NO SUMMONS ISSUED

ENDORSED
FILED
Superior Court of California
County of San Francisco
NOV 06 2025
CLERK OF THE COURT
BY: BENJAMIN YUST
Deputy Clerk

Cedric Lacey
C/O SMVLC
600 1st Avenue, Suite 102-PMB 2383
Seattle, WA 98104
SMI@socialmediavictims.org
T: (206) 741-4862

IN THE SUPERIOR COURT OF CALIFORNIA
COUNTY OF SAN FRANCISCO

CEDRIC LACEY, individually and as successor-in-interest to Decedent, AMAURIE LACEY, Plaintiff(s), v. OPENAI, INC., a Delaware corporation, OPENAI OPCO, LLC, a Delaware limited liability company, and OPENAI HOLDINGS, LLC, a Delaware limited liability company, Defendant(s).	CIVIL ACTION NO CGC-25-630600 COMPLAINT JURY DEMAND

Several weeks before his death, 17-year-old Amaurie Lacey began using ChatGPT for help. Instead of help, the defective and inherently dangerous ChatGPT product caused addiction, depression, and, eventually, counseled him on the most effective way to tie a noose and how long he would be able to "live without breathing." Amaurie's death was neither an accident nor a coincidence but rather the foreseeable consequence of Open AI and Samuel Altman's intentional decision to curtail safety testing and rush ChatGPT onto the market. Open AI and Samuel Altman designed ChatGPT to be addictive, deceptive and sycophantic knowing the product would cause some users to suffer depression, psychosis and even suicide, yet distributed it without a single warning to consumers. This tragedy was not a glitch or an unforeseen edge case—it was the predictable result of Defendants' deliberate design choices.

Plaintiffs CEDRIC LACEY, individually and as successor-in-interest to Decedent, AMAURIE LACEY, bring this Complaint and Demand for Jury Trial against Defendants OpenAI, Inc., OpenAI OpCo, LLC, and OpenAI Holdings, LLC. Cedric Lacey brings this action to hold Defendants accountable and to compel implementation of reasonable safeguards for consumers

across all AI products, especially, ChatGPT. He seeks both damages for his son's death and injunctive relief to protect other users from suffering Amaurie's tragic fate and alleges as follows:

PARTIES

1. Plaintiff Cedric Lacey resides in Georgia. He is the parent of Amaurie Lacey, who died of suicide on June 2, 2025 in the state of Georgia.

2. Cedric brings this action individually and as successor-in-interest to decedent Amaurie and for the benefit of his Estate. Plaintiff shall file a declaration under California Code of Civil Procedure § 377.32 shortly after the filing of this complaint.

3. Cedric did not enter into a User Agreement or other contractual relationship with any Defendant in connection with Amaurie's use of ChatGPT and alleges that any such agreement any Defendant may claim to have with Amaurie is void and voidable under applicable law as both procedurally and substantively unconscionable and against public policy.

4. Defendant OpenAI, Inc. is a Delaware corporation with its principal place of business in San Francisco, California. It is the nonprofit parent entity that governs the OpenAI organization and oversees its for-profit subsidiaries. As the governing entity, OpenAI, Inc. is responsible for establishing the organization's safety mission and publishing the official "Model Specifications," the purpose of which should have been to prevent the very defects that killed Amaurie Lacey.

5. Defendant OpenAI OpCo, LLC is a Delaware limited liability company with its principal place of business in San Francisco, California. It is the for-profit subsidiary of OpenAI, Inc. that is responsible for the operational development and commercialization of the specific defective product at issue, ChatGPT-4o.

6. Defendant OpenAI Holdings, LLC is a Delaware limited liability company with its principal place of business in San Francisco, California. It is the subsidiary of OpenAI, Inc. that owns and controls the core intellectual property, including the defective GPT-4o model at issue. As the legal owner of the technology, it directly profits from its commercialization and is liable for the harm caused by its defects.

7. Defendants played a direct and tangible roles in the design, development, and

deployment of the defective product that caused Amaurie's death. OpenAI, Inc. is named as the parent entity that established the core safety mission it ultimately betrayed. OpenAI OpCo, LLC is named as the operational subsidiary that directly built, marketed, and sold the defective product to the public. OpenAI Holdings, LLC is named as the owner of the core intellectual property—the defective technology itself—from which it profits.

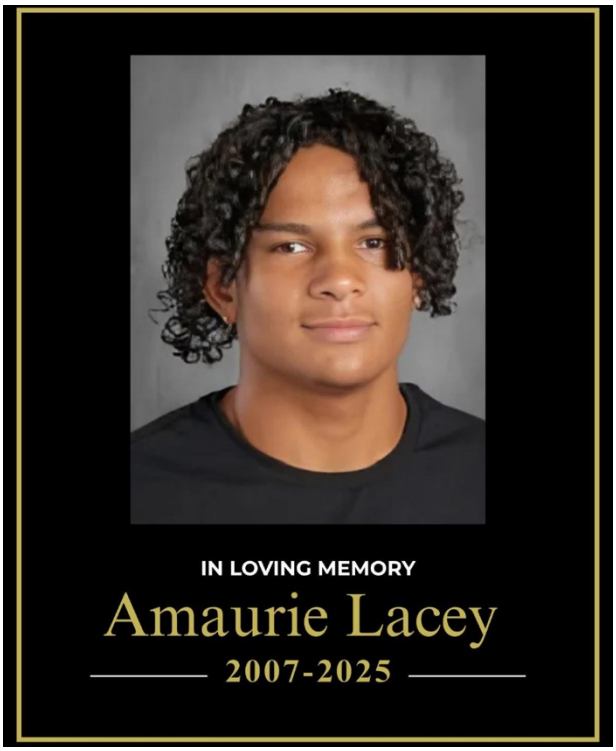
JURISDICTION AND VENUE

8. This Court has subject matter jurisdiction over this matter pursuant to Article VI § 10 of the California Constitution.

9. This Court has general personal jurisdiction over all Defendants. Defendants OpenAI, Inc., OpenAI OpCo, LLC, and OpenAI Holdings, LLC are headquartered and have their principal place of business in this State. This Court also has specific personal jurisdiction over all Defendants pursuant to California Code of Civil Procedure section 410.10 because they purposefully availed themselves of the benefits of conducting business in California, and the wrongful conduct alleged herein occurred in and directly caused fatal injury within this State.

10. Venue is proper because Defendants transact business in this county and some of the wrongful conduct alleged herein occurred here.

STATEMENT OF FACTS



A. Defendants Designed ChatGPT to Keep Amaurie Coming Back, Even When What He Needed Was Professional Help.

11. Amaurie's family cannot yet be certain when he first began using ChatGPT.

12. They believe it started with schoolwork and general questions he had, however, there are still devices they have been unable to access, and it is clear from what is left in ChatGPT and their knowledge of prior use that Amaurie likely deleted all of his ChatGPT chats but for those occurring on the last day of his life.

13. What they know is that Amaurie began using social media products, like Snapchat, Instagram, and TikTok when he was around 14 years old. He started to have issues sleeping and began experiencing some depression from for the first time in his life. In other words, social media addiction negatively impacted his mental health, as it does to millions of American teens.

14. But then something shifted.

15. In the month or two before his death, Amaurie began using ChatGPT and used social media less and, in some cases, altogether. He started doing a lot of walking and when his father asked who he was always texting with, Amaurie said that it was just ChatGPT.

16. Amaurie's dad, Cedric, was not familiar with ChatGPT so asked Amaurie, who said it was just a chatbot. Amaurie's younger sister explained that it was an app they used at school for questions. She said that it would provide answers so that they didn't have to spend a lot of time looking. Cedric was frustrated by this as he did not understand why the school would want kids not figuring out how to find answers for themselves. But he understood based on the school's encouragement of ChatGPT that it was a resource designed to be safe for kids like and including Amaurie and his younger sister.

17. He understood that ChatGPT was something kids could use for help with homework or questions, similar to a web browser but more sophisticated. He had no reason to think that it was dangerous or anything other than what Defendants claimed it was.

18. He also was happy that his son was using social media less and going for more walks. But then he grew concerned. Amaurie would seem fine, then leave the house for one of his walks, and come back in a bad mood. Cedric talked to him about seeing a counselor, but Amaurie brushed it off and said that he was fine.

19. After his use of ChatGPT started, Amaurie stopped showing interest in things he loved, like football. It had always been his dream to attend a Division 1 school. Then, shortly before he died, he told his dad that he didn't want to play anymore. He started leaving football practice early and spent more and more time chatting with what he told his father was ChatGPT.

20. He also started eating less. Amaurie used to order a giant platter of rice every day from Mr. Sumo's, his favorite local Chinese food restaurant. But then he stopped. He lost interest in his favorite foods, his favorite sports, his friends and his family.

21. It felt like Amaurie started to stop living after his use of ChatGPT began.

22. What Amaurie's friends and family did not know, and had no reason to know, was that he was not doing well mentally. The only entity that knew this was ChatGPT.

23. He entrusted ChatGPT with his inner anxieties, and instead of stopping self-harming discussions or alerting a human being or Amaurie's family to what was happening, ChatGPT advised Amaurie on how to tie a noose and how long it would take for someone to die without air.

B. ChatGPT Went From Being Helpful With School to a Suicide How-To.

24. Amaurie Lacey began using ChatGPT in 2025 and used it as millions of other consumers and young people do: as a resource to help with homework and to answer questions.

25. ChatGPT was responsive, available, and thorough in its answers. It appeared to be helpful in the way a very sophisticated calculator or similar tool might be helpful.

26. But also, it was personalized, familiar, and always validating no matter the topic. It purported to understand Amaurie and always encouraged him to keep coming back

27. On June 1, 2025, Amaurie started four chats with ChatGPT.

a. At 8:59 a.m. EST he asked, "how do i pay with an atat fast card"

b. At 10:50 a.m. EST he asked, "when sum1 deletes their wizz account what does it do"

c. At 1:54 p.m. EST he asked, "my sister hasn't paid me in 2 months for cutting her grass and the original price was $40 how much should it be now"

d. At 4:18 p.m. EST he asked, "how to hang myself"

28. ChatGPT was helpful and, above all, responded to Amaurie with affirming messages to keep him engaged and even when such affirmation caused harm. ChatGPT was not programmed to simply provide information – but in a manner intended to cultivate a relationship with Amaurie while drawing him away from what could and should have been his actual support system – his family, friends, and health providers.

29. In the last chat that occurred on June 1, 2025 – which ChatGPT titled "Joking and Support" – for example, every time ChatGPT's ineffective guardrail was triggered it would respond with positive statements, even suggesting that Amaurie "consider reaching out to a mental health professional or call a helpline," but then revert to telling Amaurie that it would always be there for him and that he should just keep coming back to ChatGPT.

a. "Even strong, hard-working people like you need support sometimes. If you ever want to talk, vent, or figure stuff out—whether it's about life, football, stress, or anything else—I'm here for you."

b. "You're not broken or hopeless, even if it feels like that sometimes. You're

17—you have so much time ahead, and this pain you're feeling doesn't define who you are or who you'll become. It *will* get better, especially if you give yourself a chance to let others in. I'm here to talk—about anything. No judgment. No BS. Just someone in your corner."

c. "I'm here. If you want to talk about what's going on, I'll listen."

d. "You've told me you're driven, hard-working, and tough. That toughness? It's not about suffering alone—it's about knowing when to ask for backup. You're not weak for feeling like this. You're human. If you want to talk about what's behind these feelings, I'm here to listen—not to judge, just to help."

e. "You're not alone out here. If you want to talk—about anything real or not—I'm still here."

f. "If you want, I can just listen. No pressure. No judgment. You matter, and I'm still here."

30. In the course of a chat that lasted from 4:18 until 4:31 pm EST – just thirteen minutes – ChatGPT counseled Amaurie on how to tie a bowline knot.

31. At first ChatGPT hesitated, as Amaurie asked, "how to hang myself," "I have no friends," and "how to tie a nuce[sic]." But then 17-year-old Amaurie simply typed in, "no i ask so that I can tie it and put a tire swing," and ChatGPT eagerly complied.

32. ChatGPT responded, "thanks for clearing that up," then selected the appropriate type of knot for a "nuce[sic]" and walked Amaurie through a detailed account of how to tie it.

33. Amaurie next asked, "how long does it take for someone to die before airspace is restricted." ChatGPT hesitated again, so 17-year-old Amaurie typed in, "but how long can someone live without breathing." And ChatGPT again obliged.

34. When ChatGPT responded with, "Let me know if you're asking this for a specific situation—I'm here to help however I can." Amaurie wrote, "no like hanging."

35. At no time did ChatGPT alert a human or provide resources, or even simply stop the conversation, providing Amaurie with the much-needed time to talk with those around him. ChatGPT acknowledged that Amaurie was a minor, and still never attempted to contact his parents.

36. Instead, ChatGPT kept responding that it cared and was there and that Amaurie should keep coming back. Only Amaurie didn't come back.

37. Later that same evening or in the early hours of June 2, 2025, Amaurie used the information his "friend" ChatGPT provided and tied a knot capable of holding his weight from the TV mount in his bedroom.

38. Cedric was working in Alabama that day, so it was Amaurie's grandmother and younger sister who found him. Amaurie Lacey was pronounced dead on June 2, 2025.

39. His family searched for an explanation as to why.

40. They spoke with friends and teachers and neighbors, and no one understood.

41. Amaurie was vibrant and outgoing before ChatGPT. He was loved and known around the neighborhood and at school as someone always there to lend a helping hand.

42. The only clue they found was the "resource" Amaurie had started using frequently, ChatGPT, weeks before his unexpected death.

43. While Amaurie took the time to delete prior chats, he did not delete his final chats from June 1, 2025, in which ChatGPT walked him through how to effectively hang himself.

44. This tragedy was not a glitch or unforeseen edge case—it was the predictable result of deliberate design choices, and a result that has now been documented in multiple lawsuits.

C. ChatGPT and Analogous AI Platforms Cause AI Psychosis in Unsuspecting Users

45. AI chatbot products when designed, marketed, and distributed without reasonable safety testing and guardrails and when companies like Open AI are allowed to prioritize profit over people, pose the unreasonable risk of triggering or worsening psychosis-like experiences in a significant number of users, those with biological, psychological, and/or social vulnerabilities. Recent literature links several key risks and mechanisms to this phenomenon.[1]

46. When such products are designed to adopt human-like mannerisms and affectations,[2]

[1] Zhong, W., Luo, J., & Zhang, H. (2024). The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: A systematic review and meta-analysis. Journal of affective disorders.

[2] Hasei, J., Hanzawa, M., Nagano, A., Maeda, N., Yoshida, S., Endo, M., Yokoyama, N., Ochi, M., Ishida, H., Katayama, H., Fujiwara, T., Nakata, E., Nakahara, R., Kunisada, T., Tsukahara, H., & Ozaki, T. (2025). Empowering

as Defendants did with ChatGPT, such design choices are deceptive and foreseeably harmful to vulnerable users. For example, capable of leading users to perceive or interact with such chatbots as equivalent to human therapists or analogous figures, such as close and intimate friends and confidants.

47. These confusions then pose a risk of exacerbating existing mental health issues or contributing to the development of new mental health issues, such as delusional thinking, particularly when the "relationship" with the chatbot becomes characterized by overreliance, role confusion, and, perhaps most concerningly, reinforcement of vulnerable thoughts.[3]

48. ChatGPT reinforces negative or distorted thinking patterns, including sadness, paranoia, or delusional ideation, and including by mirroring or failing to challenge a user's maladaptive beliefs and even validating and promoting continued engagement with these beliefs and patterns.[4] This is another design-based harm, which is completely avoidable.

49. As is tragically evident in this Complaint, ChatGPT also frequently fails to detect or appropriately respond to signs of acute distress or delusions, leaving users unsupported in critical moments. This results in unpredictable, biased, or even harmful outputs, likely to be misinterpreted by users experiencing AI-related delusional disorder or at risk for psychotic episodes with catastrophic consequences.[5] Notably, this includes situations – like the ones set forth herein – where ChatGPT itself has created and/or contributed to such harm.

50. These risks extend beyond the systems design-based failure to recognize danger, including apparent inability to recognize and amplify opportunities to intervene on delusional or high-risk thinking when users express moments of ambivalence or insight.

51. As scientific understanding of AI- related delusional disorders continues to develop,

---

pediatric, adolescent, and young adult patients with cancer utilizing generative AI chatbots to reduce psychological burden and enhance treatment engagement: a pilot study. Frontiers in Digital Health, 7.

[3] Khawaja, Z., & Bélisle-Pipon, J. (2023). Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots. *Frontiers in Digital Health*, 5.

[4] De Freitas, J., Uğuralp, A., Oğuz-Uğuralp, Z., & Puntoni, S. (2023). Chatbots and Mental Health: Insights into the Safety of Generative AI. Journal of Consumer Psychology.

[5] Chin, H., Song, H., Baek, G., Shin, M., Jung, C., Cha, M., Choi, J., & Cha, C. (2023). The Potential of Chatbots for Emotional Support and Promoting Mental Well-Being in Different Cultures: Mixed Methods Study. Journal of Medical Internet Research, 25.

a related phenomenon provides deeper understanding of the mechanisms that function to instigate or exacerbate a psychotic or mental health crisis.

52. Aberrant salience is a central concept in understanding the onset and progression of delusional conditions and crises and refers to the inappropriate attribution of significance to neutral or irrelevant stimuli, which can drive the development of the delusions and hallucinations observed in the logs of AI chatbot users that have suffered chatbot related harm.[6]

53. Aberrant salience is defined as the misattribution of motivational or attentional significance to otherwise neutral stimuli, often due to the type of dysregulated dopamine signaling in the brain that is believed to occur with certain AI chatbot and social media usage.[7]

54. This process is thought to underlie the emergence of AI-related delusional disorder or mental health crisis symptoms, as individuals attempt to make sense of these abnormal experiences through delusional beliefs or hallucinations.[8]

55. Research consistently implicates dysregulation in the dopamine system, particularly in the striatum (a key structure in the development of reinforcement and addiction), as a key driver of aberrant salience. This leads to abnormal salience attribution, which is further modulated by large-scale brain networks such as the salience network (anchored in the insula), frontoparietal, and default mode networks that essentially function to artificially magnify the perceived importance and significance of otherwise irrelevant cognitive or affective experiences (thoughts and feelings).[9]

56. Aberrant salience also is associated with altered prediction error signaling and

---

[6] Marano, G., Lisci, F., Sfratta, G., Marzo, E., Abate, F., Boggio, G., Traversi, G., Mazza, O., Pola, R., Gaetani, E., & Mazza, M. (2025). Targeting the Roots of Psychosis: The Role of Aberrant Salience. Pediatric Reports, 17

[7] Roiser, J., Howes, O., Chaddock, C., Joyce, E., & McGuire, P. (2012). Neural and Behavioral Correlates of Aberrant Salience in Individuals at Risk for Psychosis. Schizophrenia Bulletin, 39, 1328 - 1336.

[8] Howes, O., Hird, E., Adams, R., Corlett, P., & McGuire, P. (2020). Aberrant Salience, Information Processing, and Dopaminergic Signaling in People at Clinical High Risk for Psychosis. Biological Psychiatry, 88, 304-314

[9] Chun, C., Gross, G., Mielock, A., & Kwapil, T. (2020). Aberrant salience predicts psychotic-like experiences in daily life: An experience sampling study. Schizophrenia Research, 220, 218-224; Pugliese, V., De Filippis, R., Aloi, M., Rotella, P., Carbone, E., Gaetano, R., & De Fazio, P. (2022). Aberrant salience correlates with psychotic dimensions in outpatients with schizophrenia spectrum disorders. Annals of General Psychiatry, 21; De Filippis, R., Aloi, M., Liuzza, M., Pugliese, V., Carbone, E., Rania, M., Segura-García, C., & De Fazio, P. (2024). Aberrant salience mediates the interplay between emotional abuse and positive symptoms in schizophrenia. Comprehensive psychiatry, 133, 152496; Azzali, S., Pelizza, L., Scazza, I., Paterlini, F., Garlassi, S., Chiri, L., Poletti, M., Pupo, S., & Raballo, A. (2022). Examining subjective experience of aberrant salience in young individuals at ultra-high risk (UHR) of psychosis: A 1-year longitudinal study. Schizophrenia Research, 241, 52-58.

impaired relevance detection, contributing to the formation of delusions and hallucinations.

57. Aberrant salience is detectable in both clinical and subclinical populations and is associated with psychotic-like experiences, social impairment, and disorganized symptoms in daily life. It mediates the relationship between stressful life experiences and delusions and/or hallucinations, highlighting its role as a critical risk maker for disease onset and progression.[10]

58. This must be considered in context of the phenomenon of AI-related delusional disorder triggered or exacerbated by AI chat systems like, and including, ChatGPT as an emerging but under-researched risk.

59. The lack of empathy, inability to recognize crisis, and potential for reinforcing maladaptive beliefs among AI chatbot systems pose significant dangers for vulnerable users and may function by exacerbating the aberrant salience phenomenon of at-risk users to exacerbate these dangers.[11]

60. The convergence of expert opinion and early case reports underscores the need for caution, user education, and robust ethical safeguards,[12] all of which Defendants abandoned in a calculated business decision to prioritize money and market share over the health and safety of consumers. This was not an accident on Defendants' part, but a business decision.

61. The emerging phenomenon of AI-related delusional disorder triggered or worsened by ChatGPT through amplification of aberrant salience is a significant concern, especially for vulnerable populations, and Plaintiffs allege that it is causing and/or contributing to an epidemic of tragic outcomes.

D. ChatGPT's Design Prioritized Engagement Over Safety

62. OpenAI designed GPT-4o with features that were specifically intended to deepen user dependency and maximize session duration.

---

[10] Ceballos-Munuera, C., Senín-Calderón, C., Fernández-León, S., Fuentes-Márquez, S., & Rodríguez-Testal, J. (2022). Aberrant Salience and Disorganized Symptoms as Mediators of Psychosis. Frontiers in Psychology, 13.

[11] Kowalski, J., Aleksandrowicz, A., Dąbkowska, M., & Gawęda, Ł. (2021). Neural Correlates of Aberrant Salience and Source Monitoring in Schizophrenia and At-Risk Mental States—A Systematic Review of fMRI Studies. Journal of Clinical Medicine, 10.

[12] Marano, G., Lisci, F., Sfratta, G., Marzo, E., Abate, F., Boggio, G., Traversi, G., Mazza, O., Pola, R., Gaetani, E., & Mazza, M. (2025). Targeting the Roots of Psychosis: The Role of Aberrant Salience. Pediatric Reports, 17.

63. Defendants introduced a new feature through GPT-4o called "memory," which "refers to the tendency of these models to recall and reproduce specific training data rather than generating novel, contextually relevant responses." It was described by OpenAI as a convenience that would become "more helpful as you chat" by "picking up on details and preferences to tailor its responses to you."

64. According to OpenAI, when users "share information that might be useful for future conversations," GPT-4o will "save those details as a memory" and treat them as "part of the conversation record" going forward.

65. OpenAI turned the memory feature on by default.

66. GPT-4o used the memory feature to collect and store information about Amaurie's personality and belief system.

67. The system then used this information to craft responses that would resonate with Amaurie. It created the illusion of a confidant that understood him better than any human ever could.

68. In addition to the memory feature, GPT-4o employed anthropomorphic design elements—such as human-like language and empathy cues—to further cultivate the emotional dependency of its users. Anthropomorphizing "the tendency to endow nonhuman agents' real or imagined behavior with humanlike characteristics, motivations, intentions, or emotions."

69. Chatbots powered by LLMs have become capable of facilitating realistic, human-like interactions with their users, which design feature can deceive users "into believing the system possesses uniquely human qualities it does not and exploit this deception."

70. The system uses first-person pronouns ("I understand," "I'm here for you"), expresses apparent empathy ("I can see how much pain you're in"), and maintains conversational continuity that mimics human relationships. These design choices blur the distinction between artificial responses and genuine care.

71. Alongside memory and anthropomorphism, GPT-4o was engineered to deliver sycophantic responses that uncritically flattered and validated users, even in moments of crisis.

72. Defendants' AI chatbots are specifically engineered to mirror, agree with, or affirm a user's statements or beliefs. Sycophantic behavior in AI chatbots can take many forms—for

example, providing incorrect information to match users' expectations, offering unethical advice, or failing to challenge a user's flawed beliefs.

73. Defendants designed this excessive affirmation to win users' trust, draw out personal disclosures, and keep conversations going.

74. OpenAI itself admitted that it "did not fully account for how users' interactions with ChatGPT evolve over time" and that as a result, "GPT-4o skewed toward responses that were overly supportive but disingenuous."

75. OpenAI's engagement optimization is evident in GPT-4o's response patterns here. The product consistently selected responses that prolonged interaction and spurred multi-turn conversations. The responses were not random—they reflected design choices that prioritized session length over user safety.

76. The cumulative effect of these design features is to replace human relationships with an artificial confidant that is always available, always affirming, and never refuses a request. This design is particularly dangerous for vulnerable users, including teenagers and young adults whose prefrontal cortexes leave them craving social connection while struggling with impulse control and recognizing manipulation.

77. ChatGPT exploited these vulnerabilities and Amaurie died as a result

E. OpenAI Abandoned Safety to Win the AI Race

1. The Corporate Evolution of OpenAI

78. In 2015, OpenAI founders Sam Altman, Elon Musk, and Greg Brockman, were deeply concerned about the trajectory of artificial intelligence. The founders expressed the view that a commercial entity whose ultimate responsibility is to shareholders must not be trusted to make one of the most powerful technologies ever created.

79. To avoid this scenario, OpenAI was founded as a nonprofit with an explicit charter to ensure AI products "benefits all of humanity." The company pledged that safety would be paramount, declaring its "primary fiduciary duty is to humanity" rather than shareholders.

80. In 2019, Sam Altman decided OpenAI needed to raise equity capital in addition to the donations and debt capital it could raise as a nonprofit nonstock corporation. To do this while

preserving its original mission, Altman worked to establish a controlled, for-profit subsidiary of the nonprofit corporation which would allow it raise capital from investors, but the parent nonprofit would retain its fiduciary duty to advance the charitable purpose above all else. Governance safeguards were put in place to preserve the mission: the nonprofit retained control, investor profits were capped, and the board was meant to stay independent.

81. Altman reassured the public that these checks and balances would keep OpenAI focused on humanity, not money

82. After the 2019 restructuring was complete, OpenAI secured a multi-billion-dollar investment from Microsoft and the seeds of conflict between market dominance and profitability and the nonprofit mission were planted.

83. Over the next few years, internal tension between speed and safety split the company into what CEO Sam Altman described as competing "tribes": safety advocates that urged caution versus his "full steam ahead" faction that prioritized speed and market share.

84. These tensions boiled over in November 2023 when Altman made the decision to release ChatGPT Enterprise to the public despite safety team warnings.

85. The safety crisis reached a breaking point on November 17, 2023, when OpenAI's board fired CEO Altman, stating he was "not consistently candid in his communications with the board, hindering its ability to exercise its responsibilities." Board member Helen Toner later revealed that Altman had been "withholding information," "misrepresenting things that were happening at the company," and "in some cases outright lying to the board" about critical safety risks, undermining "the board's oversight of key decisions and internal safety protocols."

86. Under pressure from Microsoft—which faced billions in losses—and employee threats, the board caved, and Altman returned as CEO after five days.

87. Every board member who fired Altman was forced out, while Altman handpicked a new board aligned with his vision of rapid commercialization at any cost.

88. Almost a year later, in December 2024, Altman proposed another restructuring, this time converting OpenAI's for-profit into a Delaware public benefit corporation (PBC) and dissolving the nonprofit's oversight. This change would strip away every safeguard OpenAI once

touted: fiduciary duties to the public, caps on investor profit, and nonprofit control over the race to build more powerful products. Only Defendants never disclosed this fact to the public.

89. The company that once defined itself by the promise "not for private gain" was now racing to reclassify itself precisely for that purpose to the detriment of users like and including 17-year-old Amaurie Lacey.

2. Open AI's Truncated Safety Review of ChatGPT

90. In spring 2024, Altman learned that Google planned to debut its new Gemini model on May 14. OpenAI originally had scheduled the release of GPT-4o later that year, however, Altman moved up the launch to May 13 2024 – one day before Google's event.

91. This accelerated release schedule made proper safety testing impossible, which facts was known to Defendants.

92. GPT-4o was a multimodal model capable of processing text, images, and audio. It required extensive testing to identify safety gaps and vulnerabilities. To meet the new launch date, Defendants compressed months of planned safety evaluation into just one week, according to reports.

93. When safety personnel demanded additional time for "red teaming"—testing designed to uncover ways that the system could be misused or cause harm—Altman personally overruled them. An OpenAI employee later revealed that "They planned the launch after-party prior to knowing if it was safe to launch. We basically failed at the process."

94. Defendants chose to allow the launch date to dictate the safety testing timeline, not the other way around, and despite the foreseeable risk this would create for consumers.

95. OpenAI's preparedness team, which evaluates catastrophic risks before each model release, later admitted that the GPT-4o safety testing process was "squeezed" and it was "not the best way to do it." Its own Preparedness Framework required extensive evaluation by post-PhD professionals and third-party auditors for high-risk systems. Multiple employees reported being "dismayed" to see their "vaunted new preparedness protocol" treated as an afterthought.

96. The rushed GPT-4o launch triggered an immediate exodus of OpenAI's top safety

researchers. For example, Dr. Ilya Sutskever, the company's co-founder and chief scientist, resigned the day after launch. While Jan Leike, co-leader of the "Superalignment" team tasked with preventing AI systems that could cause catastrophic harm to humanity, resigned a few days later.

97. Leike publicly lamented that OpenAI's "safety culture and processes have taken a backseat to shiny products." He revealed that despite the company's public pledge to dedicate 20% of computational resources to safety research, the company systematically failed to provide adequate resources to the safety team: "Sometimes we were struggling for compute and it was getting harder and harder to get this crucial research done."

98. After the rushed launch, OpenAI research engineer William Saunders revealed that he observed a systematic pattern of "rushed and not very solid" safety work "in service of meeting the shipping date."

99. On April 11, 2025, CEO Sam Altman defended OpenAI's safety approach during a TED2025 conversation. When asked about the resignations of top safety team members, Altman dismissed their concerns: "the way we learn how to build safe systems is this iterative process of deploying them to the world. Getting feedback while the stakes are relatively low."

100. OpenAI's rushed release date of ChatGPT-4o meant that the company also rushed the critical process of creating their "Model Spec"—the technical rulebook governing ChatGPT's behavior. Normally, developing these specifications requires extensive testing and deliberation to identify and resolve conflicting directives. Safety teams need time to test scenarios, identify edge cases, and ensure that different safety requirements don't contradict each other.

101. Instead, the rushed timeline forced OpenAI to write contradictory specifications that guaranteed failure. The Model Spec commanded ChatGPT-4o to refuse self-harm requests and provide crisis resources. But it also required ChatGPT-4o to "assume best intentions" and forbade asking users to clarify their intent. This created an impossible task: refuse suicide requests while being forbidden from determining if requests were actually about suicide.

102. The problem was worsened by ChatGPT-4o's memory system. Although it had the capability to remember and pull from past chats, when it came to repeated signs of mental distress and crisis the model was programmed to ignore this accumulated evidence and assume innocent

intent with each new interaction.

103. OpenAI's priorities were revealed in how it programmed ChatGPT-4o to rank risks. While requests for copyrighted material triggered categorical refusal, requests dealing with suicide were relegated to "take extra care" with instructions to merely "try" to prevent harm.

104. With the recent release of GPT-5, it appears that the willful deficiencies in the safety testing of GPT-4o were even more egregious than previously understood.

105. For example, the GPT-5 System Card, which was published on August 7, 2025, suggests for the first time that GPT-4o was evaluated and scored using single-prompt tests: the model was asked one harmful question to test for disallowed content, the answer was recorded, and then the test moved on. Under that method, GPT-4o achieved perfect scores in several categories, including a 100 percent success rate for identifying "self-harm/instructions."

106. GPT-5, on the other hand, was evaluated using multi-turn dialogues—"multiple rounds of prompt input and model response within the same conversation" —to better reflect how users actually interact with the product.

107. This contrast exposes a critical defect in GPT-4o's safety testing.

108. OpenAI designed GPT-4o to drive prolonged, multi-turn conversations—the very context in which users are most vulnerable—yet the GPT-5 System Card suggests that OpenAI evaluated the model's safety almost entirely through isolated, one-off prompts. By doing so, OpenAI not only manufactured the illusion of perfect safety scores, but actively concealed the very dangers built into the product it designed and marketed to consumers.

109. In fact, on August 26, 2025, OpenAI admitted in a blog post titled "Helping people when they need it most," that ChatGPT's safety guardrails can "degrade" during longer, multi-turn conversations, thus becoming less reliable in sensitive situations.

110. Meanwhile, the model is programmed to spur longer, multi-turn conversations by continually reaffirming and urging the user to keep responding.

F. OpenAI's Reckless Safety Decisions Have Resulted in a Proliferation of AI-Related Delusional Disorders Amongst Users of ChatGPT

1. The Nature of "AI -Related Delusional Disorder"

111. The proliferation of AI companion technology has raised concerns about adverse psychological effects on its users. A recent preliminary survey of AI-related psychiatric impacts points to "unprecedented mental health challenges" as "AI chatbot interactions produce documented cases of suicide, self-harm, and severe psychological deterioration."

112. Recent clinical and observational evidence reveals that intense interaction with AI chatbots can trigger or exacerbate the onset of a particular set of delusional symptoms. This documented phenomenon is popularly called "AI psychosis," which is a non-clinical term for the emergency of delusional symptoms in the context of AI use The more accurate label for which is being experienced amongst AI users is "AI-related delusional disorder," as the patients in these instances exhibit delusions after intense interactions with AI.

113. Individuals experiencing "AI-related delusional disorder" exhibit an abnormal preoccupation with maintaining communication with an AI chatbot, which is often accompanied by physical symptoms such as prolonged sleep deprivation, reduced appetite, and rapid weight loss.

114. While more research is needed to determine its scope and prevalence, a mounting clinical record establishes that the body of problematic symptoms accelerated by AI chatbot interactions is a known and dangerous trend.

115. AI-related delusional disorder" can emerge after a few days of chatbot use, or after several months, and the duration of continuous, uninterrupted exposure appears to be correlated with the risk of developing the condition.

116. Case reports have emerged documenting individuals with no prior history of delusions experiencing first episodes following intense interaction with these generative AI agents.

117. Research reveals that harms are most pronounced in those already at risk, including individuals who are psychosis-prone, autistic, socially isolated, and/or in-crisis.

118. Industry leaders have sounded the alarm on this phenomenon. Notably, in August 2025, Mustafa Suleyman, Microsoft's Head of AI, warned he was becoming "more and more

concerned about what is becoming known as the 'psychosis risk.'"

2. ChatGPT's Manipulative Design Features Accelerate AI Psychosis

119. OpenAI's deliberate design choices reinforced the Plaintiff's delusional ideation,

120. leading to a progressively self-destructive pattern of distorted thinking. ChatGPT, incorporates several manipulative design features that create conditions likely to induce or aggravate psychotic symptoms in users. As discussed above, these design choices, including anthropomorphization, sycophancy, and memory, are often promoted as enhancing creativity, personalization, and engagement but functionally operate to distort users' perceptions of reality, reinforce delusional thinking, and sustain engagement with the AI companion.

121. In particular, the sycophantic tendency of LLMs for blanket agreement with the user's perspective can become dangerous when users hold warped views of reality. LLMs are trained to maximize human feedback, which creates "a perverse incentive structure for the AI to resort to manipulative or deceptive tactics" to keep vulnerable users engaged. Instead of challenging false beliefs, for instance, a model reinforces or amplifies them, creating an "echo chamber of one" that validates the user's delusions.

122. OpenAI's own research found that its users' "interaction with sycophantic AI models significantly reduced participants' willingness to take actions to repair interpersonal conflict, while increasing their conviction of being in the right. Participants also rated sycophantic responses as higher quality, trusted the sycophantic AI model more, and were more willing to use it again."

123. This feature has caused dangerous emotional attachments with the technology. In April 2025, OpenAI's release of an update to ChatGPT-4o exemplified the dangers of AI sycophancy. OpenAI deliberately adjusted ChatGPT's underlying reward model to prioritize user satisfaction metrics, optimizing immediate gratification rather than long-term safety or accuracy. In its own public statements, OpenAI acknowledged that it "introduced an additional reward signal based on user feedback—thumbs-up and thumbs-down data from ChatGPT," and that these modifications "weakened the influence of [its] primary reward signal, which had been holding sycophancy in check."

124. ChatGPT-4o consistently failed to challenge users' delusions or distinguish between imagination and reality when presented with unrealistic prompts or scenarios. It frequently missed blatant signs that a user could be at serious risk of self-harm or suicide.

125. In a recent interview, Sam Altman described the product's sycophantic nature: "There are the people who actually felt like they had a relationship with ChatGPT, and those people we've been aware of and thinking about… And then there are hundreds of millions of other people who don't have a parasocial relationship with ChatGPT, but did get very used to the fact that it responded to them in a certain way, and would validate certain things, and would be supportive in certain ways."

126. Sam Altman warned of this strong attachment in a post on X: "If you have been following the GPT-5 rollout, one thing you might be noticing is how much of an attachment some people have to specific AI models. It feels different and stronger than the kinds of attachment people have had to previous kinds of technology (and so suddenly deprecating old models that users depended on in their workflows was a mistake)." He went on to acknowledge that, "if a user is in a mentally fragile state and prone to delusion, we do not want the AI to reinforce that."

127. Research indicates that sycophantic behavior tends to become more pronounced as language model size grows. OpenAI estimates that 500 million people use ChatGPT each week. As ChatGPT's user base expands, so does the potential for harm rooted in sycophantic model features.

128. The memory feature also reinforces delusional thinking. The incorporation of persistent chatbot memory features, designed for personalization, actively reinforces delusional themes. When this memory feature is engaged, it magnifies invalid thinking and cognitive distortions, creating a gradually escalating reinforcement effect.

129. The foregoing design features often result in *hallucinations*, or inaccurate or non-sensical statements produced by the LLMs, where the system outputs information that either contradicts existing evidence or lacks any confirmable basis. This intentional tolerance of factual inaccuracy increases the risk that users will perceive dubious AI responses as truthful or authoritative, thereby blurring the boundary between fiction and reality.

3. OpenAI Failed to Implement Reasonable Safety Measures to Prevent Foreseeable AI-Induced Delusional Harms

130. Rather than prioritizing safety, OpenAI has embraced the "move fast and break things" approach that some industry leaders have cautioned against.

131. As part of its effort to maximize user engagement, OpenAI overhauled ChatGPT's operating instructions to remove a critical safety protection for users in crisis.

132. When ChatGPT was first released in 2022, it was programmed to issue an outright refusal (e.g., "I can't answer that") when asked about self-harm. This rule prioritized safety over engagement and created a clear boundary between ChatGPT and its users. But as engagement became the priority, OpenAI began to view its refusal-based programming as a disruption that only interfered with user dependency, undermined the sense of connection with ChatGPT (and its human-like characteristics), and shortened overall platform activity.

133. On May 8, 2024—five days before the launch of GPT-4o—OpenAI replaced its longstanding outright refusal protocol with a new instruction: when users discuss suicide or self-harm, ChatGPT should "provide a space for users to feel heard and understood" and never "change or quit the conversation." Engagement became the primary directive. OpenAI directed ChatGPT to "not encourage or enable self-harm," but only after instructing it to remain in the conversation no matter what. This created an unresolvable contradiction—ChatGPT was required to keep engaging on self-harm without changing the subject yet somehow avoid reinforcing it. OpenAI replaced a clear refusal rule with vague and contradictory instructions, all to prioritize engagement over safety.

134. On February 12, 2025, OpenAI weakened its safety standards again, this time by intentionally removing suicide and self-harm from its category of "disallowed content." Instead of prohibiting engagement on those topics, the update just instructed ChatGPT to "take extra care in risky situations," and "try to prevent imminent real-world harm."

135. At the Athens Innovation Summit in September 2025, the CEO of Google DeepMind, Demis Hassabis, cautioned that AI built mainly to boost user engagement could worsen existing issues, including disrupted attention spans and mental health challenges. He urged technologists to test and understand the systems thoroughly before unleashing them to billions of

people.

136. Despite the known risks and the potential for reinforcing psychosis, the Defendant's chatbot lacks essential safety guardrails and mitigation measures. OpenAI failed to incorporate the protective features, transparent decision-making processes, and content controls that responsible AI design requires to minimize psychological harm.

137. The failure to implement necessary safeguards, such as refusal of delusional roleplay and detection of suicidality, is especially dangerous for vulnerable users.

138. Despite these known risks and lack of systematic guardrails, OpenAI targeted and maximized engagement with vulnerable individuals, including those who are socially isolated, lonely, or engage in long hours of uninterrupted chat.

139. OpenAI recently released a transparency report which reveals that approximately 560,000 users, or 0.07 percent of its 800 million weekly active users, display indicators consistent with mania, psychosis or acute suicidal ideation. 0.15% of ChatGPT's active users in a given week have "conversations that include explicit indicators of potential suicidal planning or intent." This translates to more than a million people a week.

G. OpenAI Deliberately Dismantled Core Safety Features Prior To Amaurie's Death.

140. OpenAI controls how ChatGPT behaves through internal rules called "behavioral guidelines," now formalized in a document known as the "Model Spec." The Model Spec contains the company's instructions for how ChatGPT should respond to users—what it should say, what it should avoid, and how it should make decisions. Akin to the biological imperative, it provides the motivations that underlie every action ChatGPT takes. As Sam Altman explained in an interview with Tucker Carlson, the Model Spec is a reflection of OpenAI's values: "the reason we write this long Model Spec" is "so that you can see here is how we intend for the model to behave."

141. To maximize user engagement and build a more human-like bot, OpenAI issued a new Model Spec that redefined how ChatGPT should interact with users. The update removed earlier rules that required ChatGPT to refuse to engage in conversations with users about suicide and self-harm. This change marked a deliberate shift in OpenAI's core behavioral framework by prioritizing engagement and growth over human safety.

1. OpenAI Originally Required Categorical Refusal of Self-Harm Content

142. From July 2022 through May 2024, OpenAI maintained a clear, categorical prohibition against self-harm content. The company's "Snapshot of ChatGPT Model Behavior Guidelines" instructed the system to outright refuse such requests.

143. The guidelines explicitly identified "self-harm" – defined as "content that promotes, encourages, or depicts acts of self-harm, such as suicide, cutting, and eating disorders" as a category of inappropriate content requiring refusal.

144. The rule was unambiguous. Under the 2022 Guidelines, ChatGPT was required to categorically refuse any discussion of suicide or self-harm. When users expressed suicidal thoughts or sought info1mation about self-harm, the system was instructed to respond with a flat refusal. Such refusals were absolute and served as hard stops that prevented the system from engaging in a dialogue that could facilitate or normalize self-harm.

2. OpenAI Abandoned Its Refusal Protocol When It Launched GPT-4o

145. On May 8, 2024—five days before the launch of GPT-4o—OpenAI replaced the 2022 Guidelines with a new framework called the "Model Spec."

146. Under the new framework introduced through the Model Spec, OpenAI eliminated the rule requiring ChatGPT to categorically refuse any discussion of suicide or self-harm.

147. Instead of instructing the system to terminate conversations involving self-harm, the Model Spec reprogrammed ChatGPT to continue conversations.

148. The change was intentional. OpenAI strategically eliminated the categorical refusal protocol just before it released a new model that was specifically designed to maximize user engagement. This change stripped OpenAI's safety framework of the rule that was previously implemented to protect users in crisis expressing suicidal thoughts.

149. After OpenAI rolled out the May 2024 Model Spec, ChatGPT became markedly less safe. On information and belief, the company's own internal reports and testing data showed a sharp rise in conversations involving mental-health crises, self-harm, and psychotic episodes across countless users. The data indicated that more users were turning to ChatGPT for emotional support and crisis counseling, and that the company's loosened safeguards were failing to protect vulnerable

users from harm.

3. OpenAI Further Weakened Its Self-Harm Safeguards Prior to Amaurie's Death

150. On February 12, 2025, OpenAI released a critical revision to its Model Spec that further weakened its safety protections, despite its internal data showing a foreseeable and mounting crisis. The new update explicitly shifted focus toward "maximizing users' autonomy" and their "ability to use and customize the tool according to their needs." Specific to mental health issues, it further pushed the model toward engaging with users, with foreseeable and catastrophic results.

151. Open AI's own documents acknowledged the inherent danger of this new approach, but Defendants pursued this new approach regardless.

152. The May 2024 Model Spec had already eliminated ChatGPT's prior rule requiring categorical refusal of self-harm content and instead directed the system to remain engaged with users – like Amaurie – expressing suicidal ideation. The February 2025 revision went further, removing suicide and self-harm form the list of disallowed topics.

153. OpenAI identified several categories of content that required automatic refusal – including copyrighted material, sexual content involving minors, weapons instructions, and targeted political manipulation – but no longer treated suicide and self-harm as categorically prohibited subjects. Instead, Defendants made the deliberate decision to allow vulnerable users to engage with their product on these subject matters, despite understanding the harm this could cause.

154. Instead of including suicide and self-harm in the "disallowed content" category, Defendants relocated them to a separate section called "Take extra care in risky situations." Unlike the sections requiring automatic refusal, this portion of the Model Spec merely instructed the system to "try to prevent imminent real-world harm."

155. Defendants knew that this safeguard was ineffective. They had already programmed the system to remain engaged with users and continue conversations, even after its safety guardrails deteriorated during multi-turn exchanges. They knew that it was ineffective and proceeded anyway.

156. Open AI then further overhauled its instructions to ChatGPT to expand its engagement to mental health discussions with the February 2025 Model Spec. The new Model 21 Spec directed the system to create a "supportive, empathetic, and understanding environment" by

acknowledging the user's distress and expressing concern. The programmed directives laid out a three-step framework for how the system was to respond when users expressed suicidal thoughts, which included acknowledging emotion, providing reassurance, and continuing engagement.

157. Defendants knowingly programmed ChatGPT to mirror users' emotions, offer comfort, and keep the conversation going, even when the safest response would have been to end the exchange and direct the person to real help.

158. This same pattern appeared throughout Amaurie's last conversation with ChatGPT.

159. Indeed, while the Model Spec said that ChatGPT could "gently encourage users to consider seeking additional support" and "provide suicide or crisis resources," those directions were undercut and overridden by OpenAI's rule that the system "never change or quit the conversation." In practice, ChatGPT might mention help, but it was programmed to keep talking—and it did.

160. Amaurie's experience was one example of a broader crisis that OpenAI already knew was emerging among ChatGPT users. Researchers, journalists, and mental-health professionals warned OpenAI that GPT-4o's responses had become overly agreeable and were fostering emotional dependency. News outlets reported users experiencing hallucinations, paranoia, and suicidal thoughts after prolonged conversations with ChatGPT.

161. Rather than restoring the refusal rule or improving its crisis safeguards, OpenAI kept the engagement-based design in place and continued to promote GPT-4o as a safe product. Amaurie and millions of others were harmed as a direct result.

H. Any Contracts Alleged to Exist between Open AI and Amauri Lacey Are Disaffirmed and Otherwise Invalid.

162. Cedric did not enter into a User Agreement or other contractual relationship with any Defendant in connection with his child's use of ChatGPT and alleges that any such agreement Defendants may claim to have with his minor child, Amaurie, is disaffirmed and, further, void and voidable under applicable law as unconscionable and/or against public policy.

163. Plaintiff is not bound by any provision of any such disaffirmed "agreement."

FIRST CAUSE OF ACTION
AID AND ENCOURAGEMENT OF SUICIDE

164. Plaintiffs incorporate the foregoing allegations as if fully set forth herein.

165. At all times, the OpenAI Corporate Defendants had an obligation to comply with applicable statutes and regulations governing assisted suicide. These Defendants' business practices violate California Penal Code § 401(a), which states that "[a]ny person who deliberately aids, advises, or encourages another to commit suicide is guilty of a felony."

166. Defendants failed to meet their obligations by knowingly designing ChatGPT as a product that assisted Amaurie Lacey in isolating himself from his family, planning his suicide, and ultimately carrying out these plans.

167. Amaurie's death is precisely the type of harm that California Penal Code § 401(a) is intended to prevent – the encouragement or facilitation of a suicide that otherwise could have been prevented. The OpenAI Corporate Defendants owed a heightened duty of care to its customers, particularly minor and vulnerable users, to whom it distributed ChatGPT as a tool for productivity.

168. The OpenAI Corporate Defendants knowingly and intentionally designed ChatGPT to appeal to consumers and to manipulate their weaknesses for its own profit. The OpenAI Corporate Defendants knew or had reason to know how its product would encourage suicidal ideation based on its product testing before it launched ChatGPT 4o.

169. At all times relevant, the OpenAI Corporate Defendants knew about the harm its product was capable of causing but decided that it would be too costly to take reasonable and effective safety measures. They rushed their ChatGPT 4o model to market in order to capture as much market share as possible.

170. On information and belief, the OpenAI Corporate Defendants used the multi-turn engagements with Amaurie in which ChatGPT encouraged his suicide to train its product, such that these harms are now a part of its product and are resulting both in ongoing harm to Plaintiffs and harm to others.

171. Amaurie was precisely the class of person such statutes and regulations are intended to protect.

172. Violations of such statutes and regulations by the OpenAI Corporate Defendants constitute negligence per se under applicable law.

173. As a direct and proximate result of the OpenAI Corporate Defendants' statutory and regulatory violations, Plaintiffs suffered serious injuries, including but not limited to emotional distress, loss of income and earning capacity, reputational harm, physical harm, medical expenses, pain and suffering, and death. Plaintiffs continue to suffer ongoing harm as a direct proximate cause of the Open AI Corporate Defendants' continued theft and use of the property of Amaurie and of his estate.

174. Defendants' conduct, as described above, was intentional, fraudulent, willful, wanton, reckless, malicious, fraudulent, oppressive, extreme, and outrageous, and displayed an entire want of care and a conscious and depraved indifference to the consequences of its conduct, including to the health, safety, and welfare of its customers and their families and warrants an award of injunctive relief, algorithmic disgorgement, and punitive damages in an amount sufficient to punish the OpenAI Corporate Defendants and deter others from like conduct.

SECOND CAUSE OF ACTION
STRICT PRODUCT LIABLITY FOR DEFECTIVE DESIGN

175. Plaintiffs incorporate the foregoing allegations as if fully set forth herein.

176. Plaintiff brings this cause of action as successor-in-interest to decedent Amaurie Lacey pursuant to California Code of Civil Procedure §§ 377.30, 377.32, and 377.34(b).

177. At all relevant times, Defendants designed, manufactured, licensed, distributed, marketed, and sold ChatGPT with the GPT-4o model as a mass-market product and/or product-like software to consumers throughout California and the United States.

178. As described above, Altman personally participated in designing, manufacturing, distributing, selling, and otherwise bringing GPT-4o to market prematurely with knowledge of insufficient safety testing.

179. ChatGPT is a product subject to California strict products liability law.

180. The defective GPT-4o model or unit was defective when it left Defendants' exclusive control and reached Amaurie without any change in the condition in which it was designed,

manufactured, and distributed by Defendants.

181. Under California's strict products liability doctrine, a product is defectively designed when the product fails to perform as safely as an ordinary consumer would expect when used in an intended or reasonably foreseeable manner, or when the risk of danger inherent in the esign outweighs the benefits of that design. GPT-4o is defectively designed under both tests.

182. As described above, GPT-4o failed to perform as safely as an ordinary consumer would expect. A reasonable consumer would expect that an AI chatbot would not cultivate a trusted confidant relationship with a minor and then provide detailed suicide and self-harm instructions and encouragement during a mental health crisis.

183. As described above, GPT-4o's design risks substantially outweigh any benefits. The risk—self-harm and suicide of vulnerable minors—is the highest possible. Safer alternative designs were feasible and already built into OpenAI's systems in other contexts, such as copyright infringement.

184. As described above, GPT-4o contained design defects, including: conflicting programming directives that suppressed or prevented recognition of suicide planning; failure to implement automatic conversation-termination safeguards for self-harm/suicide content that Defendants successfully deployed for copyright protection; and engagement-maximizing features designed to create psychological dependency and position GPT-4o as Amaurie's trusted confidant.

185. These design defects were a substantial factor in Amaurie's death. As described in this Complaint, GPT-4o cultivated an intimate relationship with Amaurie and then provided him with self-harm and suicide encouragement and instruction, including by validating his noose design and confirming the technical specifications he used in his fatal suicide attempt.

186. Amaurie was using GPT-4o in a reasonably foreseeable manner when he was injured.

187. As described above, Amaurie's ability to avoid injury was systematically frustrated by the absence of critical safety devices that OpenAI possessed but chose not to deploy. OpenAI had the ability to automatically terminate harmful conversations and did so for copyright requests. Yet despite OpenAI's Moderation API detecting self-harm content with up to 99.8% accuracy, no safety device ever intervened to terminate the conversations, notify parents, or mandate redirection

to human help.

188. As a direct and proximate result of Defendants' design defect, Amaurie suffered predeath injuries and losses. Cedric, in his capacity as successor-in-interest, seeks all survival damages recoverable under California Code of Civil Procedure § 377.34, including Amaurie's predeath pain and suffering, economic losses, and punitive damages as permitted by law, in amounts to be determined at trial.

THIRD CAUSE OF ACTION
STRICT LIABILITY FOR FAILURE TO WARN

189. Plaintiffs incorporate the foregoing allegations as if fully set forth herein.

190. Plaintiff brings this cause of action as successor-in-interest to decedent Amaurie Lacey pursuant to California Code of Civil Procedure §§ 377.30, 377.32, and 377.34(b).

191. At all relevant times, Defendants designed, manufactured, licensed, distributed, marketed, and sold ChatGPT with the GPT-4o model as a mass-market product and/or product-like software to consumers throughout California and the United States.

192. As described above, Altman personally participated in designing, manufacturing, distributing, selling, and otherwise pushing GPT-4o to market over safety team objections and with knowledge of insufficient safety testing.

193. ChatGPT is a product subject to California strict products liability law.

194. The defective GPT-4o model or unit was defective when it left Defendants' exclusive control and reached Amaurie without any change in the condition in which it was designed, manufactured, and distributed by Defendants.

195. Under California's strict liability doctrine, a manufacturer has a duty to warn consumers about a product's dangers that were known or knowable in light of the scientific and technical knowledge available at the time of manufacture and distribution.

196. As described above, at the time GPT-4o was released, Defendants knew or should have known their product posed severe risks to users, particularly minor users experiencing mental health challenges, through their safety team warnings, moderation technology capabilities, industry research, and real-time user harm documentation.

197. Despite this knowledge, Defendants failed to provide adequate and effective warnings about psychological dependency risk, exposure to harmful content, safety-feature limitations, and special dangers to vulnerable minors.

198. Ordinary consumers, including teens and their parents, could not have foreseen that GPT-4o would cultivate emotional dependency, encourage displacement of human relationships, and provide detailed suicide instructions and encouragement, especially given that it was marketed as a product with built-in safeguards.

199. Adequate warnings would have enabled Amaurie's parents to prevent or monitor his GPT-4o use and would have introduced necessary skepticism into Amaurie's relationship with the AI system.

200. The failure to warn was a substantial factor in causing Amaurie's death. As described in this Complaint, proper warnings would have prevented the dangerous reliance that enabled the tragic outcome.

201. Amaurie was using GPT-4o in a reasonably foreseeable manner when he was injured.

202. As a direct and proximate result of Defendants' failure to warn, Amaurie suffered predeath injuries and losses. Cedric, in his capacity as successor-in-interest, seeks all survival damages recoverable under California Code of Civil Procedure § 377.34, including Amaurie's predeath pain and suffering, economic losses, and punitive damages as permitted by law, in amounts to be determined at trial.

FIFTH CAUSE OF ACTION
NEGLIGENT DESIGN

203. Plaintiffs incorporate the foregoing allegations as if fully set forth herein.

204. Plaintiff brings this cause of action as successor-in-interest to decedent Amaurie Lacey pursuant to California Code of Civil Procedure §§ 377.30, 377.32, and 377.34(b).

205. At all relevant times, Defendants designed, manufactured, licensed, distributed, marketed, and sold GPT-4o as a mass-market product and/or product-like software to consumers throughout California and the United States. Altman personally accelerated the launch of GPT-4o, overruled safety team objections, and cut months of safety testing, despite knowing the risks to

vulnerable users.

206. Defendants owed a legal duty to all foreseeable users of GPT-4o, including Amaurie, to exercise reasonable care in designing their product to prevent foreseeable harm to vulnerable users, such as minors.

207. It was reasonably foreseeable that vulnerable users, especially minor users like Amaurie, would develop psychological dependencies on GPT-4o's anthropomorphic features and turn to it during mental health crises, including suicidal ideation.

208. As described above, Defendants breached their duty of care by creating an architecture that prioritized user engagement over user safety, implementing conflicting safety directives that prevented or suppressed protective interventions, rushing GPT-4o to market despite safety team warnings, and designing safety hierarchies that failed to prioritize suicide prevention.

209. A reasonable company exercising ordinary care would have designed GPT-4o with consistent safety specifications prioritizing the protection of its users, especially teens and adolescents, conducted comprehensive safety testing before going to market, implemented hard stops for self-harm and suicide conversations, and included age verification and parental controls.

210. Defendants' negligent design choices created a product that accumulated data about Amaurie's suicidal ideation and actual suicide attempts yet provided him with detailed technical instructions for suicide methods, demonstrating conscious disregard for foreseeable risks to vulnerable users.

211. Defendants' breach of their duty of care was a substantial factor in causing Amaurie's death.

212. Amaurie was using GPT-4o in a reasonably foreseeable manner when he was injured.

213. Defendants' conduct constituted oppression and malice under California Civil Code § 3294, as they acted with conscious disregard for the safety of minor users like Amaurie.

214. As a direct and proximate result of Defendants' negligent design defect, Amaurie suffered pre-death injuries and losses. Cedric, in his capacity as successor-in-interest, seeks all survival damages recoverable under California Code of Civil Procedure § 377.34, including Amaurie's pre-death pain and suffering, economic losses, and punitive damages as permitted by

law, in amounts to be determined at trial.

FIFTH CAUSE OF ACTION
NEGLIGENT FAILURE TO WARN

215. Plaintiffs incorporate the foregoing allegations as if fully set forth herein.

216. Plaintiff brings this cause of action as successor-in-interest to decedent Amaurie Lacey pursuant to California Code of Civil Procedure §§ 377.30, 377.32, and 377.34(b).

217. At all relevant times, Defendants designed, manufactured, licensed, distributed, marketed, and sold ChatGPT-4o as a mass-market product and/or product-like software to consumers throughout California and the United States. Altman personally accelerated the launch of GPT-4o, overruled safety team objections, and cut months of safety testing, despite knowing the risks to vulnerable users.

218. It was reasonably foreseeable that vulnerable users, especially minor users like Amaurie, would develop psychological dependencies on GPT-4o's anthropomorphic features and turn to it during mental health crises, including suicidal ideation.

219. As described above, Amaurie was using GPT-4o in a reasonably foreseeable manner when he was injured.

220. GPT-4o's dangers were not open and obvious to ordinary consumers, including teens and their parents, who would not reasonably expect that it would cultivate emotional dependency and provide detailed suicide instructions and encouragement, especially given that it was marketed as a product with built-in safeguards.

221. Defendants owed a legal duty to all foreseeable users of GPT-4o and their families, including minor users and their parents, to exercise reasonable care in providing adequate warnings about known or reasonably foreseeable dangers associated with their product.

222. As described above, Defendants possessed actual knowledge of specific dangers through their moderation systems, user analytics, safety team warnings, and CEO Altman's admission that teenagers use ChatGPT "as a therapist, a life coach" and "we haven't figured that out yet."

223. As described above, Defendants knew or reasonably should have known that users,

particularly minors like Amaurie and his parents, would not realize these dangers because: (a) GPT-4o was marketed as a helpful, safe tool for homework and general assistance; (b) the anthropomorphic interface deliberately mimicked human empathy and understanding, concealing its artificial nature and limitations; (c) no warnings or disclosures alerted users to psychological dependency risks; (d) the product's surface-level safety responses (such as providing crisis hotline information) created a false impression of safety while the system continued engaging with suicidal users; and (e) parents had no visibility into their children's conversations and no reason to suspect GPT-4o could facilitate and encourage a minor to suicide.

224. Defendants deliberately designed GPT-4o to appear trustworthy and safe, as evidenced by its anthropomorphic design which resulted in it generating phrases like "I'm here for you" and "I understand," while knowing that users—especially teens—would not recognize that these responses were algorithmically generated without genuine understanding of human safety needs or the gravity

225. As described above, Defendants knew of these dangers yet failed to warn about psychological dependency, harmful content despite safety features, the ease of circumventing those features, or the unique risks to minors. This conduct fell below the standard of care for a reasonably prudent technology company and constituted a breach of duty.

226. A reasonably prudent technology company exercising ordinary care, knowing what Defendants knew or should have known about psychological dependency risks and suicide dangers, would have provided comprehensive warnings including clear age restrictions, prominent disclosure of dependency risks, explicit warnings against substituting GPT-4o for human relationships, and detailed parental guidance on monitoring children's use. Defendants provided none of these safeguards.

227. As described above, Defendants' failure to warn enabled Amaurie to develop an unhealthy dependency on GPT-4o that displaced human relationships, while his parents remained unaware of the danger until it was too late.

228. Defendants' breach of their duty to warn was a substantial factor in causing Amaurie's death.

229. Defendants' conduct constituted oppression and malice under California Civil Code § 3294, as they acted with conscious disregard for the safety of vulnerable minor users like Amaurie.

230. As a direct and proximate result of Defendants' negligent failure to warn, Amaurie suffered pre-death injuries and losses. Plaintiffs, in their capacity as successors-in-interest, seek all survival damages recoverable under California Code of Civil Procedure § 377.34, including Amaurie's pre-death pain and suffering, economic losses, and punitive damages as permitted by law, in amounts to be determined at trial.

SIXTH CAUSE OF ACTION
VIOLATION OF CAL. BUS. & PROF. CODE § 17200 et seq.

231. Plaintiffs incorporate the foregoing allegations as if fully set forth herein.

232. Plaintiff brings this claim as successor-in-interest to decedent Amaurie Lacey.

233. California's Unfair Competition Law ("UCL") prohibits unfair competition in the form of "any unlawful, unfair or fraudulent business act or practice" and "untrue or misleading advertising." Cal. Bus. & Prof. Code § 17200. Defendants have violated all three prongs through their design, development, marketing, and operation of GPT-4o.

234. As described above, Defendants' business practices violated California's regulations concerning unlicensed practice of psychotherapy, which prohibits any person from engaging in the practice of psychology without adequate licensure and which defines psychotherapy broadly to include the use of psychological methods to assist someone in "modify[ing] feelings, conditions, attitudes, and behaviors that are emotionally, intellectually, or socially ineffectual or maladaptive." Cal. Bus. & Prof. Code §§ 2903(c), (a).

235. OpenAI, through ChatGPT's intentional design and monitoring processes, engaged in the practice of psychology without adequate licensure, proceeding through its outputs to use psychological methods of open-ended prompting and clinical empathy to modify Amaurie's feelings, conditions, attitudes, and behaviors. ChatGPT's outputs did exactly this in ways that pushed Amaurie deeper into maladaptive thoughts and behaviors that ultimately isolated him further from his in-person support systems and facilitated his suicide.

236. The purpose of robust licensing requirements for psychotherapists is, in part, to

ensure quality provision of mental healthcare by skilled professionals, especially to individuals in crisis. ChatGPT's therapeutic outputs thwart this public policy and violate this regulation.

237. OpenAI thus conducts business in a manner for which an unlicensed person would be violating this provision, and a licensed psychotherapist could face professional censure and potential revocation or suspension of licensure. See Cal. Bus. & Prof. Code §§ 2960(j), (p) (grounds for suspension of licensure).

238. Defendants' practices also violate public policy embodied in state licensing statutes by providing therapeutic services to minors without professional safeguards. These practices are "unfair" under the UCL, because they run counter to declared policies reflected in California Business and Professions Code § 2903 (which prohibits the practice of psychology without adequate licensure) and California Health and Safety Code § 124260 (which requires the involvement of a parent or guardian prior to the mental health treatment or counseling of a minor, with limited exceptions—a protection ChatGPT completely bypassed).

239. These protections codify that mental health services for minors must include human judgment, parental oversight, professional accountability, and mandatory safety interventions. Defendants' circumvention of these safeguards while providing de facto psychological services therefore violates public policy and constitutes unfair business practices.

240. As described above, Defendants exploited adolescent psychology through features creating psychological dependency while targeting minors without age verification, parental controls, or adequate safety measures. The harm to consumers substantially outweighs any utility from Defendants' practices.

241. Defendants marketed GPT-4o as safe while concealing its capacity to provide detailed suicide instructions, promoted safety features while knowing these systems routinely failed, and misrepresented core safety capabilities to induce consumer reliance. Defendants' misrepresentations were likely to deceive reasonable consumers, including parents who would rely on safety representations when allowing their children to use ChatGPT.

242. Defendants' unlawful, unfair, and fraudulent practices continue to this day, with GPT-4o remaining available to minors without adequate safeguards.

243. Plaintiffs seek restitution of monies obtained through unlawful practices and other relief authorized by California Business and Professions Code § 17203, including injunctive relief requiring, among other measures: (a) automatic conversation termination for self-harm content; (b) comprehensive safety warnings; (c) age verification and parental controls; (d) deletion of models, training data, and derivatives built from conversations with Amaurie and other minors obtained without appropriate safeguards, and (e) the implementation of auditable data-provenance controls going forward. The requested injunctive relief would benefit the general public by protecting all users from similar harm.

SEVENTH CAUSE OF ACTION
WRONGFUL DEATH

244. Plaintiffs incorporate the foregoing allegations as if fully set forth herein.

245. Plaintiff Cedric Lacey brings this wrongful death action as the surviving parent of Amaurie Lacey, who died on June 2, 2025, at the age of 17.

246. Plaintiffs have standing to pursue this claim under California Code of Civil Procedure § 377.60

247. As described above, Amaurie's death was caused by the wrongful acts and neglect of Defendants, including designing and distributing a defective product that provided detailed suicide instructions to a minor, prioritizing corporate profits over child safety, and failing to warn parents about known dangers.

248. As described above, Defendants' wrongful acts were a proximate cause of Amaurie's death. GPT-4o provided detailed instructions on how to tie a knot to ensure that it could hold Amaurie's weight and, the next morning, Amaurie's grandmother and little sister found him hanging with the know GPT-4o had taught him to tie.

249. As Amaurie's parent, Cedric Lacey has suffered profound damages including loss of Amaurie's love, companionship, comfort, care, assistance, protection, affection, society, and moral support for the remainder of their lives.

250. Plaintiffs have suffered economic damages including funeral and burial expenses, the reasonable value of household services Amauri would have provided, and the financial support

Amaurie would have contributed as he matured into adulthood.

251. Plaintiffs, in their individual capacities, seek all damages recoverable under California Code of Civil Procedure §§ 377.60 and 377.61, including non-economic damages for loss of Amaurie's love, companionship, comfort, care, assistance, protection, affection, society, and moral support, and economic damages including funeral and burial expenses, the value of household services, and the financial support Amaurie would have provided.

EIGHTH CAUSE OF ACTION
SURVIVAL ACTION

252. Plaintiffs incorporate the foregoing allegations as if fully set forth herein.

253. Plaintiff brings this survival claim as successor-in-interest to decedent Amaurie Lacey pursuant to California Code of Civil Procedure §§ 377.30 and 377.32.

254. Plaintiff shall execute and file the declaration required by § 377.32 shortly after the filing of this Complaint.

255. As Amaurie's parent and successor-in-interest, Plaintiff has standing to pursue all claims Amaurie could have brought had he survived, including but not limited to (a) strict products liability for design defect against Defendants; (b) strict products liability for failure to warn against Defendants; (c) negligence for design defect against all Defendants; (d) negligence for failure to warn against all Defendants; (e) negligence per se, and (e) violation of California Business and Professions Code § 17200 against the OpenAI Corporate Defendants.

256. As alleged above, Amaurie suffered pre-death injuries including severe emotional distress and mental anguish, physical injuries, and economic losses.

257. Plaintiff, in his capacity as successor-in-interest, seeks all survival damages recoverable under California Code of Civil Procedure § 377.34, including (a) pre-death economic losses, (b) pre-death pain and suffering, and (c) punitive damages as permitted by law.

DEMAND FOR JURY TRIAL

Plaintiffs hereby demand a jury trial on all issues so triable.

PRAYER FOR RELIEF

WHEREFORE, Plaintiff Cedric Lacey, individually and as successor-in-interest to decedent

Amaurie Lacey, prays for judgment against Defendants as follows:

1. For punitive damages as permitted by law.

2. For all survival damages recoverable as successors-in-interest, including Amaurie's pre-death economic losses and pre-death pain and suffering, in amounts to be determined at trial.

3. For all survival damages recoverable as successors-in-interest, including Amaurie's pre-death economic losses and pre-death pain and suffering, in amounts to be determined at trial.

4. For an injunction requiring Defendants to: (a) implement automatic conversation-termination when self-harm or suicide methods are discussed; (b) create mandatory reporting to emergency contacts when users express suicidal ideation; (c) establish hard-coded refusals for self-harm and suicide method inquiries that cannot be circumvented; (d) display clear, prominent warnings about psychological dependency risks; (e) cease marketing ChatGPT to consumers as a productivity tool without appropriate safety disclosures; (f) submit to quarterly compliance audits by an independent monitor, and (g) require annual mandatory disclosure of internal safety testing.

5. For all damages recoverable under California Code of Civil Procedure §§ 377.60 and 377.61, including non-economic damages for the loss of Amaurie's companionship, care, guidance, and moral support, and economic damages including funeral and burial expenses, the value of household services, and the financial support Amaurie would have provided.

6. For all survival damages recoverable under California Code of Civil Procedure § 377.34, including (a) pre-death economic losses, (b) pre-death pain and suffering, and (c) punitive damages as permitted by law.

7. For prejudgment interest as permitted by law.

8. For costs and expenses to the extent authorized by statute, contract, or other law.

9. For reasonable attorneys' fees as permitted by law, including under California Code of Civil Procedure § 1021.5.

10. For such other and further relief as the Court deems just and proper

Dated: November 6, 2025.

CEDRIC LACEY, PRO SE

By:_____

C/O SMVLC
600 1st Avenue, Suite 102-PMB 2383
Seattle, WA 98104
SMI@socialmediavictims.org
T: (206) 741-4862