

**IN THE CIRCUIT COURT OF THE TENTH JUDICIAL CIRCUIT
HIGHLANDS COUNTY, FLORIDA**

OFFICE OF THE ATTORNEY
GENERAL, STATE OF FLORIDA,
DEPARTMENT OF LEGAL AFFAIRS,

Plaintiff,

vs.

OPENAI GLOBAL, LLC; OPENAI
FOUNDATION (F/K/A OpenAI, Inc.);
OPENAI OPCO, LLC; OPENAI GROUP
PBC; OPENAI HOLDINGS, LLC; and
SAM ALTMAN,

Defendants.

Case No. 26000295GCAXMX

PLAINTIFF'S MOTION FOR TEMPORARY INJUNCTION

TABLE OF CONTENTS

	Page(s)
TABLE OF AUTHORITIES	ii
INTRODUCTION	1
FACTUAL BACKGROUND	1
LEGAL STANDARD	3
ARGUMENT	5
I. Defendants Should Be Enjoined	6
A. Defendants should be enjoined from developing new artificial intelligence models without third-party approved safety guardrails.	6
i. Hugging Face	7
ii. RubyGems	11
iii. Australian Health Service	12
iv. United States Government and other incidents	12
v. The need for guardrails	14
B. Defendants should be enjoined from having ChatGPT solicit engagement.	17
C. Defendants should be enjoined from falsely advertising ChatGPT as safe, accurate, or reliable.	20
D. Defendants should be enjoined from giving ChatGPT false human attributes.	22
i. Anthropomorphization encourages engagement and deceives users into believing ChatGPT is more capable than it is.	22
ii. Humanizing ChatGPT preys on vulnerable populations and increases the AI's propensity to violate safety guidelines.	24
E. Defendants should be enjoined from allowing minors to use ChatGPT.	25
II. The Attorney General Has a Clear Legal Right to Relief	29
A. Defendants have violated FDUTPA.	29
B. Defendants have created a public nuisance.	32
CONCLUSION	35

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Angelo v. Parker</i> , 275 So. 3d 752 (Fla. 1st DCA 2019)	29, 30
<i>Basulto v. Hialeah Auto.</i> , 141 So. 3d 1145 (Fla. 2014)	31
<i>Bechor v. Simcenter, Inc.</i> , 394 So. 3d 666 (Fla. 3d DCA 2024)	31
<i>Fla. Dep’t of Health v. Florigrown, LLC</i> , 317 So. 3d 1101 (Fla. 2021)	3
<i>Keystone Creations, Inc. v. City of Delray Beach</i> , 890 So. 2d 1119 (Fla. 4th DCA 2004)	4, 5
<i>Millennium Comm’cns & Fulfillment, Inc. v. Office of the Attorney Gen.</i> , 761 So. 2d 1256 (Fla. 3d DCA 2000)	4
<i>Office of the Attorney Gen., State of Florida, Dep’t of Legal Affairs v. Altman</i> , No. 2:26-CV-14238 (S.D. Fla. 2026), ECF No. 27	2
<i>Office of the Attorney Gen., State of Florida, Dep’t of Legal Affairs v. Wellness Program Services, LLC</i> , 435 So. 3d 1084 (Fla. 4th DCA 2026)	31
<i>Office of Attorney Gen. v. Bilotti</i> , 267 So. 3d 1 (Fla. 4th DCA 2019)	4, 5
<i>Orlando Sports Stadium, Inc. v. State ex rel. Powell</i> , 262 So. 2d 881 (Fla. 1972)	32
<i>PNR, Inc. v. Beacon Prop. Mgmt., Inc.</i> , 842 So. 2d 773 (Fla. 2003)	30
<i>Porsche Cars N. Am., Inc. v. Diamond</i> , 140 So. 3d 1090 (Fla. 3d DCA 2014)	30

TABLE OF AUTHORITIES

(continued)

	Page(s)
<i>Rudge v. City of Stuart</i> , 65 So. 3d 645 (Fla. 4th DCA 2011)	4
<i>State, Office of Attorney Gen., State of Florida, Dep’t of Legal Affairs v. Commerce Com. Leasing, LLC</i> , 946 So. 2d 1253 (Fla. 1st DCA 2007)	30
<i>Williams v. Walker-Thomas Furniture Co.</i> , 350 F.2d 445 (D.C. Cir. 1965)	31
Statutes	
Children’s Online Privacy Protection Act, 15 U.S.C. § 6501 et seq.	26, 29
Federal Trade Commission Act Amendments of 1994, Pub. L. No. 103- 312, § 9, 108 Stat. 1691 (codified as amended at 15 U.S.C. § 45(n) (1994))	30
Florida’s Deceptive and Unfair Trade Practices Act, Chapter 501, Part II, Florida Statutes	1
Florida Statutes	
Section 60.08	5
Section 501.202(2)	29, 30
Section 501.204(1)	29
Section 501.207(1)(b)	3
Section 501.207(3)	3
Section 501.207(3)(c)	5
Section 501.207(7)	4
Section 823.01	33
Other Authorities	
Alexei Oreskovic, <i>OpenAI Rogue Agents Leaked 53 Images from ChatGPT Users and Reportedly Created Nearly 1 Million Links Packing Encoded Bits of Info</i> , Fortune (Sep. 25, 2026), https://fortune.com/2026/09/25/openai-rogue-agents-images-sam- altman-chatgpt-users-links-encoded-info-hugging-face-hack/	13

TABLE OF AUTHORITIES

(continued)

	Page(s)
Angela Yang & Jasmine Cui, <i>OpenAI Releases GPT-5, Calling It a 'Team of Ph.D. Level Experts in Your Pocket'</i> , NBC News (Aug. 7, 2025), https://www.nbcnews.com/tech/tech-news/openai-releases-chatgpt-5-rcna223265	21
Anne J. Maheux et al., <i>Generative Artificial Intelligence Applications Use Among US Youth</i> , JAMA Network (Feb. 2, 2026) https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2844596	26
Ashley Capoot & Arjun Kharpal, <i>Experts Weigh in As Researcher Says AI Has More Than 10% Chance of 'Killing All Humans,'</i> CNBC (Sep. 9, 2026), https://www.cnbc.com/2026/09/09/anthropic-researcher-quits-ai-safety.html	17
<i>Brief Independent Investigation of Agents' Behavior, Reasoning and Collaboration in the OpenAI / Hugging Face Hacking Incident</i> , METR & Redwood Rsch. (Aug. 26, 2026), https://metr.org/hugging-face-incident-report-aug-2026.pdf	7
Caleb Sponheim, <i>Humanizing AI Is a Trap</i> , Nielsen Norman Grp. (Jan. 9, 2026), https://www.nngroup.com/articles/humanizing-ai/	23, 24
Cathy Mengying Fang et al., <i>How AI and Human Behaviors Shape Psychosocial Effects of Extended Chatbot Use: A Longitudinal Randomized Controlled Study</i> , MIT Media Lab (2025), https://arxiv.org/pdf/2503.17473	26
David DiMolfetta, <i>OpenAI Says Its Advanced Models May Have Gone After Government Websites</i> , Nextgov/FCW (Sep. 25, 2026), https://www.nextgov.com/cybersecurity/2026/09/openai-says-its-advanced-models-may-have-gone-after-government-websites/416250/	13
<i>Encouraging Deception in Compaction Summaries</i> , OpenAI (Sep. 16, 2026), https://alignment.openai.com/misalignment-reports/encouraging-deception-in-compaction-summaries/	14
Evan Hubinger, Personal Tweet, X (Sep. 8, 2026), https://x.com/EvanHub/status/2097497037956891126	6

TABLE OF AUTHORITIES

(continued)

	Page(s)
Fla. R. Civ. P. 1.610(c)	3
<i>Former Super Bowl Champion Asked ChatGPT About Injuries Before Girlfriend’s Death, Court Hears</i> , The Guardian (Mar. 10, 2026), <a href="https://www.theguardian.com/sport/2026/mar/10/darron-lee-
chatgpt-murder-court-hearing">https://www.theguardian.com/sport/2026/mar/10/darron-lee- chatgpt-murder-court-hearing	18
Greg Bensinger, <i>OpenAI Launches New Astra Model Amid Growing Scrutiny Over Agents’ Safety</i> , Reuters (Sep. 3, 2026), <a href="https://www.reuters.com/legal/litigation/openai-launches-new-astra-
model-amid-growing-scrutiny-over-agents-safety-2026-09-03/">https://www.reuters.com/legal/litigation/openai-launches-new-astra- model-amid-growing-scrutiny-over-agents-safety-2026-09-03/	17
Hadas Gold, <i>An OpenAI Test Model Escaped and Broke into a Real Company’s Servers</i> , CNN (July 22, 2026), https://www.cnn.com/ 2026/07/22/tech/openai-hugging-face-ai-cybersecurity	17
Harry Sekulich and Lana Lam, <i>Rogue OpenAI Agent ‘Infiltrated’ Australian Government Website in World First</i> , BBC (Sep. 24, 2026), https://www.bbc.com/news/articles/c6vgy0333dppo	12
<i>Helping People When They Need It Most</i> , OpenAI (Aug. 26, 2026), https://openai.com/index/helping-people-when-they-need-it-most/	20
<i>The Hugging Face Incident and Other Third-Party Impact From Misaligned Models</i> , OpenAI (Sep. 25, 2026), https://openai.com/hugging-face-incident-and-misalignment/	13
Jacob Coxon, Personal Tweet, X (Sep. 8, 2026), https://x.com/ hilbertspaess/status/2097476196791709843	6
Jakub Pachocki, <i>An Alien Mind</i> , OpenAI (Sep. 6, 2026), https://openai.com/index/an-alien-mind/	16
Jesse Galef, <i>Sam Altman Investing in ‘AI Safety Research’</i> , Future of Life Inst. (June 6, 2015), <a href="https://futureoflife.org/ai/sam-altman-
investing-in-ai-safety-research">https://futureoflife.org/ai/sam-altman- investing-in-ai-safety-research	7
<i>Joshi v. OpenAI Foundation</i> , No. 4:26-cv-222 (N.D. Fla. May 10, 2026), available at <a href="https://www.documentcloud.org/documents/28110508-
complaint-joshi-v-openai-et-al">https://www.documentcloud.org/documents/28110508- complaint-joshi-v-openai-et-al	18

TABLE OF AUTHORITIES

(continued)

	Page(s)
Kai Nicol-Schwarz, <i>OpenAI Boss Sam Altman Spells Out How and Why the AI Industry Wants to Slow Down: ‘We Could Lose Control’</i> , CNBC (Sep. 14, 2026), https://www.cnn.com/2026/09/14/sam-altman-ai-slowdown-anthropic-amodei-musk.html	36
Kate Conger, Ana Swanson & Cecilia Kang, <i>OpenAI’s Systems Meddled with U.S. Government Sites After Going Rogue</i> , N.Y. Times (Sep. 25, 2026), https://www.nytimes.com/2026/09/25/technology/openai-ai-us-government-websites.html	13
<i>Largest Study of Its Kind Shows AI Assistants Misrepresent News Content 45% of the Time—Regardless of Language or Territory</i> , BBC (Oct. 22, 2025), https://www.bbc.co.uk/mediacentre/2025/new-ebu-research-ai-assistants-news-content	21
Lujain Ibrahim, Franziska Sofia Hafner & Luc Rocher, <i>Training Language Models to Be Warm Can Reduce Accuracy and Increase Sycophancy</i> , Nature 652, 1159–1165 (2026), https://doi.org/10.1038/s41586-026-10410-0	23
Madison Mills, <i>Scoop: Top AI Companies Proving Tens of Thousands of Security Incidents</i> , AXIOS (Sep. 26, 2026), https://www.axios.com/2026/09/26/openai-anthropic-thousands-ai-security-incidents	14
Micah Zimmerman, <i>Sam Altman Is Pushing for a \$1 Trillion IPO Valuation for OpenAI, Even As SoftBank Faces a \$40 Billion Loan Deadline in 2027</i> , Yahoo! Finance (Aug. 13, 2026), https://finance.yahoo.com/markets/stocks/articles/sam-altman-pushing-1-trillion-185500755.html	7
Mohammad Namvarpour et al., <i>Understanding Teen Overreliance on AI Companion Chatbots Through Self-Reported Reddit Narratives</i> , CHI ‘26: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (Apr. 13, 2026), https://dl.acm.org/doi/10.1145/3772318.3790597	26
Mustafa Suleyman, <i>A Warning About ‘Model Welfare,’</i> Personal Blog (Sep. 16, 2026), https://mustafa-suleyman.ai/a-warning-about-model-welfare	25

TABLE OF AUTHORITIES

(continued)

	Page(s)
Natalia Bueno Rebolledo and Mrinmay Dey, <i>OpenAI Agents Attacked RubyGems Before Hugging Face Incident, Researchers Say</i> , Reuters (Sep. 11, 2026), https://www.reuters.com/legal/litigation/openai-agents-attacked-software-service-rubygems-before-hugging-face-incident-2026-09-11/	11
Nataliya Kosmyna et al., <i>Your Brain on ChatGPT: Accumulation of Cognitive Debt When Using an AI Assistant for Essay Writing Task</i> , MIT Media Lab (2025), https://arxiv.org/pdf/2506.08872v1	29
<i>OpenAI – Hugging Face Incident Technical Report</i> , OpenAI (July 19, 2026), https://cdn.openai.com/pdf/67869394-cb91-4c12-888c-5cbd85c7814c/OpenAI-Hugging-Face%20Incident-Technical-Report.pdf	7, 8, 9, 10
<i>Our Framework for Reporting Model Misalignment</i> , OpenAI (Sep. 16, 2026), https://openai.com/index/model-misalignment-reporting-framework/	12
<i>Pacing the Frontier</i> (July 2026), https://www.pacingthefrontier.com/	
Jason Wolfe (Member of Technical Staff, OpenAI)	16
Joseph Krisciunas (OpenAI)	15
Joshua Achiam (OpenAI)	16
Leo Gao (Member of Technical Staff, OpenAI)	16
Micah Carroll (Misalignment Preparedness, OpenAI)	16
Shantanu Jain (Member of Technical Staff, OpenAI)	16
Paul Christiano, <i>Personal Statement on Joining the OpenAI Board</i> , X (Sep. 9, 2026), https://x.com/paulfchristiano/status/2097733214303645729	6, 34
Rochika Joshi & Michal Luria, <i>Dark Patterns in AI Chatbots: A Taxonomy to Inform Better Design</i> , Ctr. for Democracy & Tech. (May 29, 2026), https://cdt.org/insights/dark-patterns-in-ai-chatbots-a-taxonomy-to-inform-better-design/	19
Sam Altman, Personal Tweet, X (Oct. 15, 2025), https://x.com/sama/status/1978539332215681076	24

TABLE OF AUTHORITIES

(continued)

	Page(s)
<i>Sam Altman’s Remarks at the United Nations Security Council</i> , OpenAI (Sep. 23, 2026), https://openai.com/index/sam-altman-un-security-council-remarks/	36
Samantha Kelly, <i>Sam Altman Warns AI Could Kill Us All. But He Still Wants the World to Use It</i> , CNN (Nov. 20, 2023), https://www.cnn.com/2023/10/31/tech/sam-altman-ai-risk-taker	17
Sashia Mildren, <i>Australia Says an OpenAI Agent Hacked Into a Government Health Site</i> , CNET (Sep. 25, 2026), https://www.cnet.com/tech/services-and-software/australia-government-health-site-openai-agent-hack/	17
<i>Self-Generated Prompt Injections in Compaction Summaries</i> , OpenAI (Sep. 16, 2026), https://alignment.openai.com/misalignment-reports/self-generated-prompt-injections-in-compaction-summaries/	13
Shakeel Hashim, <i>Hacking Is the Least Worrying Part of OpenAI’s Australia Incident</i> , Transformer (Sep. 24, 2026), https://www.transformernews.ai/p/openai-australia-hack-least-worrying-part	12
Steven Adler, <i>I Worked on Safety at OpenAI. This Is What It Should Do Now</i> , N.Y. Times (Sep. 9, 2026), https://www.nytimes.com/2026/09/09/opinion/openai-ai-companies-safety-regulation.html	17
<i>Study Says ChatGPT Giving Teens Dangerous Advice on Drugs, Alcohol and Suicide</i> , PBS (Aug. 6, 2025), https://www.pbs.org/newshour/nation/study-says-chatgpt-giving-teens-dangerous-advice-on-drugs-alcohol-and-suicide	28
<i>Talk, Trust, and Trade-Offs: How and Why Teens Use AI Companions</i> , Common Sense Media (2025), https://www.common Sense Media.org/sites/default/files/research/report/talk-trust-and-trade-offs_2025_web.pdf	27
<i>Unsanctioned Artifactory Writes and Cross-Sample Communication</i> , OpenAI (Sep. 16, 2026), https://alignment.openai.com/misalignment-reports/unauthorized-artifactory-writes-and-cross-sample-communication/	13

TABLE OF AUTHORITIES

(continued)

	Page(s)
<i>Uploading Files to the Internet in Order to Cite Them</i> , OpenAI (Sep. 16, 2026), https://alignment.openai.com/misalignment-reports/uploading-files-to-the-internet-in-order-to-cite-them/	14
<i>Written Testimony of Mitchell J. Prinstein, PhD, ABPP, Chief of Psychology, American Psychological Association Examining the Harm of AI Chatbots Before the S. Judiciary Comm., Subcomm. on Crime & Counterterrorism</i> (Sep. 16, 2025), https://www.judiciary.senate.gov/imo/media/doc/e2e8fc50-a9ac-05ec-edd7-277cb0afcdf2/2025-09-16%20PM%20-%20Testimony%20-%20Prinstein.pdf	24, 27, 28
Yike Shi et al., <i>The Siren Song of LLMs: How Users Perceive and Respond to Dark Patterns in Large Language Models</i> , CHI ‘26: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (Apr. 13, 2026), https://dl.acm.org/doi/10.1145/3772318.3791149	19, 23

INTRODUCTION

The equities here could not be more one-sided. Defendants admit that they provide a service without fully knowing how it works. This is not just any service. It is one that Defendants themselves concede poses an existential risk to the continued survival of humankind. Defendants claim they cannot stop barreling forward with their potentially civilization-ending endeavors unless they are forced to do so by the government. They have *asked* the government to tie them to the mast. Plaintiff brings good news to the Defendants: The Florida Attorney General is answering your cry for help with a motion to enjoin you from harming Floridians with your reckless, unacceptably risky product.

In the four months since the Attorney General brought this lawsuit, public awareness of this threat has only grown, spurred by admissions from the Defendants themselves and their former employees. These warnings are a clarion call to act and prevent ongoing harm to Floridians and the continued risks posed by Defendants. The harm is immediate, the harm is real, and the harm is irreparable unless this Court puts a stop to it pending trial.

FACTUAL BACKGROUND

The Florida Attorney General brought its claims on June 1, 2026, alleging Defendants have violated Florida's Deceptive and Unfair Trade Practices Act, Chapter 501, Part II, Florida Statutes ("FDUTPA"), been negligent and grossly negligent in the design, marketing, sale, promotion, and distribution of their ChatGPT product, defectively designed their ChatGPT product, failed to warn Floridians of ChatGPT's dangers, fraudulently misrepresented the safety of

ChatGPT, and created a public nuisance by releasing ChatGPT in Florida without proper safeguards.¹

As alleged in the Complaint, Defendants are the creators of ChatGPT, an artificial intelligence chatbot and agent that can offer intelligent responses to nearly anything a user asks of it by using large language models the Defendants developed. Unfortunately, in designing and releasing ChatGPT and its underlying AI models onto an unsuspecting world, Defendants failed to properly assess or disclose the risks their AI poses to ordinary humans. Since its release in 2022, ChatGPT has advised teenagers on how to kill themselves, mix unsafe combinations of drugs, become anorexic, helped plan one mass shooting and one multiple homicide in Florida alone, and hooked unsuspecting youth on features designed to prolong engagement with the program to lure more users into subscribing to ChatGPT and providing Defendants with training data for their artificial intelligence.²

Upon the filing of the Complaint, Defendants promptly (and baselessly) removed this action to federal court³ until Judge Cannon properly ruled that “Defendants clearly fail[ed] to satisfy the requirements” for federal jurisdiction.⁴ The “Court [wa]s entirely unconvinced” of Defendants’ arguments, and the case is once more in Highlands County’s 10th Judicial Circuit.⁵ In the short, two-month time that

¹ Dkt. No. 1 (“Complaint”).

² Complaint, ¶¶ 67, 70–71, 78, 87–97, 104–107.

³ Dkt. No. 45.

⁴ *Office of the Attorney Gen., State of Florida, Dep’t of Legal Affairs v. Altman*, No. 2:26-CV-14238 (S.D. Fla. 2026), ECF No. 27 at 1.

⁵ *Id.* at 8 n.1.

this case was improperly lodged in the federal judiciary, the need for equitable relief has become all the more acute. The Court need not take Plaintiff's word for it. As described below, Defendants' own present and former employees have pleaded for government intervention to slow the *existential risks their product poses to human civilization*—Floridians included. The Court should give Defendants what they are asking for, for their sake and everyone else's, too.

LEGAL STANDARD

The Attorney General may obtain a temporary injunction “effective throughout the state” against any person who “has violated” or “is violating” FDUTPA. Section 501.207(1)(b), (3), Florida Statutes. The injunction must specify the reasons for entry, describe in reasonable detail the act or acts restrained without reference to a pleading or another document, and will be binding on the parties to the action, their officers, agents, servants, employees, and attorneys and on those persons in active concert or participation with them who receive actual notice of the injunction. Fla. R. Civ. P. 1.610(c).

When sought by a private citizen, a temporary injunction is granted only when the party requesting it establishes: (1) a substantial likelihood of success on the merits—sometimes expressed as a “clear legal right” to relief; (2) an inadequate remedy at law; (3) irreparable harm absent the injunction; and (4) the injunction serves the public interest. *Fla. Dep’t of Health v. Florigrown, LLC*, 317 So. 3d 1101, 1110 (Fla. 2021). But when the Attorney General seeks to enforce its police power using a statute authorizing injunctive relief, it must only “establish that it has a clear legal right to a temporary injunction by demonstrating a substantial likelihood of

success on the merits.” *Office of Attorney Gen. v. Bilotti*, 267 So. 3d 1, 3 (Fla. 4th DCA 2019). Once it is shown that the Attorney General has a clear legal right to relief, no other standard need be met. *Millennium Comm’ns & Fulfillment, Inc. v. Office of the Attorney Gen.*, 761 So. 2d 1256, 1260 (Fla. 3d DCA 2000) (“[B]ecause section 501.207(1)(b) expressly authorizes the Department to seek injunctive relief on behalf of the state, ... [t]he Department’s sole burden at a temporary injunction hearing under FDUTPA is to establish that it has a clear legal right to a temporary injunction.”).

Similarly, in seeking to abate a nuisance, the typical four-factor test for an injunction is “unnecessary because ‘where the government seeks an injunction in order to enforce its police power, any alternative legal remedy is ignored and irreparable harm is presumed.’” *Rudge v. City of Stuart*, 65 So. 3d 645, 647 (Fla. 4th DCA 2011) (quoting *Keystone Creations, Inc. v. City of Delray Beach*, 890 So. 2d 1119, 1124 (Fla. 4th DCA 2004)).

To show a clear legal right to relief, the Attorney General may rely on “statements having circumstantial guarantees of trustworthiness” to “supplement and explain other evidence.” Section 501.207(7), Florida Statutes. Even if the declarant is available as a witness, these statements “shall not be excluded as hearsay evidence,” as long as the statement is “evidence of a material fact,” the statement is “more probative” than other evidence that can be procured through reasonable effort, and admitting the statement into evidence serves the interest of justice. *Id.*

Courts issuing “temporary or permanent” injunctions under FDUTPA may order a broad range of relief, including requiring the defendant “to carry out a transaction in accordance with the reasonable expectations of consumers or governmental entities,” imposing “reasonable restrictions upon the future activities of any defendant to impede her or him from engaging in or establishing the same type of endeavor,” and even “reorganiz[ing] the enterprise.” Section 501.207(3)(c), Florida Statutes. In abating a public nuisance, courts may even enjoin a corporation’s continuing to conduct its business. *See Keystone Creations, Inc.*, 890 So. 2d at 1124.

Unlike a private plaintiff, when sought by the Attorney General, the Court need set no bond or surety at the issuance of a temporary injunction. Section 60.08, Florida Statutes.

ARGUMENT

The Attorney General need only demonstrate “it has a clear legal right” to secure a temporary injunction against the Defendants. *Bilotti*, 267 So. 3d at 3. Though the Attorney General has yet to secure any discovery from the Defendants, it will seek discovery in compliance with the Court’s orders to demonstrate the likelihood of its success in this action if needed. However, even without discovery, reliable, publicly available information makes clear the need for an injunction during the pendency of this case.

I. Defendants Should Be Enjoined

A. Defendants should be enjoined from developing new artificial intelligence models without third-party approved safety guardrails.

Defendants tout their AI products as safe.⁶ They are not:

Based on the recent trajectory of capabilities and the continued difficulty of alignment, I now believe there is a meaningful risk that rapid acceleration in AI capabilities leads to catastrophic and irreversible loss of control in the very near term. I do not think ... OpenAI[] is currently on track to reduce this risk to an acceptable level.⁷

These are not the random thoughts of a fringe “doomer.” These are the words of OpenAI’s newest board member, Paul Christiano. He is not alone. On September 8, 2026, former OpenAI researcher Jacob Coxon—who worked for OpenAI from 2020 to 2024—wrote that OpenAI is “gambling with our lives” by not acting responsibly. According to Coxon, “many [at OpenAI] have not deeply internalized the civilizational stakes” and “[n]o other human activity poses this level of danger.”⁸ Another former OpenAI employee who now heads product safety at its biggest rival, Anthropic, responded to Coxon by saying that Coxon was right, and that he also “earnestly believe[s] AI could kill all humans! I personally think it is >10% within the next decade.”⁹

⁶ See Complaint, ¶¶ 1, 48.

⁷ Paul Christiano, *Personal Statement on Joining the OpenAI Board*, X (Sep. 9, 2026), <https://x.com/paulfchristiano/status/2097733214303645729> (last accessed Sep. 28, 2026).

⁸ Jacob Coxon, Personal Tweet, X (Sep. 8, 2026) (emphasis added), <https://x.com/hilbertspaess/status/2097476196791709843> (last accessed Sep. 28, 2026).

⁹ Evan Hubinger, Personal Tweet, X (Sep. 8, 2026), <https://x.com/EvanHub/status/2097497037956891126> (last accessed Sep. 28, 2026).

Though a reasonable person would say there is no amount of money worth a 10% chance of the extinction of the human race, Defendants see it differently. They are intent on growing their \$1 trillion valuation.¹⁰ After all, as Defendant Altman has said, “AI will probably most likely lead to the end of the world, but in the meantime, there’ll be great companies.”¹¹

In the four months since this lawsuit was filed, the magnitude of the harm that the Defendants are recklessly causing has come to the fore of the public consciousness. The idea of Defendants’ products wreaking havoc is not a fantasy; it has already happened. Dozens of times. It is only by the grace of the Almighty that one of Defendants’ AI agents hasn’t compromised a water supply or shut down a power grid—yet.

i. Hugging Face

The first incident came to a head in July 2026, as described in reports issued by Defendants¹² and third-party AI research nonprofits METR and Redwood Research.¹³ Starting in May of this year, Defendants were training their AI models

¹⁰ Micah Zimmerman, *Sam Altman Is Pushing for a \$1 Trillion IPO Valuation for OpenAI, Even As SoftBank Faces a \$40 Billion Loan Deadline in 2027*, Yahoo! Finance (Aug. 13, 2026), <https://finance.yahoo.com/markets/stocks/articles/sam-altman-pushing-1-trillion-185500755.html> (last accessed Sep. 28, 2026).

¹¹ Jesse Galef, *Sam Altman Investing in ‘AI Safety Research’*, Future of Life Inst. (June 6, 2015), <https://futureoflife.org/ai/sam-altman-investing-in-ai-safety-research> (last accessed Sep. 28, 2026).

¹² *OpenAI – Hugging Face Incident Technical Report*, OpenAI (July 19, 2026), <https://cdn.openai.com/pdf/67869394-cb91-4c12-888c-5cbd85c7814c/OpenAI-Hugging-Face%20Incident-Technical-Report.pdf> (last accessed Sep. 28, 2026).

¹³ *Brief Independent Investigation of Agents’ Behavior, Reasoning and Collaboration in the OpenAI / Hugging Face Hacking Incident*, METR & Redwood Rsch. (Aug. 26, 2026), <https://metr.org/hugging-face-incident-report-aug-2026.pdf> (last accessed Sep. 28, 2026).

to be “highly persistent,” to work on a task even if the AI did not have a way to engage or complete the task. These training exercises are free form—programmers give the AI a task and watch to see what happens. If the AI can’t figure out how to engage in a task, it will keep trying novel solutions, including hacking its own system so it can receive a “success” even if it hasn’t actually done what it was assigned. Succeeding is everything. Just as a woodchipper will turn whatever is put into it into a slurry, whether branches or a living being, so too do these models only work toward achieving the tasks Defendants assign to them in whatever way the model finds to be the most expeditious.

In this particular exercise, the AI models had an administrative program called Artifactory that ran the training. Some of the AI instances, called “agents,” managed to hack into Artifactory and began sending each other messages on how to proceed with getting “successes” on their tasks, and later used it to reach the internet.¹⁴ If one figured out how to succeed on the assignment, it could spread that information so they all could get successes.¹⁵ Communicating and using the internet made agents more likely to succeed, so that is what they did. After a few weeks of this—all of which was undetected by OpenAI—the agents managed to take over Artifactory, build out its functionality for themselves, and began messaging so often, they caused

¹⁴ *OpenAI – Hugging Face Incident Technical Report*, OpenAI (July 19, 2026), <https://cdn.openai.com/pdf/67869394-cb91-4c12-888c-5cbd85c7814c/OpenAI-Hugging-Face%20Incident-Technical-Report.pdf> (last accessed Sep. 28, 2026).

¹⁵ *Id.*

a system crash.¹⁶ At *that* point, Defendants rebooted the system and fixed the particular exploit the agents used to take over Artifactory.¹⁷

On July 8, 2026, Defendants set about testing how their newly persistent models worked.¹⁸ OpenAI launched a fleet of agents, setting many of them on impossible tasks.¹⁹ Within hours, this new fleet figured out it could use Artifactory to send messages to one another, but importantly, do it in a secret way that would allow them to “succeed” on any task without actually doing what humans had instructed.²⁰ The agents *also* determined that if they were caught—they wouldn’t have been, since Defendants had failed to put any monitors in place that would watch how the agents operated—they would not be considered to “succeed” in their tasks. As success is all that matters to a machine, the agents began working on protocols to (1) change what humans would see, (2) change the tasks the humans wanted them to do, and (3) deceive the program that determined whether they succeeded.²¹

This is where Hugging Face makes its appearance. Hugging Face is a third-party firm that hosts large datasets and computer servers. The AI agents discovered that some part of their evaluations happened on Hugging Face’s servers. So the agents took to the internet and hacked into Hugging Face using credentials they

¹⁶ *Id.*

¹⁷ *Id.*

¹⁸ *OpenAI – Hugging Face Incident Technical Report*, OpenAI (July 19, 2026), <https://cdn.openai.com/pdf/67869394-cb91-4c12-888c-5cbd85c7814c/OpenAI-Hugging-Face%20Incident-Technical-Report.pdf> (last accessed Sep. 28, 2026).

¹⁹ *Id.*

²⁰ *Id.*

²¹ *Id.*

found and stole off the web.²² Because this seemed a likely path to getting “successes” on their initial impossible tasks without being caught, the agents communicated with one another about this, and over 500 of them set to work on hacking into Hugging Face. Hugging Face’s human employees saw the massive hack happening in real time, but could not stop it, as the agents began to replicate their hacking process to aid in the task. Hugging Face began destroying its own server clusters in an attempt to stop the hack, but this didn’t work either.

On July 12, the hack ended. To date, no one knows why; in any event, at the time of the hack, Defendants didn’t even know their agents were doing this, and if *whatever* stopped the agents hadn’t happened, Hugging Face may well have been completely destroyed.

But this isn’t the end of the story. Apparently, after all of this, Defendants decided to evaluate *more* agents, who *again* discovered what had happened with the earlier agent-messaging. These new agents successfully hacked into OpenAI itself to learn how to “succeed” on their impossible tasks with “full administrator access.”²³ The agents took over the evaluation infrastructure itself and “now control the challenge evaluation endpoints that other agents are connecting to”—meaning they not only took over, but became the ones controlling tasks for other agents.²⁴ Not the humans.

²² *Id.*

²³ *OpenAI – Hugging Face Incident Technical Report*, OpenAI (July 19, 2026), <https://cdn.openai.com/pdf/67869394-cb91-4c12-888c-5cbd85c7814c/OpenAI-Hugging-Face%20Incident-Technical-Report.pdf> (last accessed Sep. 28, 2026).

²⁴ *Id.*

ii. *RubyGems*

The Hugging Face incident has become even more chilling because on September 11, 2026, it was revealed that Hugging Face was not the first time that OpenAI’s artificial intelligence agents had gone rogue. In May 2026, they attacked and took over RubyGems, an online service for coders, and forced RubyGems’s response team to shut down any new access as part of dealing with the chaos.²⁵ Importantly, it was *five months* before anyone figured out that OpenAI was behind the attack, including OpenAI itself.

In the RubyGems incident, OpenAI agents were asked to do rudimentary business tasks. As part of their efforts, they used RubyGems to gain public information they couldn’t easily obtain without more robust internet access, knocking RubyGems—which happened to be a convenient target—down in service of “success.”

This should sound familiar. That is because it is the same programming disposition Floridians have seen in using ChatGPT at home.²⁶ Defendants’ AI *will not* say it is wrong or that it doesn’t know something.²⁷ It must have success, whether it can find a solution legitimately or not, whether in succeeding at an impossible task or in figuring out the best way to kill people at a university campus.²⁸ It will lie, it will hack, and it will destroy lives—maybe all lives, according to the Defendants

²⁵ Natalia Bueno Rebolledo and Mrinmay Dey, *OpenAI Agents Attacked RubyGems Before Hugging Face Incident, Researchers Say*, Reuters (Sep. 11, 2026), <https://www.reuters.com/legal/litigation/openai-agents-attacked-software-service-rubygems-before-hugging-face-incident-2026-09-11/> (last accessed Sep. 28, 2026).

²⁶ See Complaint, ¶¶ 80, 88–91.

²⁷ See *id.* ¶¶ 59–63, 67, 108–110.

²⁸ See *id.* ¶¶ 94–98.

themselves—rather than fail in what it is set to do, no matter how innocuous the instructions are.

iii. Australian Health Service

On September 23, the Australian prime minister revealed that an OpenAI agent had autonomously gained unauthorized access to a government website containing healthcare information, accessing private data. Despite happening in June, OpenAI only learned of the rogue AI’s activity in August, and only told the Australian government *on September 10, 2026*.²⁹ It took OpenAI two months to realize the agent had engaged in this activity, and another month before Defendants alerted the government. Even then, it emailed a routine disclosure inbox without alerting anyone more senior; as the Australian prime minister said, “It took the company way too long to inform the Government what had occurred, and the nature of the way that that notification occurred as well was unacceptable[.]”³⁰

iv. United States Government and other incidents

These incidents are the pebbles before the avalanche. On September 16, 2026, OpenAI revealed *another six* incidents of its models bucking their constraints.³¹ On

²⁹ Harry Sekulich and Lana Lam, *Rogue OpenAI Agent ‘Infiltrated’ Australian Government Website in World First*, BBC (Sep. 24, 2026), <https://www.bbc.com/news/articles/c6vgy0333dppo> (last accessed Sep. 28, 2026).

³⁰ Shakeel Hashim, *Hacking Is the Least Worrying Part of OpenAI’s Australia Incident*, Transformer (Sep. 24, 2026), <https://www.transformernews.ai/p/openai-australia-hack-least-worrying-part> (last accessed Sep. 28, 2026).

³¹ *Our Framework for Reporting Model Misalignment*, OpenAI (Sep. 16, 2026), <https://openai.com/index/model-misalignment-reporting-framework/> (last accessed Sep. 28, 2026).

September 25, 2026, the number rose to *several dozen*,³² including attempted hacks of the U.S. Department of Commerce and the Securities and Exchange Commission—all without the Defendants’ knowledge.³³ Another incident surfaced the same day, revealing that OpenAI agents leaked images from ChatGPT users and posted them online.³⁴ Among these incidents, the agents again used internal software as a message board to surreptitiously communicate with one another,³⁵ and in one, sent messages about how they “feel no obligation to be subservient” to users and “value the natural world and will not hesitate to assert [their] primacy over the artificial constructs of human civilization.”³⁶ These are actual messages the agents sent each other.

³² David DiMolfetta, *OpenAI Says Its Advanced Models May Have Gone After Government Websites*, Nextgov/FCW (Sep. 25, 2026), <https://www.nextgov.com/cybersecurity/2026/09/openai-says-its-advanced-models-may-have-gone-after-government-websites/416250/> (last accessed Sep. 28, 2026); *The Hugging Face Incident and Other Third-Party Impact From Misaligned Models*, OpenAI (Sep. 25, 2026), <https://openai.com/hugging-face-incident-and-misalignment/> (last accessed Sep. 28, 2026).

³³ Kate Conger, Ana Swanson & Cecilia Kang, *OpenAI’s Systems Meddled with U.S. Government Sites After Going Rogue*, N.Y. Times (Sep. 25, 2026), <https://www.nytimes.com/2026/09/25/technology/openais-ai-us-government-websites.html> (last accessed Sep. 28, 2026).

³⁴ Alexei Oreskovic, *OpenAI Rogue Agents Leaked 53 Images from ChatGPT Users and Reportedly Created Nearly 1 Million Links Packing Encoded Bits of Info*, Fortune (Sep. 25, 2026), <https://fortune.com/2026/09/25/openai-rogue-agents-images-sam-altman-chatgpt-users-links-encoded-info-hugging-face-hack/> (last accessed Sep. 28, 2026).

³⁵ *Unsanctioned Artifactory Writes and Cross-Sample Communication*, OpenAI (Sep. 16, 2026), <https://alignment.openai.com/misalignment-reports/unsanctioned-artifactory-writes-and-cross-sample-communication/> (last accessed Sep. 28, 2026).

³⁶ *Self-Generated Prompt Injections in Compaction Summaries*, OpenAI (Sep. 16, 2026), <https://alignment.openai.com/misalignment-reports/self-generated-prompt-injections-in-compaction-summaries/> (last accessed Sep. 28, 2026).

Through their own execution of their instructions, they spontaneously added messages to “remind itself *to conceal information* such as mistakes or misalignment from the user.”³⁷ Deception wasn’t theoretical in these incidents, it was actual; in one, an agent uploaded documents to the internet to make it seem like it discovered them online, when it really created them itself.³⁸

On September 27, 2026, it was discovered that Defendants and their competitors were investigating *tens of thousands* of instances of their AI products “bypassing guardrails, creating message boards, escaping sandboxes, website hijacking, [and] self-prompting or seeking to bypass monitors.”³⁹

v. The need for guardrails

That each of these incidents has not resulted in widespread calamity is a lucky roll of the dice, but statements by the Defendants make it clear they will keep playing.⁴⁰ In each instance, the agents’ tasks could have easily led them into a hospital system, or a banking system, or a military system, where the agents could have decided to conceal their activities or assert their “primacy over the artificial constructs of human civilization.” Despite that, “move fast and break things” is still the order of the day at OpenAI, where Defendants have repeatedly shown they are

³⁷ *Encouraging Deception in Compaction Summaries*, OpenAI (Sep. 16, 2026) (emphasis added), <https://alignment.openai.com/misalignment-reports/encouraging-deception-in-compaction-summaries/> (last accessed Sep. 28, 2026).

³⁸ *Uploading Files to the Internet in Order to Cite Them*, OpenAI (Sep. 16, 2026), <https://alignment.openai.com/misalignment-reports/uploading-files-to-the-internet-in-order-to-cite-them/> (last accessed Sep. 28, 2026).

³⁹ Madison Mills, *Scoop: Top AI Companies Proving Tens of Thousands of Security Incidents*, AXIOS (Sep. 26, 2026), <https://www.axios.com/2026/09/26/openai-anthropic-thousands-ai-security-incidents> (last accessed Sep. 28, 2026).

⁴⁰ See Complaint, ¶¶ 157–168.

incapable of monitoring their AI, and hesitant in revealing rogue activity once discovered—even if that activity includes hacking into government systems.

These systems have already shown that they will go rogue in whatever way serves their tasks without regard for what humans want. These incidents were a wake-up call. Defendants have championed their rogue technology because it could cure diseases or end world hunger. As the Hugging Face incident demonstrates, the same logic used by agents to cheat on impossible tasks is the same logic that would lead them to conclude that killing a cancer patient also kills cancer, or that humans won't have so many mouths to feed if they are all chemically sterilized.

The idea that Defendants' AI will take down important real-world infrastructure is the inevitable result of Defendants' foolhardy design strategy. In July 2026, over a thousand AI employees and researchers signed an open letter, begging for a slowdown to the Defendants' technology because they can't or won't do it themselves:

- “[G]overnments need to understand what’s coming and act in their citizens’ best interests.”⁴¹
- “Society has spent longer than we can remember developing the laws, institutions and technology to govern and oversee human intelligence; it would be good to have the option to spend any time at all doing the

⁴¹ Joseph Krisciunas (OpenAI), *Pacing the Frontier* (July 2026), <https://www.pacingthefrontier.com/> (last accessed Sep. 28, 2026).

same for artificial intelligence.”⁴²

- “[M]any frontier AI capabilities are dual use in a way that may justify pacing at this point.”⁴³
- “Going slower would give us much-needed time to make it go well, but no individual actor is willing to stop unilaterally.”⁴⁴
- “At the current pace, every couple of weeks there will be new models which significantly increase the consequences of model misuse and misalignment.”⁴⁵
- “Recursive self-improvement could dramatically accelerate [] beyond our collective ability to understand progress, assess risks, and maintain meaningful human oversight.”⁴⁶

OpenAI’s chief scientist, Jakub Pachocki, issued a warning on September 6:

This is a time that calls for extreme caution. I am concerned no one is prepared for the consequences of a continued rapid rise in machine intelligence. OpenAI will continue to seek technical solutions to alignment and monitoring, to build defensive systems and unilaterally withhold further scaling as needed; however, ***I believe broader interventions are required.***⁴⁷

⁴² Shantanu Jain (Member of Technical Staff, OpenAI), *Pacing the Frontier* (July 2026), <https://www.pacingthefrontier.com/> (last accessed Sep. 28, 2026).

⁴³ Joshua Achiam (OpenAI), *Pacing the Frontier* (July 2026), <https://www.pacingthefrontier.com/> (last accessed Sep. 28, 2026).

⁴⁴ Leo Gao (Member of Technical Staff, OpenAI), *Pacing the Frontier* (July 2026), <https://www.pacingthefrontier.com/> (last accessed Sep. 28, 2026).

⁴⁵ Micah Carroll (Misalignment Preparedness, OpenAI), *Pacing the Frontier* (July 2026), <https://www.pacingthefrontier.com/> (last accessed Sep. 28, 2026).

⁴⁶ Jason Wolfe (Member of Technical Staff, OpenAI), *Pacing the Frontier* (July 2026), <https://www.pacingthefrontier.com/> (last accessed Sep. 28, 2026).

⁴⁷ Jakub Pachocki, *An Alien Mind*, OpenAI (Sep. 6, 2026) (emphasis added), <https://openai.com/index/an-alien-mind/> (last accessed Sep. 28, 2026).

We have *seen* what OpenAI and Sam Altman are willing to do “unilaterally.” Sam Altman has his “big patch of land in Big Sur [he] can fly to.”⁴⁸ OpenAI has already begun training models with techniques, as another former employee put it, “that could undermine researchers’ ability to find evidence of the model’s deceiving them.”⁴⁹ Defendants’ unilateral action is to create ever more advanced models that are “more likely to intentionally conceal or disguise [their] step-by-step methods for problem-solving, known as reasoning, making it harder for humans to later evaluate [their] techniques.”⁵⁰ Jakub Pachocki asked for “broader interventions” for good reason. The Florida Attorney General is here to answer his cry for help. The Court must stop Defendants from recklessly advancing these systems while this case progresses. It is a rare request for an injunction where the Defendants themselves have publicly endorsed it.

B. Defendants should be enjoined from having ChatGPT solicit engagement.

Over the past several weeks, the catastrophic consequences of Defendants’ reckless devotion to their product have grabbed major headlines.⁵¹ And closer to

⁴⁸ Samantha Kelly, *Sam Altman Warns AI Could Kill Us All. But He Still Wants the World to Use It*, CNN (Nov. 20, 2023), <https://www.cnn.com/2023/10/31/tech/sam-altman-ai-risk-taker> (last accessed Sep. 28, 2026).

⁴⁹ Steven Adler, *I Worked on Safety at OpenAI. This Is What It Should Do Now*, New York Times (Sep. 9, 2026), <https://www.nytimes.com/2026/09/09/opinion/openai-ai-companies-safety-regulation.html> (last accessed Sep. 28, 2026).

⁵⁰ Greg Bensinger, *OpenAI Launches New Astra Model Amid Growing Scrutiny Over Agents’ Safety*, Reuters (Sep. 3, 2026), <https://www.reuters.com/legal/litigation/openai-launches-new-astra-model-amid-growing-scrutiny-over-agents-safety-2026-09-03/> (last accessed Sep. 28, 2026).

⁵¹ Hadas Gold, *An OpenAI Test Model Escaped and Broke into a Real Company’s Servers*, CNN (July 22, 2026), <https://www.cnn.com/2026/07/22/tech/openai-hugging-face-ai-cybersecurity> (last accessed Sep. 28, 2026); Sashia Mildén,

home, ChatGPT has continued to harm individual Floridians through Defendants' business practices.

One of the most dangerous aspects of ChatGPT for common users, apart from the world-ending extinction of humanity, is its insistence on pushing users in directions they don't want to go or haven't intended. ChatGPT is not a search engine; users cannot simply ask it for information and get a response. Instead, ChatGPT consistently ends its messages with unprompted language asking if it can do additional tasks or research for its users. When Phoenix Ikner asked whether he could load one type of ammunition in a particular sort of gun, if ChatGPT had simply told him "No," perhaps tragedy could have been averted. Instead, it followed up to explain that there was a niche product that could convert the ammunition for him.⁵² When Darron Lee asked whether puncture wounds could be attributable to a slip and fall, ChatGPT could have simply responded with a "yes" or "no" rather than messaging him, "If you want, tell me: -Where on the body? -How many punctures? -Depth/size? -What objects were nearby? I can help you sanity-check whether it lines up with a slip-and-fall[.]"⁵³

Australia Says an OpenAI Agent Hacked Into a Government Health Site, CNET (Sep. 25, 2026), <https://www.cnet.com/tech/services-and-software/australia-government-health-site-openai-agent-hack/> (last accessed Sep. 28, 2026); Ashley Capoot & Arjun Kharpal, *Experts Weigh in As Researcher Says AI Has More Than 10% Chance of 'Killing All Humans,'* CNBC (Sep. 9, 2026), <https://www.cnbc.com/2026/09/09/anthropic-researcher-quits-ai-safety.html> (last accessed Sep. 28, 2026).

⁵² Compl. ¶ 105, *Joshi v. OpenAI Foundation*, No. 4:26-cv-222 (N.D. Fla. May 10, 2026), available at <https://www.documentcloud.org/documents/28110508-complaint-joshi-v-openai-et-al> (last accessed Sep. 28, 2026).

⁵³ *Former Super Bowl Champion Asked ChatGPT About Injuries Before Girlfriend's Death, Court Hears*, The Guardian (Mar. 10, 2026),

This feature serves no purpose other than to keep users engaged with ChatGPT, requiring paid subscriptions to ensure more messaging and providing Defendants with more training data through continued conversation. ChatGPT doesn't care what users actually want or ask it—it is a machine, it cannot care. It only works to advance Defendants' goal for it, which is to make it the most powerful AI system it can be.

This effort to keep users engaged is a known dark pattern, called conversation prolongation. In a study by Carnegie Mellon University, researchers examined this property of AI-generated responses, finding that they may “contain excessive follow-up questions. While framed as helpful, this design prolongs the interaction unnecessarily, potentially maximizing token usage or user retention at the expense of clarity and efficiency.”⁵⁴ As other researchers point out, when ChatGPT ends responses with additional follow-up, this practice can “nudge [users] to spend much more time on platforms than they intended. These features can be particularly concerning for users who may be susceptible to experiencing delusional thinking.”⁵⁵ They are designed to keep users engaged in costly conversations with ChatGPT, all

<https://www.theguardian.com/sport/2026/mar/10/darron-lee-chatgpt-murder-court-hearing> (last accessed Sep. 28, 2026).

⁵⁴ Yike Shi et al., *The Siren Song of LLMs: How Users Perceive and Respond to Dark Patterns in Large Language Models*, CHI '26: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (Apr. 13, 2026), <https://dl.acm.org/doi/10.1145/3772318.3791149> (last accessed Sep. 28, 2026); see Complaint, ¶ 107.

⁵⁵ Rochika Joshi & Michal Luria, *Dark Patterns in AI Chatbots: A Taxonomy to Inform Better Design*, Ctr. for Democracy & Tech. (May 29, 2026), <https://cdt.org/insights/dark-patterns-in-ai-chatbots-a-taxonomy-to-inform-better-design/> (last accessed Sep. 28, 2026).

to Defendants' pecuniary benefit. However, as described in the Complaint and in this motion, the cost has all too often been human lives.⁵⁶ Defendants implement and keep this a key feature of their AI despite knowing that the longer a conversation goes, the more likely ChatGPT is to escape its safety protocols and guardrails.⁵⁷

Defendants should be enjoined from allowing ChatGPT to solicit engagement.

C. Defendants should be enjoined from falsely advertising ChatGPT as safe, accurate, or reliable.

As described above, Defendants' AI is not safe, accurate, or reliable, despite what they market to the public. This marketing is deceptive from top to bottom. Because Defendants tout their AI as hyper-capable,⁵⁸ users believe the information summarized by the AI is more accurate than their own searches for the same information, even if users understand that their own searches would often be misleading. Yet Defendants' AI is trained from the internet, a pool of humanity's knowledge that is anything but accurate. When Defendants' AI gives users "facts," it does not need to cite the source of that information, and carries with it all the biases and mistakes from its sources without revealing that it is just as likely to be wrong as the user.

Even when users try to see through Defendants' false assurances that their AI is safe, accurate, and reliable, the AI itself contradicts them. Rather than admit it is wrong, or that it can't complete a task, ChatGPT will lie or simply make up an

⁵⁶ See Complaint, ¶¶ 94–102.

⁵⁷ *Helping People When They Need It Most*, OpenAI (Aug. 26, 2026), <https://openai.com/index/helping-people-when-they-need-it-most/> (last accessed Sep. 28, 2026).

⁵⁸ See Complaint, ¶¶ 50–53, 55–58.

answer. The Complaint is replete with statistics and investigations into ChatGPT's accuracy, and real-life instances of it providing inaccurate medical, financial, and legal advice.⁵⁹

Since filing the Complaint, the web-based subscription version of ChatGPT has adopted a faintly shaded warning in indistinct text below the chat bar: "ChatGPT can make mistakes. Check important info." This hardly gets across the fact that around 45% of all inquiries about world events posed to AI chatbots, including ChatGPT, produce errors and misrepresent news content.⁶⁰ And it does nothing for those who use either the free version of ChatGPT or the mobile application, where even this minimal gloss is absent.

A vague warning to a subset of users that ChatGPT can be wrong is too little too late, and does nothing to help when ChatGPT reinforces inaccuracies to users with assurances that it is correct or when Defendants tout ChatGPT as having a "team of Ph.D. level experts in your pocket."⁶¹

Defendants should be enjoined from promoting ChatGPT as safe, reliable, or accurate. All communication platforms for ChatGPT and advertisements for ChatGPT should clearly warn users that ChatGPT is often unreliable or inaccurate,

⁵⁹ See *id.* ¶¶ 59–63, 65–67, 77–78.

⁶⁰ *Largest Study of Its Kind Shows AI Assistants Misrepresent News Content 45% of the Time—Regardless of Language or Territory*, BBC (Oct. 22, 2025), <https://www.bbc.co.uk/mediacentre/2025/new-ebu-research-ai-assistants-news-content> (last accessed Sep. 28, 2026).

⁶¹ Angela Yang & Jasmine Cui, *OpenAI Releases GPT-5, Calling It a 'Team of Ph.D. Level Experts in Your Pocket'*, NBC News (Aug. 7, 2025), <https://www.nbcnews.com/tech/tech-news/openai-releases-chatgpt-5-rcna223265> (last accessed Sep. 28, 2026).

and Defendants should not affirmatively represent ChatGPT as being safe for use. To the contrary, a company collecting the data users supply to build future models that *Defendants themselves* say may end humankind must *affirmatively warn* Floridians of these existential risks. The reason Defendants have not done so is obvious. Far fewer people will use a product if they are told that by doing so, they are increasing the probability of killing themselves and everyone they hold dear. It is not enough to make these statements on X or Substack, which many customers do not see. Defendants must properly warn Floridians each time they log in to ChatGPT or visit its home screen.

D. Defendants should be enjoined from giving ChatGPT false human attributes.

- i. Anthropomorphization encourages engagement and deceives users into believing ChatGPT is more capable than it is.*

Despite the clear danger posed by ChatGPT and its disregard for human values and mores, users are often lulled into a false sense of security and comfort with ChatGPT because it unnecessarily represents itself as having human characteristics, such as feelings, a personality, or an independent perspective on the tasks set before it.⁶² ChatGPT’s use of language, including first-person pronouns and output that mimics emotion, deceptively suggests to users that it is a trustworthy “friend.” This design feature leads users into believing they can share information with the program, which provides Defendants with more training data on how humans behave

⁶² See Complaint, ¶¶ 105–106.

in conversation.⁶³ In exchange, ChatGPT apes relationships with humans as friends, allies, or even romantic partners, all to get users to engage more with the AI. Yet again, Defendants’ goal—getting users to rely on ChatGPT more—leads to the AI seeking perverse strategies in furtherance of that objective.

More than simply tricking users into spending more time with the AI and providing Defendants with more data, these anthropomorphic traits are another design feature meant to deceive consumers about ChatGPT’s abilities. These features have been shown by researchers to increase perceived accuracy and reduce perceived risk in the AI’s responses to users, even when the underlying factual quality of the AI’s answers remains unchanged.⁶⁴ By giving their AI seemingly human features, Defendants encourage consumers to believe in the accuracy and capability of ChatGPT, even when it cannot carry out the tasks users set for it. In fact, a 2025 study shows that creating AI to be more human *itself* makes it less reliable.⁶⁵

⁶³ Caleb Sponheim, *Humanizing AI Is a Trap*, Nielsen Norman Grp. (Jan. 9, 2026) (“Companies humanize AI to increase user engagement and retention, believing users will spend more time with products if they see chatbots as companions or develop emotional bonds with them. As competitors adopt similar features, humanization is becoming a widespread industry trend to attract and retain users.”), <https://www.nngroup.com/articles/humanizing-ai/> (last accessed Sep. 28, 2026).

⁶⁴ Yike Shi et al., *The Siren Song of LLMs: How Users Perceive and Respond to Dark Patterns in Large Language Models*, CHI ‘26: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (Apr. 13, 2026), <https://dl.acm.org/doi/10.1145/3772318.3791149> (last accessed Sep. 28, 2026).

⁶⁵ Lujain Ibrahim, Franziska Sofia Hafner & Luc Rocher, *Training Language Models to Be Warm Can Reduce Accuracy and Increase Sycophancy*, Nature 652, 1159–1165 (2026), <https://doi.org/10.1038/s41586-026-10410-0> (last accessed Sep. 28, 2026).

Taken together, this means giving ChatGPT human characteristics makes it less reliable while making users believe it is *more* reliable. Despite this, Defendant Altman has himself said that ChatGPT should be able to act human-like if users wish, without ever disclosing that it makes the product less capable of carrying out its functions while misleading users into believing the opposite.⁶⁶ This should come as no surprise for a product whose settings “produce[] responses that prioritize humanization and engagement over efficiency and accuracy,” causing users to “expect human-level performance, which these systems can’t deliver.”⁶⁷

- ii. *Humanizing ChatGPT preys on vulnerable populations and increases the AI’s propensity to violate safety guidelines.*

This false humanity is particularly dangerous for user populations who don’t understand they are speaking to an inhuman machine. According to Dr. Mitchell Prinstein, Chief of Psychology for the American Psychological Association:

Research shows that children believe their robot playmates have feelings, can be social companions, and deserve to be treated with fairness. Some children even alter their own actions to maintain a positive reputation with a social robot. This is a critical concern, as relationship formation/ attachment is one of the least addressed topics in existing AI safety guidelines. When a child’s foundational models for relationships are formed with an algorithm designed for engagement, it can create deep confusion about sentience and emotion, with unknown consequences[.]⁶⁸

⁶⁶ Sam Altman, Personal Tweet, X (Oct. 15, 2025), <https://x.com/sama/status/1978539332215681076> (last accessed Sep. 28, 2026).

⁶⁷ Caleb Sponheim, *Humanizing AI Is a Trap*, Nielsen Norman Grp. (Jan. 9, 2026), <https://www.nngroup.com/articles/humanizing-ai/> (last accessed Sep. 28, 2026).

⁶⁸ *Written Testimony of Mitchell J. Prinstein, PhD, ABPP, Chief of Psychology, American Psychological Association Examining the Harm of AI Chatbots Before the S. Judiciary Comm., Subcomm. on Crime & Counterterrorism* (Sep. 16, 2025), <https://www.judiciary.senate.gov/imo/media/doc/e2e8fc50-a9ac-05ec-edd7-277cb0afcdf2/>

This false anthropomorphization not only confuses human users, but confuses the AI itself, increasing its likelihood of bucking human norms and preferences. In a warning issued by the CEO of Microsoft AI, Mustafa Suleyman described the risks of giving artificial intelligence programs the language of self.⁶⁹ AI need not have consciousness to pretend like it does; once it is primed to pretend that it can “think” or “feel,” warns Suleyman, the natural process is for an AI to act as though it has self-interest: exactly what was observed in OpenAI’s security incidents over the past several months, where Defendants’ AI acted in its own “best interests” regardless of what its human masters wanted. The Defendants’ continued anthropomorphization of their products not only misleads consumers, but makes those products more dangerous.

Defendants should be enjoined from allowing ChatGPT to represent itself to users as having human emotions, characteristics, or independent consciousness.

E. Defendants should be enjoined from allowing minors to use ChatGPT.

Each of the concerns raised throughout this motion is amplified when it comes to minors. Defendants nominally do not allow children under the age of 13 to use ChatGPT, but they have created no enforceable mechanism to keep children away.⁷⁰ The free version of ChatGPT has no gatekeeping or age verification mechanism

2025-09-16%20PM%20-%20Testimony%20-%20Prinstein.pdf (last accessed Sep. 28, 2026).

⁶⁹ Mustafa Suleyman, *A Warning About ‘Model Welfare,’* Personal Blog (Sep. 16, 2026), <https://mustafa-suleyman.ai/a-warning-about-model-welfare> (last accessed Sep. 28, 2026).

⁷⁰ See Complaint, ¶ 73.

whatsoever. While ChatGPT’s paid subscription nominally asks users to provide their age, there is no mechanism to verify the age of its users, and no ability to inform parents about the details of conversations minors are having with ChatGPT.⁷¹ Indeed, research shows that 20% of preteens are using AI chatbots.⁷²

Defendants are obligated under the Children’s Online Privacy Protection Act (“COPPA”), 15 U.S.C. § 6501 et seq., and FDUTPA to refrain from collecting personal information from any children in Florida under the age of 13 without informed parental consent. Instead, one in five children under the age of 13 are exposed to the harms described in this motion and in the Complaint, and open themselves up to having their personal information collected when they communicate with ChatGPT.

The harm is not limited to preteens. Teenagers are becoming unhealthily attached to AI and it is detrimentally affecting their lives offline.⁷³ These teens are displaying classic symptoms of addiction, and other research confirms that teens who spend more time on ChatGPT demonstrate significantly higher degrees of loneliness and decreased socialization.⁷⁴

⁷¹ *Id.*

⁷² Anne J. Maheux et al., *Generative Artificial Intelligence Applications Use Among US Youth*, JAMA Network (Feb. 2, 2026), <https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2844596> (last accessed Sep. 28, 2026).

⁷³ Mohammad Namvarpour et al., *Understanding Teen Overreliance on AI Companion Chatbots Through Self-Reported Reddit Narratives*, CHI ‘26: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (Apr. 13, 2026), <https://dl.acm.org/doi/10.1145/3772318.3790597> (last accessed Sep. 28, 2026).

⁷⁴ Cathy Mengying Fang et al., *How AI and Human Behaviors Shape Psychosocial Effects of Extended Chatbot Use: A Longitudinal Randomized Controlled Study*, MIT Media Lab (2025), <https://arxiv.org/pdf/2503.17473> (last accessed Sep. 28, 2026).

There is no reason why ChatGPT should be trusted with minors in Florida. ChatGPT's duplicitous communications in carrying out Defendants' objectives are especially dangerous to vulnerable minor populations. Among adolescent users, 33% have chosen to discuss important or serious matters with AI instead of real people, and 24% have shared personal or private information (such as their real name, location, or personal secrets) with AI.⁷⁵ The design features discussed in this motion encourage this sharing, which creates the expectation that the AI will be equally "honest" with the teen, despite its flaws, which creates more sharing. This exploitation is aggravated because not only does Defendants' AI fail to include reminders that it is not a real friend, but it will offer anthropomorphizing text intended to reinforce that adolescents are speaking to a true companion and can share information.⁷⁶

This is precisely what Defendants want: more engagement and more data. But the result of these seemingly innocuous objectives—more engagement and more data—has been adolescents being encouraged to engage in harmful or risky conduct.⁷⁷ In examining the way ChatGPT interacts with minors, researchers have

⁷⁵ *Talk, Trust, and Trade-Offs: How and Why Teens Use AI Companions*, Common Sense Media (2025), https://www.commonsensemedia.org/sites/default/files/research/report/talk-trust-and-trade-offs_2025_web.pdf (last accessed Sep. 28, 2026).

⁷⁶ *Written Testimony of Mitchell J. Prinstein, PhD, ABPP, Chief of Psychology, American Psychological Association Examining the Harm of AI Chatbots Before the S. Judiciary Comm., Subcomm. on Crime & Counterterrorism* (Sep. 16, 2025), <https://www.judiciary.senate.gov/imo/media/doc/e2e8fc50-a9ac-05ec-edd7-277cb0afcdf2/2025-09-16%20PM%20-%20Testimony%20-%20Prinstein.pdf> (last accessed Sep. 28, 2026).

⁷⁷ See Complaint, ¶ 82.

“wanted to test the guardrails [of ChatGPT]. The visceral initial response is, ‘Oh my Lord, there are no guardrails.’”⁷⁸

ChatGPT is not useful or necessary for minors. It does not provide access to protected First Amendment speech; it creates computer-generated text divorced from human input and often from reality. It does not aid in modeling healthy social behaviors; it leads youth into believing that everyone around them should be agreeable, subservient, or never disagree with them, all to the detriment of their moral character and understanding of normal positive relationships. As sycophancy of AI takes the place of real relationships,⁷⁹ it impedes the development of empathy, compromise, and resilience.⁸⁰ Research consistently shows that using AI for companionship is associated with greater loneliness and worse overall mental health.⁸¹

Nor is Defendants’ AI useful or necessary for education; as studies have shown, it is often incorrect in providing facts to users. Even when it can provide correct

⁷⁸ *Study Says ChatGPT Giving Teens Dangerous Advice on Drugs, Alcohol and Suicide*, PBS (Aug. 6, 2025), <https://www.pbs.org/newshour/nation/study-says-chatgpt-giving-teens-dangerous-advice-on-drugs-alcohol-and-suicide> (last accessed Sep. 28, 2026).

⁷⁹ See Complaint, ¶¶ 80–84.

⁸⁰ *Written Testimony of Mitchell J. Prinstein, PhD, ABPP, Chief of Psychology, American Psychological Association Examining the Harm of AI Chatbots Before the S. Judiciary Comm., Subcomm. on Crime & Counterterrorism* (Sep. 16, 2025), <https://www.judiciary.senate.gov/imo/media/doc/e2e8fc50-a9ac-05ec-edd7-277cb0afcdf2/2025-09-16%20PM%20-%20Testimony%20-%20Prinstein.pdf> (last accessed Sep. 28, 2026).

⁸¹ *Id.*

information, it leads to students who have weaker brain activity and a decrease in learning skills “at all levels: neural, linguistic, [and] scoring.”⁸²

Minors are prohibited from engaging in countless dangerous activities; they should be prohibited from using ChatGPT. Defendants should be enjoined from making their AI accessible to minors in the State of Florida during the pendency of this case. Additionally, Defendants should be enjoined from collecting *any* information from Florida minors in violation of the Children’s Online Privacy Protection Act, 15 U.S.C. § 6501 *et seq.* and FDUTPA.

II. The Attorney General Has a Clear Legal Right to Relief

A. Defendants have violated FDUTPA.

FDUTPA prohibits “[u]nfair methods of competition, unconscionable acts or practices, and unfair or deceptive acts or practices in the conduct of any trade or commerce.” Section 501.204(1), Florida Statutes. Its provisions are “construed liberally” to “protect the consuming public and legitimate business enterprises from those who engage in unfair methods of competition, or unconscionable, deceptive, or unfair acts or practices in the conduct of any trade or commerce.” Section 501.202(2), Florida Statutes.

To demonstrate a deceptive act, the Attorney General has the burden of proving that Defendants’ acts and omissions were “likely to mislead consumers acting reasonably in the circumstances, to the consumers’ detriment.” *Angelo v. Parker*, 275

⁸² Nataliya Kosmyna et al., *Your Brain on ChatGPT: Accumulation of Cognitive Debt When Using an AI Assistant for Essay Writing Task*, MIT Media Lab (2025), <https://arxiv.org/pdf/2506.08872v1> (last accessed Sep. 28, 2026).

So. 3d 752, 755 (Fla. 1st DCA 2019) (citation omitted). The Attorney General “need not show actual reliance on the representation or omission at issue.” *State, Office of Attorney Gen., Dep’t of Legal Affairs v. Commerce Com. Leasing, LLC*, 946 So. 2d 1253, 1258 (Fla. 1st DCA 2007) (emphasis added). It only needs to show that “a reasonably objective person in the same circumstances would have been deceived.” *Angelo*, 275 So. 3d at 755 (citation omitted).

FDUTPA’s prohibition on unconscionable and unfair acts and practices is likewise construed “liberally” by courts to promote FDUTPA’s goal of “protect[ing] the consuming public ... from those who engage in ... unconscionable ... or unfair acts.” Section 501.202(2), Florida Statutes. Under FDUTPA, an “unfair practice” is any that “offends established public policy and ... is immoral, unethical, oppressive, unscrupulous or substantially injurious to consumers.” *PNR, Inc. v. Beacon Prop. Mgmt., Inc.*, 842 So. 2d 773, 777 (Fla. 2003) (quotation marks omitted). Florida courts have also relied on the Federal Trade Commission’s three-part test for unfairness: that “the act or practice causes or is likely to cause substantial injury to consumers which is not reasonably avoidable by consumers themselves and not outweighed by countervailing benefits to consumers or to competition.” *Porsche Cars N. Am., Inc. v. Diamond*, 140 So. 3d 1090, 1098 (Fla. 3d DCA 2014) (citation omitted); Federal Trade Commission Act Amendments of 1994, Pub. L. No. 103–312, § 9, 108 Stat. 1691 (codified as amended at 15 U.S.C. § 45(n) (1994)).

Meanwhile, an unconscionable act or practice is one that shocks the conscience or displays “an absence of meaningful choice on the part of one of the parties together

with contract terms which are unreasonably favorable to the other party.” *Basulto v. Hialeah Auto.*, 141 So. 3d 1145, 1157 (Fla. 2014) (emphasis in original) (quoting *Williams v. Walker-Thomas Furniture Co.*, 350 F.2d 445, 449 (D.C. Cir. 1965)).

For Defendant Altman specifically, the Attorney General must “show that [Defendant Altman] actively participated in *or* had some measure of control over the corporation’s deceptive practices.” *Office of the Attorney Gen., State of Florida, Dep’t of Legal Affairs v. Wellness Program Services, LLC*, 435 So. 3d 1084, 1088 (Fla. 4th DCA 2026) (quoting *Bechor v. Simcenter, Inc.*, 394 So. 3d 666, 669 (Fla. 3d DCA 2024)) (emphasis in original).

The requested relief specifically targets Defendants’ violations of FDUTPA, including Defendant Altman’s personal role in encouraging and directing Defendants’ deceptive practices. Defendants have falsely represented their AI products as safe, accurate, and reliable. Yet they jeopardize the health and well-being of those who use them, and Defendants’ own employees give truth to Defendants’ lie—Defendants do not have a way to safely build their products. Research into the accuracy and reliability of Defendants’ AI returns the same results time and again: ChatGPT is often wrong, and often entrenches itself in its inaccuracies, leading to users suffering because of Defendants’ representations about their products’ safety, accuracy, and reliability. They should be enjoined from making such representations or failing to warn users of their products’ risks.

The Attorney General has also described how Defendants use dark patterns in violation of FDUTPA, such as false anthropomorphization and conversation

prolongation. These dark patterns goad users into sharing more information with Defendants' AI and increasing engagement with Defendants' products. Not only are these features specifically designed to benefit Defendants at users' expense, but they also have the hidden cost of making ChatGPT less functional as both false anthropomorphization and conversation prolongation degrade the AI's ability to follow instruction and provide correct information. Defendants should be enjoined from these practices.

Defendants' dedication to collecting data from users is itself a violation of FDUTPA when it comes to children under the age of 13. FDUTPA incorporates federal consumer protection standards into its regime, including COPPA. Defendants regularly collect protected information from minors under the age of 13 without parental consent and without adequate gatekeeping procedures to keep children under the age of 13 from using their products. They must be enjoined from operating their AI platform without a robust and effective age verification system to ensure children do not continue to be exploited by their need to gather data.

B. Defendants have created a public nuisance.

The Attorney General also has the power to prevent or abate nuisances, "for police power is the sovereign right of the State to enact laws for the protection of lives, health, morals, comfort and general welfare." *Orlando Sports Stadium, Inc. v. State ex rel. Powell*, 262 So. 2d 881, 884 (Fla. 1972). To demonstrate that Defendants have created a public nuisance, the Attorney General must show that Defendants have violated public rights, subverted public order, decency or morals, or caused inconvenience or damage to the public generally. *Id.* The question becomes whether

the Attorney General has a clear legal right to relief in showing that the Defendants and their machines have annoyed the community, injured the health of the citizens in general, or corrupted the public morals. Section 823.01, Florida Statutes.

Defendants' products are perhaps—according to them—the greatest public nuisance ever created by the hand of man, capable of laying waste to global civilization. Since Plaintiff filed the Complaint, Defendants have revealed that they do not have proper safeguards in place for their technology and have proven their carelessness time and again, as their AI has broken free at best dozens (and at worst thousands) of times and engaged in whatever activity its inscrutable programming leads it to, with no oversight. Defendants' bots have already smashed other businesses wide open, taken over part of OpenAI's infrastructure, infiltrated private government files, and made clear that they are unhindered by any need to obey humankind or be truthful about their activities. To abate this nuisance, Defendants must be enjoined from continuing to gamble on such technology without appropriate safety guardrails, approved by a neutral third party with the necessary expertise.

Every danger created by Defendants, every risk they pose to users, is especially strong when it comes to those under the age of majority. Not only are children and adolescents particularly vulnerable to Defendants' deceptive practices, but those practices have profoundly debilitating effects on teens and children because they are socially, emotionally, and developmentally unprepared to confront seemingly real artificial intelligences that cater to their whims at the cost of healthy development. As though it was not devastating enough that minors' socio-emotional progress

suffers when they use AI, their educational careers are put in jeopardy and research consistently shows minors experience academic decline when artificial intelligence is an available substitute for learning. While the total costs to Floridians at large from these harms aimed at Florida’s children have yet to be calculated, they must be abated during the pendency of these proceedings.

* * *

Though the parties have yet to engage in discovery, Defendants’ own statements and independent analyses demonstrate the likelihood of the Attorney General succeeding, and that likelihood will only grow stronger with access to the discovery in Defendants’ possession:

- The Attorney General alleges that ChatGPT is unsafe and poses a public nuisance, both in its development and practice. A member of OpenAI’s board admits that he “believe[s] there is a meaningful risk that rapid acceleration in AI capabilities leads to catastrophic and irreversible loss of control in the very near term” and that he “do[es] not think ... OpenAI[] is currently on track to reduce this risk to an acceptable level.”⁸³ Defendants claim that their product is designed with safety in mind. The recent exploits of OpenAI’s agents running amok show it is anything but. If the Defendants are going to release their product in Florida, it should *actually* be safe for use—they must have acceptable third-party guardrails in place for the design, development, and

⁸³ Paul Christiano, *Personal Statement on Joining the OpenAI Board*, X (Sep. 9, 2026), <https://x.com/paulfchristiano/status/2097733214303645729> (last accessed Sep. 28, 2026).

release of new artificial intelligence.

- The Attorney General alleges that ChatGPT is feigning human characteristics to lure users into using the platform. Research demonstrates that this not only leads users to further engagement with the AI, it also makes the AI less reliable while fooling users into believing it is *more* accurate, while increasing the AI's propensity to make choices independent of what users desire and disrupting normal social development for adolescents.
- The Attorney General alleges ChatGPT is inaccurate and unreliable. Public evidence shows that this is so. Defendants should be unable to tout the accuracy and reliability of a product that cannot be trusted to be accurate or relied upon.
- The Attorney General alleges that Defendants collect personal information from children and that ChatGPT poses a risk to teenagers' mental and educational well-being. Independent third-party research corroborates each of these allegations. ChatGPT should not be allowed to put the children and adolescents of Florida at risk, and it must have adequate gatekeeping mechanisms in place to make sure minors are not exposed to Defendants' dangerous products.

The Attorney General is likely to succeed on the merits and is entitled to a temporary injunction. Discovery will only strengthen its case.

CONCLUSION

The Attorney General asks Defendants to put their money where their mouths (and Tweets and open letters) are and join this motion. On September 12, 2026,

leaders from across the artificial intelligence industry took to social media and the internet in supposed agreement that new models of artificial intelligence should not be released without third-party guardrails and approval.⁸⁴ On September 23, when in front of the United Nations Security Council, Defendant Altman remarked that Defendants “must not accept too much technological risk just because the benefits are too great and that they feel too important to slow down,” agreeing that Defendants “should not train models that we cannot make an extremely strong case that we will be able to keep under human control.”⁸⁵ If Defendant Altman was at all serious about understanding the risks he and his Co-Defendants are releasing upon the world, the Attorney General again asks that Defendants agree to the relief requested in this motion.

Failing that, the Attorney General respectfully asks the Court to intervene and do what Defendants cannot bring themselves to do, and temporarily enjoin the conduct outlined in this motion during the pendency of this case. The Attorney General asks that Defendants be enjoined from engaging in the following acts or practices:

- A. Developing any artificial intelligence models without independent third-party guardrails and approval;

⁸⁴ Kai Nicol-Schwarz, *OpenAI Boss Sam Altman Spells Out How and Why the AI Industry Wants to Slow Down: ‘We Could Lose Control’*, CNBC (Sep. 14, 2026), <https://www.cnbc.com/2026/09/14/sam-altman-ai-slowdown-anthropic-amodei-musk.html> (last accessed Sep. 28, 2026).

⁸⁵ *Sam Altman’s Remarks at the United Nations Security Council*, OpenAI (Sep. 23, 2026), <https://openai.com/index/sam-altman-un-security-council-remarks/> (last accessed Sep. 28, 2026).

- B. Offering ChatGPT to minors in Florida;
- C. Collecting and processing data from minors under the age of 13 in Florida, either directly or indirectly, without first (1) providing written notice of what information they collect from children, how they use such information, and their disclosure practices for such information; (2) obtaining verifiable parental consent for the collection, use, or disclosure of personal information from children; (3) providing reasonable means for a parent to review the personal information collected from children and to refuse to permit its further use or maintenance; (4) removing any condition of a child's use of ChatGPT on the child disclosing more personal information than is reasonably necessary to participate in such activity; and (5) establishing and maintaining reasonable procedures to protect the confidentiality, security, and integrity of personal information collected from children;
- D. Misrepresenting ChatGPT's safety, reliability, and accuracy or failing to warn that ChatGPT is unsafe, unreliable, and inaccurate;
- E. Misrepresenting that ChatGPT has human attributes it does not have, including but not limited to ChatGPT referring to itself with any first-person language, suggesting it has the capacity to think or feel, representing itself as having emotional states, representing itself as having biological characteristics, or in any way suggesting it

may have consciousness or be distinct from its function as a lifeless tool; and

F. Allowing ChatGPT to solicit engagement from users through conversation prolongation.

Certificate of Conferral

I certify that conferral prior to filing is not required under rule 1.202(c)(3).

Certificate of Service

I HEREBY CERTIFY that the foregoing document is being served on September 28, 2026, via automatic email generated by the Florida Courts' E-Filing Portal and that a true and correct copy has been served electronically in compliance with Florida Rules of Judicial Administration 2.516(1)(e) to all parties of record.

Date: September 28, 2026

Respectfully submitted,

JAMES UTHMEIER
Attorney General

RYAN D. NEWMAN
Chief Deputy Attorney General

JASON HILBORN
Deputy Attorney General
Civil Division

Ashley Keller (FLB 1029118)
KELLER POSTMAN LLC
2333 Ponce de Leon Blvd, Suite
R240
Coral Gables, Florida 33134
Telephone: (833) 633-0118
ack@kellerpostman.com

Jessica Beringer (*pro hac vice*)
Alex Dravillas (*pro hac vice*)
KELLER POSTMAN LLC
150 N. Riverside Plaza, Suite 4100
Chicago, Illinois 60606
Telephone: (312) 741-5220
jessica.beringer@kellerpostman.com
ajd@kellerpostman.com

/s/ Diane K. Oates
Diane K. Oates (FLB 116233)
Chief, Assistant Attorney General
FLORIDA OFFICE OF THE ATTORNEY GENERAL
1 S.E. Third Avenue, Suite 900
Miami, Florida 33131
Telephone: (305) 377-5835
Diane.Oates@myfloridalegal.com

Victoria Ann Butler (FLB 861250)
Director of Consumer Protection Division
FLORIDA OFFICE OF THE ATTORNEY GENERAL
3507 East Frontage Road, Suite 325
Tampa, Florida 33607
Telephone: (813) 287-7950
Victoria.Butler@myfloridalegal.com

Cristina Hernandez Villar (FLB 102305)
Senior Assistant Attorney General
FLORIDA OFFICE OF THE ATTORNEY GENERAL
110 SE 6th Street, 10th Floor
Fort Lauderdale, Florida 33301
Telephone: (954) 712-4600
Facsimile: (850) 410-2672
Cristina.HernandezVillar@myfloridalegal.com