that consumers and researchers have easy access to A.I. model and system information, including when significant adverse incidents occur or failures in A.I. cause harms."[136] This guideline for a licensing scheme closely parallels the proposal in this Note in that it requires certain AI systems to be licensed with a federal agency so that if those AI systems cause harm, information about them might be obtained to redress such harms. Although the framework and the proposal in this Note might use similar methods, they address different problems. The framework proposes a licensing scheme for *companies* "developing sophisticated general-purpose A.I. models"[137] but not for the *users* that deploy AI agents based on such AI models. And the framework's licensing scheme does not focus on

---

[132] Blumenthal & Senator Hawley, *Bipartisan Framework for U.S. AI Act* (Oct. 2023) [hereinafter *Senate AI Framework*], https://www.blumenthal.senate.gov/imo/media/doc/09072023bipartisanaiframework.pdf [https://perma.cc/4DT7-F3Q8]; Varun Aggarwal, *Senate Leaders Propose New Bipartisan Framework for AI Regulation*, HARVARD J. L. & TECH (Oct. 12, 2023), https://jolt.law.harvard.edu/digest/senate-leaders-propose-new-bipartisan-framework-for-ai-regulation [https://perma.cc/2TWM-XC6D].

[133] *Senate AI Framework*, *supra* note 132.

[134] *Compare Senate AI Framework*, *supra* note 132, *with* Biden AI Order, *supra* note 92.

[135] *See Senate AI Framework*, *supra* note 132.

[136] *See id.*

[137] *Id.*

identification at all, instead it focuses on storing information from companies—which are *known* entities.[138] In fact, the licensing section of the framework uses the example of GPT-4.[139] Far from being anonymous, GPT-4 is one of the fastest growing and most popular AI systems ever.[140] Another difference is that the framework uses the information collected by the licensing scheme as a means for diagnosing problems in known AI systems when they cause harm[141]—not for identifying which AI system caused the harm or who deployed the system. In fact, this requirement for storing information about AI systems to subsequently diagnose them when they cause harm presupposes that the harm-causing AI system can be identified. In other words, it is not possible to use system-specific diagnostic information to find the cause of a harm when you do not know which system committed the harm. Therefore, the identification laws proposed in this Note would complement, not duplicate, the licensing scheme proposed in the framework because it would prevent anonymous AI agents from skirting the scheme's purpose to connect AI-enabled harms to known companies.

Second, the framework suggests that "users should have a right to an affirmative notice that they are interacting with an A.I. model or system."[142] This guideline, designed to prevent AI systems from pretending that they are human, invokes anonymous AI agents because the guideline suggests that the user is unclear about the identification of the party with whom they are interacting.[143] Notifying the user that they are interacting with an AI system is a positive initial step, but it is not sufficient to hold parties accountable for harms caused by the AI agents they deploy. As this Note proposes, these user notification laws need to include the identification of the party that deployed the AI agent and the identification of the party that created the underlying AI system because without such identification no liability can be assigned when AI agents operating anonymously cause harm. Harmed parties will find hollow consolation in knowing that they were harmed by an AI if they are unable to identify the party responsible for the AI.

Third, the framework suggests several guidelines for holding parties accountable for the actions of their AIs including "oversight body enforcement,"[144] "private rights of action,"[145] and the exclusion of AI from the Section 230[146] protections that have traditionally shielded tech

---

[138] *See id.* ("Companies developing sophisticated general-purpose A.I. models (e.g., GPT-4) or models used in high-risk situations (e.g., facial recognition) should be required to register with an independent oversight body.").

[139] *See id.* ("companies developing sophisticated general-purpose A.I. models (e.g., GPT-4)").

[140] *See, e.g.*, Jon Porter, *ChatGPT Continues to be One of the Fastest-Growing Services Ever*, THE VERGE (Nov. 6, 2023), https://www.theverge.com/2023/11/6/23948386/chatgpt-active-user-count-openai-developer-conference [https://perma.cc/VQ6S-XJRW].

[141] *See Senate AI Framework*, *supra* note 132 ("easy access to A.I. model and system information, including when significant adverse incidents occur or failures in A.I. cause harms").

[142] *See id.*

[143] *Id.*

[144] *See id.* ("Congress should ensure that A.I. companies can be held liable through oversight body enforcement and private rights of action. . . .").

[145] *See id.* ("Congress should ensure that A.I. companies can be held liable through oversight body enforcement and private rights of action. . . .").

[146] *See id.* ("Congress should ensure that enforcers and victims can take companies and perpetrators to court, including clarifying that Section 230 does not apply to A.I."); 47 U.S.C. § 230.

companies from certain liabilities.[147] For these measures to be effective if implemented, harmed parties must have the ability to identify the parties that harmed them—something that might not be possible when the harm was caused by an anonymous AI agent. The proposal in this Note could allow the framework's proposed new liability laws to be more easily applied to responsible parties by providing for the identification of an AI agent's responsible parties.

Fourth, the framework refers to election interference only briefly,[148] and President Biden's executive order[149] makes no mention of elections or election interference at all.[150] This lack of emphasis on election interference might indicate that Congress and the President are primarily reacting to problems that have proven to be most exacerbated by AI in the past rather than predicting where AI will be most problematic in the future. Notably, DHS has taken a more forward-looking approach and is concerned that AI could be used by the nation's adversaries to "undermin[e] trust in government institutions, our social cohesion, and democratic processes."[151] Additionally, the National Security Commission on Artificial Intelligence, an independent commission of experts tasked by the federal government to consider the effects of AI on national security, warned that "AI is deepening the threat posed by cyberattacks and disinformation campaigns that Russia, China, and others are using to infiltrate our society, steal our data, and interfere in our democracy."[152]

In 2024, the United States experienced first-hand the effects of the latest AI technologies, including AI agents, on its election process. The U.S Department of the Treasury sanctioned entities in Iran and Russian that attempted to interfere with U.S. elections.[153] These sanctions specifically alleged that the Russian GRU "used generative AI tools to quickly create disinformation that would be distributed across a massive network of websites designed to imitate legitimate news outlets to create false corroboration between the stories, as well as to obfuscate their Russian origin."[154] Foreign actors were not the only culprits using AI to interfere with the election.[155] "AI-generated robocalls were sent to voters . . . with a fake version of President Joe

---

[147] *See, e.g.*, Doe #1 v. Twitter, Inc., No. 22-15103, No. 22-15104, 2023 WL 3220912, at *1 (9th Cir. May 3, 2023) ("Our court has held that section 230 'provides broad immunity' for claims against interactive computer service providers "for publishing content provided primarily by third parties.") (citing Carafano v. Metrosplash.com, Inc., 339 F.3d 1119, 1123 (9th Cir. 2003)); *but see* Lemmon v. Snap, Inc., 995 F.3d 1085 (9th Cir. 2021) (holding that Section 230 protections do not apply when a social media company's feature, not a third party's post on the platform, is responsible for a car accident killing three teenagers).

[148] *See Senate AI Framework*, *supra* note 132 ("Congress must take steps to directly prohibit harms that are already emerging from A.I., such as non-consensual explicit deepfake imagery of real people, production of child sexual abuse material from generative A.I., and *election interference*.") (emphasis added).

[149] *See supra* 25.

[150] *See* Biden AI Order, *supra* note 92 (making no reference to elections).

[151] *Homeland Threat Assessment 2024*, *supra* note 12.

[152] National Security Commission on Artificial Intelligence, *Final Report*, at 7 (2012), https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf [https://perma.cc/LGV6-NXRF].

[153] *See* Press Release, U.S. Dep't of the Treasury, Treasury Sanctions Entities in Iran and Russia That Attempted to Interfere in the U.S. 2024 Election (Dec. 31, 2024), https://home.treasury.gov/news/press-releases/jy2766 [https://perma.cc/LDF9-H5M].

[154] *Id.*

[155] Jacqui Goddard, *Russian Lies May Incite US Election Violence, Security Chiefs Say*, TIMES (Nov. 5, 2024), https://www.thetimes.com/world/us-world/article/russian-lies-may-incite-us-election-violence-security-chiefs-say-vvt9qdcbs [https://perma.cc/Q9M3-2NPY] ("Foreign influence is not the only element of the 'disinformation nightmare' that Americans fear could affect the election results.")

Biden's voice urging people not to vote in the election."[156] Some of these interference activities used AI Agents in the form of bots that "operate like people and cannot be distinguished" and can "like, comment, share, give more fire to extreme views."[157] These events demonstrate a need to enact election-specific AI laws,[158] as well as a need to prevent AI agents from acting anonymously in public forums.

Finally, the framework also provides guidelines for preventing harm caused by deepfakes by requiring AI-generated content to contain a specific watermark.[159] As discussed above,[160] these watermarking requirements could help to mitigate some of the risks caused by AI agents but fall short of providing accountability for anonymous AI agents when not paired with identification laws.

Using the license plate analogy, the framework, like the executive order, focuses on the companies that manufacture automobiles as well as the general prevention of the manufacture of harmful automobiles. It does not address a requirement to identify individual drivers of automobiles on public roadways. Thus, the laws proposed in the framework would benefit from a licensing scheme that makes AI agents identifiable.

4. State Laws

Several states, including California, Colorado, Utah, and Illinois have passed laws that address issues related to AI.[161] Like President Biden's Executive Order and the congressional proposal for a Bipartisan Framework for AI, these laws primarily focus on large generative AI models, disclosures that systems use AI, AI discrimination, and AI deepfakes,[162] but they do little to address the specific harms caused by AI agents. Specifically, these laws do not require that AI agents identify their deployers, so long as they identify that they are AI and not human or contain deepfaked information. Accordingly, these laws potentially leave room for anonymous AI agents to break laws with little chance of being having their deployers held accountable.

i. California

California recently enacted a host of laws that affect AI—over ten in September of 2024 alone.[163] These laws include Senate Bill 942: California AI Transparency Act ("California AI

---

[156] Aimee Picchi, *AI Chatbots Are Serving Up Wildly Inaccurate Election Information, New Study Says*, CBS NEWS (Feb. 27, 2024), https://www.cbsnews.com/news/ai-chatbots-inaccurate-election-information-proof-news [https://perma.cc/3XF5-ZG5X]. The FCC enforced civil penalties against the source of the calls. *See* FBI Press Release, *supra* note 61.

[157] Carine Harb, *Did AI Influence The Election Result? Experts' Verdicts*, NEWSWEEK (Nov. 20, 2024), https://www.newsweek.com/did-ai-influence-election-result-experts-verdicts-1987628 [https://perma.cc/99L9-7EGK].

[158] Some states have enacted AI-specific election laws, such as California's Assembly Bill 2839 discussed in Section III.A.4.i.

[159] *See Senate AI framework*, *supra* note 132.

[160] *See supra* 13 (discussing harms from AI-generated content); *see also supra* 25 (discussing how President Biden's executive order addresses watermarking of AI-generated content).

[161] *See infra* III.A.4.i (California); *infra* III.A.4.ii (Colorado); *infra* III.A.4.iii (Utah); *infra* III.A.4.iv (Illinois).

[162] *See e.g.*, The California AI Transparency Act in Section III.A.4.i (targeting systems with over 1,000,000 monthly visitors); Utah Senate Bill 149 in Section III.A.4.iii (requiring disclosure that a system is AI and not human); Colorado Senate Bill 24-205 in Section III.A.4.ii (addressing AI discrimination); California Assembly Bill 2655 in Section III.A.4.i (targeting deepfakes).

[163] Hope Anderson, Nick Reem & Sara Tadayyon, *Raft of California AI Legislation Adds to Growing Patchwork of US Regulation*, WHITE

Transparency Act"),[164] Assembly Bill 2013: Generative AI: Training Data Transparency Act ("California AI Training Act"),[165] Assembly Bill 3030: Health Care Services: Artificial Intelligence Act ("California AI Health Care Act"),[166] Assembly Bill 2839: Elections: deceptive media in advertisements ("California AI Elections Law"),[167] Assembly Bill 1008—which amends California's consumer privacy laws ("California AI Privacy Amendment"),[168] Assembly Bill 2602: Contracts against public policy: personal or professional services: digital replicas ("California AI Replicas Law I"),[169] Assembly Bill 1836: Use of likeness: digital replica ("California AI Replicas Law II"),[170] Assembly Bill 2655: Defending Democracy from Deepfake Deception Act of 2024 ("California AI Deepfake Act"),[171] Senate Bill 896: Generative Artificial Intelligence Accountability Act ("California AI Accountability Act"),[172] and Assembly Bill 2885: Artificial Intelligence ("California AI Definition Law").[173]

To help illuminate how California's spate of new AI laws might affect AI agents, this Note will address them in three categories: laws that regulate specific use cases for AI, laws that modify definitions in existing laws to apply them to AI, and laws targeted at governing AI systems specifically.

Among the laws in the first category, laws that regulate specific use cases for AI, are the California AI Health Care Act, the California AI Replicas Law I, and the California AI Replicas Law II. These laws target specific use-cases for AI in certain industries. The California AI Health Care Act requires healthcare facilities to obtain permission when using generative AI for certain types of health care activities.[174] The California AI Replicas Law I renders unenforceable contracts where the "use of a digital replica of the voice or likeness of an individual [is delivered] in lieu of the work of the individual"[175]—apparently in an effort to protect California's entertainment industry from becoming overcome by AI replicas. In a similar effort to protect the entertainment industry, the California AI Replicas Law II makes it illegal to produce a "digital replica of a deceased personality's voice or likeness" without prior permission from "any injured party,"[176] such as a performer's estate.

Another law in the first category is the California AI Accountability Act, which primarily

---

& CASE (Oct. 10, 2024), https://www.whitecase.com/insight-alert/raft-california-ai-legislation-adds-growing-patchwork-us-regulation [https://perma.cc/6AK3-V9CD] (summarizing some of the September 2024 California AI legislation).

[164] S.B. 942 2023–24 Leg. (Cal. 2024) [hereinafter Cal. AI Transparency Act] (approved by Governor Sept. 19, 2024).

[165] A.B. 2013 2023–24 Leg. (Cal. 2024) [hereinafter Cal. AI Training Act] (approved by Governor Sept. 28, 2024).

[166] A.B. 3030 2023–24 Leg. (Cal. 2024) [hereinafter Cal. AI Health Care Act] (approved by Governor Sept. 28, 2024).

[167] A.B. 2839 2023–24 Leg. (Cal. 2024) [hereinafter Cal. AI Elections Law] (approved by Governor Sept. 17, 2024).

[168] A.B. 1008 2023–24 Leg. (Cal. 2024) [hereinafter Cal. AI Privacy Amendment I] (approved by Governor Sept. 28, 2024); S.B. 1223 2023–24 Leg. (Cal. 2024) [hereinafter Cal. AI Privacy Amendment II] (approved by Governor Sept. 28, 2024).

[169] A.B. 2602 2023–24 Leg. (Cal. 2024) [hereinafter Cal. AI Replicas Law I] (approved by Governor Sept. 17, 2024).

[170] A.B. 1836 2023–24 Leg. (Cal. 2024) [hereinafter Cal. AI Replicas Law II] (approved by Governor Sept. 17, 2024).

[171] A.B. 2655 2023–24 Leg. (Cal. 2024) [hereinafter Cal. AI Deepfake Act] (approved by Governor Sept. 17, 2024).

[172] S.B. 896 2023–24 Leg. (Cal. 2024) [hereinafter Cal. AI Accountability Act] (approved by Governor Sept. 29, 2024).

[173] A.B. 2885 2023-24 Leg. (Cal. 2024) [hereinafter Cal. AI Definition Law] (approved by Governor Sept. 28, 2024).

[174] Cal. AI Health Care Act, *supra* note 166.

[175] Cal. AI Replicas Law I, *supra* note 169.

[176] Cal. AI Replicas Law II, *supra* note 170.

focuses on the use of AI by state agencies and government actors.[177] Although it is conceivable that such agencies and actors might use AI agents, governments are not a significant source of the type of harm contemplated by this Note.

Laws regulating specific use cases for AI, such as entertainment or government, could benefit from making their deployers readily identifiable through the law proposed in this Note because such deployers could not escape enforcement by operating anonymously.

Yet one specific use case for AI—influencing elections—affects AI agents more significantly because they are often deployed in ways that can affect elections.[178] The California AI Elections Law[179] and the California AI Deepfake Act both address AI's influence on elections. The California AI Elections Law is intended to prohibit "knowingly distributing an advertisement or other election communication . . . that contains certain materially deceptive content."[180] Similarly, the California AI Deepfake Act would require some platforms "to block the posting of materially deceptive content related to elections in California."[181] AI agents, such as social media bots, deployed to affect elections could potentially distribute deceptive content covered by either of these laws. Enforcing the California AI Deepfake Act would not require identifying that the content's source was an AI agent because a platform could identify that content is deceptive and block it without identifying the specific person that posted it. Conversely, the knowingly requirement in the California AI Elections Law would require a finding of the content creator's intent. Whether courts would consider the AI agent or its deployer to be the creator for the purposes of establishing such intent remains to be seen as courts apply the law. Yet for such a law to be enforceable against a deployer, that person would need to be identifiable—a task made substantially easier by the law proposed in this Note.

Among the laws in the second category, laws that modify definitions in existing laws to apply them to AI, is the California AI Definition Law. This law modifies the "the term 'artificial intelligence' [in existing state laws, including social media regulation] to mean an engineered or machine-based system that *varies in its level of autonomy* and that can, for *explicit or implicit objectives*, infer from the input it receives how to generate outputs that can *influence physical or virtual environments*."[182] The introduction of the terms "varies in its level of autonomy," "explicit or implicit objectives," and "influence physical or virtual environments" are likely to be applied to include AI agents in the application of the laws they affect because the terms are broader than the terms in the definitions they replace in ways that could encompass AI agents. For example,

---

[177] Cal. AI Accountability Act, *supra* note 172.

[178] *See generally supra* 13.

[179] There is some question as to whether the California AI Elections Law will survive judicial review. One lower California court has already found it unconstitutional, stating that it "does not pass constitutional scrutiny because the law does not use the least restrictive means available for advancing the State's interest here." *Kohls v. Bonta*, 2024 WL 4374134, at *1 (E.D. Cal. Oct. 2, 2024).

[180] Cal. AI Elections Law, *supra* note 167.

[181] Cal. AI Deepfake Act, *supra* note 171.

[182] Cal. AI Definition Law, *supra* note 173 (emphasis added). This definition is similar to a Colorado law that preceded it, as well as EU AI Act. *See* Arsen Kourinian et. al, *Colorado Governor Signs Comprehensive AI Bill*, MAYER BROWN (June 4, 2024), https://www.mayerbrown.com/en/insights/publications/2024/06/colorado-governor-signs-comprehensive-ai-bill [https://perma.cc/3Y3L-GM8S]; S.B. 24-205 2024 Leg. (Colo. 2024).

existing California law "defines an 'automated decision system' as a computational process derived from machine learning, statistical modeling, data analytics, or artificial intelligence that issues simplified output, including a score, classification, or recommendation, that is used to assist or replace human discretionary decisionmaking and materially impacts natural persons."[183] AI agents are unlikely to meet this previous definition of "automated decision systems" because they do not often merely issue simplified output like a "score, classification, or recommendation" and they might not be replacing "human discretionary decisionmaking"—a term likely implemented in an era where automated systems were replacing discrete human tasks and not acting autonomously. The new definition for AI, on the other hand, is expressly broad concerning levels of autonomy and includes both implicit and explicit objectives. Additionally, the law moved from the narrower reference to only effects that "materially impact[] natural persons" to the much broader reference to all effects that "influence physical or virtual environments." Therefore, the introduction of this new definition within existing California laws could allow courts to apply existing laws to AI systems, including AI agents, in new ways.

In addition to enacting the California AI Definition Law to broadly change the definition of AI in California laws, California also enacted the California AI Privacy Amendment to specifically alter California's existing privacy laws to ensure that AI is included in the laws' definitions of certain protected information. It expands the definitions of protected personal information to include "abstract digital formats" like "artificial intelligence systems."[184] These laws could affect AI agents that interact with the types of private information the laws protect.

Yet if AI agents are operating anonymously, enforcement of these against the AI agents' deployers would be almost impossible. The enforcement of the California AI Definition Law and the California AI Privacy Amendment against violations committed through AI agents would be more practical if the laws were combined with the law proposed in this Note because responsible parties could be identified and held accountable.

The license plate analogy illustrates how both laws applying to specific use cases for AI and laws extending definitions in existing laws to apply to AI could benefit from the identification law proposed in this Note. If governments were passing new laws to address specific activities that might be performed in a vehicle as well as updating existing laws to include vehicles, such laws could eliminate the problems associated with enforcement against anonymous parties if vehicles were required to be registered to responsible parties and identified by license plates.

The final category, laws targeted at governing AI systems specifically, includes the California AI Transparency Act and the California AI Training Act. Both of these acts have the potential to impact AI agents by regulating the generative AI systems on which they depend.

The California AI Transparency Act requires "covered provider[s]" of "generative artificial intelligence system[s]" to follow several requirements, including a watermarking requirement for

---

[183] Cal. AI Definition Law, *supra* note 173.
[184] Cal. AI Privacy Amendment I, *supra* note 168.

identifying certain content as being AI generated and detection requirement for providing a tool for detecting AI generated content.[185] At first glance, this law appears to not apply to AI agents because it defines a "content provider" that is subject to its requirements as someone who "produces a generative artificial intelligence system that has over 1,000,000 monthly visitors."[186] Unlike websites that provide AI services, AI agents do not often have monthly visitors. However, over time, courts might interpret the term "monthly visitors" broadly. For example, they might consider a social media bot whose posts are viewed more than 1,000,000 times per month to be a "covered provider" and require that bot to comply with the law's requirements. Like the watermarking requirements in President Bidens' Executive Order on AI, the watermarking requirement here would notify the public that content was generated by AI but would not require that a responsible party be identified. The comprehensive AI agent identification law proposed in this Note, on the other hand, would provide both notification that content was AI generated and notify the public of the party responsible for the content.

The California AI Training Act primarily focuses on requiring developers of "generative artificial intelligence system[s]" to post documentation "regarding the data used to train" those AI systems.[187] AI agents generally rely on existing generative AI systems, such as LLMs, for the underlying intelligence that would require the types of training datasets addressed by this new law.[188] So, like President Biden's Executive Order, the California AI Transparency Act will likely affect the generative AI systems that support AI agents without having much impact directly on AI agents or their identification.

These laws in the final category will help make the generative AI systems used by AI agents to be safer but fall short of requiring identification of parties that deploy AI agents. In the license plate analogy, these laws are analogous to laws that govern car manufacturers. Those laws help make cars on the road less likely to be involved in accidents but do little to identify parties responsible for accidents that do occur. Although California has passed the most sweeping set of AI laws, other states have also recently added new AI laws.

ii. Colorado

In May of 2024, Colorado enacted Senate Bill 24-205 ("Colorado AI Law").[189] The Colorado AI Law is primarily targeted at prohibiting algorithmic discrimination by AI systems against its citizens.[190] Algorithm discrimination occurs when important decisions made by an

---

[185] Cal. AI Transparency Act, *supra* note 164.

[186] *Id.*

[187] Cal. Civ. Code §§ 3110-3111 (Deering 2024).

[188] *See generally supra* Section II.A.

[189] COLO. REV. STAT. §§ 6-1-1701 - 6-1-1707 (2024); *see also* Arsen Kourinian et al., *Colorado Governor Signs Comprehensive AI Bill*, MAYER BROWN (June 4, 2024), https://www.mayerbrown.com/en/insights/publications/2024/06/colorado-governor-signs-comprehensive-ai-bill [https://perma.cc/8ABU-MSPX].

[190] For example, the first two main sections of the bill are titled "Developer duty to avoid algorithmic discrimination - required documentation" and "Deployer duty to avoid algorithmic discrimination - risk management policy and program." COLO. REV. STAT. §§ 6-1-1702 - 6-1-1703 (2024).

algorithm, such as an AI system, are biased against a particular group.[191] For example, employment decisions made by AI systems are often biased against groups that might be underrepresented in the historical hiring data on which the systems are trained, such as racial minorities or people with disabilities.[192] Although AI agents could be deployed to make these types of decisions, they would typically not be operating anonymously in such situations because algorithmic discrimination is generally related to activities like employment or housing decisions where the person or organization operating the AI system is obvious. For example, in an employment decision it is obvious that the potential employer is the deployer of the AI system. So even though AI agents might be used to perform tasks that result in algorithmic discrimination, the use of AI agents in such a role is outside the scope of this Note because the AI agents are not acting anonymously.

The Colorado AI Law also contains a disclosure requirement that could apply to AI agents. The law requires that AI systems[193] that interact with consumers shall disclose "that the consumer is interacting with an artificial intelligence system."[194] This provision might help to prevent consumers from being deceived into thinking they are interacting with a human when they are actually interacting with an AI system. However, this law falls short of requiring the deployer to identify themself, so long as they disclose "that the consumer is interacting with an artificial intelligence system."[195] So, unlike the law proposed in this Note, the Colorado AI Law does not prevent AI agents from being deployed anonymously or help connect their harms to a responsible party.

iii. Utah

In March of 2024, Utah enacted Senate Bill 149, the Utah Artificial Intelligence Policy Act.[196] This law primarily focuses on a disclosure requirement that requires AI systems to "clearly and conspicuously disclose . . . that the person is interacting with generative artificial intelligence and not a human."[197] Like the Colorado AI Law's disclosure requirement, the Utah Artificial Intelligence Policy Act requires that AI systems disclose to consumers that the AI systems are not human but the law does not require the AI system to identify its deployer. So this law too falls short of preventing AI agents from being deployed in an anonymous manner or helping to connect their harms to a responsible party.

---

[191] *See, e.g.*, *Colorado is First in Nation to Pass Legislation Tackling Threat of AI Bias in Pivotal Decisions*, CBS NEWS (May 24, 2024), https://www.cbsnews.com/news/ai-colorado-law-algorithms-bias-antidiscrimination [https://perma.cc/82PS-Q3GG].

[192] *See e.g.*, Sarah Roth & Becca Delbos, *AI is Causing Massive Hiring Discrimination Based on Disability*, HILL (Apr. 8, 2024), https://thehill.com/opinion/technology/4576649-ai-is-causing-massive-hiring-discrimination-based-on-disability/ [https://perma.cc/739Q-UBSM].

[193] The Colorado law uses a broader definition of AI, similar to the EU and the new California laws. COLO. REV. STAT. § 6-1-1701 (2024). ("Artificial intelligence system' means any machine-based system that, for any explicit or implicit objective, infers from the inputs the system receives how to generate outputs, including content, decisions, predictions, or recommendations, that can influence physical or virtual environments."); *see also* Kourinian et al., *supra* note 182.

[194] S.B. 24-205, § 6-1-1704, 2024 Leg. (Colo. 2024).

[195] *Id.*

[196] S.B. 149, § 13-2-12, 2024 Leg. (Utah 2024); *see also* Scott F. Young & Jordan C. Hilton, *Utah Enacts AI-Focused Consumer Protection Bill*, MAYER BROWN (May 13, 2024), https://www.mayerbrown.com/en/insights/publications/2024/05/utah-enacts-ai-focused-consumer-protection-bill [https://perma.cc/TH2G-NLHN].

[197] S.B. 149, § 13-2-12, 2024 Leg. (Utah 2024).

iv. Illinois

In August of 2024, Illinois enacted House Bill 3773, amending the Illinois Human Rights Act to address the use of AI in employment decisions.[198] This law is similar to the Colorado AI Law's prohibition on algorithmic discrimination, although it is tailored more towards employment than the general decision making activities covered by the Colorado law. As discussed above, laws prohibiting algorithmic discrimination address different harms than those addressed in this Note and such harms are outside the scope of this Note.

5. AI alignment

AI alignment is one way that AI agents might be controlled through non-legal means. AI alignment efforts work to create AIs that are designed to align with human values; they try to "keep[] AI systems working for humans, no matter how powerful the technology becomes."[199] These efforts are not necessarily specific to AI agents, instead they generally focus on overarching AI safety issues, such as creating "ethical guardrails [inside] the AI" that will prevent AI systems from operating "beyond preset ethical boundaries."[200] In fact, AI alignment is often targeted at preventing what is arguably the largest overarching concern about AI—the "extinction risks to humanity."[201] But not all AI alignment efforts focus on such existential threats; many work to resolve more targeted problems like AI bias and misuse.[202]

These initiatives are important and could help prevent AI from being misused in the first place. Yet, they are not a substitute for effective laws that work to prevent the harms that result from AI when it is misused. The law proposed in this Note would support AI alignment efforts because it could help to identify when and how AI is misused. Such identification of misuse is important to AI alignment because it allows changes in AI designs that might prevent future occurrences of the types of misuse that are identified.

Using the license plate analogy, AI alignment efforts focus on making automobiles generally safer—primarily by encouraging safety practices by automobile manufacturers. Although safer automobiles are better, they will still be involved in accidents or used intentionally in criminal activities. No matter how safe an automobile might be, it should still not be operated

---

[198] H.B. 3773 103 Gen. Assemb. (Ill. 2024); Arsen Kourinian, Kyla Miller & Charles E. Harris, *Illinois Passes Artificial Intelligence (AI) Law Regulating Employment Use Cases*, MAYER BROWN (Sept. 9. 2024), https://www.mayerbrown.com/en/insights/publications/2024/09/illinois-passes-artificial-intelligence-ai-law-regulating-employment-use-cases [https://perma.cc/8N5J-8HJS].

[199] Daniel Langkilde, *Why Business Leaders Should Understand AI Alignment*, FORBES (Oct. 6, 2023), https://www.forbes.com/sites/forbestechcouncil/2023/10/06/why-business-leaders-should-understand-ai-alignment/ [https://perma.cc/3WJJ-5H7Z].

[200] *See* Lance Eliot, *Implanting Legal Reasoning Into AI Could Smartly Attain Human-Value Alignment Says AI Ethics And AI Law*, FORBES (Dec. 1, 2022), https://www.forbes.com/sites/lanceeliot/2022/12/01/implanting-legal-reasoning-into-ai-could-smartly-attain-human-value-alignment-says-ai-ethics-and-ai-law/ [https://perma.cc/YV3A-37UV] (quoting Lance Eliot, *Crafting Ethical AI That Monitors Unethical AI And Tries To Deter Bad AI From Acting Up*, FORBES (Mar. 28, 2022)).

[201] Eliza Strickland, *OpenAI's Moonshot: Solving the AI Alignment Problem*, IEEE SPECTRUM (Aug. 31, 2023), https://spectrum.ieee.org/the-alignment-problem-openai [https://perma.cc/3YCL-XP8H].

[202] Stephen Khan, *What is 'AI Alignment'? Silicon Valley's Favourite Way to Think about AI Safety Misses the Real Issues*, CONVERSATION (July 11, 2023), https://theconversation.com/what-is-ai-alignment-silicon-valleys-favourite-way-to-think-about-ai-safety-misses-the-real-issues-209330 [https://perma.cc/44C8-XWNW].

anonymously in public. Similarly, AI alignment efforts will be good for AI as a whole, but individual AI agents should still be identifiable.

B. Efforts to regulate specific AI agents

The laws discussed so far were designed to address AI generally but were not targeted specifically at AI agents. However, a few states have passed laws that are specific to AI agents. These laws are narrowly tailored to address the problem of social media bots.

1. Bot laws

California, Maryland, and New Jersey have enacted so-called bot laws that apply a disclosure requirement to AI agents operating as fake users of online accounts, e.g., social media bots.[203] The California law requires bots, which it defines as "an automated online account" that is not operated by a person,[204] to "inform persons with whom the bot communicates or interacts that it is a bot,"[205] but only when it is used to "incentivize a . . . commercial transaction or to influence a vote in an election."[206] The New Jersey law is similar and applies to activities such as the "sale or advertisement of any merchandise or real estate or to solicit support for any candidate."[207] The Maryland bot law is specific to political campaigns and does not apply to commercial transactions.[208]

All of these laws are limited to "online accounts," such as social media accounts.[209] Although regulating social media bots is an important part of regulating AI agents, social media is not the only place that AI agents might cause harm.[210] Additionally, like the AI disclosure guidelines discussed above,[211] these laws are not sufficient to hold users of AI agents accountable because they merely require an AI agent to disclose its presence but does not require it to disclose the human parties responsible for its acts.[212]

2. S. 2125

In 2019, Senator Diane Feinstein proposed S. 2125, 116th Cong. § 4(a)(1)(A) (2019), a bill

---

[203] CAL. BUS. & PROF. CODE §§ 17940–17943; MD. CODE ANN., ELEC. LAW § 13-401; N. J. Rev. Stat. § 56:18-2 (2023).

[204] CAL. BUS. & PROF. CODE § 17940(a) ("'Bot' means an automated online account where all or substantially all of the actions or posts of that account are not the result of a person.").

[205] CAL. BUS. & PROF. CODE § 17941(b) ("The disclosure required by this section shall be clear, conspicuous, and reasonably designed to inform persons with whom the bot communicates or interacts that it is a bot.").

[206] CAL. BUS. & PROF. CODE § 17941(a) (defining conditions when a bot must identify itself).

[207] N. J. Rev. Stat. § 56:18-2 (2023).

[208] MD. ELECTION L. CODE § 13-401.1 ("section applies to any candidate, campaign finance entity").

[209] In fact, the California and Maryland laws have character-for-character identical definitions of the term bot. *Compare* CAL. BUS. & PROF. CODE § 17940(a) ("'Bot' means an automated online account where all or substantially all of the actions or posts of that account are not the result of a person.") *with* MD. ELECTION L. CODE § 13-401.1(b)(2) ("'Bot' means an automated online account where all or substantially all of the actions or posts of that account are not the result of a person."). New Jersey's law reflects the same language. N.J. REV. STAT. § 56:18-1 (2023) ("'Bot' means an automated online account where all or substantially all of the actions or posts of that account are not directly generated by a live natural person.").

[210] *See supra* 13.

[211] *See supra* 29.

[212] Distributing political material might inherently disclose a responsible party but not necessarily—AI agents can also distribute political material about an opponent.

that would have created a law requiring that (a) social media websites disclose when a user's account is being operated by an automated process (e.g., a social media bot), and (b) political candidates and organizations cannot use any software "intended to impersonate or replicate human activity online."[213] Put more simply, social media platforms would have to identify social media bots and political organizations could not make use of them.

Like the bot laws discussed above, this legislation was targeted at identifying non-human activities and restricting the use of social media bots in elections, but *not* at identifying deployers of such AI agents to tie a responsible party to their harms.[214]

Using the license plate analogy, both the bot laws and S. 2125 regulate identification for very specific automobiles in very specific circumstances. Such laws might be comparable to making license plates required only on certain roads or when automobiles are driven for certain organizations. Although such laws work toward making some automobiles identifiable, they fall short of making all automobiles identifiable all of the time. Similarly, the bot laws and S. 2125 fall short of requiring all AI agents to be identifiable all of the time.

IV. PROPOSAL

A. Proposed law

This Note proposes a federal law requiring a person or an entity that deploys autonomous AI agents ("Deployer") to register with a regulatory agency ("Agency") and provide identification for each AI agent ("Agent") tied to the registration ("Registration"). An Agent that must register will be defined as an artificial intelligence process that is (a) operating autonomously and (b) initiates communications with third parties.

This legal definition is slightly different from the non-legal definition in the technical background above, where an AI agent was defined as an "AI system that runs autonomously to perform a service."[215] The primary difference is that the broader term, "perform a service," has been narrowed to include only those AI agents that "initiate[] communications with third parties." As discussed below, this distinction exempts from regulation those AI agents that are not deployed to interact with entities other than their Deployer or are simply responding to communications initiated by other entities.

The term "artificial intelligence" will be defined broadly because the scope of this law is designed to be limited by the definitions of "operating autonomously" and "initiates communications with third parties," as discussed below. Using a broad definition for artificial intelligence—combined with setting limitations through these other terms—avoids the problems

---

[213] S. 2125, 116th Cong. § 4(a)(1)(A) (2019); S. 2125, 116th Cong. § 325(a)(1)(A) (2019).

[214] The requirements in section § 4(c) require social media platforms to "provide clear and conspicuous notice of the automated program," and take other actions such as removing posts made by those programs—but there is no requirement to identify a deployer. S. 2125, 116th Cong. § 4(c) (2019).

[215] *See supra* Section II.A.

associated with defining artificial intelligence[216] and the confusion created by the natural overlap between the terms "artificial intelligence" and "autonomous."

The term "autonomously" will be defined through rules promulgated by the Agency. As discussed above,[217] autonomy is not a binary condition; instead, levels of autonomy exist along a spectrum. The definition of such a spectrum and its application to Agents are both too complex and too likely to change over time to be defined in the statute. Therefore, this definition should be left to a more fluid rulemaking procedure that considers input from subject matter experts inside and outside of the Agency.

Unlike the definitions in the bot laws that focus on particular technologies by using terms like "online accounts,"[218] the definition of autonomy promulgated in the Agency's rules must be designed to reach a broader set of forms of communication, such as internet communications, text messaging, video conferencing, the metaverse, or communications mechanisms not yet invented. Additionally, this definition of communications should encapsulate in-person communications between humans and AI agents in the form of humanoid robots—a form of communication that will likely exist in the foreseeable future.[219]

The term "initiates communications with third parties" exempts two important categories of Deployers. The first category includes Deployers that merely receive, as opposed to "initiate," communications. For example, a web server might be powered by an autonomous AI system but does not initiate communications, instead it merely processes communications requests initiated by others. In another example, a customer service AI might be interacting with customers who initiated communications to receive assistance; such an AI would be exempted from a licensing requirement so long as the AI is only handling incoming requests and not initiating unwanted communications. The second category includes Deployers that do not "communicate with third parties." Communicating with third parties will be defined as communications where the communication reaches a destination entity that is not the Deployer. For example, if an Agent sends an email to its Deployer, then such communication is exempt. But if the Agent emails a third-party then the communication falls within the proposed registration law and would require a license.

Additionally, Agents that merely collect data will be specifically exempted from the definition of communications. Some AI agents are designed to collect data from external sources without otherwise interacting with such sources. For example, AI agents that "screen scrape"—a term used to describe automatically retrieving data from the internet—would not be considered to be communicating under the definition of the proposed law.[220] This exemption is undergirded by

---

[216] *See supra* note 26.

[217] *See supra* 13 (discussing autonomy).

[218] *See supra* 45 (discussing the bot laws).

[219] *See, e.g.*, Camilo Fonseca, *Elon Musk Welcomes Competition from Humanoid Robot Rivals: 'Bring It On'*, BUSINESS INSIDER (Mar. 1, 2024), https://businessinsider.com/elon-musk-comments-figure-ai-rival-openai-robot-machines-2024-3 [https://perma.cc/2LUZ-WUE8].

[220] Alexander S. Gillis, *Screen Scraping*, TECHTARGET, https://www.techtarget.com/searchdatacenter/definition/screen-scraping [https://perma.cc/ZH4A-B4CK] ("Screen scraping is a data collection method used to gather information shown on a display to use for

the policy rationale that harms from AI agents are generally caused when they are producing information and not just passively consuming it. For example, interfering with elections requires actively providing information, such as through a social media post. Likewise, the fraud schemes discussed above require the two-way exchange of data, such as providing stolen credit card information, and not its mere passive collection.[221]

One issue faced by this data collection exemption is the fact that all interactions on the Internet require two-way communications. For example, when a user visits a website, the user's computer must create and send a request to a server that hosts the website.[222] To avoid this technical question of whether data is actually exchanged, the law might use an intent-based clause like those used in the bot laws. For example, the law might include language exempting Agents whose Deployers "intend to merely passively collect data." Such an exemption would minimize the regulatory burden on the common practice of automated data collection.

The proposed law will require a registration process whereby Deployers can register and receive a license from the Agency to deploy their Agents. This Agency might be the FTC, FCC, or a new government entity designed to regulate AI specifically.[223] This task should be handled by a single one of these entities, as opposed to creating a scheme that allows duplicate enforcement. Determining which of these agencies should be the Agency defined in this law is beyond the scope of this Note.

Unlike the licensing requirements discussed by President Biden's AI Executive Order, this license will be issued to Deployers of AI agents—those that launch AI agents into the public—not the companies that develop the AI technology. This focus on Deployers instead of developers regulates how AI agents are *used* rather than how they are *made*.

Registered Deployers must identify themselves as the Deployer and also identify the developer or manufacturer ("Developer") of the Agent. This is comparable to how vehicle registration requires the name of the registered car owner in addition to the make and model of the vehicle. This requirement allows for liabilities to be apportioned correctly when harms occur. With vehicles, liability for a vehicle-caused harm might be attributed to the vehicle's manufacturer, its driver, or both. Likewise, liability for an Agent-caused harm might be attributed to the Agent's Developer, its Deployer, or both. So, under this proposed law, harmed parties can use the Registration information to identify both the Deployer and the Developer of the Agent.

---

another purpose. Typically, screen scraping is used to collect data from one application to then translate it to another—or, more controversially, to steal data.").

[221] Some completely passive data collection could still be nefarious, such as passively collecting stolen credit card information. However, broadening the scope of this proposed law to include such activity is unnecessary because these activities are likely already illegal under existing laws. *See generally* DANIEL J. SOLOVE & PAUL M. SCHWARTZ, PRIVACY LAW FUNDAMENTALS 43–47 (7th ed. 2024) (discussing the Electronic Communications Privacy Act's prohibitions on intercepting electronic communications); *id.* at 167 ("The Computer Fraud and Abuse Act (CFAA) prevents unauthorized access to computers.").

[222] *See generally*, *An Overview of HTTP*, MDN WEB DOCS, https://developer.mozilla.org/en-US/docs/Web/HTTP/Overview [https://perma.cc/A5HZ-8UDE] ("To display a Web page, the browser sends an original request [to the server] to fetch the HTML document that represents the page.").

[223] This Note recommends that Congress either create a new entity to regulate AI or properly fund the FTC or FCC if additional regulatory burdens are placed on them.

After Registration is completed, licensed Deployers can deploy their Agents. Deployed Agents must comply with an identification protocol that allows the agents to be identified in a secure manner. This protocol could be modeled after the HTTPS protocol that secures web traffic and identifies it source.[224] Such a protocol could ensure that Registrations are not faked by malicious AI agents.

In addition to the underlying identification protocol, rules would need to be promulgated for how and when protocol verifications are triggered. For example, in a social media environment, the underlying social media platform might process the proper identification protocol and notify its users that a certain account is an Agent and disclose the Deployer and Developer associated with the Registration. In other circumstances the underlying platform might not be capable of such verification. For example, when an Agent makes a telephone call, the underlying network might not be capable of running the identification protocol. In these cases, a rule might be promulgated by the Agency that defines keywords or other signals that trigger an identification. In the telephone example, an agent might have to respond to a question like "please provide your agent identification" by providing the registration information to the called party.

This proposed law could enable a simple regulatory mechanism that drastically curbs the harms associated with autonomous and anonymous AI agents by making their deployers identifiable and therefore accountable.

B. Counterarguments

1. Enforcement Feasibility

Enforcement will be critical to the effectiveness of this proposed law. As with any law, the agency tasked with the enforcement of this law will need to receive the funding required to police communications activities by AI agents and take actions against noncompliant deployers. One unique problem faced by the regulation of AI agent activities is the need to separate the activities of an AI agent from human activities prior to enforcement for noncompliance. In other words, noncompliant AI agents might simply pretend to be human to avoid compliance. However, this issue can be overcome using automated AI agent identification in public communications as well as reporting mechanisms in private communications. In public communications, such as public social media posts, AI-driven algorithms can identify posts that likely come from AI agents.[225] In private communications, such as phone calls or direct messages, where government AI-analysis would be inappropriate or illegal, the agency tasked with enforcement might rely on reporting

---

[224] Rahul Awati, *Hypertext Transfer Protocol Secure (HTTPS)*, TECHTARGET, HTTPS://WWW.TECHTARGET.COM/SEARCHSOFTWAREQUALITY/DEFINITION/HTTPS [https://perma.cc/2E5Z-WFBX] ("HTTPS is based on the TLS encryption protocol, which secures communications between two parties.").

[225] Orabi et al., *supra* note 47 (discussing techniques for identifying social media bot accounts through AI and other detection mechanisms). Although these techniques would be helpful in detecting potential violators of the registration law, they are not a suitable replacement identification system because once an offending IP address is identified then a subpoena is often required to learn the true identity of the deployer. *See, e.g.*, Arista Records, LLC v. Does 1-19, 551 F. Supp. 2d 1, 5-6 (D.D.C. 2008) (discussing obtaining a subpoena for an IP address associated with unauthorized sharing of copyrighted material).

mechanisms, just as the FTC does for violations of many consumer protection laws.[226]

In addition to agency enforcement, Congress could also include a private right of action. However, such an addition could slow the progress of AI technology with overburdensome regulation. For example, private rights of action can have substantial negative consequences for unsophisticated small businesses when strict liability for unintentional violations can be combined with high per-violation penalties.[227] So, while a private right of action would be an option for Congress, this Note suggests that such a private right of action should only be applicable to extreme cases that involve serious violations and not directed towards small businesses or individual developers. However, such limitations would not apply to agency enforcement actions where violations by small businesses or individual developers might be appropriately regulated by a neutral agency. For example, the agency tasked with enforcing the law might fine an individual developer for failing to license an AI agent, even when such a developer is exempted from private civil litigation.

2. Control over foreign actors

As discussed above, many of the bad actors responsible for AI-based fraud and election interference are based in foreign countries. For example, the virtual kidnappings discussed above were based in Mexico[228] and significant portions of election interference are thought to be driven by state actors.[229] This presents problems for both enforcement and jurisdiction. However, in many cases, the communications platforms used by these overseas actors are both under the jurisdiction of the U.S., and the U.S. is capable of enforcing the proposed law at the platform level. Consider the example above where social media networks would provide information about Agents and their Deployers. Such information could stop state actors from activities like election interference because their presence on the social media platform would be immediately identifiable. Also, as mentioned above, enforcement is possible using AI-driven algorithms to identify Agents that do not comply. In fact, the use of the proposed law and the suggested enforcement mechanisms could help government officials more readily identify sources of foreign crime and election interference because source identifiers, like internet IP addresses, controlled by these offending entities would be frequently flagged as noncompliant by the enforcement mechanisms discussed above.

---

[226] Consumers use an online portal to report violations of certain consumer protection laws directly to the FTC, allowing the agency to be aware of these violations without actively detecting them. *See, e.g.*, *Report to Help Fight Fraud!*, FED. TRADE COMM'N (2024), https://reportfraud.ftc.gov. For example, illegal communications, such as "robocalls," are reported through these mechanisms. *See National Do Not Call Registry*, FED. TRADE COMM'N (2024), https://www.donotcall.gov/report.html [https://perma.cc/994L-LJXY].

[227] *Cf.* Shari A. Kolnicki, *The Telephone Consumer Protection Act and Its Burden on Small Business: An Evaluation of the Law and Its Ramifications on Telecommunication Advances*, 28 CAP. U. L. REV. 223, 237 (1999) (discussing the burden of TCPA regulations on unsophisticated small businesses); Stuart L. Pardau, *Good Intentions and the Road to Regulatory Hell: How the TCPA Went from Consumer Protection Statute to Litigation Nightmare*, 2018 U. ILL. J.L. TECH. & POL'Y 313, 321 (2018) ("the damages exposure to defendants in TCPA litigation can skyrocket quickly, even for small and medium-sized businesses").

[228] *See* CNN, *supra* note 71.

[229] *See* DHS, *supra* note 12.

3. Preventing a stifling effect on innovation

Any regulation on technology brings with it concerns about stifling innovation. However, the requirement to identify a responsible party is a very low bar. Unlike many of the other invasive regulatory frameworks proposed by legislative bodies, this law can be implemented relatively easily and is unlikely to hinder legal activities. Deployers with legitimate reasons to deploy Agents need only register and make their Agents identifiable. Just as license plates on vehicles likely did not hinder the development of the automobile, registration for AI agents is unlikely to hinder AI innovation. In fact, it might help to increase public trust in AI innovations by removing the harms of many bad actors from the public discourse about AI.

V. CONCLUSION

Autonomous and anonymous AI agents pose threats to free elections, cybersecurity, consumer protection, and fraud prevention. The law proposed by this Note would make AI agents and their deployers clearly identifiable to the public. A relatively simple registration scheme could prevent bad actors from acting anonymously without negatively impacting AI innovation. Additionally, this registration scheme is inclusive enough to address many of the foreseeable advances in AI agent technologies.