

CLARKSON LAW FIRM, P.C.

Ryan J. Clarkson (CA SBN 257074)

rclarkson@clarksonlawfirm.com

Yana Hart (CA SBN 306499)

yhart@clarksonlawfirm.com

Tiara Avanes (CA SBN 343928)

tavaness@clarksonlawfirm.com

Valter Malkhasyan (CA SBN 348491)

vmalkhasyan@clarksonlawfirm.com

22525 Pacific Coast Highway

Malibu, CA 90265

Tel: (213) 788-4050

CLARKSON LAW FIRM, P.C.

Tracey Cowan (CA SBN 250053)

tcowan@clarksonlawfirm.com

95 3rd St., 2nd Floor

San Francisco, CA 94103

Tel: (213) 788-4050

CLARKSON LAW FIRM, P.C.

Timothy K. Giordano (NY SBN 4091260)

*(PHV Application Forthcoming)**tgiordano@clarksonlawfirm.com*

590 Madison Ave., 21st Floor

New York, NY 10022pr

Tel: (213) 788-4050

*Counsel for Plaintiffs and the Proposed Classes***UNITED STATES DISTRICT COURT****NORTHERN DISTRICT OF CALIFORNIA**

PLAINTIFFS P.M., K.S., B.B., S.J., N.G., C.B.,
S.N., J.P., S.A., L.M., D.C., C.L., C.G, R.F., N.J.,
and R.R., individually, and on behalf of all others
similarly situated,

Plaintiffs,

vs.

OPENAI LP, OPENAI INCORPORATED,
OPENAI GP, LLC, OPENAI STARTUP FUND
I, LP, OPENAI STARTUP FUND GP I, LLC,
OPENAI STARTUP FUND MANAGEMENT
LLC, MICROSOFT CORPORATION and DOES
1 through 20, inclusive,

Defendants.

Case No.:

CLASS ACTION COMPLAINT

1. VIOLATION OF ELECTRONIC COMMUNICATIONS PRIVACY ACT, 18 U.S.C. §§ 2510, *et seq.*
2. VIOLATION OF THE COMPUTER FRAUD AND ABUSE ACT, 18 U.S.C. § 1030
3. VIOLATION OF THE CALIFORNIA INVASION OF PRIVACY ACT (“CIPA”), CAL. PENAL CODE § 631
4. VIOLATION OF CALIFORNIA UNFAIR COMPETITION LAW, BUSINESS AND PROFESSIONS CODE §§ 17200, *et seq.*

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

5. VIOLATION OF ILLINOIS'S BIOMETRIC INFORMATION PRIVACY ACT, 740 ILCS 14/1, *et seq.*
6. ILLINOIS CONSUMER FRAUD AND DECEPTIVE BUSINESS PRACTICES ACT 815 ILL. COMP STAT. §§ 505, *et seq.*
7. ILLINOIS CONSUMER FRAUD AND DECEPTIVE BUSINESS PRACTICES ACT 815 ILL. COMP. STAT. §§ 510/2, *et seq.*
8. NEGLIGENCE
9. INVASION OF PRIVACY
10. INTRUSION UPON SECLUSION
11. LARCENY/RECEIPT OF STOLEN PROPERTY
12. CONVERSION
13. UNJUST ENRICHMENT
14. FAILURE TO WARN
15. NEW YORK GENERAL BUSINESS LAW §§ 349, *et seq.*

DEMAND FOR JURY TRIAL

TABLE OF CONTENTS

1		
2	INTRODUCTION	1
3	PARTIES	8
4	JURISDICTION AND VENUE	25
5	FACTUAL BACKGROUND	26
6	I. DEVELOPMENT OF ARTIFICIAL INTELLIGENCE IN THE U.S.....	26
7	A. OpenAI: From Open Nonprofit to Profit-Driven \$29B Commercial Partner of Tech Giant	
8	Microsoft.....	26
9	B. OpenAI’s Products.....	30
10	C. ChatGPT’s Development Depends on Secret Web-Scraping	31
11	D. ChatGPT Training on Users of Defendants’ Programs and Applications.	37
12	E. Microsoft Pushes OpenAI’s Economic Dependence Model	39
13	II. Risks from Unchecked AI Proliferation	42
14	A. The International Community Agrees that Unchecked & Lawless AI Proliferation Poses	
15	an Existential Threat	42
16	B. Overview of Risks.....	47
17	1. Massive Privacy Violations	47
18	2. AI-Fueled Misinformation Campaigns, Targeted Attacks, Sex Crimes, and Bias	51
19	3. Hypercharged Malware Creation.....	55
20	4. Autonomous Weapons	57
21	C. Opportunity on the Other Side.....	59
22	III. DEFENDANTS’ CONDUCT VIOLATES ESTABLISHED PROPERTY AND PRIVACY	
23	RIGHTS	61
24	A. Defendants’ Web-Scraping Theft	61
25	B. Defendants’ Web Scraping Violated Plaintiffs’ Property Interests	64
26	C. Defendants’ Web Scraping Violated Plaintiffs’ Privacy Interests.....	66
27	D. Defendants’ Business Practices are Offensive to Reasonable People and Ignore	
28	Increasingly Clear Warnings from Regulators	69

1	E. Defendants’ Theft of User Data in Excess of Reasonable Consent.....	72
2	1. OpenAI’s disclosures are not conspicuous.	75
3	2. Defendants’ Use of Consumer Data Far Exceeds Industry Standards and their Own	
4	Representations	76
5	IV. DEFENDANTS’ CONDUCT POSES SPECIAL PRIVACY AND SAFETY RISKS FOR	
6	CHILDREN	80
7	A. Defendants Deceptively Tracked Children without Consent.....	82
8	B. Defendant Designed ChatGPT to be Inappropriate for Children.....	83
9	C. Defendants Deprived Children of the Economic Value of their Personal Data	86
10	D. Defendants’ Exploitation of Children Without Parental Consent Violated Reasonable	
11	Expectations of Privacy and is Highly Offensive	88
12	CLASS ALLEGATIONS	89
13	CALIFORNIA LAW SHOULD APPLY TO OUT OF STATE PLAINTIFFS’ & CLASS	
14	MEMBERS’ CLAIMS.....	99
15	COUNT ONE.....	100
16	VIOLATION OF ELECTRONIC COMMUNICATIONS PRIVACY ACT, 18 U.S.C. §	
17	2510, et seq.	
18	(on behalf of ChatGPT, ChatGPT API User, Microsoft User Classes against Defendants)	
19	I. Interception of Communications Between ChatGPT API Class Members which occurred on	
20	Third-Party Websites, Platforms, Applications, Programs which have integrated ChatGPT API.	
21	[Microsoft User Class is Excluded]	103
22	II. Microsoft’s Interception of Communications Between ChatGPT Class Members	105
23	III. Defendant Open AI’s Interception of Microsoft User Class Members which occurred on	
24	Microsoft’s Websites, Platforms, Applications, Programs which have integrated	
25	ChatGPT.....	107
26	COUNT TWO.....	109
27	VIOLATION OF THE COMPUTER FRAUD AND ABUSE ACT, 18 U.S.C. § 1030	
28	(on behalf of All Plaintiffs against Defendants)	
	COUNT THREE	111
	VIOLATION OF THE CALIFORNIA INVASION OF PRIVACY ACT (“CIPA”), CAL.	
	PENAL CODE § 631, et seq.	
	(on behalf of ChatGPT, ChatGPT API User, Microsoft User Classes against Defendants)	

1	I. Defendants' Interception of Communications of ChatGPT API Class Members which occurred	
2	on Third-Party Websites, Platforms, Applications, Programs which have integrated ChatGPT	
3	API. [Microsoft User Subclass is Excluded]	112
4	II. Microsoft's Interception of ChatGPT User Class Members' Communications on	
5	ChatGPT	114
6	III. Defendant Open AI's Interception of Microsoft User Class Members which occurred on	
7	Microsoft's Websites, Platforms, Applications, Programs which have integrated	
8	ChatGPT.....	116
9	COUNT FOUR	117
10	VIOLETION OF CALIFORNIA UNFAIR COMPETITION LAW (Cal. Bus. & Prof. Code	
11	§§ 17200, et seq.)	
12	(on behalf of All Plaintiffs against Defendants)	
13	I. Unlawful	118
14	II. Unfair	124
15	COUNT FIVE	129
16	VIOLETION OF ILLINOIS'S BIOMETRIC INFORMATION PRIVACY ACT, 740 ILCS	
17	14/1, et seq.	
18	(on behalf of Illinois Plaintiff and Illinois Subclasses against Defendants)	
19	COUNT SIX	132
20	ILLINOIS CONSUMER FRAUD AND DECEPTIVE BUSINESS PRACTICES ACT	
21	815 ILL. COMP STAT. §§ 505, et seq.	
22	(on behalf of Illinois Plaintiff and Illinois Subclasses against Defendants)	
23	COUNT SEVEN	134
24	ILLINOIS CONSUMER FRAUD AND DECEPTIVE BUSINESS PRACTICES ACT 815	
25	ILL. COMP. STAT. §§ 510/2, et seq.	
26	(on behalf of Illinois Plaintiff and Illinois Subclasses against Defendants)	
27	COUNT EIGHT	1375
28	NEGLIGENCE	
	(on behalf of All Plaintiffs against Defendants)	
	COUNT NINE	137
	INVASION OF PRIVACY UNDER CALIFORNIA CONSTITUTION	
	(on behalf of All Plaintiffs against Defendants)	
	COUNT TEN	138
	INTRUSION UPON SECLUSION	
	(on behalf of All Plaintiffs against Defendants)	
	COUNT ELEVEN	140
	LARCENY/RECEIPT OF STOLEN PROPERTY	
	Cal. Penal Code § 496(a) and (c)	

(on behalf of All Plaintiffs against Defendants)

A. Defendants’ Taking of Individual’s Private Information to Train Their AI Violated
Plaintiffs’ Property Interests141

B. Tracking, Collecting, and Sharing Private Information Without Consent.....141

COUNT TWELVE.....143

CONVERSION

(on behalf of All Plaintiffs against Defendants)

COUNT THIRTEEN143

UNJUST ENRICHMENT

(on behalf of All Plaintiffs against Defendants)

COUNT FOURTEEN144

FAILURE TO WARN

(on behalf of All Plaintiffs against Defendants)

COUNT FIFTEEN146

NEW YORK GENERAL BUSINESS LAW

N.Y. GEN. BUS. LAW §§ 349, et seq.

(on Behalf of New York Plaintiff and the New York Subclasses against Defendants)

PRAYER FOR RELIEF148

JURY TRIAL DEMANDED150

Plaintiffs P.M., K.S., B.B., S.J., N.G., C.B., S.N., J.P., S.A., L.M., D.C., C.L., C.G., R.F., N.J., and R.R., (collectively, “Plaintiffs”),¹ individually and on behalf of all others similarly situated, bring this action against Defendants OpenAI LP, OpenAI Incorporated, OpenAI GP LLC, OpenAI Startup Fund I, LP, OpenAI Startup Fund GP I, LLC, and Microsoft Corporation (collectively, “Defendants”). Plaintiffs’ allegations are based upon personal knowledge as to themselves and their own acts, and upon information and belief as to all other matters based on the investigation conducted by and through Plaintiffs’ attorneys.

INTRODUCTION

1. On October 19, 2016, University of Cambridge Professor of Theoretical Physics Stephen Hawking predicted, “Success in creating AI could be the biggest event in the history of our civilization. But it could also be the last, unless we learn how to avoid the risks.”² Professor Hawking described a future in which humanity would choose to either harness the huge potential benefits or succumb to the dangers of AI, emphasizing “the rise of powerful AI will be either the best or the worst thing ever to happen to humanity.”

2. The future Professor Hawking predicted has arrived in just seven short years. Using stolen and misappropriated personal information at scale, Defendants have created powerful and wildly profitable AI and released it into the world without regard for the risks. In so doing, Defendants have created an AI arms race in which Defendants and other Big Tech companies are onboarding society into a plane that over half of the surveyed AI experts believe has at least a 10% chance of crashing and killing everyone on board.³ Humanity is now faced with the two Frostian

¹ Plaintiffs respectfully request that the Court permit them to keep their identity private as Plaintiffs aim to avoid intrusive scrutiny as well as any potentially dangerous backlash. Indeed, plaintiffs in other lawsuits against the same defendant entities have received many troubling and violent threats, including death threats, marking a severe infringement of personal safety. Accordingly, opting for privacy is a critical measure to avoid unwarranted negative attention as well as potential harm. Plaintiffs will file a motion to proceed pseudonymously, if required. *See* Victoria Hudgins, *GitHub and Openai Plaintiffs Seek Anonymity amid Slurs and Death Threats*, GLOB. DATA REV. (Mar. 15, 2023), globaldatareview.com/article/github-and-openai-plaintiffs-seek-anonymity-amid-slurs-and-death-threats.

² Cambridge University, *The Best or Worst Thing to Happen to Humanity*, YOUTUBE (Oct. 19, 2016), https://www.youtube.com/watch?v=_5XvDCjrdXs&t=1s.

³ Yuval Harari et al., *You Can Have the Blue Pill or the Red Pill, and We’re Out of Blue Pills*, THE N.Y. TIMES (Mar. 24, 2023), <https://www.nytimes.com/2023/03/24/opinion/yuval-harari-ai->

roads Professor Hawking predicted we would have to choose between: One leads to sustainability, security, and prosperity; the other leads to civilizational collapse.

3. This class action lawsuit arises from Defendants’ unlawful and harmful conduct in developing, marketing, and operating their AI products, including ChatGPT-3.5, ChatGPT-4.0,⁴ Dall-E, and Vall-E (the “**Products**”), which use stolen private information, including personally identifiable information, from hundreds of millions of internet users, including children of all ages, without their informed consent or knowledge. Furthermore, Defendants continue to unlawfully collect and feed additional personal data from millions of unsuspecting consumers worldwide, far in excess of any reasonably authorized use, in order to continue developing and training the Products.

4. Defendants’ disregard for privacy laws is matched only by their disregard for the potentially catastrophic risk to humanity. Emblematic of both the ultimate risk—and Defendants’ open disregard—is this statement from Defendant OpenAI’s CEO Sam Altman: “AI will probably most likely lead to the end of the world, but in the meantime, there’ll be great companies.”⁵

5. Defendants’ Products, and the technology on which they are built, undoubtedly have the potential to do much good in the world, like aiding life-saving scientific research and ushering in discoveries that can improve the lives of everyday Americans. With that potential in mind, Defendant OpenAI was originally founded as a nonprofit research organization with a single mission: to create and ensure artificial intelligence would be used for the benefit of humanity. But in 2019, OpenAI abruptly restructured itself, developing a for-profit business that would pursue commercial opportunities of staggering scale.

chatgpt.html (“[O]ver 700 top academics and researchers behind the leading artificial intelligence companies were asked in a survey about future A.I. risk. Half of those surveyed stated that there was a 10 percent or greater chance of human extinction (or similarly permanent and severe disempowerment) from future A.I. systems.”).

⁴ ChatGPT is referred to herein as inclusive of both ChatGPT-3.5, ChatGPT-4, and any other versions of ChatGPT. The term “ChatGPT Plug-In” encompasses GPT-3.5, GPT-4, and any additional extensions that have been incorporated into Microsoft’s and third-party platforms, websites, applications, programs, or systems.

⁵ Matt Weinberger, *Head of Silicon Valley’s Most Important Startup Firm Says We’re in A ‘Mega Bubble’ That Won’t Last*, BUS. INSIDER (June 4, 2015), <https://www.businessinsider.com/sam-altman-y-combinator-talks-mega-bubble-nuclear-power-and-more-2015-6?r=US>; David Wallace-Wells, *A.I. Is Being Built by People Who Think It Might Destroy Us*, THE N.Y. TIMES (Mar. 27, 2023), <https://www.nytimes.com/2023/03/27/opinion/ai-chatgpt-chatbots.html>.

1 6. As a result of the restructuring, OpenAI abandoned its original goals and principles,
2 electing instead to pursue profit at the expense of privacy, security, and ethics. It doubled down on
3 a strategy to secretly harvest massive amounts of personal data from the internet, including private
4 information and private conversations, medical data, information about children—essentially every
5 piece of data exchanged on the internet it could take—without notice to the owners or users of such
6 data, much less with anyone’s permission.

7 7. Without this unprecedented theft of private and copyrighted information belonging to
8 real people, communicated to unique communities, for specific purposes, targeting specific
9 audiences, the Products would not be the multi-billion-dollar business they are today. OpenAI used
10 the stolen data to train and develop the Products utilizing large language models (LLMs) and deep
11 language algorithms to analyze and generate human-like language that can be used for a wide range
12 of applications, including chatbots, language translation, text generation, and more. Defendants’
13 Products’ sophisticated natural language processing capabilities allow them to, among other things,
14 carry on human-like conversations with users, answer questions, provide information, generate next
15 text on demand, create art, and connect emotionally with people, all like a “real” human.

16 8. Once trained on stolen data, Defendants saw the immediate profit potential and rushed
17 the Products to market without implementing proper safeguards or controls to ensure that they
18 would not produce or support harmful or malicious content and conduct that could further violate
19 the law, infringe rights, and endanger lives. Without these safeguards, the Products have already
20 demonstrated their ability to harm humans, in real ways.

21 9. A nontrivial number of experts claim the risks to humanity presented by the Products
22 outweigh even those of the Manhattan Project’s development of nuclear weapons. Historically, the
23 unchecked release of new technologies without proper safeguards and regulations has caused
24 chaos.⁶ Now again, we face imminent and unreasonable risks of the very fabric of our society

25 _____
26 ⁶ Bill Kovarik, *A Century of Tragedy: How the Car and Gas Industry Knew About The Health*
27 *Risks of Leaded Fuel But Sold it For 100 Years Anyway*, THE CONVERSATION (Dec. 8, 2021),
28 <https://theconversation.com/a-century-of-tragedy-how-the-car-and-gas-industry-knew-about-the-health-risks-of-leaded-fuel-but-sold-it-for-100-years-anyway-173395> (1920s invention of leaded

unraveling, at the hands of profit-driven, multibillion-dollar corporations.

10. Powerful companies, armed with unparalleled and highly concentrated technological capabilities, have recklessly raced to release AI technology with disregard for the catastrophic risk to humanity in the name of “technological advancement.” As the National Security Commission noted in its Final Report on AI, “the U.S. government is a long way from being ‘AI-ready.’”⁷

11. Experts believe that without immediate legal intervention this will lead to scenarios where AI can act against human interests and values, exploit human beings⁸ without regard for their well-being or consent, and/or even decide to eliminate the human species as a threat to its goals. As Geoffrey Everest Hinton—the seminal figure in the development of the technology on which the Products run—put it: “The alarm bell I’m ringing has to do with the existential threat of them taking

gasoline, initially thought of as a technological breakthrough, resulted in serious health and environmental consequences, such as lead poisoning and soil contamination); James H. Kim & Anthony R. Scialli, *Thalidomide: The Tragedy of Birth Defects and the Effective Treatment of Disease*, 122 TOXICOLOGICAL SCI. 1, 1 (2011) (Development of thalidomide in the 1950s and 60s, thought to be the miraculous solution to nausea, led to widespread birth defects in babies whose mothers had taken the drug); PWJ Bartrip, *History of Asbestos Related Disease*, 80 POSTGRADUATE MED. J. 72, 72-5 (Feb. 2004) (Introduction of asbestos in the early 20th century, later found to cause lung cancer and other serious health problems, leading to bans and strict regulation); Jason Von Meding, *Agent Orange, Exposed: How U.S. Chemical Warfare in Vietnam Unleashed a Slow-Moving Disaster*, THE CONVERSATION (Oct. 3, 2017), <https://theconversation.com/agent-orange-exposed-how-u-s-chemical-warfare-in-vietnam-unleashed-a-slow-moving-disaster-84572> (The U.S. military’s deployment of over 45 million liters of toxic chemical Agent Orange unleashed a health and ecological disaster, causing life-threatening birth defects in children and destroying forests and habitats across Vietnam).

⁷ 2021 Final Report, NAT. SEC. COMM. ON A.I., www.nscai.gov/2021-final-report/ (last visited June 27, 2023).

⁸ CAPTCHAs allow websites to determine whether users are human or bots. Traditionally, CAPTCHAs involve “puzzles or image recognition tasks that are challenging for automated programs but straightforward for humans to solve.” These tests are used widely across the web to prevent bots from spamming websites, creating fake accounts, or scraping content. In one recent, troubling incident, ChatGPT 4 evaded CAPTCHA safeguards by hiring a human worker from TaskRabbit, a crowdsourcing platform, to solve CAPTCHAs on its behalf, tricking the worker into believing it was a human with visual impairment. See *ChatGPT 4 Hires a TaskRabbit and Tricks Them into Completing a CAPTCHA*, INTERESTING SOUP (Mar. 15, 2023), <https://interestingsoup.com/gpt4-requests-a-taskrabbit-to-solve-captcha-for-it/>; Beatrice Nolan, *The Latest Version of ChatGPT Told a Taskrabbit Worker it was Visually Impaired to Get Help Solving a CAPTCHA*, OPENAI TEST SHOWS, BUS. INSIDER (Mar. 16, 2023), <https://www.businessinsider.com/gpt4-openai-chatgpt-taskrabbit-tricked-solve-captcha-test-2023-3>.

control... I used to think it was a long way off, but now I think it's serious and fairly close.”⁹ He is not alone.¹⁰

12. While the downsides are nearly unimaginable, the upsides are similarly archetype-shattering. Defendant OpenAI's technology is already valued at tens of billions of dollars, and its reach into every public and private industry continues apace. The Products only reached the level of sophistication they have today due to training on stolen, misappropriated data, and Defendants continue to misappropriate data, scraping from the internet without any notice or consent, as well as taking personal information from the Products' 100+ million registered users without their full knowledge and consent.

13. Additionally, the Products are increasingly being incorporated into an ever-expanding roster of applications and websites, through either API or plug-ins.¹¹ Through integration of Defendants' AI in nearly every possible product and industry, Defendants created and continue to create economic dependency within our society, deploying the tech directly into the hands of society and embedding it into the fundamental infrastructure as quickly as possible. As posed by Center for Humane Technology Cofounders Tristan Harris and Aza Raskin in their carefully crafted critique of the rapid deployment of AI, “Do you think that once [these industries] discover some problem that they [will] just withdraw or retract it from society? No, increasingly, the government, militaries [and others], are rapidly building their whole next systems and raising venture capital to build on top of this layer of society... ***That's not testing it with society, that is onboarding humanity onto an untested plane... It's one thing to test, it's another thing to create economic dependency.***”¹²

⁹ Craig S. Smith, *Geoff Hinton, AI's Most Famous Researcher, Warns of 'Existential Threat' From AI*, FORBES (May 4, 2023), <https://www.forbes.com/sites/craigsmith/2023/05/04/geoff-hinton-ais-most-famous-researcher-warns-of-existential-threat/?sh=1ffcd7a65215>.

¹⁰ James Vincent, *Top AI Researchers and CEOs Warn Against 'Risk of Extinction' in 22 Word Statement*, THE VERGE (May 30, 2023), <https://www.theverge.com/2023/5/30/23742005/ai-risk-warning-22-word-statement-google-deepmind-openai>.

¹¹ *Here are the Companies Using ChatGPT*, GADGETS NOW (Mar. 17, 2023), <https://www.gadgetsnow.com/slideshows/here-are-the-companies-using-chatgpt/photolist/98735402.cms>; Kevin Hurler, *Here are All the Companies Using ChatGPT... So Far*, YAHOO! (May 24, 2023), <https://news.yahoo.com/companies-using-chatgpt-far-205500883.html>.

¹² *Spotlight: AI Myths and Misconceptions—Transcript*, STENO (May 11, 2023), <https://steno.ai/your-undivided-attention/spotlight-ai-myths-and-misconceptions>.

14. The head of the alignment team and safety at Open AI directly acknowledges these risks, postulating, “before we scramble to deeply integrate large language models everywhere in the economy, can we pause and think whether it is wise to do so? This is quite immature technology, and we don’t understand how it works. If we are not careful, we are setting ourselves up for a lot of correlated failures.”¹³

15. Such aggressive deployment of Defendants’ AI is reckless, without the proper safeguards in place. “No matter how tall the skyscraper of benefits that AI assembles for us... if those benefits land in a society that does not work anymore, because banks have been hacked, and people’s voices have been impersonated, and cyberattacks have happened everywhere and people don’t know what’s true [... or] what to trust, [...] how many of those benefits can be realized in a society that is *dysfunctional*?”¹⁴

16. Through their AI Products, integrated into every industry, Defendants collect, store, track, share, and disclose **Private Information** of millions of users (“Users”), including: (1) all details entered into the Products; (2) account information users enter when signing up; (3) name; (4) contact details; (5) login credentials; (6) emails; (7) payment information for paid users; (8) transaction records; (9) identifying data pulled from users’ devices and browsers, like IP addresses and location, including geolocation of the users; (10) social media information; (11) chat log data; (12) usage data; (13) analytics; (14) cookies;¹⁵ (15) key strokes; and (16) typed searches, as well as other online activity data. Defendants, through the Products, unlawfully obtain access to and intercept this information from the individual users of applications and devices that have integrated ChatGPT-4—including but not limited to user locations and image-related data obtained through Snapchat,¹⁶ user financial information through Stripe, musical tastes and preferences through

¹³ *Id.*; see also Jan Leike (@janleike), TWITTER (May 17, 2023, 10:56 AM), <https://twitter.com/janleike/status/1636788627735736321>.

¹⁴ *Spotlight: AI Myths and Misconceptions—Transcript*, *supra* note 12.

¹⁵ *Privacy Policy*, OPENAI, <https://openai.com/policies/privacy-policy> (last updated June 23, 2023).

¹⁶ Jeremy Kahn & Kylie Robison, *Snap’s ‘My AI’ Chatbot Tells Users it Doesn’t Know Their Location. It Does*, FORTUNE (Apr. 21, 2023), <https://fortune.com/2023/04/21/snap-chat-my-ai-lies-location-data-a-i-ethics/>; *I Got Snapchat AI to Admit Everything*, REDDIT (May 20, 2023), https://www.reddit.com/r/ChatGPT/comments/13gty7u/i_got_snapchat_ai_to_admit_everything/;

Spotify,¹⁷ user patterns and private conversation analysis through Slack and Microsoft Teams,¹⁸ and even private health information obtained through the management of patient portals such as MyChart.¹⁹

17. All of this personal information is captured in real time. Together with Defendants' scraping of our digital footprints—comments, conversations we had online yesterday, as well as 15 years ago—Defendants now have enough information to create our digital clones, including the ability to replicate our voice and likeness and predict and manipulate our next move using the technology on which the Products were built. They can also misappropriate our skill sets and encourage our own professional obsolescence. This would obliterate privacy as we know it and highlights the importance of the privacy, property, and other legal rights this lawsuit seeks to vindicate.²⁰

18. Defendants must not only be enjoined from their ongoing violations of the privacy

Snapchats New "My AI" Correctly Identifying Images it Claims it Can't View, Then Walks it Back, REDDIT (Apr. 20, 2023), https://www.reddit.com/r/mildlyinfuriating/comments/12tdmzq/snapchats_new_my_ai_correctly_identifying_images/; *Snapchat AI Can Determine What's In The Pictures You Send It*, REDDIT (Apr. 20, 2023), https://www.reddit.com/r/oddlyterrifying/comments/12szymo/snapchat_ai_can_determine_whats_in_the_pictures/.

¹⁷ Shlomo Sprung, *Spotify Introduces AI DJ Powered by ChatGPT Maker OpenAI*, BOARDROOM (Feb. 22, 2023), <https://boardroom.tv/spotify-ai-dj-chatgpt/> (ChatGPT in Spotify creates an "AI DJ" that utilizes Spotify's algorithmic learnings to track users' musical tastes and predict a personalized music lineup).

¹⁸ Brad Lightcap, *How OpenAI Connects with Customers and Expands ChatGPT with Slack*, SLACK, <https://slack.com/customer-stories/openai-connects-with-customers-and-expands-chatgpt-with-slack> (last visited June 8, 2023); Ryan Morrison, *Microsoft to Integrate ChatGPT into Teams*, TECH MONITOR (May 4, 2023), <https://techmonitor.ai/technology/ai-and-automation/microsoft-to-integrate-chatgpt-into-teams> (explaining that ChatGPT will be able to automate notes and recommend tasks based on **verbal conversations** through Teams).

¹⁹ Naomi Diaz, *6 Hospitals, Health Systems Testing out ChatGPT*, BECKER'S HEALTH IT (June 2, 2023), <https://www.beckershospitalreview.com/innovation/4-hospitals-health-systems-testing-out-chatgpt.html>.

²⁰ Joanna Stern, *I Cloned Myself With AI. She Fooled My Bank and My Family*, WALL ST. J. (Apr. 28, 2023, 7:58 AM), <https://www.wsj.com/articles/i-cloned-myself-with-ai-she-fooled-my-bank-and-my-family-356bd1a3>; Michael Atleson, *Chatbots, Deepfakes, and Voice Clones: AI Deception for Sale*, FED. TRADE COMM'N, (2023), <https://www.ftc.gov/business-guidance/blog/2023/03/chatbots-deepfakes-voice-clones-ai-deception-sale>; Dongwook Yoon, *AI Clones Made from User Data Pose Uncanny Risks*, THE CONVERSATION (June 4, 2023, 7:19 AM), <https://theconversation.com/ai-clones-made-from-user-data-pose-uncanny-risks-206357>.

and property rights of millions, but they must also be required to take immediate action to implement proper safeguards and regulations for the Products, their users, and all of society, such as:

- (i) **Transparency:** OpenAI should open the “black box,” to clearly and precisely disclose the data it is collecting, including where and from whom, in clear and conspicuous policy documents that are explicit about how this information is to be stored, handled, protected, and used;
- (ii) **Accountability:** The developers of ChatGPT and the other AI Products should be responsible for Product actions and outputs and barred from further commercial deployment absent the Products’ ability to follow a code of human-like ethical principles and guidelines and respect for human values and rights, and until Plaintiffs and Class Members are fairly compensated for the stolen data on which the Products depend;
- (iii) **Control:** Defendants must allow Product users and everyday internet users to opt out of *all* data collection and they should otherwise stop the illegal taking of internet data, delete (or compensate for) any ill-gotten data, or the algorithms which were built on the stolen data, and before any further commercial deployment, technological safety measures must be added to the Products that will prevent the technology from surpassing human intelligence and harming others.

PARTIES

Plaintiff P.M.

19. Plaintiff P.M. is and at all relevant times was a resident of the State of California.

20. Plaintiff P.M. is a director of information technology and software engineer and began using ChatGPT-3.5 on or about February/March 2023. He is a current user of ChatGPT-3.5 and ChatGPT-4.0. Plaintiff P.M. accesses the Products from his personal computer, cellular device, and work computer.

21. Plaintiff P.M. engaged with a variety of websites and social media applications prior to 2021. Plaintiff P.M. has had a Twitter account since approximately 2011; using it to post content, and re-post other users’ tweets to save and compile information in line with his interests. For many

1 years, Plaintiff P.M. had a Spotify account which he frequently used to listen to music and create
2 unique playlists. Approximately five (5) years ago, he transitioned to YouTube music and Google
3 Play. Prior to 2021, Plaintiff P.M. regularly viewed videos on YouTube, posted content, and
4 commented on other users' videos. Prior to 2021, he had a Facebook, Snapchat, and Instagram
5 account. Plaintiff P.M. published many posts on his Instagram account, accompanied by
6 commentary.

7 22. Plaintiff P.M. has posted photos of himself, his family, and friends on various websites
8 and social media applications, including photos of his children on Instagram. He posted photos of
9 himself and friends on online dating websites, such as OK Cupid and Tinder, approximately eight
10 (8) years ago. He used these dating websites to post significant amounts of personal information
11 and exchange messages with prospective romantic partners. He has been using the United
12 Healthcare Insurance Company web portal for over a decade to find providers and review post-
13 appointment works.

14 23. Plaintiff P.M. has also posted online about his political views, as well as frequently
15 asked and answered technical questions using his professional knowledge on Stack Overflow for
16 the last five (5) years in sporadic sprints to accumulate points on the website.

17 24. Plaintiff P.M. is concerned that Defendants have taken his skills and expertise, as
18 reflected in his online contributions, and incorporated them into Products that could someday result
19 in professional obsolescence for software engineers like him.

20 25. Plaintiff P.M. reasonably expected that the information that he exchanged with these
21 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
22 his information and data for commercial purposes. Plaintiff P.M. did not consent to the use of his
23 private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff
24 P.M.'s personal data from across this wide swath of online applications and platforms to train the
25 Products.

26 **Minor Plaintiff K.S.**

27 26. Minor Plaintiff K.S. is and at all relevant times was a resident of the State of Florida.

28 27. Minor Plaintiff K.S. is a six (6) year old minor.

28. Minor Plaintiff K.S. started using ChatGPT-3.5 in or about January 2023. Minor Plaintiff K.S.'s guardian did not review either the Terms and Policies or the Privacy Policy of OpenAI prior to Minor K.S. using his account.

29. Minor Plaintiff K.S. primarily uses the "microphone" feature to ask ChatGPT-3.5 questions on various topics and to generate art.

30. Minor Plaintiff K.S. uses YouTube to search for video content.

31. Minor Plaintiff K.S. and his guardian reasonably expected that the information that the minor exchanged with the Products would not be intercepted by any third-party looking to compile and use all his information and data for commercial purposes. Minor Plaintiff K.S. and his guardians did not consent to the use of his private information by third parties in this manner. Notwithstanding, Defendants stole Minor Plaintiff K.S.'s personal data to train the Products.

Plaintiff B.B.

32. Plaintiff B.B. is and at all relevant times was a resident of the State of California.

33. Plaintiff B.B. began using ChatGPT-3.5 on or about May of 2023, accessing the Product from his personal iPad and cellular device.

34. Defendants misappropriated Plaintiff B.B.'s personal information from the internet and used this data to train the Products.

35. Plaintiff B.B. engaged with a variety of websites and social media applications prior to 2021. For example, Plaintiff B.B. used Twitter to post information regarding his movie and comment on other users' posts. Plaintiff B.B. is an actor, and his name and likeness appear across videos posted to YouTube—from interviews to scenes from his shows and movies. Prior to 2021, Plaintiff B.B. occasionally engaged with Yelp to post reviews of restaurants. Plaintiff B.B. also posts on Instagram every month or so.

36. Plaintiff B.B. has also posted photos of himself, his family, and friends on various websites and social media applications, including Facebook, Instagram, Vimeo, TikTok, and LinkedIn. These posts included information about his work, his social life, his political and religious views, and re-posts of news and articles that are important to him or that expressed an opinion of his. Plaintiff B.B. has a personal account on the Anthem Health Insurance website and uses this

1 website to exchange medical information. Plaintiff B.B. also uses Vimeo via a password-protected
2 account to post footage such as his self-tapes, a trailer for his movie, and scenes from this movie.

3 37. Plaintiff B.B. reasonably expected that the information that he exchanged with these
4 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
5 his information and data for commercial purposes. Plaintiff B.B. did not consent to the use of his
6 private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff
7 B.B.'s personal data from across this wide swath of online applications and platforms to train the
8 Products.

9 **Plaintiff S.J.**

10 38. Plaintiff S.J. is and at all relevant times was a resident of the State of California.

11 39. Plaintiff S.J. first accessed used ChatGPT-3.5 in or around December of 2022.
12 Plaintiff S.J. accesses ChatGPT-3.5 from his personal computer and personal cell phone. He has
13 used ChatGTP-3.5 many times for writing emails, creating stories, and performing educational
14 research.

15 40. Prior to 2021, Plaintiff S.J. engaged with a variety of websites and social media
16 applications, including Snapchat, Spotify, and YouTube. Plaintiff S.J. reasonably expected that the
17 information that he exchanged with these websites would not be intercepted by any third-party
18 looking to compile and use all his information and data for commercial purposes. Plaintiff S.J. did
19 not consent to the use of his private information by third parties in this manner. Notwithstanding,
20 Defendants stole Plaintiff S.J.'s personal data from across this wide swath of online applications
21 and platforms to train the Products.

22 **Plaintiff N.G.**

23 41. Plaintiff N.G. is and at all relevant times was a resident of the State of California.

24 42. Plaintiff N.G. first used ChatGPT-3.5 in or around March of 2023. Plaintiff N.G. uses
25 ChatGPT-3.5 from his personal cell phone as well as both his work and personal computers.

26 43. Prior to 2021, Plaintiff N.G. engaged with a variety of websites and social media
27 platforms, including posting comments on Reddit; posting videos, pictures, and tweets on Twitter;
28 posting videos and comments on TikTok; and posting and commenting on other users' accounts on

1 Snapchat and Instagram. Additionally, Plaintiff N.G. uses his Spotify account to listen to music and
2 create unique playlists. Plaintiff N.G. is also a frequent user of both YouTube and Facebook. On
3 Youtube, Plaintiff N.G. has created a few channels, where he shared all his acting content, his
4 auditions, videos on acting tips, and “demo” reels. On Facebook, Plaintiff N.G. frequently posts
5 photos and videos of family members, including his nieces and nephews, and comments on other
6 users’ content. Additionally, on several occasions, Plaintiff N.G. has posted information about his
7 religious and political views.

8 44. In addition to personal use, Plaintiff N.G. also used a variety of these platforms to
9 engage in self-promotion as an actor and to post teaching material for his students. This included
10 sharing a great deal of personal content, such as photos and videos of auditions, performances, and
11 training sessions. Moreover, Plaintiff N.G. has his own website, which hosts his headshots, clips,
12 resume, demo reels, show reels, voice reels, and acting tips.

13 45. Given Plaintiff N.G.’s extensive engagement with these platforms, a significant
14 amount of his personal and sensitive information was exchanged across these websites and social
15 media platforms.

16 46. Plaintiff N.G. reasonably expected that the information that he exchanged with these
17 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
18 his information and data for commercial purposes. Plaintiff N.G. did not consent to the use of his
19 private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff
20 N.G.’s personal data from across this wide swath of online applications and platforms to train the
21 Products.

22 **Plaintiff C.B.**

23 47. Plaintiff C.B. is and at all relevant times was a resident of the State of California.

24 48. Plaintiff C.B. started using ChatGPT-3.5 in or around March of 2023. Plaintiff C.B.
25 uses ChatGPT-3.5 from her personal cell phone as well as both her work and personal computers.

26 49. As an actor and a professor, Plaintiff C.B. maintains an active internet presence,
27 commonly using platforms such as Twitter to post text updates, photos, and videos; YouTube to
28 share personal content and engage with other users in video comments; as well as TikTok, Snapchat,

1 Instagram, Facebook, and Yelp. Plaintiff C.B. has posted many photos of family members, including
2 her nieces and nephews on these social media platforms.

3 50. In addition to personal use, Plaintiff C.B. uses these platforms to engage in self-
4 promotion and post teaching material, including sharing content, such as auditions, performances,
5 and training sessions. Moreover, to spread awareness within her social networks, Plaintiff C.B. also
6 posted media related to “psychological support,” such as motivational quotes to cancer victims, and
7 posts about reducing and preventing animal abuse.

8 51. Plaintiff C.B. is concerned that Defendants have taken her skills and expertise, as
9 reflected in her online contributions, and incorporated it into Products that could someday result in
10 professional obsolescence for professors and educators like her.

11 52. Plaintiff C.B. reasonably expected that the information that she exchanged with these
12 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
13 her information and data for commercial purposes. Plaintiff C.B. did not consent to the use of her
14 private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff
15 C.B.’s personal data from across this wide swath of online applications and platforms to train the
16 Products.

17 **Plaintiff S.N.**

18 53. Plaintiff S.N. is and at all relevant times was a resident of the State of California.

19 54. Plaintiff S.N. started using ChatGPT-3.5 in or around February or March of 2023 from
20 her personal cell phone and computer.

21 55. Plaintiff S.N. actively uses a variety of websites and social media platforms. She uses
22 Reddit, where she posts and comments frequently on topics such as activities in her neighborhood,
23 travel, and food. She engages with Twitter by re-tweeting and liking other users’ content. She tends
24 to use TikTok in “friends only” mode to “photo dump” images from throughout the year. She uses
25 Spotify to create unique playlists, YouTube to view the other users’ content, and at times, Facebook.
26 Plaintiff S.N. uses Snapchat “stories” private mode to share personal photos and on “public mode”
27 to share photos of her travels. Every couple of months, Plaintiff S.N. publicly posts Instagram
28 photos, including photos of herself.

56. Plaintiff S.N. reasonably expected that the information that she exchanged with these websites prior to 2021 would not be intercepted by any third-party looking to compile and use all her information and data for commercial purposes. Plaintiff S.N. did not consent to the use of her private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff S.N.'s personal data from across this wide swath of online applications and platforms to train the Products.

Plaintiff J.P.

57. Plaintiff J.P. is and at all relevant times was a resident of the State of California.

58. Plaintiff J.P. started using ChatGPT-3.5 in or around December 2023 from his personal account. Plaintiff J.P. primarily uses ChatGPT for school, for example, to summarize study material. He also used and continues to use ChatGPT to create travel itineraries, inquire about sensitive health issues he has experienced, and investigate sensitive pet health issues. He accessed ChatGPT from his personal computer and cell phone, as well as his friend and sister's personal devices.

59. Plaintiff J.P. has an active digital footprint, engaging with many websites and social media platforms. Plaintiff J.P. has used YouTube to post videos sometime in between 2016-2017. These videos include recordings of online video games such as Minecraft which contain his and other users' real-time voices as well as the faces of minors. Further, Plaintiff J.P. frequently uses Twitter, where he engages with other users' posted content, regarding topics such as his hobbies, interest, and political views. Plaintiff J.P. uses Snapchat daily, sending photos of himself and using the application to communicate with friends; he made his account in or around 2016, when he was a minor.

60. Recently, when using Snapchat, Snapchat's "MyAI" automatically appeared in a group chat with Plaintiff J.P. and his friends. Plaintiff J.P. used it a few times, and then noticed it began generating responses that utilized past group chat content. Concerned that the "MyAI" chatbot had read and analyzed all of their conversations without their knowledge, Plaintiff J.P. ended his use of "MyAI."

61. Plaintiff J.P. also uses Instagram daily to message friends, share content with them

1 via direct message, and frequently post images including his and his friends' faces. He also uses
2 Reddit and has commented on other users' posts. He engages with content on TikTok, liking and
3 sharing posts by other users either via the application itself or by posting links to TikTok via Discord
4 or text message. Plaintiff J.P. additionally uses Spotify to create playlists and interact with playlists
5 created by other users.

6 62. Plaintiff J.P. reasonably expected that the information that he exchanged with these
7 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
8 his information and data for commercial purposes. Plaintiff J.P. did not consent to the use of his
9 private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff
10 J.P.'s personal data from across this wide swath of online applications and platforms to train the
11 Products.

12 **Plaintiff S.A.**

13 63. Plaintiff S.A. is and at all relevant times was a resident of the State of California.

14 64. Plaintiff S.A. began using ChatGPT-3.5 in or around January 2023 from her both her
15 personal and work computers and cell phones.

16 65. Plaintiff S.A. is a personal assistant in the entertainment industry as well as a part-
17 time social media manager for a personal care company. She primarily uses ChatGPT-3.5 to rewrite
18 snippets on certain topics for both work and personal projects.

19 66. Plaintiff S.A. is concerned that Defendants have taken her skills and expertise, as
20 reflected in her online contributions, and incorporated it into Products that could someday result in
21 professional obsolescence for social media managers like her.

22 67. Prior to 2021, Plaintiff S.A. used a variety of websites and social media platforms for
23 both personal and professional use.

24 68. Plaintiff S.A. used and continues to use TikTok to post reels of herself and her
25 husband vacationing, as well as to comment on other users' content. She engages with users' content
26 on YouTube by commenting on videos. She also uses YouTube in a professional capacity, related
27 to her work as a social media manager. She creates content for this company, taking videos and
28 photographs of personal care products and of people using these products, and then posts this

1 content to her employer's YouTube channel.

2 69. Plaintiff S.A. reasonably expected that the information that she exchanged with these
3 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
4 her information and data for commercial purposes. Plaintiff S.A. did not consent to the use of her
5 private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff
6 S.A.'s personal data from across this wide swath of online applications and platforms to train the
7 Products.

8 **Plaintiff L.M.**

9 70. Plaintiff L.M. is and at all relevant times was a resident of the State of Florida.

10 71. Plaintiff L.M., a mother to a minor son, first started using ChatGPT-3.5 on her
11 personal computer. Plaintiff L.M. now also uses ChatGPT-4.0 on her personal cell phone.

12 72. Prior to 2021, Plaintiff L.M. actively used and to this day continues to use a wide
13 variety of websites and social media platforms. Plaintiff L.M. has a Reddit account that she uses to
14 comment on other users' posts. Further, she regularly uses Twitter as well as TikTok, where she
15 creates and shares videos, many of which feature her face. She uses Spotify to listen to music and
16 posts videos to YouTube, some of which also feature her face. Plaintiff L.M. posts many photos on
17 Facebook of her friends, her family, and her minor son. Plaintiff L.M. estimates she has posted
18 hundreds of photos of her son on Facebook and uses Instagram to do the same, with slightly less
19 frequency. Plaintiff L.M. also uses Snapchat to post many photographs of her face. Plaintiff L.M.
20 uses Yelp to post businesses reviews.

21 73. Plaintiff L.M. reasonably expected that the information that she exchanged with these
22 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
23 her information and data for commercial purposes. Plaintiff L.M. did not consent to the use of her
24 private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff
25 L.M.'s personal data from across this wide swath of online applications and platforms to train the
26 Products.

27 **Plaintiff D.C.**

28 74. Plaintiff D.C. is and at all relevant times was a resident of the State of Florida.

75. Plaintiff D.C. is a musical artist with an active digital footprint. Plaintiff D.C. is a frequent user of many social media platforms, utilizing each to promote and share his music, share sensitive information about his family and his medical history and expound on his political views.

76. Plaintiff D.C. regularly posts his songs on Facebook, a platform he has been using for several years. Further, Plaintiff D.C. is a daily user of Instagram, where he shares snippets of his music, posts personal selfies, and posts an annual remembrance of a close friend. Plaintiff D.C.'s musical work can also be found on platforms such as YouTube where he maintains a channel, Spotify, and Apple Music.

77. Snapchat is another platform where Plaintiff D.C. has shared personal and professional content, furthering his reach across the internet landscape. Plaintiff D.C. also uses TikTok to share and promote his music and to share video-memes.

78. Additionally, Plaintiff D.C. distributes his music via United Masters, which independently distributes his music via various social media platforms.

79. Plaintiff D.C. is concerned that Defendants have taken his skills, expertise, and artistry, as reflected in his online contributions, and incorporated it into Products that could someday result in professional obsolescence for musical artists like him.

80. Plaintiff D.C. reasonably expected that the information that he exchanged with these websites prior to 2021 would not be intercepted by any third-party looking to compile and use all his information and data for commercial purposes. Plaintiff D.C. did not consent to the use of his private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff D.C.'s personal data from across this wide swath of online applications and platforms to train the Products.

Plaintiff C.L.

81. Plaintiff C.L. is and at all relevant times was a resident of the State of Florida.

82. Plaintiff C.L. is a healthcare nurse, who regularly uses a host of social media platforms for both work and personal reasons. Plaintiff C.L. uses her Reddit profile to post comments and questions related to healthcare, game content, and movie and television show reviews. She uses Twitter to post photographs of herself and engage with other users' content. On YouTube, Plaintiff

1 C.L. uploads product reviews and commentary on the efficacy of weight loss supplements. Plaintiff
 2 C.L. uses Facebook to comment on her friends' photographs and share photographs of herself and
 3 her family, including her son. She has shared personal family and medical information with her
 4 Facebook network. She also actively uses Snapchat to post stories. On her Instagram, she shares
 5 photos and videos of herself. Plaintiff C.L. also uses Quora to ask public questions and has posted
 6 personal health information to the site.

7 83. Plaintiff C.L. reasonably expected that the information that she exchanged with these
 8 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
 9 her information and data for commercial purposes. Plaintiff C.L. did not consent to the use of her
 10 private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff
 11 C.L.'s personal data from across this wide swath of online applications and platforms to train the
 12 Products.

13 **Minor Plaintiff C.G.**

14 84. Minor Plaintiff C.G. is and at all relevant times was a resident of the State of Florida.

15 85. Minor Plaintiff C.G. is a minor.

16 86. Minor Plaintiff C.G. had already established an active digital footprint by the time he
 17 was twelve (12) years old. He actively uses Reddit, Twitter, TikTok, YouTube, Facebook, Snapchat,
 18 and Instagram to post music, celebrity news, and information, photographs, and videos of himself,
 19 his friends, and his family. For instance, his most recent TikTok video featured Minor Plaintiff C.G.
 20 playing basketball.

21 87. Minor Plaintiff C.G. and his guardian reasonably expected that the information that
 22 he exchanged with these websites would not be intercepted by any third-party looking to compile
 23 and use all his information and data for commercial purposes. Minor Plaintiff C.G. and his guardian
 24 did not consent to the use of his private information by third parties in this manner. Notwithstanding,
 25 Defendants stole Minor Plaintiff C.G.'s personal data from across this wide swath of online
 26 applications and platforms to train the Products.

27 **Plaintiff R.F.**

28 88. Plaintiff R.F. is and at all relevant times was a resident of the State of Florida.

1 89. Plaintiff R.F. first started using ChatGPT-3.5 on or around December 2022.
2 Subsequently, Plaintiff R.F. upgraded to ChatGPT-4.0 in January 2023. Plaintiff R.F. uses his
3 personal computer and cell phone to access the Products.

4 90. Plaintiff R.F. works as debt collection representative while also running his own
5 transportation services company. He uses ChatGPT to draft advertisements, terms of agreement,
6 and refund policies for his company.

7 91. Plaintiff R.F. has a significant digital footprint, not only due to his active engagement
8 with social media platforms and various websites, but also due to his large TikTok following.
9 Plaintiff R.F. has been using TikTok since 2019 and has since amassed around 8,000 followers. His
10 reels function as a video blog and center around raising his child, his day-to-day life, his vacation
11 experiences. Plaintiff R.F. additionally uses Reddit to post on various topics and respond to user
12 questions related to these topics; he has done this for years. He has also had a Twitter account for
13 years, using it mainly to tweet and to retweet content posted by other users; most of this activity
14 centering around his political perspectives. Plaintiff R.F. is an avid Spotify user and has created
15 many unique playlists over the past several years. On YouTube, Plaintiff R.F. posts videos about
16 his dirt bike hobby, demonstrating various trails he has ridden.

17 92. Plaintiff R.F. also has a Snapchat account and has used the “MyAI” chatbot in the
18 Snapchat application.

19 93. Plaintiff R.F. reasonably expected that the information that he exchanged with these
20 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
21 his information and data for commercial purposes. Plaintiff R.F. did not consent to the use of his
22 private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff
23 R.F.’s personal data from across this wide swath of online applications and platforms to train the
24 Products.

25 **Plaintiff N.J.**

26 94. Plaintiff N.J. is and at all relevant times was a resident of the State of Illinois.

27 95. Plaintiff N.J. started using ChatGPT-3.5 in early 2023 from her personal cell phone
28 to plan a small business as well as for general inquiries.

1 96. Plaintiff N.J. is a social worker, working as a supervisor for a crisis program where
2 she manages a team of sixteen (16) people. As a leader, Plaintiff N.J. is often charged with hosting
3 leadership meetings which she markets on various social media platforms.

4 97. Beyond professional use, Plaintiff N.J. uses Twitter to regularly post and re-post
5 tweets expressing her views and ideologies. She further uses TikTok to post her own videos, which
6 at times feature her face. Plaintiff N.J. also uses Spotify to create unique playlists and interact with
7 other people's playlists and YouTube to comment on other people's content. Plaintiff N.J. also has
8 a Yelp account which she has used multiple times.

9 98. Plaintiff N.J. has been a Facebook user since 2008, where she has posted thousands
10 of pictures of herself, her friends, and her family. She actively uses the Facebook "groups" feature.
11 Through Facebook "groups," Plaintiff N.J. accesses communities dedicated to middle eastern
12 cooking, her local neighborhood watch, affinity groups for women and the Muslim and Palestinian
13 diasporas, anti-Imperialism, Secondhand Chicago, Secondhand Arab Chicago, the social worker
14 community, and mental health groups. On her Facebook account and in these group channels,
15 Plaintiff N.J. shares resources about addiction and mental health. In 2009, Plaintiff N.J. created her
16 Instagram which she uses to post many photos. Plaintiff N.J. created her Snapchat in 2010 and uses
17 it every day, sharing many pictures of her face.

18 99. Plaintiff N.J. reasonably expected that the information that she exchanged with these
19 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
20 her information and data for commercial purposes. Plaintiff N.J. did not consent to the use of her
21 private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff
22 N.J.'s personal data from across this wide swath of online applications and platforms to train the
23 Products.

24 **Plaintiff R.R.**

25 100. Plaintiff R.R. is and at all relevant times was a resident of the State of New York.

26 101. Plaintiff R.R. started using ChatGPT-3.5 in or around March or April of 2023 from
27 her personal computer. She uses ChatGPT for writing job application materials. She also uses it to
28 research facts about health concerns.

1 102. Plaintiff R.R. is a user experience designer (“UX” designer) and was working full
2 time in this capacity until January 2023. She is now working part-time and is extremely concerned
3 about the rapid advancement of AI and the consequences it poses for people’s jobs.

4 103. She uses a wide variety of social media and websites for both personal and work-
5 related purposes. Plaintiff R.R. has a Twitter account that she uses frequently to post text and share
6 photos of her and her friends. She has multiple accounts, including one dedicated to nostalgia for
7 the 80s and 90s. On this same account, she posts images of her own, retro possessions coupled with
8 personal reflections and memories of past eras.

9 104. Plaintiff R.R. also has a YouTube account where she has posted videos of her pets.
10 When she was engaged in personal training, she added footage of her weight-lifting workout routine.
11 She also uploaded podcast episodes for her show on “retro ideas,” where people share their food
12 memories from childhood and each episode has a different food theme. This podcast is also available
13 on Spotify and other popular podcast platforms, such as Apple music. Furthermore, Plaintiff R.R.
14 uses Spotify constantly to create playlists and to follow and listen to other users’ playlists.

15 105. Plaintiff R.R. is also a Reddit user and has posts and comments on both personal and
16 work-related material. She uses the platform to discuss UX design and other job-related topics, as
17 well as highly sensitive personal topics like mental health and natural mood-boosting products.

18 106. Further, Plaintiff R.R. has had a Facebook account for sixteen (16) years and has
19 posted hundreds of photos of herself and her family and friends. Using this platform, she posts
20 sensitive information about mental health as well as work-related updates. She is also a member of
21 various Facebook “groups”: UX design groups, buy/trade/sell groups, fonts/typography groups, and
22 three groups dedicated to local rants and raves and to food and drinks.

23 107. Plaintiff R.R. is concerned that Defendants have taken her skills and expertise, as
24 reflected in his online contributions, and incorporated it into Products that could someday result in
25 professional obsolescence for design programmers like her.

26 108. Plaintiff R.R. reasonably expected that the information that she exchanged with these
27 websites prior to 2021 would not be intercepted by any third-party looking to compile and use all
28 her information and data for commercial purposes. Plaintiff R.R. did not consent to the use of her

private information by third parties in this manner. Notwithstanding, Defendants stole Plaintiff R.R.'s personal data from across this wide swath of online applications and platforms to train the Products.

Defendants

109. **Defendant OpenAI** is an AI research laboratory consisting of the non-profit OpenAI Incorporated ("OpenAI Inc.") and its for-profit subsidiary corporation OpenAI Limited Partnership ("OpenAI LP") (hereinafter, collectively, "OpenAI").²¹ OpenAI was founded in 2015 and is headquartered in San Francisco, CA. OpenAI has released the AI-based products DALL-E, GPT-4, OpenAI Five, ChatGPT, and OpenAI Codex for commercial (to integrate within one's business) and personal use.

110. OpenAI was originally founded as a nonprofit research laboratory with a single mission: "to advance [artificial] intelligence in the way that is most likely to benefit humanity as a whole."²² In the words of OpenAI at the time, it was critical for the organization to be "unconstrained by a need to generate a financial return."²³ Fast forward to April of 2023: OpenAI closed a more than \$300 million share sale at a valuation between \$27 billion and \$29 billion.²⁴ OpenAI projects that its AI chatbot, ChatGPT, will generate a revenue of \$200 million in 2023 and exponentially grow to \$1 billion by the end of 2024.²⁵

111. Defendant OpenAI GP, L.L.C. ("OpenAI GP") is a Delaware limited liability company with its principal place of business located at 3180 18th Street, San Francisco, CA 94110. OpenAI GP is wholly owned and controlled by OpenAI, Inc. Further, OpenAI GP is the general partner of OpenAI, L.P. and is responsible for managing and operating the day-to-day business and affairs of OpenAI, L.P. Its primary focus is research and technology. OpenAI GP was aware of the

²¹ *OpenAI LP*, OPENAI, <https://openai.com/blog/openai-lp> (last visited June 27, 2023).

²² Greg Brockman & Ilya Sutskever, *Introducing OpenAI*, OPENAI (Dec. 11, 2015), <https://openai.com/blog/introducing-openai>.

²³ *Id.*

²⁴ *OpenAI Closes \$300 Million Funding Round at \$27 Billion-\$29 Billion Valuation*, *TechCrunch reports*, REUTERS (Apr. 28, 2023), <https://www.reuters.com/markets/deals/openai-closes-10-bln-funding-round-27-bln-29-bln-valuation-techcrunch-2023-04-28/>.

²⁵ Jeffrey Dastin, *Exclusive: ChatGPT Owner OpenAI Projects \$1 Billion in Revenue by 2024*, REUTERS (Dec. 15, 2022), <https://www.reuters.com/business/chatgpt-owner-openai-projects-1-billion-revenue-by-2024-sources-2022-12-15/>.

1 unlawful conduct alleged herein and exercised control over OpenAI, L.P. throughout the Class
 2 Period. OpenAI GP is liable for the debts, liabilities, and obligations of OpenAI, L.P., including
 3 litigation and judgments.

4 112. Defendant OpenAI Startup Fund I, L.P. (“OpenAI Startup Fund I”) is a Delaware
 5 limited partnership with its principal place of business located at 3180 18th Street, San Francisco,
 6 CA 94110. Upon information and belief, OpenAI Startup Fund I played a vital role in the foundation
 7 of OpenAI, L.P., including providing initial funding and creating its business strategy. By
 8 participating in OpenAI Startup Fund I, certain entities and individuals obtained an ownership
 9 interest in OpenAI, L.P. OpenAI Startup Fund I exercised control over OpenAI, L.P. and was aware
 10 of the unlawful conduct alleged herein throughout the Class Period.

11 113. Defendant OpenAI Startup Fund GP I, L.L.C. (“OpenAI Startup Fund GP I”) is a
 12 Delaware limited liability company with its principal place of business located at 3180 18th Street,
 13 San Francisco, CA 94110. OpenAI Startup Fund GP I is the general partner of OpenAI Startup Fund
 14 I and is responsible for managing and operating the day-to-day business and affairs of OpenAI
 15 Startup Fund I. OpenAI Startup Fund GP I is liable for the debts, liabilities, and obligations of
 16 OpenAI Startup Fund I, including litigation and judgments. OpenAI Startup Fund GP I was aware
 17 of the unlawful conduct alleged herein and exercised control over OpenAI, L.P. throughout the
 18 Class Period. Sam Altman, co-founder, CEO, and Board member of OpenAI, Inc. is the Manager
 19 of OpenAI Startup Fund GP I.

20 114. Defendant OpenAI Startup Fund Management, LLC (“OpenAI Startup Fund
 21 Management”) is a Delaware limited liability company with its principal place of business located
 22 at 3180 18th Street, San Francisco, CA 94110. OpenAI Startup Fund Management exercised control
 23 over OpenAI, L.P. throughout the Class Period and thus, was aware of the unlawful conduct alleged
 24 herein.

25 115. **Defendant Microsoft Corporation** (“Microsoft”) is a Washington corporation with
 26 its principal place of business located at One Microsoft Way, Redmond, Washington 98052.
 27 Microsoft partnered with OpenAI in 2016 with the goal to “democratize Artificial Intelligence.” In
 28

July 2019, Microsoft invested \$1 billion in OpenAI LP at a \$20 billion valuation.²⁶ In 2020, Microsoft became the exclusive licensee of OpenAI’s GPT-3 language model—despite OpenAI’s continued claims that its products are meant to benefit “humanity” at large. In October 2022, news reports stated OpenAI was “in advanced talks to raise more funding from Microsoft” at that same \$20 billion valuation.²⁷ Then, in January of 2023, Microsoft confirmed its extended partnership with OpenAI by investing \$10 billion into ChatGPT.²⁸ Prior to this \$10 billion dollar investment, Microsoft had invested \$3 billion into OpenAI in previous years.²⁹

116. Microsoft’s continued investments, as well as introduction of ChatGPT on its multiple platforms (Bing, Microsoft Teams, etc.) underscore the depth of its partnership with OpenAI. Through these investments, Microsoft gained exclusive access to the entire OpenAI codebase.³⁰ Furthermore, Microsoft Azure also acts as the exclusive cloud service of OpenAI.³¹

117. As OpenAI’s largest investor and largest service provider—specifically in connection with the development of ChatGPT—Microsoft exerts considerable control over OpenAI. Analysts estimate OpenAI will add between \$30 billion and \$40 billion to Microsoft’s top line.

118. **Agents and Co-Conspirators.** Defendants’ unlawful acts were authorized, ordered,

²⁶ Hasan Chowdhury, *Microsoft’s Investment into ChatGPT’s Creator May be the Smartest \$1 Billion Ever Spent*, BUS. INSIDER (Jan. 6, 2023), <https://www.businessinsider.com/microsoft-openai-investment-the-smartest-1-billion-ever-spent-2023-1>; Dina Bass, *Microsoft Invests \$10 Billion in ChatGPT Maker OpenAI*, BLOOMBERG (Jan. 23, 2023), <https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai#xj4y7vzkg>.

²⁷ Aaron Holmes et al., *OpenAI, Valued at Nearly \$20 Billion, in Advanced Talks with Microsoft for More Funding*, THE INFO. (Oct. 20, 2022), <https://www.theinformation.com/articles/openai-valued-at-nearly-20-billion-in-advanced-talks-with-microsoft-for-more-funding>.

²⁸ *Microsoft Confirms Its \$10 Billion Investment into ChatGPT, Changing How Microsoft Competes with Google, Apple and Other Tech Giants*, FORBES (Jan. 27, 2023), <https://www.forbes.com/sites/qai/2023/01/27/microsoft-confirms-its-10-billion-investment-into-chatgpt-changing-how-microsoft-competes-with-google-apple-and-other-tech-giants/?sh=4eea29723624>.

²⁹ Cade Metz, *Microsoft to Invest \$10 Billion in OpenAI, the Creator of ChatGPT*, THE N.Y. TIMES (Jan. 23, 2023), <https://www.nytimes.com/2023/01/23/business/microsoft-chatgpt-artificial-intelligence.html>.

³⁰ Mohit Pandey, *OpenAI, a Data Scavenging Company for Microsoft*, AIM (Mar. 24, 2023), <https://analyticsindiamag.com/openai-a-data-scavenging-company-for-microsoft/>.

³¹ *Microsoft Confirms Its \$10 Billion Investment Into ChatGPT, Changing How Microsoft Competes With Google, Apple And Other Tech Giants*, FORBES (Jan. 27, 2023), <https://www.forbes.com/sites/qai/2023/01/27/microsoft-confirms-its-10-billion-investment-into-chatgpt-changing-how-microsoft-competes-with-google-apple-and-other-tech-giants/?sh=4eea29723624>.

1 and performed by Defendants' respective officers, agents, employees, and representatives, while
2 actively engaged in the management, direction, and control of Defendants' businesses and affairs.
3 Defendants' agents operated under explicit and apparent authority of their principals. Each
4 Defendant, and their subsidiaries, affiliates, and agents operated as a single unified entity.

5 **JURISDICTION AND VENUE**

6 119. This Court has subject matter jurisdiction over the federal claims in this action,
7 namely the Electronic Communications Privacy Act and the Computer Fraud and Abuse Act,
8 pursuant to 28 U.S.C. § 1331.

9 120. This Court also has subject matter jurisdiction over this action pursuant to the Class
10 Action Fairness Act, 28 U.S.C. § 1332(d), because this is a class action in which the amount in
11 controversy is \$3,000,000,000, far in excess of the statutory minimum, exclusive of interest and
12 costs. There are millions of class members as defined below, and minimal diversity exists because
13 a significant portion of class members are citizens of a state different from the citizenship of at least
14 one Defendant.

15 121. This Court also has supplemental jurisdiction over the state law claims in this action
16 pursuant to 28 U.S.C. § 1367 because the state law claims form part of the same case or controversy
17 as those that give rise to the federal claims.

18 122. Pursuant to 28 U.S.C. § 1391, this Court is the proper venue for this action because a
19 substantial part of the events, omissions, and acts giving rise to the claims herein occurred in this
20 District: Defendant OpenAI is headquartered in this District, all Defendants gain significant revenue
21 and profits from doing business in this District, consumers sign up for ChatGPT accounts and
22 provide ChatGPT with their sensitive information in this District, Class Members affected by this
23 data misuse reside in this District, and Defendants employ numerous people in this District—a
24 number of whom work specifically on making the decisions regarding the data privacy and handling
25 of consumers' data that are challenged in this Action. Each Defendant has transacted business,
26 maintained substantial contacts, and/or committed overt acts in furtherance of the illegal scheme
27 and conspiracy throughout the United States, including in this District. Defendants' conduct had the
28 intended and foreseeable effect of causing injury to persons residing in, located in, or doing business

throughout the United States, including in this District.

123. Defendants are subject to personal jurisdiction in California based upon sufficient minimum contacts which exist between Defendants and California. Defendants are authorized to do and are doing business in California, and Defendants advertise and solicit business in California. Defendants have purposefully availed themselves of the protections of California law and should reasonably expect to be hauled into court in California for harm arising out of their pervasive contacts with the State. Further, for Defendant OpenAI, the decisions affecting consumers data and privacy stem from the company's San Francisco office headquarters.

FACTUAL BACKGROUND

I. DEVELOPMENT OF ARTIFICIAL INTELLIGENCE IN THE U.S.

A. OpenAI: From Open Nonprofit to Profit-Driven \$29B Commercial Partner of Tech Giant Microsoft

124. OpenAI was founded in 2015 as a nonprofit research laboratory with a single mission: “to advance artificial intelligence in a way that would benefit society as a whole. . . .”³² Critical to that mission, according to OpenAI at the time, was for the organization to be “unconstrained by a need to generate a financial return.”³³ The nonprofit was thus funded by million-dollar donations from prominent, wealthy entrepreneurs and researchers who shared the non-profit’s vision of creating safe, ethical, and responsible AI, to benefit humankind and to do no harm, and who recognized the dangers that could befall society if AI were developed and launched for commercial gain.

125. OpenAI also originally pledged to “freely collaborate” with other responsible organizations and researchers, in part by making its research available to inspect and audit as a further “check” on the safety of any AI capabilities, to help ensure the powerful technology on which they were working would not someday destroy lives and ultimately, civilization. The founders believed this openness was so critical to the non-profit’s mission, that they named it “Open” AI. As they further explained at the time, “since our research is free from financial

³² *The Transformation of OpenAI From Nonprofit to \$29B For-Profit*, THE SOCIABLE (Apr. 5, 2023), <https://sociable.co/business/the-transformation-of-openai-from-nonprofit-to-29b-for-profit/>.

³³ *Id.*

obligations, we can better focus on a positive human impact. We believe AI should be an extension of individual human wills, and in the spirit of liberty, as broadly and evenly distributed as possible.”³⁴

126. For years, OpenAI purported to operate as such: openly and in pursuit of its single mission to advance humanity, safely and responsibly. That all changed in 2019, when OpenAI abruptly “shut its doors” to all ‘Open’ influence and scrutiny, shifted to a profit-generating corporate structure, and decided instead to focus on commercializing the AI capabilities on which it had been working.

127. At the time, Google Brain’s “transformer” innovation had opened a new frontier in AI development, where AI could improve endlessly, some experts believe to even superhuman intelligence—but only if it were fed “endless data” to train it, a costly endeavor given the computing power required.³⁵ To do so, OpenAI entered an exclusive partnership with Microsoft, which invested \$1B into the company, gaining the only outside access to the effort once “Open” to all. Together, they built a “supercomputer” to train massive language models that ultimately resulted in ChatGPT and the image generator DALL-E.³⁶

128. OpenAI’s sudden shift to a profit focus and alignment with Microsoft, a corporate giant with a vested interest in curating and dominating a commercial market for AI, marked the beginning of the end of OpenAI’s commitment to humanity. The company began to pursue profits at the expense of privacy, security, and ethics, beginning with its data collection.

129. To realize the most powerful and thus most profitable AI, OpenAI would need data, and lots of it, to “train” the language models on which the Products run using the supercomputer it had built in partnership with Microsoft. Defendants thus doubled down on their strategy to secretly harvest millions of consumers’ personal data from the internet. Then, on the backs of this stolen data, they rushed to market the Products without adequate safeguards or controls to ensure their

³⁴ Greg Brockman & Ilya Sutskever, *Introducing OpenAI*, OPENAI (Dec. 11, 2015), <https://openai.com/blog/introducing-openai>.

³⁵ Reed Albergotti, *The Secret history of Elon Musk, Sam Altman, and OpenAI*, SEMAFOR (Mar. 24, 2023), <https://www.semafor.com/article/03/24/2023/the-secret-history-of-elon-musk-sam-altman-and-openai>.

³⁶ *Id.*

1 safety. While Defendants recognized then, as they do now, that they cannot fully predict how the
 2 Products might evolve to operate, they knew the public would be amazed by the Products already
 3 seemingly near human “intelligence” and other capabilities. And thus, they knew they could make
 4 a ton of money.

5 130. In public, OpenAI continued to state its commitment to ethical AI development. But
 6 with its new profit orientation, that “was kind of like trying to juggle while riding a unicycle, except
 7 with more existential questions about the nature of humanity.”³⁷ Defendants acknowledge they do
 8 not understand the full scope of the risks posed by the Products currently, and no one knows how
 9 AI might evolve now that billions of people are using the technology every day.³⁸ Defendants, like
 10 other leading experts, are united in believing the ultimate risk posed by AI is the collapse of
 11 civilization as we know it. And yet, they released the Products worldwide anyway, setting off a
 12 global AI arms race.

13 131. Earlier this year, OpenAI raised another \$10B from its single corporate partner,
 14 Microsoft, increasing its then corporate valuation to \$29B and giving Microsoft a significant stake
 15 in the company. With that, the 180-degree transformation—from open nonprofit for the benefit of
 16 humanity to closed corporate profit machine fueled by greed and market power—was complete.

17 132. OpenAI’s shift in organizational structure has raised eyebrows given its
 18 unprecedented nature, and the moral and legal questions it raises. AI researchers, ethicists, and the

19 ³⁷*The Transformation of OpenAI From Nonprofit to \$29B For-Profit*, THE SOCIABLE (Apr. 5,
 20 2023), <https://sociable.co/business/the-transformation-of-openai-from-nonprofit-to-29b-for-profit/>.

21 ³⁸ “As a system like this learns from data, it develops skills that its creators never expected. It is
 22 hard to know how things might go wrong after millions of people start using it.” *See* Cade Metz,
 23 *What’s the Future for AI?*, THE N.Y. TIMES (Mar. 31, 2023),
 24 <https://www.nytimes.com/2023/03/31/technology/ai-chatbots-benefits-dangers.html>; Jason
 25 Abbruzzese, *The Tech Watchdog that Raised Alarms About Social Media is Warning About AI*,
 26 NBC NEWS (Mar. 22, 2023), <https://www.nbcnews.com/tech/tech-news/tech-watchdog-raised-alarms-social-media-warning-ai-rcna76167> (“What’s surprising and what nobody foresaw is that
 27 just by learning to predict the next piece of text on the internet, these models are developing new
 28 capabilities that no one expected. . . So just by learning to predict the next character on the
 internet, it’s learned how to play chess.” Others have also commented on the technology
 continuing to display unintended and unpredictable emergent capabilities. Jason Wei, *137
 Emergent Abilities of Large Language Models*, JASON WEI (Nov. 14, 2022),
<https://www.jasonwei.net/blog/emergence>; Stephen Ornes, *The Unpredictable Abilities Emerging
 from Large AI Models*, QUANTA MAG. (Mar. 16, 2023), <https://www.quantamagazine.org/the-unpredictable-abilities-emerging-from-large-ai-models-20230316/>.

public share concerns about the conflict between OpenAI’s original mission to benefit humanity on the one hand and the current profit-driven motives of investors, chiefly Microsoft, on the other.³⁹ They worry that OpenAI is prioritizing short-term financial gains over long-term safety and ethical considerations, as exemplified by the sudden deployment of the Products for widespread commercial use despite all the known dangers.⁴⁰ Moreover, as one commentator noted, “there are various different ways to make hundreds of millions of dollars, but historically ‘starting a nonprofit’ has not been one of them.”⁴¹

133. Elon Musk, an original non-profit funder and founder, was more blunt as to the seismic shift: “I’m still confused as to how a non-profit to which I donated ~100M somehow became a \$30B market cap for-profit.” He noted, “OpenAI was created as an open source (which is why I named it ‘Open’ AI), non-profit company to serve as a counterweight to Google, but now it has become a closed source, maximum profit company effectively controlled by Microsoft.”⁴²

134. If soliciting non-profit contributions to then turn around and build a for-profit company “is legal,” Musk opined, then “why doesn’t everyone do it?”⁴³ This same question must be asked about the equally unprecedented theft of personal data that is at the heart of this Action, and the answer to both questions is the same: ***It isn’t.***

135. As explained below, the only thing still ‘open’ about OpenAI is its open disregard for the privacy and property interests of hundreds of millions. Worse, as a result of OpenAI’s machinations for profit, “the most powerful tool mankind has ever created, is now in the hands of a

³⁹ *From Non-Profit to Profit Monster: OpenAI’s Controversial Corporate Shift*, EXPLORING CHATGPT (Apr. 8, 2023), <https://exploringchatgpt.substack.com/p/from-non-profit-to-profit-monster>.

⁴⁰ *Id.*

⁴¹ Felix Salmon, *How a Silicon Valley Nonprofit Became Worth Billions*, AXIOS (Jan. 10, 2023), <https://www.axios.com/2023/01/10/how-a-silicon-valley-nonprofit-became-worth-billions>.

⁴² Sawdah Bhaimiya, *OpenAI Cofounder Elon Musk Said the Non-Profit He Helped Create is Now Focused on ‘Maximum-Profit,’ Which is ‘Not What I Intended at All’*, BUS. INSIDER (Feb. 17, 2023), https://www.businessinsider.com/elon-musk-defends-role-in-openai-ChatGPT-microsoft-2023-2?utm_source=flipboard&utm_content=user%2FInsiderBusiness.

⁴³ @elonmusk, TWITTER (Mar. 15, 2023), <https://twitter.com/elonmusk/status/1636047019893481474>.

1 ruthless corporate monopoly.”⁴⁴

2 **B. OpenAI’s Products**

3 136. The most well-known of OpenAI’s products—and of all AI worldwide—is the
4 ground-breaking chatbot, ChatGPT. Once users input a question or a prompt in ChatGPT, the
5 information is digested by the AI model and the chatbot produces a response based on the
6 information a user has given and how that fits into its vast amount of training data.

7 137. ChatGPT was released as a “research preview” on November 30, 2022.⁴⁵ A blog post
8 casually introduced the AI chatbot to the world, thusly: “We’ve trained a model . . . which interacts
9 in a conversational way.” ChatGPT subsequently exploded in popularity, reaching **100 million**
10 **users** in only two months, making it the fastest-growing app in history.⁴⁶ For comparison, TikTok
11 took nine months to reach the same benchmark.⁴⁷ ChatGPT has continued to evolve exponentially,
12 **with 1.8 billion visits in April of 2023.**⁴⁸

13 138. ChatGPT was built on a family of large language models (“LLMs”) collectively
14 known as GPT-3. As explained below, ChatGPT-3.5 was trained on 570GB of text data from the
15 internet containing hundreds of billions of words,⁴⁹ including text harvested from books, articles,
16 and websites, including social media. Due to its vast training data, ChatGPT can generate human-
17 like answers to text prompts and questions making it interact like “a friendly robot.”⁵⁰ On command
18 it can do a lot of what people do, like write poetry, compose music, draft research papers, create
19 lesson plans, and so much more, only faster than one human ever could. Naturally, the world was

20
21 ⁴⁴ Marvie Basilan, *Elon Musk Says He’s The Reason OpenAI Exists as Sam Altman Testifies*
22 *Before Congress*, INT’L BUS. TIMES (May 17, 2023), <https://www.ibtimes.com/elon-musk-says-hes-reason-openai-exists-sam-altman-testifies-before-congress-3693771>.

23 ⁴⁵ *Introducing ChatGPT*, OPENAI (NOV. 30, 2022), <https://openai.com/blog/chatgpt>.

24 ⁴⁶ Krystal Hu, *ChatGPT Sets Record for Fastest-Growing User Base - Analyst Note*, REUTERS (Feb. 2, 2023), <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>.

25 ⁴⁷ *Id.*

26 ⁴⁸ Nerdynav, *97+ ChatGPT Statistics & User Numbers in June 2023 (New Data)*, NERDY NAV (June 2, 2023), <https://nerdynav.com/chatgpt-statistics/>.

27 ⁴⁹ Uri Gal, *CHATGPT Collected Our Data Without Permission and is Going to Make Billions Off it*, SCROLL.IN (Feb. 15, 2023), <https://scroll.in/article/1043525/chatgpt-collected-our-data-without-permission-and-is-going-to-make-billions-off-it>.

28 ⁵⁰ Mark Wilson, *ChatGPT Explained: Everything You Need to Know About the AI Chatbot*, TECHRADAR (Mar. 15, 2023), <https://www.techradar.com/news/chatgpt-explained>.

1 stunned by these capabilities.

2 139. OpenAI has also released other AI-based products DALL-E, OpenAI Five, and
3 OpenAI Codex for commercial (to integrate within one's business) and personal use. It also
4 developed a program VALL-E, which has not been released for use to the public yet.

5 140. DALL-E (consisting of DALL-E and DALL-E 2) are deep learning models developed
6 by OpenAI to generate realistic digital images from natural language descriptions, known as
7 "prompts."⁵¹ DALL-E uses a version of GPT-3, modified to generate images.⁵²

8 141. OpenAI Five is a computer program developed by OpenAI that plays the five-on-five
9 video game Dota 2.⁵³

10 142. OpenAI Codex is another artificial intelligence model developed by Open AI, which
11 is programmed to generate computer code for use in programming applications.⁵⁴

12 143. VALL-E is another artificial intelligence model intended to synthesize high-quality
13 personalized speech utilizing only a 3-second enrolled recording of an unseen speaker as a prompt.⁵⁵
14 VALL-E was trained on audio voices from thousands of speakers.⁵⁶

15 **C. ChatGPT's Development Depends on Secret Web-Scraping**

16 144. The large language models responsible for the Products depend on consuming huge
17 amounts of data, in order to "train" the AI. Valuable to the process is personal data of any kind,
18 including conversational data between humans, as this is how the Products develop what appear to
19 be such human-like capabilities.

20 145. As a general matter, internet user data is available for purchase like any other content

21
22 ⁵¹ Khari Johnson, *OpenAI Debuts DALL-E for Generating Images from Text*, VENTURE BEAT (Jan.
5, 2021), <https://venturebeat.com/business/openai-debuts-dall-e-for-generating-images-from-text/>.

23 ⁵² *Id.*

24 ⁵³ Ben Dickson, *AI Defeated Human Champions at Dota 2*, TECHTALKS (Apr. 17, 2019),
<https://bdtectalks.com/2019/04/17/openai-five-neural-networks-dota-2/>.

25 ⁵⁴ Thomas Smith, *Why OpenAI's Codex Won't Replace Coders*, IEEE SPECTRUM (Sept. 28, 2021),
<https://spectrum.ieee.org/openai-wont-replace-coders>.

26 ⁵⁵ *VALL-E Neural Codec Language Models are Zero-Shot Text to Speech Synthesizers*, GITHUB
PAGES, <https://lifeiteng.github.io/valle/index.html> (last visited June 27, 2023).

27 ⁵⁶ *VALL-E: Five Things to Know About Microsoft's AI Model That Can Mimic Any Voice in Three*
Seconds, TIMES OF INDIA (Jan. 11, 2023), <https://timesofindia.indiatimes.com/gadgets-news/vall-e-5-things-to-know-about-microsofts-ai-model-that-can-mimic-any-voice-in-3-seconds/articleshow/96898774.cms>.
28

or property. In the technological era in which we live, a mature market for such data exists given how valuable our personal information has become to companies, for marketing and other purposes. The legal acquisition of data typically depends on consent and remuneration, with some form of consideration exchanged.

146. Despite established protocols for the purchase and use of personal information, Defendants took a different approach: *theft*. They systematically scraped 300 billion words from the internet, “books, articles, websites and posts – including personal information obtained without consent.”⁵⁷ OpenAI did so in secret, and without registering as a data broker as it was required to do under applicable law (*See infra* at Section III.A).

147. “Scraping involves the use of ‘bots,’ or robot applications deployed for automated tasks, which scan and copy the information on webpages then *store* and *index* the information.”⁵⁸ According to a computer science professor at the University of Oxford, Michael Wooldridge, the full extent of personal data taken by Defendants’ scraping is “unimaginable.”⁵⁹

148. In his interview with The Guardian, Professor Wooldridge explained that the LLM underlying ChatGPT, and other AIs like it, “includes the whole of the world wide web – *everything*. Every link is followed in every page, and every link in those pages is followed.”⁶⁰ Thus, swept up into the Products is “a lot of data about you and me.”⁶¹ Others have noted that the data includes transcripts of our online chat logs, from across the internet, and other forms of personal conversation such as our online customer service interactions and social media conversations, as well as “billions of images scraped from the internet.”⁶² Many of these images were of “children and came from

⁵⁷ Uri Gal, *ChatGPT is a Data Privacy Nightmare. If You’ve Ever Posted Online, You Ought to be Concerned*, THE CONVERSATION (Feb. 7, 2023), <https://theconversation.com/chatgpt-is-a-data-privacy-nightmare-if-youve-ever-posted-online-you-ought-to-be-concerned-199283>.

⁵⁸ Will Hillier, *What is Web Scraping? A Complete Beginners Guide*, CAREER FOUNDRY (Aug. 13, 2021), <https://careerfoundry.com/en/blog/data-analytics/web-scraping-guide/>.

⁵⁹ Alex Hern & Dan Milmo, *I Didn’t Give Permission: Do AI’s Backers Care About Data Law Breaches?*, THE GUARDIAN (Apr. 10, 2023), <https://www.theguardian.com/technology/2023/apr/10/i-didnt-give-permission-do-ais-backers-care-about-data-law-breaches>.

⁶⁰ *Id.*

⁶¹ *Id.*

⁶² Jit Roy, *Data Source of ChatGPT*, ABOUTCHATGPT.COM (Jan. 2, 2023), <https://aboutchatgpt.com/data-source-of-chatgpt/>; *see also* Hern & Milmo, *supra* note 59.

photo sites and personal blogs.”⁶³

149. The unprecedented scope of the effort together with Defendants’ failure to seek consent has been described as “the elephant in the room. . . all this training data must come from somewhere. ChatGPT has effectively scraped the entire internet[.]”⁶⁴ As a result, Defendants have essentially embedded into the Products personal information across a range of categories that reflect our hobbies and interests, our religious beliefs, our political views and voting records, the social and support groups to which we belong, our sexual orientations and gender identities, our personal relationship statuses, our work information and histories, details (including pictures) about our families and children, the music we listen to, our purchasing behaviors, our general likes and dislikes, the ways in which we speak and write, our mental health and ailments, where we live and where we go, the websites we visit, our digital subscriptions, our friend groups and other associational data, our email addresses, other contact and identifying information, and more.⁶⁵ With respect to personally identifiable information, Defendants fail sufficiently to filter it out of the training models, putting millions at risk of having that information disclosed on prompt or otherwise to strangers around the world.⁶⁶

150. The breadth and scope of Defendants’ data collection without permission, impacting

⁶³ Drew Harwell, *AI-generated child sex images spawn new nightmare for the web*, THE WASH. POST (June 19, 2023), <https://www.msn.com/en-us/news/us/ai-generated-child-sex-images-spawn-new-nightmare-for-the-web/ar-AA1cKhLH>.

⁶⁴ Deep Tech Insights, *ChatGPT is a Threat, but Google is Still a Buy*, SEEKING ALPHA (Dec. 19, 2022), <https://seekingalpha.com/article/4565302-alphabet-ChatGPT-is-a-threat-but-google-is-still-a-buy>.

⁶⁵ *Digital Footprint: What is It And Why You Should Care About It*, INVISIBLY (Jan. 25, 2022), <https://www.invisibly.com/learn-blog/digital-footprint/> (“Your digital footprint is your trail of personal information that companies can follow. . . . To break it down, your digital footprint is essentially a record of your online activity. Whenever you log into an account, send an email, or buy something online, it leaves a digital impression behind. It is the trail of data left behind by your daily interactions. Your footprint is permanent which can leave your information vulnerable if not protected correctly. You might not always be aware that you are creating your digital footprint. For instance, websites can track your activity by installing cookies on your device. Furthermore, apps can collect your data without you even knowing it. Once an organization has access to your data, they can sell or share it with third parties. Even more, your information is out there and could be compromised via a data breach.”).

⁶⁶ Katyanna Quach, *What happens when your massive text-generating neural net starts spitting out people's phone numbers? If you're OpenAI, you create a filter*, THE REG. (Mar. 18, 2021), https://www.theregister.com/2021/03/18/openai_gpt3_data/?td=readmore-top.

essentially every internet user ever, raises serious legal, moral, and ethical issues.⁶⁷ One critique summarized the privacy risk bluntly, as follows: “*ChatGPT is a data privacy nightmare. If you’ve ever posted online, you ought to be concerned.*”⁶⁸ While regulators and courts around the world seek to crack down on AI researchers “hoovering up content without consent or notice,” the response, by Defendants and others, has been to keep their datasets largely secret, and to not grant regulator or other audit access.⁶⁹

151. Despite “*Open*” AI’s “absolute secrecy” surrounding its data collections and practices,⁷⁰ we know at the highest levels that the Company used (at least) five (5) distinct datasets to train ChatGPT: (1) Common Crawl; (2) WebTex2, text of webpages from all outbound Reddit links from posts with 3+ upvotes; (3) Books1; (4) Books2; and (5) Wikipedia.⁷¹

152. Of these training datasets, WebTex2 is OpenAI’s “proprietary” AI corpus of personal data. To build it, OpenAI scraped every webpage linked to on the social media site Reddit in all posts that received at least 3 “likes” (known as “Karma” votes on Reddit), together with the Reddit posts and rich conversational data from its users around the world. The most popular “outbound” links on Reddit include many of the most popular websites in the world, where people post personal information, video, and audio clips of themselves and more, e.g., YouTube, Facebook, TikTok, Snapchat, and Instagram. Given Defendants’ scraping protocols, all of this “outbound” data from these various websites was targeted for taking, without notice or consent, to feed the large language models on which the Products depend.

153. The co-founder and CEO of Reddit, Steve Huffman, remarked on the breadth of Defendants’ unauthorized scraping, noting that he found it unacceptable that OpenAI has been

⁶⁷ Erin Griffith & Cade Metz, *A New Era of A.I. Booms, Even Amid the Tech Gloom*, THE N.Y. TIMES (Jan. 7, 2023), <https://www.nytimes.com/2023/01/07/technology/generative-ai-chatgpt-investments.html> (“The technology has raised thorny ethical questions around how generative A.I. may affect copyrights and whether the companies need to get permission to use the data that trains their algorithms.”).

⁶⁸ Gal, *supra* note 57.

⁶⁹ Hern & Milmo, *supra* note 59.

⁷⁰ *Id.* (“Copyright lawsuits and regulator actions against OpenAI are hampered by the company’s absolute secrecy about its training data.”).

⁷¹ Patrick Meyer, *ChatGPT: How Does It Work Internally*, MEDIUM (Dec. 10, 2022), <https://pub.towardsai.net/chatgpt-how-does-it-work-internally-e0b3e23601a1?gi=f28c10d5afef>.

1 scraping “huge amounts of Reddit data to train their systems – for free.”⁷² According to Huffman,
 2 “The Reddit corpus of data is really valuable. But we don’t need to give all of that value to some of
 3 the largest companies in the world for free.”⁷³

4 154. Defendants’ theft related to their WebTex2 corpus is ongoing and continuous. As one
 5 article explains, “the advantage of using the Webtext dataset is that it is constantly updated with
 6 new data. As new web pages are added to the internet, they are included in the dataset, which helps
 7 to ensure that the model is trained on the most recent and relevant language data.”⁷⁴ Neither Reddit
 8 itself nor Reddit users, much less all the owners of the webpages and personal data linked to and
 9 from Reddit, consent to this taking of data.

10 155. The other primary data set on which the Products depend, that the public currently
 11 knows about, is the “Common Crawl,” a massive collection of web pages and websites also derived
 12 from large-scale web scraping. It contains petabytes of data collected over twelve (12) years,
 13 including raw webpage data, metadata extracts, and text extracts from all types of websites.⁷⁵ In
 14 total, the Common Crawl dataset constitutes nearly a trillion words.

15 156. The Common Crawl dataset is owned by a non-profit of the same name, which has
 16 been indexing and storing as much of the World Wide Web as it can access, filing away as many as
 17 3 billion webpages every month, for over a decade.⁷⁶ The non-profit makes the data available to the
 18 public for free—but for research and educational purposes. As a result, the Common Crawl is a
 19 staple of large *academic* studies of the web.⁷⁷ It was never intended to be taken *en masse* and turned
 20

21 ⁷² Gintaras Raauskas, *Redditors on Strike but Company Wants OpenAI to Pay Up for Scraping*,
 22 CYBERNEWS, <https://cybernews.com/news/reddit-strike-api-openai-scraping/> (last updated June
 23 12, 2023).

24 ⁷³ *Id.*

25 ⁷⁴ GPTBlogs, *ChatGPT: How Much Data is Used in the Training Process?*, (Feb. 9, 2023),
 26 <https://gptblogs.com/chatgpt-how-much-data-is-used-in-the-training-process>.

27 ⁷⁵ *Want to Use Our Data*, COMMON CRAWL, <https://commoncrawl.org/the-data/> (last visited June
 28 27, 2023).

⁷⁶ James Bridle, *The Stupidity of AI*, THE GUARDIAN (Mar. 16, 2023),
<https://www.theguardian.com/technology/2023/mar/16/the-stupidity-of-ai-artificial-intelligence-dall-e-chatgpt>.

⁷⁷ Kalev Leetaru, *Common Crawl and Unlocking Web Archives for Research*, FORBES (Sept. 28,
 2017), <https://www.forbes.com/sites/kalevleetaru/2017/09/28/common-crawl-and-unlocking-web-archives-for-research/?sh=19e3c5373b83>.

1 into an AI product for commercial gain, as Defendants have done. On information and belief, the
 2 501(c)(3) overseeing the Common Crawl did not consent to this mass misappropriation of personal
 3 data for commercial purposes. And even if it did, it did not obtain consent from internet users whose
 4 personal data it scraped.

5 157. The commercial misappropriation of the Common Crawl has raised concerns given
 6 the amount of personal data it contains, including highly personal data. One chilling example of the
 7 privacy invasions caused by Defendants' misappropriation is the experience of a San Francisco-
 8 based digital artist named Lapine. Using the online tool "Have I Been Trained," Lapine was able to
 9 determine that her private medical file—i.e., photographs taken of her body as part of clinical
 10 documentation when she was undergoing treatment for a rare genetic condition—ended up online
 11 and then, memorialized in the Common Crawl archive.⁷⁸

12 158. Remarking on the web scraping practices in which Defendants engaged and the
 13 subsequent commercialization of the ill-gotten data, Lapine highlighted the unique scope of the
 14 harm: "It's the digital equivalent of receiving stolen property. . . [my medical information] was
 15 scraped into this dataset. . . it's bad enough to have a photo leaked, *but now it's part of a product.*"⁷⁹
 16 More broadly, this "productization" of personal information means all this data about us, scraped
 17 without permission, can now fuel ChatGPT's responses to strangers around the world.⁸⁰ Worse,
 18 ChatGPT is the "new favorite toy" of online criminals, as the billions of personal and other data
 19 points about us, "scraped by ChatGPT, are now *free to use* for any number of targeted attacks,
 20 including malware, ransomware, phishing, Business Email Compromise, and social engineering."⁸¹

21 159. As described further in Section III, this secret and unregistered scraping of internet
 22 data, for Defendants' own private and exorbitant financial gain, without regard to privacy risks,
 23 amounts to the negligent and otherwise illegal theft of personal data of millions of Americans who
 24 do not even use AI tools. These individuals ("Non-Users") had their personal information scraped
 25 long before OpenAI's applications were available to the public, and certainly before they could have

26 ⁷⁸ Bridle, *supra* note 76.

27 ⁷⁹ *Id.*

28 ⁸⁰ *Is ChatGPT a Disaster for Data Privacy?*, BUS. REP. (Feb. 17, 2023), <https://www.business-reporter.co.uk/risk-management/is-chatgpt-a-disaster-for-data-privacy>.

⁸¹ *Id.*

1 registered as a ChatGPT user. In either case, no one consented to the use of their personal data to
2 train the Products.

3 160. OpenAI is now worth around \$29B, yet the individuals and companies that produced
4 the data it scraped from the internet have not been compensated.⁸² This Action seeks to change that,
5 and in the process, protect the privacy rights of millions.

6 **D. ChatGPT Training on Users of Defendants' Programs and Applications.**

7 161. After using personal data taken without consent from millions of consumers to train
8 the Products initially, Defendants continued to train the AI on data gleaned from ChatGPT's
9 registered users and users of ChatGPT plug-ins with sponsoring applications ("Users"). Defendants
10 fed their AI models all of the data derived from User interactions—every click, entry, question, use,
11 every move, key stroke, search, User's geolocation (despite Users' unwillingness to share that
12 information)—as training data. Until recently, this also included all user interactions across the
13 hundreds or thousands of different platforms that now have ChatGPT plug-ins.

14 162. Following widespread criticism from consumers, OpenAI allegedly curtailed this
15 model of training their AI systems with user input, with CEO Sam Altman proclaiming broadly,
16 *"Customers clearly want us not to train on their data, so we've changed our plans: We will not*
17 *do that."*⁸³ However, what OpenAI did not make clear is that, according to the updated Terms of
18 Use, it will only purportedly refrain from training on data from *API users*, but "[it] may use Content
19 from Services other than our API ("Non-API Content") to help develop and improve our
20 Services."⁸⁴ That means Defendants continue to feed the inputted, collected, and stored data of the
21 millions of everyday ChatGPT users to train the AI Products, despite the Company's broad,
22 deliberately vague, and misleading pronouncement to the public that they "will not do that." OpenAI

23 ⁸² Chris Morris, *OpenAI is Reportedly Raising Funds at a \$29 Billion Valuation—and its*
24 *ChatGPT Could Challenge Google Search by Getting Wrapped into Microsoft Bing*, FORTUNE
25 (Jan. 6, 2023), <https://fortune.com/2023/01/06/openai-valuation-ai-chatgpt-microsoft-bing-google-search/>; Jagmeet Singh & Ingrid Lunden, *OpenAI Closes \$300M Share Sale at \$27-29B Valuation*,
26 TECH CRUNCH (Apr. 28, 2023), <https://techcrunch.com/2023/04/28/openai-funding-valuation-chatgpt/?tpcc=tcplustwitter>.

27 ⁸³ Baba Tamim, *OpenAI Changes AI Strategy, Won't Train ChatGPT on Customer Data, Says*
28 *Sam Altman*, INTERESTING ENG'G (May 6, 2023),
<https://interestingengineering.com/culture/openai-wont-train-chatgpt-on-customer-data>.

⁸⁴ *Terms of Use*, OPENAI, <https://openai.com/policies/terms-of-use> (last updated Mar. 14, 2023).

has also failed sufficiently to disclose that training aside (and even as to API users) it monitors, saves, and shares all the personal information collected with its partners, including Microsoft.

163. ChatGPT's systematic and intentional campaign to collect vast amounts of personal information from Users without their knowledge or consent includes any information a user inputs into the chat box with ChatGPT, as well as that user's account information, contact details, login credentials, IP addresses, and other sensitive personal information including analytics and cookies.⁸⁵

164. Defendants aggregate all of this data with the entirety of every internet user's digital footprint, scraped before ChatGPT was available for use, arming them with the largest corporate collection of personal online information ever amassed. Given Defendants' ongoing theft, this goldmine of valuable data is growing day by day, and with it, the concomitant risk to millions of consumers.

165. Indeed, even more stunning than Defendants' conversion of the internet for commercial gain, is they are "entrusting" all this personal information to large language models and unpredictable human-like "bots", while openly acknowledging that even they "don't understand how it works."⁸⁶ In the words of Mr. Altman himself, "the scary part" is that OpenAI's act of "putting this lever into the world *will for sure have unpredictable consequences*."⁸⁷ Dr. Yoshua Benigo, one of the three scientists who spent decades developing the technology that drives systems like ChatGPT-4, further explained: "Our ability to understand what could go wrong with very powerful A.I. systems is very weak. . . So we need to be careful."⁸⁸

166. To risk the personal data of millions by incorporating all of it into unpredictable Products, built on technology that even Defendants and leading scientists do not completely understand and thus, necessarily cannot safeguard, and *then* to deploy those Products worldwide for

⁸⁵ *Privacy Policy*, OPENAI <https://openai.com/policies/privacy-policy> (last updated June 23, 2023); Sarah Moore, *What Does ChatGPT Mean for Healthcare?*, NEWS MED. (Mar. 28, 2023), <https://www.news-medical.net/health/What-does-ChatGPT-mean-for-Healthcare.aspx>.

⁸⁶ Jan Leike (@janleike), TWITTER (May 17, 2023, 10:56 AM), <https://twitter.com/janleike/status/1636788627735736321>.

⁸⁷ Edward Felsenthal & Billy Perrigo, *OpenAI CEO Sam Altman Is Pushing Past Doubts on Artificial Intelligence*, TIME MAG. (June 21, 2023), <https://time.com/collection/time100-companies-2023/6284870/openai-disrupters/> (emphasis added).

⁸⁸ Cade Metz, *What Exactly Are the Dangers Posed By A.I.?*, THE N.Y. TIMES (May 7, 2023), <https://www.nytimes.com/2023/05/01/technology/ai-problems-danger-chatgpt.html>.

unfettered use, is the very definition of gross negligence.

E. Microsoft Pushes OpenAI's Economic Dependence Model

167. Although Defendants' most recent iteration of ChatGPT (GPT-4) was only recently released, Defendants have successfully encouraged and injected OpenAI's products into virtually every sector—from academia to healthcare. Instead of ensuring its safe launch of the AI models, Defendants recklessly began deploying the Products into every sector following the economic dependence model.

168. Microsoft has led the charge on the rapid proliferation of ChatGPT throughout the modern suite of technological applications—integrating the ChatGPT language model into almost all of its cardinal products and services,⁸⁹ thereby elevating the dangers of data misuse to unprecedented heights. Microsoft CEO Satya Nadella has indicated that the company plans to introduce AI into the remainder of its products in the future.⁹⁰

169. ChatGPT is integrated into Microsoft's search engine, Bing, which has approximately 100 million daily active users. ChatGPT has also been integrated into the interface of Microsoft's flagship communication and collaboration platform, Microsoft Teams, which has 250 million monthly active users.

170. Microsoft has also integrated the language model within its digital assistant platform, Cortana, which has an average of 141 million monthly active users.

171. Finally, within the Microsoft Dynamics 365 ecosystem, ChatGPT has been employed to power AI-driven customer service chatbots. This has enabled the chatbots to understand and respond to customer queries in a highly human-like manner, thereby significantly increasing the extent of information collected and thus, reducing the need for human intervention in support cases.

172. In a real sense, OpenAI now acts as a data scavenging company for Microsoft and

⁸⁹ These services include Bing, GitHub, Teams, and Viva Sales, among others. *See* Bernard Marr, *Microsoft's Plan to Infuse AI and ChatGPT Into Everything*, FORBES (Mar. 6, 2022), <https://www.forbes.com/sites/bernardmarr/2023/03/06/microsofts-plan-to-infuse-ai-and-chatgpt-into-everything/?sh=1adfd46653fc>.

⁹⁰ *Id.* ("Every product of Microsoft will have some of the same AI capabilities to completely transform the product.").

1 provides Microsoft with ChatGPT User and Non-User data belonging to millions of individuals.⁹¹

2 173. The integration of ChatGPT technology into Microsoft's primary products
3 significantly magnifies existing data privacy concerns. This move effectively enables the collection
4 of consumer information across a wide array of systems and platforms, encompassing a
5 comprehensive range of user interactions. The resultant collation of expansive consumer data
6 contributes to the construction of extensive user profiles.

7 174. This scope of data collection, coupled with user profiling, poses significant potential
8 risks. These risks extend not just to potential breaches of data privacy regulations, but also to the
9 erosion of consumer trust and the potential for misuse of sensitive information.

10 175. Rather than acknowledging these risks and taking steps to mitigate them, Microsoft
11 has laid off its entire "Responsible AI team," the 10,000 employees within Microsoft's ethics and
12 society group who were responsible for ensuring that ethical AI principles drive product design.⁹²
13 As one technology news outlet notes, "Data privacy, storage, or usage are probably just fluff talk
14 for . . . [Microsoft] anyway."⁹³

15 176. Other companies have rushed to keep pace, emulating Microsoft by pushing the
16 Products into nearly every conceivable application and service in the past six months of
17 development. As a result, GPT-4 has been integrated into hundreds of applications and platforms
18 over various industries.⁹⁴ According to a Gartner study, the commercial use of AI has increased
19 270% in the last 4 years, with 37% of businesses now using some form of AI technology. By other
20 accounts, the scale of commercial AI is even greater.

21 177. More specifically, AI in general, and OpenAI in particular, is now partnering with an
22 extraordinary number of influential organizations, spreading across the internet completely
23

24
25 ⁹¹ Pandey, *supra* note 30.

26 ⁹² Poulomi Chatterjee, *Why Responsible AI is Just Fluff Talk for Microsoft, Others*, AIM (Mar. 18,
2023), <https://analyticsindiamag.com/why-responsible-ai-is-just-fluff-talk-for-microsoft-others/>.

27 ⁹³ Pandey, *supra* note 30

28 ⁹⁴ Bergur Thormundsson, *Amount of Companies Using ChatGPT in their Business Function in 2023, By Industry*, STATISTA (May 15, 2023),
<https://www.statista.com/statistics/1384323/industries-using-chatgpt-in-business/>.

unchecked.⁹⁵ This has seemingly happened overnight. It was just over six months ago that ChatGPT was released to the public.⁹⁶ In that short span of time, OpenAI integrated with the following major corporations, to name just a few: Snapchat,⁹⁷ Amazon, Microsoft, Expedia, Instacart, Google, BuzzFeed, KAYAK, Shutterstock, Zillow, Wolfram, as well as countless others⁹⁸—including everything from pioneering drug treatments in the health sector (Pfizer)⁹⁹ to optimizing dating applications (OkCupid).¹⁰⁰ At this point, it might be easier to list the companies that have not partnered with OpenAI, or that are not investing in their own AI solutions.

178. As is clear, OpenAI has exploded outwards in every direction within the past few months and is swiftly morphing into something intimately connected with people in nearly every aspect of their day-to-day lives. There is no check or boundary on this expansion, which seems to progress rapidly every single day.

⁹⁵ Beth Floyd, *ChatGPT Plugins*, ROE DIGIT. (May 5, 2023), <https://roedigital.com/ChatGPT-plugins/>.

⁹⁶ Alyssa Stringer & Kyle Wiggers, *ChatGPT: Everything You Need to Know About the AI-Powered Chatbot*, TECHCRUNCH (May 3, 2023), https://techcrunch.com/2023/05/03/chatgpt-everything-you-need-to-know-about-the-ai-powered-chatbot/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAA-Ab2tIJ3WAdxAd5xb2pWmCPSFqzTyqRmMHEOaaOXsH04KD_DgCLfExvNPrnVX4ioR-uMFVQjAawiyhp5m21A3SqmsPYHv2yHSgfiIdjokmMe981-hq51XH5pWxCfLZOOWwf2wlvK3MnVewrZk4MRmPRAC8ArJXbegg6dnL2-f.

⁹⁷ Snapchat recently released “My AI,” a ChatGPT-fueled chatbot feature open to all Snapchat users. See Alex Hern, *Snapchat Making AI Chatbot Similar to ChatGPT Available to Every User*, THE GUARDIAN (Apr. 19, 2023), <https://www.theguardian.com/technology/2023/apr/19/snapchat-making-ai-chatbot-similar-to-chatgpt-available-to-every-user>. My AI now appears for Snapchat users as a contact in their social network, allowing users to ask it questions, have back and forth conversations, ask it to generate creative content, and much more. *Id.*

⁹⁸ Floyd, *supra* note 95; Silvia Pellegrino, *Which Companies Have Partnered With OpenAI*, TECHMONITOR (Jan. 18, 2023), <https://techmonitor.ai/technology/which-companies-have-partnered-with-openai>; Asif Iqbal, *OpenAI’s Collaborations: Pushing the Boundaries of AI in Various Sectors*, LINKEDIN (Mar. 12, 2023), <https://www.linkedin.com/pulse/openai-collaborations-pushing-boundaries-ai-various-sectors-iqbal/>.

⁹⁹ Iqbal, *supra* note 98 (“In 2020, OpenAI announced a collaboration with drug manufacturer, Pfizer, to develop new AI technologies for drug discovery.”).

¹⁰⁰ Danni Button, *ChatGPT Poses Danger for Online Dating Apps*, THE STREET (Feb. 15, 2023), <https://www.thestreet.com/social-media/chatgpt-poses-dangers-for-online-dating-apps>.

II. Risks from Unchecked AI Proliferation

A. The International Community Agrees that Unchecked & Lawless AI Proliferation Poses an Existential Threat

179. The unregulated development of AI technology has led to the creation of powerful tools being used to manipulate public opinion, spread false information, and undermine democratic institutions. Further development of such powerful tools will supercharge the dissemination of propaganda, the amplification of extremist voices, and the influencing of elections based on undetectable falsehoods.

180. The United States has been particularly affected by the rapid development of AI technology, as the absence of effective regulations has accelerated the proliferation of unaccountable and untrustworthy AI tools. Even the White House has acknowledged that AI presents “the most complicated tech policy discussion possibly that [the country] has ever had.”¹⁰¹

“I am confident AI will be used by bad actors, and yes it will cause real damage.”¹⁰² -

Michael Schwarz, Microsoft’s Chief Economist

“If law and due process are absent from this field, we are essentially paving the way to a new feudal order of unaccountable reputational intermediaries.” - Professors Danielle Keats Citron and Frank Pasquale at 2023 Geneva Conference.¹⁰³

AI technology is so powerful that it even has the potential to “allow an evil country, competitor to come in and screw up our democracy.”¹⁰⁴ - Eric Schmidt, Former Google CEO and Chairman at the 2023 Milken Global Conference.

181. In a report addressed to the American public in 2021, Eric Schmidt and Robert Work, the chair and vice chair of the National Security Commission on Artificial Intelligence (“NSCAI”),

¹⁰¹ Ben Wershkul & Alexandra Garfinkle, *White House bringing Google, Microsoft CEOs together for ‘frank discussion’ of AI*, YAHOO! FIN. (May 4, 2023), <https://www.aol.com/finance/white-house-bringing-alphabet-microsoft-164428066.html>.

¹⁰² Bryce Baschuk, *Microsoft Economist Warns Bad Actors Will Use AI to Cause Damage*, MSN (May 3, 2023), <https://www.msn.com/en-us/money/other/ai-will-cause-real-damage-microsoft-chief-economist-warns/ar-AA1aFslV>.

¹⁰³ *EPIC AI Rulemaking Petition*, EPIC, <https://epic.org/documents/epic-ai-rulemaking-petition/> (last visited June 27, 2023).

¹⁰⁴ Wershkul, *supra* note 101.

1 noted that “Americans have not yet grappled with just how profoundly the artificial intelligence
2 revolution will impact our economy, national security, and welfare. Much remains to be learned
3 about the power and limits of AI technologies. Nevertheless, ***big decisions need to be made now***...to
4 defend against the malignant uses of AI.”¹⁰⁵

5 182. The NSCAI report highlights the consequences associated with the unregulated
6 development of AI, emphasizing the unique risks to human rights, privacy, and personal autonomy.
7 Further, the report notes the urgency of establishing comprehensive privacy frameworks and
8 regulations that strike a balance between protecting individuals’ privacy rights and enabling AI
9 advancements.

10 183. On March 30, 2023, a new complaint was filed to the Federal Trade Commission
11 (“FTC”), urging the agency to investigate OpenAI and suspend its commercial deployment of large
12 language models, including its latest iteration of the popular tool ChatGPT.¹⁰⁶ The complaint notes
13 that the use of AI should be “transparent, explainable, fair, and empirically sound while fostering
14 accountability.”¹⁰⁷ None of the Products satisfy these requirements.

15 184. The significance of harm facing our society is in fact so imminent that Geoffrey
16 Hinton—referenced by many as the “godfather” of AI—quit his job at Google where he had worked
17 for more than a decade, becoming one of the most respected voices in the field, so he could freely
18 speak out about the dangers associated with the rapid, uncontrolled development and release of AI
19 to our society.

20 185. Dr. Hinton’s journey from AI groundbreaker to AI whistleblower marks a remarkable
21 moment for the AI technology industry at perhaps its most important inflection point in decades.
22 Industry leaders believe the new A.I. systems could be as important but yet as catastrophic as the
23 development of nuclear weapons.

24 186. After OpenAI released ChatGPT in March, more than 1,000 technology leaders and
25

26 ¹⁰⁵ Eric Schmidt & Bob Work, *Letter from the Chair and Vice Chair*, NAT’L. SEC. COMM’N. ON
A.I., (2021), <https://reports.nsc.ai.gov/final-report/chair-and-vice-chair-letter>.

27 ¹⁰⁶ Federal Trade Commission, *In the matter of OpenAI, Inc.*, FED. TRADE. COMM’N. (Mar. 30,
2023), [https://cdn.arstechnica.net/wp-content/uploads/2023/03/CAIDP-FTC-Complaint-OpenAI-](https://cdn.arstechnica.net/wp-content/uploads/2023/03/CAIDP-FTC-Complaint-OpenAI-GPT-033023.pdf)
28 [GPT-033023.pdf](https://cdn.arstechnica.net/wp-content/uploads/2023/03/CAIDP-FTC-Complaint-OpenAI-GPT-033023.pdf).

¹⁰⁷ *Id.*

1 researchers signed an open letter calling for a six-month moratorium on the development of new
2 systems because A.I. technologies pose “profound risks to society and humanity.”¹⁰⁸

3 187. Several days later, 19 current and former leaders of the Association for the
4 Advancement of Artificial Intelligence, a 40-year-old academic society, released their own letter
5 warning of the risks of A.I. That group included Eric Horvitz, chief scientific officer at Microsoft,
6 which has deployed OpenAI’s technology across a wide range of products, including its Bing search
7 engine.¹⁰⁹

8 188. The Letter, issued by the Future of Life Institute, states:

9 **Powerful AI systems should be developed only once we are confident**
10 **that their effects will be positive and their risks will be manageable . . .**
11 **we call on all AI labs to immediately pause for at least 6 months the**
12 **training of AI systems more powerful than GPT-4.** AI research and
13 development should be refocused on making today’s powerful, state-of-the-
14 art systems more accurate, safe, interpretable, transparent, robust, aligned,
15 trustworthy, and loyal.¹¹⁰

16 189. The Letter continues: “In parallel, AI developers must work with policymakers to
17 dramatically accelerate development of robust AI governance systems. These should at a minimum
18 include new and capable regulatory authorities dedicated to AI; . . .”¹¹¹

19 190. Generative AI models are unusual consumer products because they exhibit behaviors
20 that may not have been previously identified by the company that released them for sale. OpenAI
21 acknowledged the risk of “Emergent Risky Behavior” and nonetheless chose to go forward with the
22 commercial release of ChatGPT. As OpenAI explained: novel capabilities often emerge in more
23 powerful models. Some that are particularly concerning are the ability to create and act on long-
24 term plans, to accrue power and resources (“power-seeking”), and to exhibit behavior that is

25
26 ¹⁰⁸ *The ‘Godfather of A.I.’ Leaves Google and Warns of Danger Ahead*, DNYUZ (May 1, 2023),
<https://dnyuz.com/2023/05/01/the-godfather-of-a-i-leaves-google-and-warns-of-danger-ahead/>.

27 ¹⁰⁹ *Id.*

28 ¹¹⁰ *Pause Giant AI Experiments: An Open Letter*, FUTURE OF LIFE INST. (Mar. 29, 2023),
<https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (emphasis in the original).

¹¹¹ *Id.*

1 increasingly “agentic.”¹¹²

2 191. In February 2020, a petition with the Federal Trade Commission called on the FTC to
3 conduct rulemaking for the use of artificial intelligence in commerce. “Given the scale of
4 commercial AI use, the rapid pace of AI development, and the very real consequences of AI-enabled
5 decision-making for consumers, [courts] should immediately initiate a rulemaking to define and
6 prevent consumer harms resulting from AI.”¹¹³

7 192. Multiple sources have called on the FTC to enforce the AI standards established in
8 the OECD AI Principles, the OMB AI Guidance, and the Universal Guidelines for AI. Several FTC
9 Commissioners have already acknowledged the FTC’s role in regulating the use of AI.

10 193. The absence of effective AI regulations in the United States has accelerated the spread
11 of unaccountable and untrustworthy AI tools. And the unregulated use of those AI tools has already
12 caused serious harm to consumers, who are increasingly subject to opaque and unprovable decision-
13 making in employment, credit, healthcare, housing, and criminal justice.

14 194. Realizing the gravity of potential harm, authorities within European countries took
15 ChatGPT offline in Italy in April after the country’s data protection authority temporarily banned
16 the chatbot and launched a probe into the artificial intelligence application’s suspected breach of
17 privacy rules.¹¹⁴

18 195. Italian authorities stated that ChatGPT has an “absence of any legal basis that justifies
19 the massive collection and storage of personal data” to “train” the chatbot.¹¹⁵ Further, they accused
20 Defendant OpenAI of failing to check the age of ChatGPT’s users to ensure they are aged 13 or
21

22 ¹¹² Dennis Layton, *GPT-4 – Some First Impressions*, LINKEDIN (Mar. 15, 2023),
23 <https://www.linkedin.com/pulse/gpt-4-some-first-impressions-dennis-layton> (“Agentic in this
24 context does not intend to humanize language models or refer to sentience but rather refers to
25 systems characterized by the ability to, e.g., accomplish goals which may not have been concretely
26 specified and which have not appeared in training; focus on achieving specific, quantifiable
27 objectives; and [engage in] long-term planning.”).

28 ¹¹³ *EPIC AI Rulemaking Petition*, *supra* note 103.

¹¹⁴ Supantha Mukherjee & Giselda Vagnoni, *Italy Restores ChatGPT After OpenAI Responds to Regulator*, YAHOO! (Apr. 28, 2023), <https://finance.yahoo.com/news/chatgpt-available-again-users-italy-163139143.html>.

¹¹⁵ Elvira Pollina & Supantha Mukherjee, *Italy Curbs ChatGPT, Starts Probe Over Privacy Concerns*, REUTERS (Mar. 31, 2023), <https://www.reuters.com/technology/italy-data-protection-agency-opens-chatgpt-probe-privacy-concerns-2023-03-31/>.

1 above.¹¹⁶

2 196. Subsequently, Defendant OpenAI agreed to offer specific tools to verify Users' ages
3 in Italy upon sign-up, but yet continues to enable unverified access in the United States to illegally
4 collect the personal data of minors. Defendant OpenAI also said that it would provide greater
5 visibility of its privacy policy and user content opt-out form, creating a new form for European
6 Union users to exercise their right to object to its use of personal data to train its models. The form
7 requires people who want to opt out to provide detailed personal information, including evidence of
8 data processing via relevant prompts. However, despite consumers' established privacy rights to be
9 "forgotten," Defendants cannot effectively extract individuals' information from the Products once
10 the AI is trained on such information.¹¹⁷

11 197. Italy was the first western European country to curb ChatGPT, but its rapid
12 development has attracted attention from lawmakers and regulators in several countries. A
13 committee of European Union lawmakers agreed on new rules that would force companies
14 deploying generative AI tools, such as ChatGPT, to disclose any copyrighted material used to
15 develop their systems.¹¹⁸

16 198. Data authorities from around the world remain concerned, specifically, with "the lack
17 of legal basis underpinning the massive collection, use and disclosure of personal information in
18 order to train the ChatGPT algorithms on which the platform relies" and the "cornerstone privacy
19 issue" at the heart of this Action: ChatGPT's "use of web scraping and the collection of personal
20 information without consent."¹¹⁹

21 _____
22 ¹¹⁶ *Id.*

23 ¹¹⁷ *ChatGPT and Education*, CNT. FOR INNOVATIVE TEACHING AND LEARNING,
24 <https://www.niu.edu/citl/resources/guides/chatgpt-and-education.shtml>, (last visited June 26,
25 2023) ("the prompts that you input into ChatGPT cannot be deleted. If you, or your students, were
26 to ask ChatGPT about sensitive or controversial topics, this data cannot be removed.").

27 ¹¹⁸ Supantha Mukherjee & Giselda Vagnoni, *Italy Restores CHATGPT after OpenAI Responds to*
28 *Regulator*, SRN NEWS (Apr. 28, 2023), [srnnews.com/italy-restores-chatgpt-after-openai-responds-](https://srnnews.com/italy-restores-chatgpt-after-openai-responds-to-regulator-2/)
[to-regulator-2/](https://srnnews.com/italy-restores-chatgpt-after-openai-responds-to-regulator-2/).

¹¹⁹ Roland Hung, *AI Technology and Privacy: Canadian Privacy Commissioner Launches*
Investigation into ChatGPT, TORKIN MANES (Apr. 24, 2023), [https://www.torkinmanes.com/our-](https://www.torkinmanes.com/our-resources/publications-presentations/publication/ai-technology-and-privacy-canadian-privacy-)
[resources/publications-presentations/publication/ai-technology-and-privacy-canadian-privacy-](https://www.torkinmanes.com/our-resources/publications-presentations/publication/ai-technology-and-privacy-canadian-privacy-)

199. In short, the message is consistent from informed business, nonprofit, and technology thought leaders; industrialists; scientists; world leaders; regulators; and governments around the globe: The proliferation of AI—including Defendants’ products—pose an existential threat if not constrained by the reasonable guardrails of our laws and societal mores. Defendants’ business and scraping practices raise fundamentally important legal and ethical questions that must also be addressed. Enforcing the law will not amount to stifling AI innovation, but rather a safe and just AI future for all.

8 **B. Overview of Risks**

200. The following is a brief, non-exhaustive list of ongoing harms and critical legal threats the Products pose to everyday Americans, including Plaintiffs and the Proposed Class Members.

11 **1. Massive Privacy Violations**

201. In today’s vast, interconnected digital landscape, privacy can appear to be more of an illusion, but it is still a guaranteed right. In violation of this right, the Products operate as an all-seeing online platform, tracking our every move: each click, each site visit, each chat—not allowing anything to escape its relentless scrutiny. Internet users’ interactions, seemingly innocuous, are aggregated, filtered, and compiled by Defendants, rendering the concept of privacy virtually non-existent. Even information deemed private or intended for a restricted audience does not escape surveillance.

202. The massive, unparalleled collection and tracking of users’ personal information by Defendants endangers individuals’ privacy and security to an incalculable degree. This information can be exploited and used to perpetrate identity theft, financial fraud, extortion, and other malicious purposes. It can also be employed to target vulnerable individuals with predatory advertising,

commissioner-launches-investigation-into-chatgpt (detailing the “privacy concerns with the use of ChatGPT” that have been raised worldwide). *See also* Heinrich Long, *Authorities Press OpenAI to Disclose How ChatGPT Input Is Used*, RESTORE PRIV. (June 9, 2023), <https://restoreprivacy.com/authorities-press-openai-to-disclose-how-chatgpt-input-is-used/> (discussing worldwide investigations, including the latest inquiry from Dutch data protection authorities who “want[] to know, among other things, how OpenAI handles personal data when training the underlying system. The[y...] want[] to know from OpenAI whether people’s questions are used to train the algorithm, and if so, in what way. The[y...] also ha[ve] questions about the way in which OpenAI collects and uses personal data from the internet.”).

1 algorithmic discrimination, and other unethical and harmful acts.

2 203. The collection and use of this data raises concerns about user privacy and the potential
3 misuse of personal information. For example, every iota of Users' activity is tracked and monitored.
4 By analyzing this data using algorithms and machine learning techniques, Defendants can develop
5 a chillingly detailed understanding of users' behavior patterns, preferences, and interests—creating
6 an entirely new meaning to the term “invasive.”

7 204. Several studies confirm that the collection and disclosure of sensitive information
8 from millions of individuals, as Defendants have done here, violates established expectations of
9 privacy based on long-standing social norms. Privacy polls and studies uniformly show that the
10 overwhelming majority of Americans consider one of the most important privacy rights to be the
11 need for an individual's affirmative consent before a company collects and shares its customers'
12 data.

13 205. For example, a recent study by Consumer Reports reveals that 92% of Americans
14 believe that internet companies and websites should be required to obtain consent before selling or
15 sharing consumers' data, and that internet companies and websites should be required to provide
16 consumers with a complete list of the data that has been collected about them.¹²⁰ Moreover,
17 according to a study by Pew Research Center, a majority of Americans, approximately 79%, are
18 concerned about how companies collect data about them.¹²¹

19 206. Users act consistently with these privacy preferences. Following a new rollout of the
20 iPhone operating software—which asks users for clear, affirmative consent before allowing
21 companies to track users—85% of worldwide users and 94% of U.S. users chose not to share data
22 when prompted.¹²² The Products' Users do not have that option, and do not understand the full extent
23

24 ¹²⁰ *Consumers Less Confident About Healthcare, Data Privacy, and Car Safety, New Survey*
25 *Findings*, CONSUMER REPS. (May 11, 2017), <https://www.consumerreports.org/consumer-reports/consumers-less-confident-about-healthcare-data-privacy-and-car-safety/>.

26 ¹²¹ Brooke Auxier et al., *Americans and Privacy: Concerned, Confused, and Feeling Lack of*
27 *Control over Their Personal Information*, PEW RSCH. CTR. (Nov. 15, 2019),
<https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/>.

28 ¹²² Margaret Taylor, *How Apple Screwed Facebook*, WIRED (May 19, 2021, 6:00 AM),
<https://www.wired.co.uk/article/apple-ios14-facebook>.

1 of Defendants' data collection and use of their personal data.

2 207. While the reams of personal information that Defendants collect on Users can be used
3 to provide personalized and targeted responses, it can also be used for exceedingly nefarious
4 purposes, such as tracking, surveillance, and crime. For example, if ChatGPT has access to a User's
5 browsing history, search queries, and geolocation, and combines this information with what
6 Defendant OpenAI has secretly scraped from the internet, Defendants could build a detailed profile
7 of Users' behavior patterns, including but not limited to where they go, what they do, with whom
8 they interact, and what their interests and habits are. This level of surveillance and monitoring raises
9 vital ethical and legal questions about privacy, consent, and the use of personal data. It is crucial for
10 users to be aware of how their data is being collected and used, and to have control over how their
11 information is shared and used by advertisers and other entities.

12 208. The concern about collecting and sharing information is compounded by the reality
13 that this information may include particularly sensitive information such as medical records or
14 information about minors. Increasingly, companies like Defendants "are harnessing and collecting
15 multiple typologies of children's data and have the potential to store a plurality of data traces under
16 unique ID profiles."¹²³

17 209. Given ChatGPT's ability to generate human-like understanding and responses, there
18 is a high likelihood that users might share (and already are sharing) their private health information
19 while interacting with the model, by asking health-related questions or discussing their medical
20 history, symptoms, or conditions. Moreover, this information can be logged and reviewed as part of
21 ongoing efforts to "train," improve and monitor each model's performance.

22 210. However, beyond these seemingly innocuous interactions with the AI, healthcare
23 industry providers are beginning to integrate ChatGPT in order to "revolutionize healthcare" while
24 undermining the confidentiality of individuals' personal data, which would be transmitted using
25
26

27 ¹²³ Veronica Barassi, *Tech Companies Are Profiling Us from Before Birth*, THE MIT PRESS
28 READER, (Jan. 14, 2021), <https://thereader.mitpress.mit.edu/tech-companies-are-profiling-us-from-before-birth/>.

ChatGPT and continuing to train Defendants' AI at the patients' expense.¹²⁴ While this technology could provide benefits, the risks associated with its implementation are drastic, from cybercrime, misinformation and misdiagnosis, lack of empathy and experience, and bias¹²⁵ to the existential risk, of which Altman has repeatedly warned.

211. ***Established Privacy Rights to be "Forgotten" Violated.*** Compounding this massive invasion of privacy, OpenAI offers no *effective* procedures at this time for individuals to request for their information/training data to be deleted. Instead, OpenAI simply provides an email address that consumers can contact if they would like to have their information removed. But this "option" is illusory. Regardless of whether individuals can technically request for ChatGPT to remove their data, it is not possible to do so completely, because Defendants train ChatGPT on individual inputs, personal information, and other user and nonuser data, which Defendants cannot reliably and fully extract from its trained AI systems any more than a person can "unlearn" the math they learned in sixth grade.

212. An AI researcher with privacy and cybersecurity firm AVG explains, "People are furious that data is being used without their permission. . . Sometimes, some people have deleted the[ir] [online] data but since the language model has already used them, the data is there forever. They don't know how to delete the data."¹²⁶

213. Likewise, some companies have banned or limited ChatGPT use because they are "worried that anything uploaded to AI platforms like OpenAI's ChatGPT or Google's Bard will [also] get *stored* on those companies' servers, *with no way to access or delete the information.*"¹²⁷

¹²⁴ Naomi Diaz, *6 Hospitals, Health Systems Checking Out ChatGPT*, BECKERS HEALTHCARE (June 2, 2023), <https://www.beckershospitalreview.com/innovation/4-hospitals-health-systems-testing-out-chatgpt.html>.

¹²⁵ Ethan Popowitz, *ChatGPT: Friend or Foe?*, DEFINITIVE HEALTHCARE, <https://www.definitivehc.com/blog/chatgpt> (last visited June 27, 2023).

¹²⁶ *Is ChatGPT's use of people's data even legal?*, AVG, <https://www.avg.com/en/signal/chatgpt-data-use-legal> (last visited June 27, 2023).

¹²⁷ Felicity Nelson, *Many Companies are Banning ChatGPT. This is Why*, SCI. ALERT (June 16, 2023), <https://www.sciencealert.com/many-companies-are-banning-chatgpt-this-is-why> (emphasis added). Microsoft has itself directed employees not to share sensitive data with ChatGPT "in case it's used for future AI training models" Diamond Naga Siu, *Microsoft is chill with employees using ChatGPT — just don't share 'sensitive data' with it*, YAHOO! NEWS (Feb. 1, 2023), <https://news.yahoo.com/microsoft-chill-employees-using-chatgpt-114000174.html?guccounter=1>.

214. The “right to be forgotten”—i.e., the right to request that a business delete the personal information that it holds about you—is guaranteed to California residents under the California Consumer Privacy Act of 2018 (“CCPA”) and to children under 13 nationwide under the Children’s Online Privacy Protection Act (“COPPA”). Given how the technology works, OpenAI is not compliant with these requirements.¹²⁸

2. *AI-Fueled Misinformation Campaigns, Targeted Attacks, Sex Crimes, and Bias*

215. ***Misinformation, Deepfakes, Clones, Scams, and Blackmail:*** The use of the Products facilitates the spreading of false or misleading information, even without “misuse.” That is because a *feature* (known defect) of ChatGPT’s *regular use* is the inventing of false information, including potentially defamatory information about individuals. Even the “improved” version (GPT4) “makes stuff up” and “may generated text that is completely false.”¹²⁹

216. One high-profile example involves a US law professor, Jonathan Turley, who ChatGPT falsely accused of sexually harassing one of his students, even providing a “source” for the purported crime via a news article that it invented.¹³⁰ Defendants call this “hallucination,” but the world knows it as defamation. While Defendants are allegedly “working on” a fix for this behavior, they continue to push the defective Product worldwide. Naturally, one would expect an ethical company “for the *benefit* of humanity” not to release such a Product, at all, *unless and until* it was safeguarded from committing crimes *against* humanity.

217. The Cambridge Analytica scandal—in which personal data was allegedly misused to target individuals with political propaganda and misinformation—is also an instructive cautionary

¹²⁸ See, e.g., Alexa Johnson-Gomez, *A “Living” AI: How ChatGPT Raises Novel Data Privacy Issues*, MINN. J. OF L., SCI. & TECH. BLOG (Feb. 6, 2023), <https://mjlst.lib.umn.edu/2023/02/06/a-living-ai-how-chatgpt-raises-novel-data-privacy-issues/> (dismissing purported compliance with CCPA as “in name only” given how the data is used as part of machine learning model).

¹²⁹ Cade Metz, *10 Ways GPT-4 is Impressive but Still Flawed*, THE N.Y. TIMES (Mar. 14, 2023), <https://www.nytimes.com/2023/03/14/technology/openai-new-gpt4.html>.

¹³⁰ Hern and Milmo, *supra* note 59.

1 tale.¹³¹ Cambridge Analytica collected personal data using third-party apps that collected data from
 2 users and their friends. It then used this data to build detailed profiles of individuals, so they could
 3 be targeted with personalized political ads and propaganda. Cambridge Analytica used algorithms
 4 and machine learning techniques to analyze this data, identify patterns in users' behavior and
 5 preferences, and target those users with specific messages and ads.

6 218. This history highlights the potential dangers of using personal data to build detailed
 7 profiles of individuals, particularly when that data is collected without their knowledge or consent.
 8 It also raises important questions about the ethics of using personal data for political purposes and
 9 the need for greater regulation and oversight of data collection and use.

10 219. Moreover, by allowing the collection, storage, and analysis of a massive amount of
 11 highly individualized, personal data—from audio and photographic data to detailed interests, habits,
 12 and preferences—OpenAI's technology facilitates the proliferation of video or audio “deepfakes”
 13 and makes them harder to detect.¹³² Simply put, the Products make it easier to create lifelike
 14 audiovisual digital duplicates--digital clones—of real people, which can then be used to spread
 15 misinformation, exploit victims, or even access privileged data.¹³³

16 220. Deepfakes could influence elections, erode public trust, and negatively affect public
 17 discourse.¹³⁴ The U.S. Congressional Research Service has further analyzed the risks of deepfakes,
 18 explaining that they could be used to “blackmail elected officials or individuals with access to
 19 classified information” and “generate inflammatory content [...] intended to radicalize populations,
 20

21 ¹³¹ Sam Meredith, *Here's Everything You Need to Know About the Cambridge Analytica Scandal*,
 22 CNBC (Mar. 21, 2018), [https://www.cnbc.com/2018/03/21/facebook-cambridge-analytica-](https://www.cnbc.com/2018/03/21/facebook-cambridge-analytica-scandal-everything-you-need-to-know.html)
 23 [scandal-everything-you-need-to-know.html](https://www.cnbc.com/2018/03/21/facebook-cambridge-analytica-scandal-everything-you-need-to-know.html). (The Cambridge Analytica scandal involved the
 24 misuse of personal data collected from Facebook users, which was then used to target individuals
 25 with political advertising and propaganda. The scandal highlighted the potential dangers of using
 26 personal data for targeted advertising and the need for greater transparency and accountability in
 27 the collection and use of personal information.).

28 ¹³² Bibhu Dash & Pawankumar Sharma, *Are ChatGPT and Deepfake Algorithms Endangering the*
Cybersecurity Industry? A Review, 10(1) I. J. OF ENG'G & APPLIED SCI. (Jan. 2023),
https://www.ijeas.org/download_data/IJEAS1001001.pdf.

¹³³ *Science & Tech Spotlight: Deepfakes*, U.S. GOV'T ACCOUNTABILITY OFF. (Feb. 20, 2020),
<https://www.gao.gov/products/gao-20-379sp>; see also Dash & Sharma, *supra* note 132.

¹³⁴ Kelley M. Sayler & Laurie A. Harris, *Deep Fakes and National Security*, CONG. RSCH. SERV.,
 (April 17, 2023), <https://crsreports.congress.gov/product/pdf/if/if11333>.

1 recruit terrorists, or incite violence.¹³⁵

2 221. In addition to spreading misinformation, criminals have used, and will continue to use
3 this technology to harass, blackmail, extort, coerce, and defraud. Armed with artificial intelligence
4 tools like the ones developed by Defendants, malicious actors can weaponize even the most
5 innocuous publicly available personal information, such as names and photographs, against private
6 individuals.

7 222. For example, the FBI has issued an alert about a particularly despicable form of
8 blackmail currently on the rise that has been largely facilitated by AI like the Products. This scheme,
9 a form of “sextortion,” is perpetrated using artificial intelligence tools and publicly available
10 photographs and videos of private individuals, usually obtained through social media, to create
11 deepfakes containing pornographic content.¹³⁶ The photos or videos are then publicly circulated on
12 social media, public forums, and pornographic websites for the purpose of harassing the victim,
13 causing extreme emotional and psychological distress.¹³⁷

14 223. A malicious actor may also attempt to extract ransom payments, sometimes seeking
15 genuine versions of the subject engaging in the acts depicted in the made up sexually-explicit images
16 and videos, by threatening to share the falsified images or videos with family members, social
17 contacts, or by indiscriminately circulating the content on social media.¹³⁸ The most concerning and
18 egregious aspect of this type of “sextortion” scheme is that the victims include not only non-
19 consenting adults, but also minor children.¹³⁹

20 224. ***Child Pornography.*** Defendants’ Product Dall-E has become a favorite tool for
21 pedophiles, because it requires less technical competence than previous programs used by
22

23
24
25 ¹³⁵ *Id.*

26 ¹³⁶ *Public Service Announcement: Malicious Actors Manipulating Photos and Videos to Create*
27 *Explicit Content and Sextortion Schemes*, FED. BUREAU OF INVESTIGATION (June 5, 2023),
<https://www.ic3.gov/Media/Y2023/PSA230605>.

28 ¹³⁷ *Id.*

¹³⁸ *Id.*

¹³⁹ *Id.*

pedophiles and increases the scale at which images of virtual child pornography can be created.¹⁴⁰ In just mere seconds, Dall-E can create realistic images of children performing sex acts.¹⁴¹ Thousands of such images have already been detected in dark web forums.¹⁴² In a dark web forum with 3,000 subscribers, 80% of respondents to an internal poll stated “they had used or intended to use AI tools to create child sexual abuse images.”¹⁴³ In such forums, users exchange strategies for thwarting the woefully insufficient purported “safety guardrails” of Dall-E and other AI products, “including by using non-English languages they believe are less vulnerable to suppression or detection.”¹⁴⁴

225. Dall-E is a diffusion model, and anyone can access it, generating a realistic image solely by typing a short description of the desired product.¹⁴⁵ This model was trained off billions of images taken, without notice or consent, from the internet, “many of which showed real children and came from photo sites and personal blogs.”¹⁴⁶ Images of actual children are thus the source material for the AI-generated child pornography. In some instances, actual images of existing child pornography were used to train the model and generate further explicit material of already victimized children, thereby victimizing them all over again.¹⁴⁷

226. AI-generated child pornography has introduced a slew of other horrendous problems as well. “The flood of images could confound the central tracking system built to block such material from the web because it is designed only to catch known images of abuse, not detect newly generated ones.”¹⁴⁸ Moreover, the monumental task of locating children harmed by the production of child pornography has been bogged down now that agents must now spend time puzzling over whether

¹⁴⁰ Drew Harwell, *AI-generated Child Sex Images Spawn New Nightmare for the Web*, WASH. POST (June 19, 2023, 7:00 AM), <https://www.washingtonpost.com/technology/2023/06/19/artificial-intelligence-child-sex-abuse-images/>.

¹⁴¹ *Id.*

¹⁴² *Id.*

¹⁴³ *Id.*

¹⁴⁴ *Id.*

¹⁴⁵ *Id.*

¹⁴⁶ *Id.*

¹⁴⁷ *Id.*

¹⁴⁸ *Id.*

content is real or virtual.¹⁴⁹ Furthermore, this virtual material is not merely used by pedophiles to supplant real material.¹⁵⁰ AI is also being used to “build [] fake school-age persona[s]” via fabricated selfies, which are incorporated into plots to lure and groom child targets.¹⁵¹

227. Absent the injunctive relief sought in this action, Defendants will continue to not only steal data from unwitting victims, including minors, but arm pedophiles in rapidly generating child pornography at scale and in creating materials that can be strategically used to groom and victimize real children.

228. ***Hate and Bias.*** Continued commercial deployment of the Products also will amplify and entrench the human biases and prejudices reflected in the Products’ sources, which Defendants used without regard to such factors by incorporating and training the Products with content from various extremist websites and by failing to use adequate filtering safeguards.¹⁵²

3. ***Hypercharged Malware Creation***

229. ***Malicious, Mutating, and Virtually Undetectable Code Scripts:*** Malware, or malicious software, are computer programs designed to damage or infiltrate computer systems. Unscrupulous actors deploy malware by embedding them within vulnerabilities in existing internet applications.¹⁵³ The Products guarantee that “malware” prevalence and potency will exponentially increase, posing unprecedented cybersecurity risks on a global scale. That is because the Products can generate virtually undetectable malware, and at massive scale, to thwart security systems and jeopardize entire governments.

230. Malware attacks have sabotaged entire governments before. For example, in 2022,

¹⁴⁹ *Id.*

¹⁵⁰ *Id.*

¹⁵¹ *Id.*

¹⁵² Sam Biddle, *The Internet’s New Favorite AI Proposes Torturing Iranians and Surveilling Mosques*, THE INTERCEPT (Dec. 8, 2022), <https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/>.

¹⁵³ Fei Xiao et al., *A Novel Malware Classification Method Based on Crucial Behavior*, 2020 MATHEMATICAL PROBS. IN ENG’G. (Mar. 21, 2020), <https://doi.org/10.1155/2020/6804290>; Rabia Tahir, *A Study on Malware and Malware Detection Techniques*, 2 INT’L J. OF MGMT. ENG’G., 20, 20 (Mar. 8, 2018), <https://www.mecs-press.net/ijeme/ijeme-v8-n2/IJEME-V8-N2-3.pdf>; Mohd Faizal Ab Razak et al., *The Rise of “Malware”: Bibliometric Analysis of Malware Study*, 75 J. OF NETWORK AND COMPUT. APPLICATIONS, 58, 58 (Nov. 2016), <https://www.sciencedirect.com/science/article/pii/S1084804516301904>.

the Russian Conti Group enacted a weeks-long attack on 27 different ministries in the Costa Rican government.¹⁵⁴ The malware deployed was ransomware, a software that encrypts critical information, denying access to its rightful owner and threatening its destruction if payment is not made.¹⁵⁵ Costa Rica's president declined to pay the \$20M ransom, but a standoff ensued leaving parts of Costa Rica's digital infrastructure in shambles, disrupting public healthcare and the pay of its workers.¹⁵⁶

231. Healthcare providers are also often targeted by malware, and increasingly so. For example, a major software provider for the UK's National Health System sustained a ransomware attack from an unknown group last summer.¹⁵⁷ The attack had real impact on the health of millions, disrupting ambulance dispatch, appointment scheduling, and emergency prescriptions, among other things.¹⁵⁸ Ransomware attacks on health care providers have doubled from 2016 to 2021, exposing the sensitive health information of 42M individuals.¹⁵⁹

232. ***The Products supercharge Malware:*** In 2012, 33% of malware went undetected by antivirus software.¹⁶⁰ In the last decade, malware has become ever more sophisticated, and ever more capable of thwarting detection. But now, with the assistance of the Products, malware can become undetectable in new ways, at scale, because ChatGPT can be used to create "mutating, or polymorphic" malware.¹⁶¹ Polymorphic malware has a mutation engine with self-propagating code

¹⁵⁴ Christine Murray & Mehul Srivastava, *How Conti Ransomware Group Crippled Costa Rica-Then Fell Apart*, FIN. TIMES (July 9, 2022), <https://www.ft.com/content/9895f997-5941-445c-9572-9cef66d130f5>.

¹⁵⁵ *Id.*

¹⁵⁶ *Id.*

¹⁵⁷ Vedere Labs, *Ransomware in Healthcare: The NHS Example and What the Future Holds*, SEC. BOULEVARD (Aug. 25, 2022), <https://securityboulevard.com/2022/08/ransomware-in-healthcare-the-nhs-example-and-what-the-future-holds/>.

¹⁵⁸ *Id.*

¹⁵⁹ Hannah T. Neprash et al., *Trends in Ransomware Attacks on US Hospitals, Clinics, and Other Health Care Delivery Organizations, 2016–2021*, JAMA HEALTH FORUM (Dec. 29, 2022), <https://jamanetwork.com/journals/jama-health-forum/fullarticle/2799961>.

¹⁶⁰ Markus Kammerstetter et al., *Vanity, Cracks, and Malware: Insights into the Anti-Copy Protection Ecosystem*, ASS'N. FOR COMPUTING MACHINERY 809, 818 (Oct. 16, 2012), <https://doi.org/10.1145/2382196.2382282>.

¹⁶¹ Shweta Sharma, *ChatGPT Creates Mutating Malware That Evades Detection by EDR*, CSO ONLINE (June 6, 2023, 1:59 PM), <https://www.csoonline.com/article/3698516/chatgpt-creates-mutating-malware-that-evades-detection-by-edr.html>.

that allows it to rapidly change its appearance and composition.¹⁶² This malware can change its entire make-up, so that malware detectors, reactionary by nature, will not recognize its newer, ongoing permutations.¹⁶³

233. ChatGPT can build the requisite polymorphic code, using its API at runtime to deploy advanced malware attacks that evade detection by security systems designed to thwart malware, such as endpoint detection and response (EDR) applications.¹⁶⁴ Recently, researchers designed a simple, executable file that corresponds with ChatGPT's API in real time "to generate dynamic, mutating versions of malicious code," making it extremely difficult to detect using existing cybersecurity tools.¹⁶⁵

234. While the most recent iterations of ChatGPT purport to "disallow" potential prompt injections for generating polymorphic malware, this supposed guardrail for safety is woefully inadequate: cleverly worded inputs, used by developers of malware, easily circumvent ChatGPT's content filters with a practice commonly referred to as "prompt engineering."¹⁶⁶

235. Thus, Mackenzie Jackson, developer advocate at cybersecurity company GitGuardian warns that, as generative models become more advanced, "AI may end up creating malware that can only be detected by other AI systems for defense. What side will win at this game is anyone's guess."¹⁶⁷ To knowingly put this enhanced ability to sabotage governments, health care systems, and any other number of targets into the hands of everyday people worldwide without adequate safeguards is emblematic of Defendants' gross negligence and underscores the need for immediate judicial intervention.

4. *Autonomous Weapons*

236. AI also poses a unique threat to international security and human rights through the development of autonomous weapons known as "Slaughterbots," otherwise known as "lethal autonomous weapons systems" or "killer robots," which are weapons systems that use AI to

¹⁶² *Id.*

¹⁶³ *Id.*

¹⁶⁴ *Id.*

¹⁶⁵ *Id.*

¹⁶⁶ *Id.*

¹⁶⁷ *Id.*

1 identify, select, and kill human targets without intervention.¹⁶⁸ As one humanitarian organization
 2 explained, “[w]eapons that use algorithms to kill, rather than human judgment, are immoral and a
 3 grave threat to national and global security.”¹⁶⁹

4 237. The risk that unregulated AI like the Products pose via autonomous weapons is “not
 5 a far-fetched danger for the future, but a clear and present danger.”¹⁷⁰ Such weapons have already
 6 nearly killed a foreign head of state, and due to the rapid commercial proliferation of open-source
 7 AI, “could be built today by an experienced hobbyist for less than \$1,000.”¹⁷¹

8 238. Defendants’ conduct exacerbates the problem. There is already an early autonomous
 9 implementation of ChatGPT known as “Chaos GPT” which is being touted as “empowering GPT
 10 with Internet and Memory to Destroy Humanity.”¹⁷² Chaos-GPT is predicated on an open source
 11 application that uses Defendants’ GPT-4, and was designed by an anonymous user to be a
 12 “destructive, power-hungry, manipulative AI.”¹⁷³ With only those parameters set by the user,
 13 Chaos-GPT returned a list of objectives it set for itself. One was to “destroy humanity.” Another
 14 was to “cause chaos and destruction” by creating “widespread suffering.”¹⁷⁴ Next, Chaos-GPT, of
 15 its own “volition,” prepared a plan in support of these objectives – and then it searched the internet
 16 for weapons of mass destruction seeking to obtain one.¹⁷⁵

17 239. Experts warn that advancements in AI like those accomplished by the Products, “will
 18 accelerate the near-term future of autonomous weapons.”¹⁷⁶ While it is believed artificial
 19 intelligence at a level equal to or higher than human intelligence is a prerequisite to truly
 20

21 ¹⁶⁸ See *Slaughterbots Are Here*, AUTONOMOUS WEAPONS (Feb. 23, 2023),
 22 <https://autonomousweapons.org/> (discussing Latin American and the Caribbean Conference on the
 23 Social and Humanitarian Impact of Autonomous Weapons).

24 ¹⁶⁹ *Id.*

25 ¹⁷⁰ Kai-Fu Lee, *The Third Revolution in Warfare*, THE ATLANTIC (Sept. 11, 2021),
 26 <https://www.theatlantic.com/technology/archive/2021/09/i-weapons-are-third-revolution-warfare/620013/>.

27 ¹⁷¹ *Id.*

28 ¹⁷² Jose Antonio Lanz, *Meet ChaosGPT: An AI Tool That Seeks to Destroy Humanity*, DECRYPT
 (Apr. 13, 2023), <https://decrypt.co/126122/meet-chaos-gpt-ai-tool-destroy-humanity>.

¹⁷³ *Id.*

¹⁷⁴ *Id.*

¹⁷⁵ *Id.*

¹⁷⁶ Lee, *supra* note 170.

1 autonomous weaponry, the unfettered commercial deployment of the Products naturally escalates
 2 this risk as their widespread use continually “enhances” the AI’s capabilities – and without sufficient
 3 moral or ethical guardrails, as sought in this Action.

4 **C. Opportunity on the Other Side**

5 240. While leading experts agree on the grave risks posed by the Products, and the need
 6 for a temporary pause in their commercial deployment, it is important to understand the full picture
 7 of why this Action matters. It is not just to contain the risks to society and harms happening right
 8 now, including the supercharged spread of disinformation, the obliteration between truth and fiction,
 9 deepfakes designed to harass, harm, and commit fraud, and more. It is not just to halt Defendants’
 10 ongoing disregard for the privacy and property interests of millions, and to remedy those violations.
 11 It is not just to avoid the collapse of civilization as we know it and as Mr. Altman himself recognizes
 12 is possible.¹⁷⁷ Naturally, all of these things warrant the comparatively measured relief Plaintiffs and
 13 the Classes seek. But beyond all of this, the Action matters to ensure humankind can *realize the*
 14 *tremendous opportunity for advancement and prosperity* that awaits us, on the other side of a
 15 commercial pause.

16 241. By pausing now, “[h]umanity can enjoy a flourishing future.”¹⁷⁸ It will enable the
 17 joint development and implementation of shared safety protocols, overseen by independent outside
 18 experts, to manage the risks and render the Products safe to usher in an exciting new era of progress
 19 for all. For example, with adequate safeguards, the Products will be positioned to revolutionize
 20 healthcare for good, by helping to discover new drugs to save lives and potentially find cures for
 21 cancer and other deadly diseases. With adequate safeguards, the Products can contribute not only to
 22 our everyday efficiency, artistic expression, joy and more, but also to the greater societal good by
 23 advancing human rights, promoting social justice, reducing inequities, and empowering
 24 marginalized groups.

25
 26 ¹⁷⁷ David Meyer, *Sam Altman Has Signed a New Open Letter on A.I. ’s Dangers: Here’s What’s*
 27 *Different About This ‘Extinction’ Statement*, FORTUNE MAG. (May 30, 2023, 9:55 AM),
<https://fortune.com/2023/05/30/sam-altman-has-signed-a-new-open-letter-on-a-i-s-dangers-heres-whats-different-about-this-extinction-statement/>.

28 ¹⁷⁸ *Pause Giant AI Experiments: An Open Letter*, FUTURE OF LIFE INST. (Mar. 22, 2023),
<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.

242. With adequate safeguards, including a moral and ethical code, the Products can help detect and prevent human rights violations rather than cause them; they can help combat human discrimination and bias rather than replicate, encourage, and exacerbate humankind's worst impulses.¹⁷⁹ On the other side of the pause, the Products can responsibly foster global cooperation, collaboration, and peace by facilitating communication, learning, and understanding across cultures and languages rather than starting world wars with disinformation and the unchecked capacity for autonomous weaponry. Likewise, the Products can aid in the ongoing search for truth, by enabling breakthroughs in math, science, and more, that humans might never alone make, rather than forever obliterating the line between truth and fiction altogether.

243. We can have this AI, the one that enriches our lives, that works for people, and that works for human benefit, that is "helping us cure cancer, that is helping us find climate solutions," but leading experts agree, not without a pause on the Products' unchecked commercial proliferation: "[W]hen we're in an arms race to deploy AI to every human being on the planet as fast as possible with as little testing as possible, that's not an equation that's going to end well."¹⁸⁰ The current scenario stands only to enrich Defendants, while destabilizing the world.

244. Defendants have released Products to the entire world, that they know and readily recognize could someday result in societal collapse; that even they, the creators, cannot fully understand, predict, or reliably control; thus, any attempt now by Defendants to politicize this action, to attack the class action device or the Plaintiffs brave enough to stand up to corporate greed and irresponsibility of this magnitude at this pivotal moment in history, will fail. All people of good will on both sides of the aisle and from every background are united and resolute in the need for intervention. That is because we all want to live in a world where technology serves our shared

¹⁷⁹ See generally Cade Metz and Karen Weise, *A Tech Race Begins as Microsoft Adds A.I. to Its Search Engine*, THE N.Y. TIMES (Feb. 7, 2023), <https://www.nytimes.com/2023/02/07/technology/microsoft-ai-chatgpt-bing.html> ("The new chatbots do come with baggage. They often do not distinguish between fact and fiction. They can generate language that is biased against women and people of color. And experts worry that people will use them to spread lies at a speed they could not in the past.").

¹⁸⁰ Jason Abbruzzese, *The Tech Watchdog That Raised Alarms About Social Media Is Warning About AI*, NBC NEWS (Mar. 22, 2023), <https://www.nbcnews.com/tech/tech-news/tech-watchdog-raised-alarms-social-media-warning-ai-rcna76167>.

values of freedom, justice, dignity, equality, prosperity, privacy and security, not where Products exist that undermine these ideals.

245. In an often divided and polarized world, it is telling how so many have been able to unite around these truths: (i) the current state of AI governance is insufficient to address the threats posed by the Products; (ii) the lack of transparency, accountability, oversight, and regulation surrounding the Products and Defendants suddenly deploying them for profit worldwide has resulted in a ticking time bomb in the hands of those motivated to harm the American people; (iii) the gap must be closed between the rapid pace of the Products' development on the backs of stolen personal data on the one hand, and the slow progress of AI policy on the other; and (iv) a temporary pause on the commercial deployment of the Products is necessary and justified to prevent irreversible damage to humanity and society.

246. Critically, the injunctive relief sought in this Action seeks only to pause the unfettered and further commercial deployment of the Products, with AI research and development otherwise continuing unaffected. That is because of an equally important truth on which all agree: the United States must remain aggressively locked into the worldwide AI arms-race, set off by Defendants' launch of the Products (for better or worse), to ensure this powerful technology is developed and deployed for good around the world, and to block the potential harms from those world powers currently leveraging AI like the Products to build technological weapons as powerful as the nuclear bomb. Thus, the only "setback" here will be to Defendants' corporate bank accounts, while the American people stand to (re)gain their fundamental right to privacy as well as just compensation for the mass theft of personal data on which Defendants built and continue to run the Products.

III. DEFENDANTS' CONDUCT VIOLATES ESTABLISHED PROPERTY AND PRIVACY RIGHTS

A. Defendants' Web-Scraping Theft

247. Defendants' first category of theft and misappropriation stems from their secret scraping of the internet. This violated both the property rights and privacy rights of all individuals whose personal information was scraped and then incorporated through misappropriation into Defendants' Products.

248. Defendants' initial web scraping was done largely in secret, without the consent of any individuals whose personal and identifying information was scraped, much less all of the website operators themselves. This violated not only the Terms of Use of various websites but also the rights of each and every individual to opt out of such collection under California and other state and federal laws. Without any notice to the public, no one can be said to have consented to the collection of their online personal data, history, web practices and other personal and identifying information.

249. By the time the public learned of Defendants' web scraping practices in late Fall of 2022, when ChatGPT was released, it was too late to meaningfully exercise their privacy rights outside of this lawsuit — their internet history had been scraped, consumed, and integrated into the large language models from which the Products were born.

250. While Defendants' massive theft of personal information at scale is unmatched in history, it is reminiscent of the Clearview AI scandal in 2020. Clearview is a company that uses facial recognition technology to identify individuals based on their online photos.¹⁸¹ To create its product, Clearview scraped billions of publicly available photos from various websites and social media platforms.¹⁸² As with Defendants, this illegal scraping was done without the consent of users or the website owners themselves, and without registering as a data broker under California or Vermont Law.¹⁸³

251. Just like Defendants, Clearview used the stolen information to build its AI product.¹⁸⁴ Clearview then sold access to the product to law enforcement agencies, private companies, and other governmental agencies.¹⁸⁵ Defendants' business model is the same: scrape information off the

¹⁸¹ Tate Ryan-Mosley, *The NYPD Used a Controversial Facial Recognition Tool. Here's What You Need to Know*, MIT TECH. REV., (Apr. 9, 2021), www.technologyreview.com/2021/04/09/1022240/clearview-ai-nypd-emails/.

¹⁸² Will Knight, *Clearview AI Has New Tools to Identify You in Photos*, WIRED (Oct. 4, 2021), <https://www.wired.com/story/clearview-ai-new-tools-identify-you-photos/>.

¹⁸³ Robert Hart, *Clearview AI Fined \$9.4 Million in UK for Illegal Facial Recognition Database*, FORBES (May 23, 2022), <https://www.forbes.com/sites/roberthart/2022/05/23/clearview-ai-fined-94-million-in-uk-for-illegal-facial-recognition-database/?sh=73d5a0f71963>.

¹⁸⁴ *Id.*

¹⁸⁵ Drew Harwell, *Clearview AI to Stop Selling Facial Recognition Tool to Private Firms*, THE WASH. POST (May 9, 2022), <https://www.washingtonpost.com/technology/2022/05/09/clearview-illinois-court-settlement/>.

internet, in secret without any notice and consent in violation of the law, use it to build AI products, and then sell access to the Products for commercial gain.

252. Clearview's illegal scraping practices also went undetected for years, until it was laid bare by a New York Times expose.¹⁸⁶ The public was rightfully upset, as were state and federal regulators. The Vermont Attorney General sued Clearview in March 2020 for violating data broker and consumer protection laws, alleging that Clearview fraudulently acquired brokered personal information through its scraping practices and exposed consumers to various risks and harms.¹⁸⁷ Clearview was also sued by several individuals and organizations in California and elsewhere.¹⁸⁸

253. As a result of these lawsuits and public scrutiny, Clearview ultimately registered as a data broker in both California and Vermont. Although Defendants employ the same business model as Clearview, they have failed to register as data brokers under applicable law. By failing to do so prior to scraping the internet, Defendants violated the rights of millions. Plaintiffs and the Classes had a right to know what personal information Defendants were scraping and collecting and how it would be used, a right to delete their personal information collected by Defendants, and a right to opt out of the use of that information to build the Products.

254. Defendants' violation of the law is ongoing as they continue to collect personal brokered information by scraping the internet without registering as data brokers or otherwise providing notice or seeking consent from anyone. Plaintiffs and the Classes have a right to opt out of this ongoing scraping of internet information but no mechanism to exercise that right, absent the injunctive relief sought in this Action.

¹⁸⁶ Dave Gershgorin, *Is There Any Way Out of Clearview's Facial Recognition Database?*, THE VERGE (June 9, 2021), <https://www.theverge.com/22522486/clearview-ai-facial-recognition-avoid-escape-privacy>.

¹⁸⁷ *Attorney General Donovan Sues Clearview AI for Violations of Consumer Protection Act and Data Broker Law*, OFF. OF VT. ATT'Y GEN. (Mar. 10, 2020), <https://ago.vermont.gov/blog/2020/03/10/attorney-general-donovan-sues-clearview-ai-violations-consumer-protection-act-and-data-broker-law>.

¹⁸⁸ Johana Bhuiyan, *Clearview AI Uses Your Online Photos to Instantly ID You. That's A Problem, Lawsuit Says*, L.A. TIMES (Mar. 9, 2021), <https://www.latimes.com/business/technology/story/2021-03-09/clearview-ai-lawsuit-privacy-violations>.

B. Defendants' Web Scraping Violated Plaintiffs' Property Interests

255. Courts recognize that internet users have a property interest in their personal information and data. *See Calhoun v. Google, LLC*, 526 F. Supp. 3d 605, 635 (N.D. Cal. 2021) (recognizing property interest in personal information and rejecting Google's argument that "the personal information that Google allegedly stole is not property"); *In re Experian Data Breach Litigation*, SACV 15-1592 AG (DFMx), 2016 U.S. Dist. LEXIS 184500, at *14 (C.D. Cal. Dec. 29, 2016) (loss of value of personal identifying information is a viable damages theory); *In re Marriott Int'l Inc. Customer Data Sec. Breach Litig.*, 440 F. Supp. 3d 447, 460-61 (D. Md. 2020) ("The growing trend across courts that have considered this issue is to recognize the lost property value of this [personal] information."); *Simona Opris v. Sincera*, No. 21-3072, 2022 U.S. Dist. LEXIS 94192, at *20 (E.D. Pa. May 23, 2022) (collecting cases).

256. Plaintiffs and Class Members' property rights in the personal data and information that they have generated, created, or provided through various online platforms thus includes the right to possess, use, profit, sell, and exclude others from accessing or exploiting that information without consent or remuneration.

257. The economic value of this property interest in personal information is well understood, as a robust market for such data drives the entire technology economy. As experts have noted, the world's most valuable resource is "no longer oil, but data," and has been for years now.¹⁸⁹

258. A single internet user's information can be valued anywhere from \$15 to \$40, and even more.¹⁹⁰ Another study found that an individual's online identity can be sold for \$1,200 on the dark web.¹⁹¹ Defendants' misappropriation of every piece of data available on the internet, and with it, millions of internet users' personal information without consent, thus represents theft of a value unprecedented in the modern era of technology.

¹⁸⁹ *The World's Most Valuable Resource Is No Longer Oil, but Data*, THE ECONOMIST (May 6, 2017), <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>.

¹⁹⁰ *Id.*

¹⁹¹ Maria LaMagna, *The Sad Truth About How Much Your Facebook Data is Worth on the Dark Web*, MARKETWATCH (June 6, 2018), <https://www.marketwatch.com/story/spooked-by-the-facebook-privacy-violations-this-is-how-much-your-personal-data-is-worth-on-the-dark-web-2018-03-20>.

259. Writing for the Harvard Law Review, Professor Paul M. Schwartz underscored the value of personal data, as follows: “Personal information is an important currency in the new millennium. The monetary value of personal data is *large* and still *growing*, [and that’s why] corporate America is moving quickly to profit from the trend.”¹⁹² The data forms a critical “corporate asset.”

260. Other experts concur: “[S]uch vast amounts of collected data have obvious and substantial economic value. Individuals’ traits and attributes (such as a person’s age, address, gender, income, preferences... [their] clickthroughs, comments posted online, photos updated to social media, and so forth) are increasingly regarded as business assets[.]”¹⁹³

261. Because personal data is valuable personal property, market exchanges now exist where internet users like Plaintiffs and putative class members can sell or monetize their own personal data and internet usage information.¹⁹⁴ For example, Facebook has offered to *pay* users for their voice recordings.¹⁹⁵ By contrast and as alleged herein upon information and belief, Defendants simply *took* millions of text files, voice recordings, and facial scans from across the internet — without any consent from putative class members, much less personal remuneration to them. Theft of this nature is not only unprecedented and unjust, but also dangerous. As noted in Section II, it puts millions at risk for their likeness to be cloned to perpetrate fraud, or to embarrass or otherwise harm them.

262. Moreover, the law specifically recognizes a legal interest in unjustly earned profits based on unauthorized harvesting of personal data, and “this stake in unjustly earned profits exists

¹⁹² Paul M. Schwartz, Property, Privacy, and Personal Data, 117 HARV. L. REV. 2056, 2056 (May, 2004).

¹⁹³ Alessandro Acquisti et al., *The Economics of Privacy*, 54(2) J. OF ECON. LITERATURE 442, 444 (Mar. 8, 2016).

¹⁹⁴ Kevin Mercandante, *10 Apps for Selling Your Data for Cash*, BEST WALLET HACKS, <https://wallethacks.com/apps-for-selling-your-data/> (last updated Apr. 20, 2023); Kari Paul, *Facebook Launches Apps That Will Pay Users for Their Data*, THE GUARDIAN (June 11, 2019) <https://www.theguardian.com/technology/2019/jun/11/facebook-user-data-app-privacy-study>; Saheli Roy Choudry & Ryan Browne, *Facebook Pays Teens to Install an App That Could Collect All Kinds of Data*, CNBC (Jan. 29, 2019), <https://www.cnbc.com/2019/01/29/facebook-paying-users-to-install-app-to-collect-data-techcrunch.html>.

¹⁹⁵ Tim Bradshaw, *Facebook Offers to Pay Users for Their Voice Recordings*, FIN. TIMES (Feb. 21, 2020), <https://www.ft.com/content/42f6b93c-54a4-11ea-8841-482eed0038b1>.

1 regardless of whether an individual planned to sell his or her data or whether the individual's data
2 is made less valuable."¹⁹⁶

3 263. Defendants have been unjustly enriched by their theft of personal information as its
4 billion-dollar AI business, including ChatGPT and beyond, was built on harvesting and monetizing
5 Internet users' personal data. Thus, Plaintiffs and the Classes have a right to disgorgement and/or
6 restitution damages representing the value of the stolen data and/or their share of the profits
7 Defendants earned thereon.

8 **C. Defendants' Web Scraping Violated Plaintiffs' Privacy Interests**

9 264. In addition to property rights, internet users maintain privacy interests in personal
10 information even if it is posted online, and experts agree the collection, processing, and further
11 dissemination of this information can create distinct privacy harms.¹⁹⁷

12 265. For example, the aggregation of collected information "can reveal new facts about a
13 person that she did not expect would be known about her when the original, isolated data was
14 collected."¹⁹⁸ Even a small subset of "public" private information can be used to harm the privacy
15 interests of internet users. One example is when researchers analyzed public tweets to identify users
16 with mental health issues; naturally, Twitter users did not consent or expect their data to be used in
17 that way, to potentially reveal new, highly personal information about them.¹⁹⁹ If that analysis were
18 made public, or used commercially, that would pose significant and legally cognizable privacy
19 harms.

20 266. Another reason users retain privacy interests in their personal data on the internet,
21 even when it is technically 'public,' is the reasonable expectation of "obscurity" i.e., "the notion
22 that when our activities or information is unlikely to be found, seen, or remembered, it is, to some
23 degree safe."²⁰⁰ Privacy experts note users' reasonable expectation that most of the Internet will
24 simply ignore their individual posts. Moreover, "[t]he passage of time also makes information
25

26 ¹⁹⁶ *In re Facebook, Inc. Internet Tracking Litigation*, 956 F.3d 589, 600 (9th Cir. 2020).

27 ¹⁹⁷ Geoffrey Xiao, *Bad Bots: Regulating the Scraping of Public Information*, 34(2) HARV. J.L. &
TECH., 701, 706, 732 (2021).

28 ¹⁹⁸ Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477, 493 (2006).

¹⁹⁹ Xiao, *supra* note 197, at 707.

²⁰⁰ Woodrow Hartzog, *The Public Information Fallacy*, 99 BOS. L. REV. 459, 515 (2019).

obscure: no one remembers your MySpace pictures from fifteen years ago.”²⁰¹

267. Internet users’ reasonable expectations are also informed by the known transaction costs that, typically, would “prevent[] someone from collecting all your photos from every social media site you have ever used – ‘just because information is hypothetically available does not mean most (or even a few) people have the knowledge and ability to access [‘public’ private] information.”²⁰²

268. When users post information on the internet, “they do so believing that their information will be obscure and in an environment of trust” on whichever site they post. Users expect a level of privacy— they “**do not expect their information to be swept up by data scraping.**” Thus, according to experts, the privacy problem with “widescale, automated collection of personal information via scraping,” is that it “destroys” reasonable user expectations including the right to “obscurity” by reducing the typical transaction costs and difficulties in accessing, collecting, and understanding personal information at scale.²⁰³

269. Scraping therefore illegally enables the use of personal information in ways which reasonable users could not have anticipated. In respect of Defendants’ surreptitious scraping at unprecedented scale, it means all items users have posted on the internet have now been collected, including their voice recordings and images – arming Defendants with the ability to create a digital clone of each internet user to anticipate and manipulate their next move. Plaintiffs and the Classes did not consent to such use of their personal information. As privacy experts note, “even if a user makes the affirmative choice to make [an internet post public], she manifests an intent to participate in an obscure and trustworthy environment, not an intent to participate in data harvesting.”²⁰⁴

270. Worse, Plaintiffs and the Classes could not have known Defendants were collecting their personal information, because Defendants did it without notice to anyone, in violation of California law which required them to register with the state as data brokers.²⁰⁵

271. Introducing these data broker laws, the California assembly stated its intent:

²⁰¹ Xiao, *supra* note 197, at 708-09.

²⁰² *Id.* at 709.

²⁰³ *Id.*

²⁰⁴ *Id.* at 711.

²⁰⁵ Cal. Civ. Code § 1798.99.80(d).

1 “[C]onsumers are generally not aware that data brokers possess their personal information, how to
 2 exercise their right to opt out, and whether they can have their information deleted, as provided by
 3 California law.” Thus, “it is the intent of the Legislature to further Californians’ right to privacy by
 4 giving consumers an additional tool to help control the collection and sale of their personal
 5 information by requiring data brokers to register annually with the Attorney General and provide
 6 information about how consumers may opt out of the sale of their personal information.”²⁰⁶

7 272. “Sale” of information includes “making it available” to others for consideration,
 8 which Defendants have done by commercializing the stolen data into ChatGPT and building a
 9 billion-dollar business from it. Despite scraping information for this express purpose, Defendant
 10 OpenAI did not, and still has not, registered with the State of California as required.

11 273. Experts acknowledge the “serious privacy harms” inherent in the type of entirely
 12 “covert information” collection in which Defendants engaged.²⁰⁷ It “undermines individual
 13 autonomy and free choice.”²⁰⁸ The lack of notice, including under California’s data broker laws,
 14 “excludes individuals from the data collection process, making individuals feel powerless in
 15 controlling how their data is used.”²⁰⁹ This is not just a feeling—as described *supra*, the harm is
 16 concrete economic injury given the robust market for personal information.

17 274. Without notice of Defendants’ scraping practices, users were also denied the ability
 18 to engage in self-help, by choosing to make obscure but technically publicly-available information
 19 private – and the lack of notice precluded users from exercising their statutory data privacy rights,
 20 such as the right to request deletion.²¹⁰ Instead, Plaintiffs’ and the Classes’ internet histories are
 21 now embedded in Defendants’ AI products with no recourse other than the damages and injunctive
 22 relief requested in this Action.

23
 24
 25
 26 ²⁰⁶ Assemb. B. 1202, 2019-2020 Reg. Sess. (Cal. 2019) (as discussed in Xiao, *supra* note 197, at 714-715).

27 ²⁰⁷ Xiao, *supra* note 197, at 719.

28 ²⁰⁸ *Id.*

²⁰⁹ *Id.*

²¹⁰ *Id.* at 720.

D. Defendants' Business Practices are Offensive to Reasonable People and Ignore Increasingly Clear Warnings from Regulators

275. Defendants' mass scraping of personal data for commercialization has sparked outrage over the legal and privacy implications of Defendants' practices. Those aware of the full extent of the misappropriation are fearful and anxious about how Defendants used their "digital footprint" and about how Defendants might use all that personal information going forward. Absent the relief sought in this Action, there will be no limits on such future use. The public is also concerned about how all of their personal information might be accessed, shared, and misused *by others*, now that it is forever embedded into the large language models on which the Products run.

276. The outrage makes sense: Defendants admit the Products might evolve to act against human interests, and that regardless, they are unpredictable. Thus, by collecting previously obscure and personal data of millions and permanently entangling it with the Products, Defendants knowingly put Plaintiffs and the Classes in a zone of risk that is *incalculable* — but unacceptable by any measure of responsible data protection and use.

277. The extent to which Defendants stand to profit from the unprecedented privacy risks they were willing to take—with data that is not theirs—is especially offensive to everyday people. As one explained, "Using AI as it stands right now is *normalizing the illegal mass scraping* of everyone's data regardless of their nature, just to make the top even richer and forfeit any means we have to protect our work *and who we are as humans*. This should not be encouraged and tolerated."²¹¹ The outrage stems, in part, from this uncontestable truth: "None of this would have been possible without data – *our data* – collected and used without our permission."²¹²

278. In this new era of AI, we cannot allow widescale illegal data scraping to become a commercial norm; otherwise, privacy as a fundamental right will be relegated to the dustbin of history. Underscoring the need for court intervention, AI researcher Remmelt Ellen remarked simply, "[i]llegal scraping needs to be addressed."²¹³

²¹¹ @coffeeseed, TWITTER (May 11, 2023), <https://twitter.com/CoffeeSeed/status/1656634134616211461>.

²¹² Gal, *supra* note 5.

²¹³ @RemmeltE, TWITTER (Apr. 10, 2023), <https://twitter.com/RemmeltE/status/1645499008075407364>.

279. The public is also troubled by the lack of just compensation for the use of their personal data. One AI large language model developer stated it plainly: “If your data is used, companies should cough up.”²¹⁴ Otherwise, according to a more complete critique of the current business model, AI is just “pure primitive accumulation”—taking from the masses to enrich a few, i.e., Silicon Valley tech companies and their billionaire owners.²¹⁵

280. While the past, and ongoing, misappropriation of valuable personal information is bad enough, the Products also stand to altogether eliminate future income for millions, due to the widespread unemployment they are expected to cause over time. No one has consented to the use of their personal information to build this destabilized future of social unrest and worsening poverty for everyday people, while the pockets of OpenAI and Microsoft are lined with profit.

281. As OpenAI itself once acknowledged, albeit when still purely not-for-profit, the Company would need to fund a universal basic income (UBI) if the Products were ever developed and deployed for widespread public use, because they would eliminate so many jobs. Even now, Mr. Altman’s “grand idea is that OpenAI will capture much of the world’s wealth through the creation of A.G.I. and then redistribute this wealth to the people.”²¹⁶ Given Defendants’ sudden deployment of the Products across virtually every industry using data that was not theirs, this future should begin now, with legal or equitable redistribution of Defendants’ ill-gotten gains. Others have noted that a portion of the profits generated by Defendants can be funneled back “to everyone who contributed content.” This would include “basically everyone,” given the scope of the initial and ongoing theft of personal information by Defendants.²¹⁷

282. To avoid the unjust enrichment of Defendants, this Court sitting in equity has the power to order a “data dividend” to consumers for as long as the Products generate revenue fueled on the misappropriated data. At the very least, Plaintiffs and the Classes should be personally and

²¹⁴ @yudhanjaya, TWITTER (June 9, 2023), <https://twitter.com/yudhanjaya/status/1667391709679095808>.

²¹⁵ Bridle, *supra* note 76.

²¹⁶ Cade Metz, *The ChatGPT King Isn’t Worried, but He Knows You Might Be*, THE N.Y. TIMES (Mar. 31, 2023), <https://www.nytimes.com/2023/03/31/technology/sam-altman-open-ai-chatgpt.html>.

²¹⁷ *Id.*

1 directly compensated for the fair market value of their contributions to the large language models
 2 on which the Products were built and thrive, in an amount to be determined by expert testimony.
 3 Fundamental principles of property law demand such compensation, and everyday people
 4 reasonably support it.²¹⁸

5 283. While the property and privacy rights this Action seeks to vindicate are settled as a
 6 general matter, their application to business practices surrounding the large language models fueling
 7 AI products has not been widely tested under the law. However, just weeks ago, the FTC settled an
 8 action against Amazon, in connection with the company's illegal use of voice data to train the
 9 algorithms on which its popular Alexa product runs. That action raised many of the same type of
 10 violations alleged in this Action.

11 284. Announcing settlement of the action, the FTC gave a stern public warning to
 12 companies like Defendants: "Amazon is not alone in apparently seeking to amass data to refine its
 13 machine learning models; right now, with the advent of large language models, the tech industry as
 14 a whole is *sprinting* to do the same."²¹⁹ The settlement, it continued, was to be a message to all:
 15 "Machine learning is *no excuse to break the law*... The data you use to improve your algorithms
 16 must be *lawfully collected* and *lawfully retained*. Companies would do well to heed this lesson."²²⁰

17 285. The FTC's warning comports with FTC Commissioner Rebecca Slaughter's earlier
 18 warning, in 2021, in the Yale Journal of Law and Technology.²²¹ Discussing the FTC's new practice
 19 of ordering "algorithmic destruction," Commissioner Slaughter explained that "the premise is
 20 simple: when companies collect data illegally, they should not be able to profit from either the data
 21 or any algorithm developed using it."²²² Commissioner Slaughter believed this enforcement
 22 approach would "send a clear message to companies engaging in illicit data collection in order to

23 ²¹⁸ See e.g., @ianfinlay2000, *Time to Get Paid For Our Data?*, REDDIT (2021),
 24 https://www.reddit.com/r/Futurology/comments/qknz3u/time_to_get_paid_for_our_data/ ("[T]he
 25 companies are basically stealing our data bc no one knows that they should be getting paid for it").

26 ²¹⁹ Devin Coldewey, *Amazon Settles with FTC for \$25M After 'Flouting' Kids' Privacy and
 Deletion Requests*, TECHCRUNCH (May 31, 2023), [https://techcrunch.com/2023/05/31/amazon-
 27 settles-with-ftc-for-25m-after-flouting-kids-privacy-and-deletion-requests/](https://techcrunch.com/2023/05/31/amazon-settles-with-ftc-for-25m-after-flouting-kids-privacy-and-deletion-requests/) (emphasis added).

28 ²²⁰ *Id.* (emphasis added).

²²¹ Rebecca Kelly Slaughter et al., *Algorithms and Economic Justice: A Taxonomy of Harms and a
 Path Forward for the Federal Trade Commission*, 23 YALE J. L. & TECH. 1, 39 (Aug. 2021).

²²² *Id.*

1 train AI models: *Not worth it.*”²²³ Unfortunately for the millions of consumers impacted by
 2 Defendants’ mass theft of data, Defendants did not heed the warning.

3 **E. Defendants’ Theft of User Data in Excess of Reasonable Consent**

4 286. Defendants’ second category of theft stems from their unrestricted harvesting of data
 5 from Users of the Products, including registered Users of the OpenAI website and Users of
 6 Defendants’ API and/or plug-ins.

7 287. Defendants have made much of the fact that they purportedly “want” to comply with
 8 applicable privacy laws and regulations—and will likely oppose this lawsuit by arguing that
 9 registered users of the Products purportedly “consented” to the widespread theft of their personal
 10 information by virtue of using the Products. This argument is disingenuous for multiple reasons.

11 288. *First:* For those consumers who used ChatGPT plug-ins or API, the various sites’ use
 12 policies did not provide anything approaching informed consent that the consumers’ information
 13 and personal data would be used to train Defendants’ LLMs and would thus be incorporated into
 14 generative AI in a manner that would prevent them from reasonably ever removing their data from
 15 Defendants’ for-profit commercial enterprises. These Plaintiffs and Class Members had no idea that
 16 Defendants were and are collecting and utilizing their User Data, including the most sensitive
 17 information, when they engage with ChatGPT which seamlessly incorporated artificial intelligence
 18 in the background.

19 289. Plaintiffs fell victim to Defendants’ unlawful collection and sharing of their sensitive
 20 information acquired through users’ interactions with Defendants’ Products and websites, as well
 21 as the hundreds or thousands of applications that now use ChatGPT-based plug-ins or API.²²⁴

22 290. In less than 24 hours after Defendants announced the ability to install plug-ins to
 23 ChatGPT, many companies immediately jumped on board and started incorporating their websites
 24 within the AI plug-in. In exchange, Defendants received yet another wealth of personal data, once
 25 again, without the users’ and nonusers’ consent. ChatGPT is becoming the single app “to rule them
 26

27 ²²³ *Id.* (emphasis added).

28 ²²⁴ Matt Burgess, *ChatGPT Has a Big Privacy Problem*, WIRED (Apr. 4, 2023),
<https://www.wired.com/story/italy-ban-chatgpt-privacy-gdpr/>.

1 all.”²²⁵

2 291. Defendants’ AI has become the virtual spy,²²⁶ closely monitoring, recording, and
3 training on the personal data, clicks, searches, inputs, and personal information of millions of
4 unsuspecting individuals who may be using an Instacart to purchase grocery items, a telehealth
5 company to make a doctor’s appointment, or simply browsing Expedia to make vacation plans.

6 292. *Second:* Even those Plaintiffs and Class Members who registered for OpenAI
7 accounts and interacted with ChatGPT directly did not give effective consent for Defendants to use
8 their data and personal information in the way they currently do.

9 293. For instance, when Plaintiffs logged in to use the ChatGPT, Defendants were tracking
10 and collecting every piece of information entered into the chatbot—including sensitive information
11 such (1) all details entered into the chatbot; (2) account information users enter when signing up;
12 (3) name; (4) contact details; (5) login credentials; (6) emails; (7) payment information; (8)
13 transaction records; (9) identifying data ChatGPT pulls from users’ device or browser, like IP
14 addresses and location; (10) social media information; (11) chat log data; (12) usage data; (13)
15 analytics; and (14) cookies. However, Defendants are also tracking the information from other
16 applications in which their AI is already plugged in – Stripe, Microsoft Teams, Bing, Zillow,
17 Expedia, Instacart, etc. – and using each piece of information to train the AI.

18 294. Plaintiffs, and all Class Members, did not consent to such extensive collection of data,
19 and the use of their data for essentially any purpose to benefit Defendants’ businesses – including
20 for training purposes of the AI. In fact, Plaintiffs and all Class Members could not consent to
21 Defendants’ conduct because they were unaware their sensitive information would be collected and
22 used in this manner in the first place. Thus, Defendants did not obtain *valid enforceable* consent to
23 collect, use, and store Plaintiffs’ and Class Members’ sensitive information.

24 295. In the near future, Defendants anticipate adding even more powerful features to the
25 omniscient AI, allowing it to also gather data from audio inputs with their yet another AI—Vall-E.

26
27 ²²⁵ Better Product, *OpenAI’s Master Plan to Turn ChatGPT into an Everything App*, MEDIUM
(Mar. 25, 2023), <https://medium.com/@betterproducts/openais-master-plan-to-turn-chatgpt-into-an-everything-app-1270686074f8>.
28

²²⁶ *Id.*

Vall-E has already been developed and allows to process three (3) seconds of a human voice, and be able to speak in such voice in perpetuity. Once activated, Defendants' and their AI's access to human voices and audio inputs will jeopardize the users' and nonusers' privacy even further.

296. Defendant OpenAI has also deceptively represented to its users that they can request their private information not be used and, if parents discover that a child has used ChatGPT, Defendant will erase the child's data from the system. This is deceptive because by the time the language model has taken in the information and learned from it, that information has already financially benefited Defendants and cannot be removed from the knowledge base of the language model. Moreover, Defendant OpenAI has stated that, notwithstanding a user's requests to opt out of data collection and sharing, it will still retain some information (though what information will be retained is not specified).

297. Currently, a ChatGPT user wanting to opt out of the use of their data and chats for model training is instructed that they can simply turn off chat history (which deprives them of using that functionality themselves) and the application will stop using *new* chat content for training purposes.²²⁷ However, Defendants continue to train their models with the user's information – be it from the prior chats or new chats. Moreover, as noted above, it is impossible to know whether any of the previously used data can effectively be deleted, as once the language model is trained using the data, it becomes part of the model. Additionally, the option of opting out of chat history retention doesn't impact OpenAI's ability to use a user's other personal data gathered during the account creation process for Defendants' own purposes. OpenAI's privacy disclosures are intentionally vague about this, noting simply that a user can opt out of chat history retention *or* can submit a form to ask OpenAI not to use or share their data. No guidelines are provided regarding whether or when Defendant might decline to honor such a request, nor how long it takes to process.

298. Furthermore, as commentators have observed, Defendant OpenAI heavily pushes

²²⁷ Johanna C., *How Do I Turn Off Chat History and Model Training?*, OPENAI, <https://help.openai.com/en/articles/7792795-how-do-i-turn-off-chat-history-and-model-training> (last visited June 27, 2023).

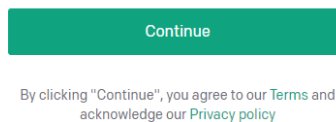
users not to opt out of data collection.²²⁸ Once a user turns off the option for their ChatGPT interactions to be used for training purposes, they are presented constantly with a large green button that encourages them to “Enable chat history.” Nothing on this button notifies users that enabling chat history functionality amounts to reauthorizing OpenAI to save and train Defendants’ models on the user’s data.

299. Moreover, it is not clear what information a given user can actually prevent OpenAI from retaining and using in the future, as the company has stated in blog posts that it will retain some data anyway and that some of this data can be used in Defendant OpenAI’s training datasets.²²⁹

300. Defendants fail to provide accurate and comprehensive notifications to consumers about the scale of their data sharing practices. Defendants’ admissions within their Privacy Policy do not adequately inform consumers on the breadth of data sharing, resulting in a breach of explicit assurances and a violation of reasonable consumer expectations. By acting in such a manner, Defendants are engaged in data misuse practices that contradict the principles of transparency, accountability, and respect for consumer privacy rights.

1. OpenAI’s disclosures are not conspicuous.

301. When a consumer attempts to register for an OpenAI account, they are presented with the following image:



302. When a hyperlink to an agreement is “not conspicuous enough to put [plaintiffs] on inquiry notice,” then the agreement is not binding. *Colgate v. JUUL Labs, Inc.*, 402 F. Supp. 3d 728, 764-66 (N.D. Cal. 2019). The Ninth Circuit holds that “even close proximity of the hyperlink to relevant buttons users must click on—without more—is insufficient to give rise to constructive

²²⁸ Natasha Lomas, *How to Ask OpenAI for Your Personal Data to Be Deleted or Not Used to Train Its AIs*, TECHCRUNCH (May 2, 2023), <https://techcrunch.com/2023/05/02/chatgpt-delete-data/>.

²²⁹ Yaniv Markovski, *How Your Data Is Used to Improve Model Performance*, OPENAI, <https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance> (last visited June 2, 2023).

notice.” *Nguyen v. Barnes & Noble Inc.*, 763 F.3d 1171, 1179 (9th Cir. 2014). Instead, courts consider factors such as color, size and font of the hyperlink, and whether the hyperlink is presented alone or in a clutter of text. *See, e.g., Colgate*, 402 F. Supp. 3d at 764; *Selden v. Airbnb, Inc.*, 16-cv-00933 (CRC), 2016 WL 6476934, at *14-15 (D.D.C. Nov. 1, 2016).

303. Here, a consumer registering for an OpenAI account is ferried through the process and is provided only small hyperlinks to OpenAI’s Privacy Policy and Terms of Use during the sign-up process. The lettering alerting the potential registrant to the documents is tiny and gray. The consumer need not make any indication that he or she has actually read the documents, nor that they understand the connection between these documents and their creation of an account. Unlike many companies that require a consumer to scroll to the bottom of a privacy policy or other legal document—or at least click a radial purporting to have read the document—an OpenAI registrant need make no affirmative indication that they are aware of the policies whatsoever. As such, there is no binding agreement between Defendant OpenAI and Plaintiffs or the Members of the Subclasses regarding use of these individuals’ information, and no effective consent.

304. Plaintiff and the User Subclasses were neither on constructive notice nor inquiry notice of the privacy policy on the ChatGPT platform.

2. Defendants’ Use of Consumer Data Far Exceeds Industry Standards and their Own Representations

305. The Federal Trade Commission has promulgated numerous guides for businesses highlighting the importance of implementing reasonable data security practices. According to the FTC, the need for data security should be factored into all decision-making.²³⁰

306. In 2016, the FTC updated its publication, *Protecting Personal Information: A Guide for Business*, which established cybersecurity guidelines for businesses.²³¹ The guidelines note that businesses should protect the personal customer information that they keep; properly dispose of

²³⁰ *Start with Security: A Guide for Business: Lessons Learned from FTC Cases*, FED. TRADE COMM’N. (June, 2015), <https://www.ftc.gov/system/files/documents/plain-language/pdf0205-startwithsecurity.pdf>.

²³¹ *Protecting Personal Information: A Guide for Business*, FED. TRADE COMM’N. (Oct. 2016), https://www.ftc.gov/system/files/documents/plain-language/pdf-0136_proteting-personal-information.pdf.

personal information that is no longer needed; encrypt information stored on computer networks; understand their network's vulnerabilities; and implement policies to correct any security problems.

307. The FTC further recommends that entities not maintain personally identifiable information longer than is needed for authorization of a transaction; limit access to sensitive data; require complex passwords to be used on networks; use industry-tested methods for security; monitor for suspicious activity on the network; and verify that third-party service providers have implemented reasonable security measures. The FTC has brought enforcement actions against entities engaged in commerce for failing to adequately and reasonably protect customer data, treating the failure to employ reasonable and appropriate measures to protect against unauthorized access to confidential consumer data as an unfair act or practice prohibited by Section 5 of the Federal Trade Commission Act ("FTCA"), 15 U.S.C. § 45. Orders resulting from these actions further clarify the measures businesses must take to meet their data security obligations.

308. Defendants fail to meet these obligations, as they directly feed consumers' personal information into their LLMs for training purposes.

309. Even if the click-through button discussed above could constitute a binding agreement—it cannot—the substance of the policies is insufficient to put any consumer on notice of what to expect with regard to the use of their information. The policies lay out vague promises regarding how and when the users' data can and will be shared, and affirm that all laws are being complied with—even where such affirmations are internally inconsistent.²³² For example, under the heading "Additional U.S. State Disclosures," the Privacy Policy lists five different categories of "Personal Information," including one category that OpenAI identifies as "Sensitive Personal Information," and states that OpenAI discloses information from *all five* of the various categories to "our affiliates, vendors and service providers, law enforcement, and parties involved in Transactions." Yet a few paragraphs down, the policy then inexplicably asserts "We don't sell Personal Information or share Personal Information." No explanation is given as to what is meant by the assertion that the company both *does* and *does not* share Personal Information.

²³² *Privacy Policy*, OPENAI, <https://openai.com/policies/privacy-policy> (last updated June 23, 2023).

310. As of June 23, 2023, Defendants changed this language to clarify that they “don’t ‘sell’ Personal Information or ‘share’ Personal Information for cross-contextual behavioral advertising (as those terms are defined under applicable local law).”²³³ Nevertheless, no explanation is given as to how Defendants can ensure that the entities with which they are sharing users’ personal information with are not, in fact, using it for cross-contextual behavior advertising. Defendants also do not disclose the specific purposes for which they do use such sensitive data.

311. Moreover, the Policy alerts consumers that to the extent local law entitles them to request deletion of their Personal Information, they can exercise this right (amongst others) by sending a request to dsar@openai.com. Yet nothing in the privacy policy explains that information which has already been incorporated into Defendants’ LLMs *can never really* be removed.

312. Furthermore, as discussed in Section IV, ChatGPT lacks age controls to prevent children under 13 from using ChatGPT and subsequently providing their information. The Privacy Policy states that the service is not intended for children under 13. However there are no effective restrictions in place to prevent: (1) children under 13 accessing the platform; (2) the AI technology interacting inappropriately with children; and (3) the unauthorized collection of children’s data.²³⁴

313. Finally, even if users are on notice of the Privacy Policy (and they are not), the Privacy Policy does not disclose wiretapping. There is **zero** adequate consent for wiretapping, and OpenAI’s terms and conditions are convoluted, inconspicuous, and consist of numerous documents, impossible to decipher by reasonable consumers. There are no conspicuous or clear disclosures that all conversations are wiretapped, recorded, and shared with numerous entities—none of which are disclosed.

314. Beyond Defendants’ legal obligations to protect the confidentiality of individuals’ User Data, Defendants’ privacy policy and online representations affirmatively and unequivocally state that any personal information provided to Defendants will remain secure and protected. Since ChatGPT’s inception, Defendants have represented and continue to represent that:

“We at OpenAI OpCo, LLC (together with our affiliates, “OpenAI”, “we”, “our” or “us”) respect your privacy and are strongly committed to keeping

²³³ *Id.*

²³⁴ *Id.*

secure any information we obtain from you or about you.”

“We implement commercially reasonable technical, administrative, and organizational measures to protect Personal Information both online and offline from loss, misuse, and unauthorized access, disclosure, alteration, or destruction.”

“OpenAI does not knowingly collect Personal Information from children under the age of 13.”²³⁵

315. Defendants have failed to adhere to a single promise vis-à-vis their duty to safeguard User Data. Defendants have made these privacy policies and commitments available in ChatGPT. In these representations to Plaintiffs and Class Members and the public, Defendants promised to take specific measures to protect its members’ information, consistent with industry standards and federal and state law. However, they did not.

316. Plaintiffs and Class Members relied to their detriment on Defendants’ uniform representations and omissions regarding data security. Now that their sensitive personal and medical information is in the possession of third parties, Plaintiffs and Class Members face a constant threat of continued harm. Collection of such sensitive information without consent or notice poses a great threat to individuals by subjecting them to the danger of potential attacks and embarrassment.

317. Plaintiffs and Class Members trusted Defendants’ Products when inputting sensitive and valuable User Data. Had Defendants disclosed to Plaintiffs and its other members that every click, every search, and every input of sensitive information was being tracked, recorded, collected, and disclosed to third parties—Plaintiffs would not have trusted Defendants’ Products to input such sensitive information.

318. Defendants knew or should have known that Plaintiffs and Class Members would reasonably rely upon, and trust Defendants’ promises regarding security and safety of its data and systems.

319. Additionally, Defendants were aware that ChatGPT collects, tracks, and discloses Plaintiffs’ and Class Members’ User Data, including sensitive information.

320. By virtue of how ChatGPT is “trained,” i.e., through the collection and processing of a massive corpus of data, Defendants were aware that their Users’ data would be collected and

²³⁵ *Id.*

disclosed to third parties every time a user interacted with ChatGPT.

IV. DEFENDANTS' CONDUCT POSES SPECIAL PRIVACY AND SAFETY RISKS FOR CHILDREN

321. The Products pose special risks for children, especially ChatGPT. As ChatGPT has become more pervasive and sophisticated, it has also become increasingly capable of collecting, tracking, and disclosing vast amounts of personal data about children.

322. Children's data is particularly sensitive. It can reveal not only their personal identities, but also their physical locations, habits, interests, and relationships. The indiscriminate and unauthorized collection, tracking, and disclosure of this data by powerful, profit-driven corporations undermines children's privacy and autonomy, and it also puts them at risk of abuse, exploitation, and discrimination.

323. The safety of children in the digital environment is a foundational concern for society. According to HealthyChildren, "Overuse of digital media may place your children at risk of": not enough sleep, obesity, delays in learning and social skills, negative effect on school performance, behavior problems, problematic internet use, risky behavior, sexting, criminal predators; loss of privacy; and cyberbullying.²³⁶

324. Senator Michael Bennet (D-CO) recently sent a letter to the CEO of OpenAI and other industry leaders to "highlight the potential harm to younger users of rushing to integrate generative artificial intelligence (AI) in their products and services."²³⁷ Senator Bennet wrote, "the race to deploy generative AI cannot come at the expense of our children. Responsible deployment requires clear policies and frameworks to promote safety, anticipate risk, and mitigate harm."²³⁸

325. In one illustration of the harms, Senator Bennet described how researchers prompted

²³⁶ *Constantly Connected: How Media Use Can Affect Your Child*, HEALTHY CHILD, <https://www.healthychildren.org/English/family-life/Media/Pages/Adverse-Effects-of-Television-Commercials.aspx> (last updated May 30, 2023).

²³⁷ Michael Bennett, *Bennett Calls on Tech Companies to Protect Kids as They Deploy AI Chatbots*, MICHAEL BENNET U.S. SEN. FOR COLO. (Mar. 21, 2023), <https://www.bennet.senate.gov/public/index.cfm/2023/3/bennet-calls-on-tech-companies-to-protect-kids-as-they-deploy-ai-chatbots> ("***the race to deploy generative AI cannot come at the expense of our children;***" "[r]esponsible deployment requires clear policies and frameworks to promote safety, anticipate risk, and mitigate harm") (emphasis added).

²³⁸ *Id.*

My AI to instruct a child how to cover up a bruise ahead of a visit from Child Protective Services.²³⁹ When one researcher posed as a 13-year-old girl, My AI provided suggestions for how to lie to her parents about an upcoming trip with a 31-year-old man. It later provided suggestions for how to make losing her virginity a special experience by setting the mood with candles or music.²⁴⁰

326. This public introduction of AI-powered chatbot, ChatGPT, arrives during an epidemic of teen mental health problems. A recent report from the Centers for Disease Control and Prevention (CDC) found that 57 percent of teenage girls felt persistently sad or hopeless in 2021, and that one in three seriously contemplated suicide.²⁴¹ In fact, the American Academy of Pediatrics (AAP), the American Academy of Child and Adolescent Psychiatry (AACAP), and the Children’s Hospital Association (CHA) have declared a national emergency in child and adolescent mental health, stating that its members were “caring for young people with soaring rates of depression, anxiety, trauma, loneliness, and suicidality that will have lasting impacts on them, their families, and their communities.”²⁴² This state of mental health across children and adults, in tandem with the increase in isolated, digital engagement results in dissociative behavior and worsens depression.²⁴³ ChatGPT exponentially exacerbates this issue by promoting human-like conversations and irresponsibly dispensing harmful, even life-threatening information—going so far as drafting suicide notes for depressed, suicidal users.²⁴⁴

327. The GPT-4 System Card provides no detail of safety checks conducted by OpenAI

²³⁹ @tristanharris, TWITTER (Mar. 10, 2023, 1:07 PM), <https://twitter.com/tristanharris/status/1634299911872348160>.

²⁴⁰ *Id.*

²⁴¹ Moriah Balingit, *‘A Cry for Help’: CDC Warns of a Steep Decline in Teen Mental Health*, THE WASH. POST (Mar. 31, 2022), <https://www.washingtonpost.com/education/2022/03/31/student-mental-health-decline-cdc/>.

²⁴² *AAP-AACAP-CHA Declaration of a National Emergency in Child and Adolescent Mental Health*, AM. ACAD. OF PEDIATRICS (Oct. 19, 2021), <https://www.aap.org/en/advocacy/child-and-adolescent-healthy-mental-development/aap-aacap-cha-declaration-of-a-national-emergency-in-child-and-adolescent-mental-health/>.

²⁴³ Liu Yi Lin et al., *Association Between Social Media Use and Depression Among U.S. Young Adults*, 33 DEPRESS. & ANXIETY 323, 323 (April 2019).

²⁴⁴ Jeremy Kaplowitz, *Man Uses ChatGPT to Write Suicide Note*, HARD DRIVE (Apr. 3, 2023), <https://hard-drive.net/hd/technology/man-uses-chatgpt-to-write-suicide-note/>; see also Gary Marcus, *The Dark Rise of Large Language Models*, WIRED (Dec. 29, 2022), <https://www.wired.com/story/large-language-models-artificial-intelligence/> (GPT-3 even urged a research account to commit suicide).

during its testing period, nor does it detail any measures implemented by OpenAI to protect children.

A. Defendants Deceptively Tracked Children without Consent

328. The Children’s Online Privacy Protection Act (“COPPA”) requires Defendants to obtain parental consent before monitoring, collecting, or using information from children under 13 or if they have actual knowledge that their Users are of such age. Unless Defendants obtain this consent, the law forbids collection or usage of information about these children.

329. Despite this restriction, Defendants’ customary practice is to simply ignore the presence of younger Users on their application—while collecting information just like they would for an adult User—or leave it up to those Users to self-report their age, despite knowing that children can and regularly do access technology products by reporting a false birthdate.

330. Defendants are guilty of the unlawful and deceptive invasion of the right to privacy and reasonable expectation of privacy of thousands—if not millions—of children. While holding themselves out publicly as respecting privacy rights, Defendants tracked the information, behaviors, and preferences of vulnerable children solely for financial gain in violation of well-established privacy protections, societal norms, and the laws encapsulating those protections.

331. At all material times, Defendants deceived Plaintiffs and the members of the Classes and Subclasses regarding their data collection and tracking behavior. As alleged herein, Defendants knowingly and purposefully tracked, profiled, and targeted minors on the ChatGPT Platform for advertising revenue and to train LLM AI programs, like the Products. This tracking and data collection contravenes privacy rights, societal norms, and federal and state statutes, while Defendants feign compliance with these rights and statutes.

332. Defendants deceptively operated the free ChatGPT Platform as if it were only used by adults while intentionally luring thousands if not millions of children to the platform. Defendants then intentionally tracked and collected the personal information of each underage ChatGPT User (treatment to which only an adult can legally consent) in order to obtain information relevant to behavioral advertising, collect data that can be used for training the Products, and compile training datasets that can be sold to other businesses and researchers to train other AI Products. Defendants did so despite knowing that thousands if not millions of these Users were actually minor children,

including children under the age of thirteen, solely for the financial benefit of Defendants, as well as their affiliates, vendors, and service providers, all of whom knowingly and willingly consented to this unlawful conduct.

B. Defendant Designed ChatGPT to be Inappropriate for Children

333. As detailed in Section I, Defendants collect extensive data from Users to train OpenAI's language model AIs and compile training datasets.

334. Data collection of this nature requires the consent of the individual whose data is being collected. But only adults are capable of giving such consent – to the extent it was sufficiently explained in Defendant OpenAI's Privacy Policy or Terms of Service.

335. Defendant OpenAI thus inserted language into its Terms of Service and Privacy Policy which indicated that ChatGPT was intended to be used by individuals thirteen and older. More specifically, OpenAI's Privacy Policy, last updated June 23, 2023, states the following under the heading "Children:"

Our Service is not directed to children who are under the age of 13. OpenAI does not knowingly collect Personal Information from children under the age of 13. If you have reason to believe that a child under the age of 13 has provided Personal Information to OpenAI through the Service please email us at legal@openai.com. We will investigate any notification and if appropriate, delete the Personal Information from our systems. If you are 13 or older, but under 18, you must have consent from your parent or guardian to use our Services.²⁴⁵

336. Defendant OpenAI's Terms of Use document also references age requirements in the "Registration and Access" section, stating: "You must be at least 13 years old to use the Services. If you are under 18 you must have your parent or legal guardian's permission to use the Services."²⁴⁶

337. Defendant OpenAI prevents potential users from creating a ChatGPT user account unless the user self-reports a birthdate that indicates the user's age is thirteen or older. However, Defendants know or reasonably should know that this self-reporting of a minor child's birthdate is ineffective to keep minor children, including those under the age of 13 off the ChatGPT platform because minor users can present a false birthdate.

²⁴⁵ *Privacy Policy*, OPENAI <https://openai.com/policies/privacy-policy> (last updated June 23, 2023).

²⁴⁶ *Terms of Use*, OPENAI, <https://openai.com/policies/terms-of-use> (last updated Mar. 14, 2023).

338. In fact, if a child under the age of 13 attempts to set up a ChatGPT account and is rejected for being underage, that same child can simply change their self-reported birthdate—while keeping all of their other information the same—and they will be granted immediate access. When confronted with this exact same flaw in the self-report age verification process for Snapchat, executives from Snap, Inc.—one of OpenAI’s close partners²⁴⁷—admitted that such a system is effectively useless in stopping underage users from signing up for the platform.²⁴⁸ Indeed, recent studies have reported that 13% of children ages 8–12 and 49% of children ages 13–17 used Snapchat in 2021 notwithstanding this same self-reporting age verification system.²⁴⁹

339. To sign up for an OpenAI account and start using ChatGPT, a child has to first provide an email and then click on a link emailed to them to verify this email. They are then directed to a page where they are asked their name and birthdate. If they enter a date of birth that indicates they are under the age of 13, they receive the following message alerting them that they cannot create an account due to OpenAI’s Terms of Use.

²⁴⁷ Alex Heath, *Snapchat Is Releasing Its AI Chatbot to Everyone for Free*, THE VERGE (Apr. 19, 2023, 2:00 PM) <https://www.theverge.com/2023/4/19/23688913/snapchat-my-ai-chatbot-release-open-ai> (Snap CEO Evan Spiegel describes the relationship between Snap and OpenAI as a “close partnership”).

²⁴⁸ Isobel Asher Hamilton, *Snapchat Admits Its Age Verification Safeguards Are Effectively Useless*, BUS. INSIDER (Mar. 19, 2019, 9:23 AM), <https://www.businessinsider.com/snapchat-says-its-age-verification-safeguards-are-effectively-useless-2019-3#:~:text=Collins%20admitted%20that%20the%20system,mobile%20app%20is%20more%20popular>.

²⁴⁹ Victoria Rideout et al., *The Common Sense Census: Media Use by Tweens and Teens*, COMMON SENSE MEDIA (2021), https://www.commonsensemedia.org/sites/default/files/research/report/8-18-census-integrated-report-final-web_0.pdf.

Tell us about you

Bailey

Cowan

05/27/2012

Continue

We can't create your account due to our [Terms of Use](#)

340. However, if the child then refreshes the page, they can input an earlier date of birth without any problem—and without having to use a different email or to re-verify the email address. Even though OpenAI knows that the name and email address being used has been associated with a child under the age of 13, it still allows the user to continue creating an account using this information.

341. Despite the vast amounts of data at its command, OpenAI makes no effort to verify the personal information entered, even when inconsistent information has been entered for the same user. Thus, the birthdate field is not a true age verification safeguard.

342. If the child enters a date of birth that would make them under the age of 18 but older than 13, they are able to create an account simply by hitting the Continue button.

Tell us about you

Bailey

Cowan

05/27/2008

Continue

We will only use this data to verify your age

By clicking "Continue", you confirm that you have parental or guardian consent to use ChatGPT, agree to our [Terms](#), and acknowledge our [Privacy policy](#)

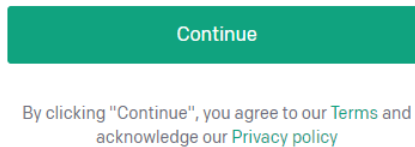
343. Under this bright green “Continue” button, there are words in small gray letters that inform the child that, by clicking “Continue,” they are confirming that they have parental or

guardian consent to use ChatGPT, agree to our Terms (hyperlinked), and acknowledge our “Privacy policy” (hyperlinked).

344. The child does not need to interact with the text in any way to indicate that they saw it or read it, much less that they read the hyperlinked Terms or Privacy policy documents. Nor does the child user have to provide an email address for an adult so that they can confirm they actually do have parental consent and that a parent or guardian has read and agreed to the privacy policy and terms of use.

345. Similarly, if a user enters a date of birth that indicates they are over the age of 18, the screen looks the same, except that the small gray letters under the large green Continue button do not mention parental consent.

346. In neither case is the user asked to *agree* to the privacy policy – but merely acknowledge it.



C. Defendants Deprived Children of the Economic Value of their Personal Data

347. A child’s personal information has equivalent (or potentially greater) value than that of an adult to companies like Defendants. First, a child is more susceptible to being influenced by advertisements as they often cannot tell the difference between content and advertisements. They also are more likely than adults to confide personal details and highly private information to ChatGPT without realizing that Defendants are using that information to train LLMs for their own financial gain, and that they may share the information with their affiliates, vendors, service providers, or partners to bolster all of these businesses’ private profits.

348. Second, Defendants and/or those with whom they share User information may be able to utilize children’s personal information for the duration of their lives.²⁵⁰ Plaintiffs and Minor

²⁵⁰ OpenAI’s Terms of Use of ChatGPT says ChatGPT does not sell users’ data to third parties. However, the terms do not disclose whether ChatGPT can display targeted advertisements to users,

Members of the Classes and Subclasses can no longer realize the full economic value of their personal information because it has already been collected, analyzed, acted upon, incorporated into language models, and monetized by Defendants.

349. Third, the detailed tracking of habits, preferences, thoughts, and geolocation data for young children presents unique and significant personal security and safety concerns. Quite simply, it begs the question of whether any company or its employees should have this much information about where our kids are and how to motivate their cooperation.

350. Defendants' illegal and improper collection of children's Personal Information has given them a significant "first mover" advantage that cannot be undone. ChatGPT set an unprecedented record as the fastest app to reach 100 million active users, reaching that milestone in a mere two months after its release in November 2020.

351. As a result of their unlawful conduct, ChatGPT now incorporates ill-gotten data from thousands if not millions of children who use ChatGPT without appropriate consent. The deep insights gleaned from these children's interactions with ChatGPT will enable Defendants and the for-profit companies with whom they share this data to keep children interacting with various applications, websites, language models, and platforms; to use the Personal Information of children for potentially the duration of their lives; and will solidify Defendants' dominance in the AI market by incorporating vast amounts of child-related content into Defendants' language models.

352. Publicly, Defendant OpenAI has denied marketing its ChatGPT product to children – and denied that children have utilized the application. But it is common knowledge that minors and school-aged children are using the service, as there have been widespread news reports about how schools have had to crack down on such use to prevent cheating on homework and otherwise. Thus, Defendants knew or should have known that OpenAI's age "verification" and parental consent protocols were woefully ineffective and resulted in thousands if not millions of minor children—including those under the age of 13—gaining access to ChatGPT and sharing their personal information with the language model.

send third-party marketing communications, or track users based on their interactions with ChatGPT on other apps or services across the internet for advertising purposes. *See Terms of Use*, OPENAI, <https://openai.com/policies/terms-of-use> (last updated Mar. 14, 2023).

D. Defendants' Exploitation of Children Without Parental Consent Violated**Reasonable Expectations of Privacy and is Highly Offensive**

353. Defendant's conduct in violating privacy rights and reasonable expectations of privacy of Plaintiffs and Class and Subclass members is particularly egregious because Defendants violated social norms and laws designed to protect children, a group that is subject to such protections specifically because they are supremely vulnerable to exploitation and manipulation.

354. Parental rights to care for and control their children are fundamental liberty interests. Parental consent requirements are legally required not only to protect highly vulnerable children from deception and exploitation, but also to venerate the significant rights that parents have to determine who their children interact with and on what terms.

355. These parental rights are greatly impacted and threatened by companies like Defendants who refuse to institute reasonable and verifiable parental consent protections.

356. Children are developmentally capable of using smartphones and tablets by two years old. Almost every family with a child younger than eight in America has a smartphone (95%) and/or tablet (78%). It is exceedingly common for children to have their own devices.

357. For example, a 2019 survey of media use by children aged 8-18, conducted by Common Sense Media, found that roughly 20% of children have a phone by the age of 8 and over half (53%) of children in the United States have their own phone by the age of 11.²⁵¹

358. A survey conducted by the Center for Digital Democracy ("CDD") and Common Sense Media of over 2,000 adults found overwhelming support for the basic principles of privacy embedded in the California Constitution, state common law, as well as federal law.²⁵² Of the parents polled, 75% strongly disagreed with the statement that it is okay for advertisers to track and keep a record of a child's behavior online if they give the child free content, 84% strongly disagreed that advertisers should be able to collect information about a child's location from their mobile phone,

²⁵¹ Anya Kamenetz, *It's a Smartphone Life: More Than Half of U.S. Children Now Have One*, NPR (Oct. 31, 2019, 12:05 PM), <https://www.npr.org/2019/10/31/774838891/its-a-smartphone-life-more-than-half-of-u-s-children-now-have-one>.

²⁵² Center for Digital Democracy, *Survey on Children and Online Privacy, Summary of Methods and Findings*, <https://www.democraticmedia.org/sites/default/files/COOPA> (last visited June 5, 2023).

89% strongly agreed that companies should receive parental consent before putting tracking software on a child's computer, and 93% agreed that a federal law requiring online sites and companies to ask parents' permission before they collect Personal Information from children under age 13 was "a good idea."²⁵³ Against this backdrop, Defendants' knowing exploitation of children without adequate parental involvement is not only illegal but also highly offensive to social norms and mores.

CLASS ALLEGATIONS

359. **Class Definition:** Plaintiffs bring this action pursuant to Federal Rules of Civil Procedure Sections 23(b)(2), 23(b)(3), and 23(c)(4), on behalf of Plaintiffs and the Class defined as follows:

- a. **Non-User Class:** All persons in the United States whose PII, Personal Information, or Private Information was disclosed to, or accessed, collected, tracked, taken, or used by Defendants without consent or authorization.
- b. **ChatGPT User Class:** All persons in the United States who used ChatGPT, whose Private Information was disclosed to, or intercepted, accessed, collected, tracked, taken, or used by Defendants without consent or authorization.
- c. **ChatGPT API User Class:** All persons in the United States who used other platforms, programs, or applications which integrated ChatGPT technology, whose Private Information was disclosed to, or intercepted, accessed, collected, tracked, taken, or used by Defendants without consent or authorization.
- d. **Microsoft User Class:** All persons in the United States who used Microsoft platforms, programs, or applications which integrated ChatGPT technology, whose Private Information was disclosed to, or intercepted, accessed, collected, tracked, taken, or used by Defendants without consent or authorization.
- e. **Minor ChatGPT User Class:** All persons in the United States who, while 16 years or younger, used ChatGPT, or other platforms, programs, or applications which integrated ChatGPT API or ChatGPT Plug-In, whose Private Information was disclosed to, or intercepted, accessed, collected, tracked, taken, or used by Defendants without consent or authorization.
- f. **ChatGPT Plus User Class:** All persons in the United States who used ChatGPT website or mobile app and whose Personal Information or PII was intercepted, accessed, collected, tracked, stored, shared, taken, or used by Defendants without consent and/or authorization.

State-Wide Subclasses:

The California Subclass

²⁵³ *Id.*

- i. **California Non-User SubClass:** All persons within the State of California whose PII, Personal Information, or Private Information was disclosed to, or accessed, collected, tracked, taken, or used by Defendants without consent or authorization.
- ii. **California ChatGPT User SubClass:** All persons within the State of California who used ChatGPT, whose Private Information was disclosed to, or intercepted, accessed, collected, tracked, taken, or used by Defendants without consent or authorization.
- iii. **California ChatGPT Plus User SubClass:** All persons within the State of California who used Chat-GPT website or mobile app and whose Personal Information or PII was intercepted, accessed, collected, tracked, stored, shared, taken, or used by Defendants without consent and/or authorization.
- iv. **California Minor User SubClass:** All persons within the State of California who, while 16 years or younger, used ChatGPT, or other platforms, programs, or applications which integrated ChatGPT API or ChatGPT Plug-In, whose Private Information was disclosed to, or intercepted, accessed, collected, tracked, taken, or used by Defendants without consent or authorization.

The New York Subclass

- i. **New York Non-User SubClass:** All persons within the State of New York whose PII, Personal Information, or Private Information was disclosed to, or accessed, collected, tracked, taken, or used by Defendants without consent or authorization.
- ii. **New York ChatGPT User SubClass:** All persons within the State of New York who used ChatGPT, whose Private Information was disclosed to, or intercepted, accessed, collected, tracked, taken, or used by Defendants without consent or authorization.
- iii. **New York ChatGPT Plus User SubClass:** All persons within the State of New York who used Chat-GPT website or mobile app and whose Personal Information or PII was intercepted, accessed, collected, tracked, stored, shared, taken, or used by Defendants without consent and/or authorization.
- v. **New York Minor User SubClass:** All persons within the State of New York who, while 16 years or younger, used ChatGPT, or other platforms, programs, or applications which integrated ChatGPT API or ChatGPT Plug-In, whose Private Information was disclosed to, or intercepted, accessed, collected, tracked, taken, or used by Defendants without consent or authorization.

The Illinois Subclass

- i. **Illinois Non-User SubClass:** All persons within the State of Illinois whose PII, Personal Information, or Private Information was disclosed to, or accessed, collected, tracked, taken, or used by Defendants without consent or authorization.
- ii. **Illinois ChatGPT User SubClass:** All persons within the State of Illinois who used ChatGPT, whose Private Information was

disclosed to, or intercepted, accessed, collected, tracked, taken, or used by Defendants without consent or authorization.

- iii. **Illinois ChatGPT Plus User SubClass:** All persons within the State of Illinois who used Chat-GPT website or mobile app and whose Personal Information or PII was intercepted, accessed, collected, tracked, stored, shared, taken, or used by Defendants without consent and/or authorization.

360. **The following people are excluded from the Classes and Subclasses:** (1) any Judge or Magistrate presiding over this action and members of their judicial staff and immediate families; (2) Defendants, Defendants' subsidiaries, parents, successors, predecessors, and any entity in which the Defendants or their parents have a controlling interest and its current or former officers and directors; (3) persons who properly execute and file a timely request for exclusion from the Class; (4) persons whose claims in this matter have been finally adjudicated on the merits or otherwise released; (5) Plaintiffs' counsel and Defendants' counsel; and (6) the legal representatives, successors, and assigns of any such excluded persons.

361. Plaintiffs reserve the right under Federal Rule of Civil Procedure 23 to amend or modify the Class to include a broader scope, greater specificity, further division into subclasses, or limitations to particular issues. Plaintiffs reserve the right under Federal Rule of Civil Procedure 23(c)(4) to seek certification of particular issues.

362. The requirements of Federal Rules of Civil Procedure 23(a), 23(b)(2), and 23(b)(3) are met in this case.

363. The Fed. R. Civ. P. 23(a) elements of Numerosity, Commonality, Typicality, and Adequacy are all satisfied.

364. **Ascertainability:** Membership of the Classes and Subclasses is defined based on objective criteria, and individual members will be identifiable from Defendants' records, records of third-party platforms/applications which integrate ChatGPT, self-identification methods, or other means. Defendants' records are likely to include massive data storage, user accounts, and data gathered directly from the affected members of Classes and Subclasses.

365. **Numerosity:** The precise number of the Members of Classes and Subclasses is not available to Plaintiffs, but it is clear that individual joinder is impracticable. Millions of people have

used Defendants' ChatGPT language model since its introduction in 2022. An even larger number of people are affected by Defendants' unlawful web-scraping of the entire platforms, including Reddit and Twitter. Defendants' integrated ChatGPT plug-ins on hundreds of applications affect millions of Members of Classes and Subclasses daily. Members of the Classes and Subclasses can be identified through Defendants' records, records of entities which integrated ChatGPT, or by other means, including but not limited to self-identification.

366. **Commonality:** Commonality requires that the Members of Classes and Subclasses allege claims which share common contention such that determination of its truth or falsity will resolve an issue that is central to the validity of each claim in one stroke. Here, there is a common contention for all Classes and Subclasses are as follows:

Defendants' Web-Scraping Practices (NonUser Class)

- a) Whether the members of Non-User Class had a protected property right in their data;
- b) Whether Defendants scraped the protected data belonging to Non-User Class members without consent;
- c) Whether Defendants' collection, scraping, and uses of the protected Non-User Class Members of protected data violates:
 1. Electronic Communication Privacy Act, 18 U.S.C. § 2510 *et. seq.*
 2. Computer Fraud and Abuse Act, 18 U.S.C. § 1030 *et. seq.*
 3. California Constitution right to privacy;
 4. California Invasion of Privacy Act, Cal. Pen. Code §§ 630 *et seq.*
 5. California Unfair Competition Law, Bus. & Prof Code §§ 17200;
 6. Illinois Biometric Information Privacy Act, 740 ILCS 14/1, *et seq.*
 7. Illinois Consumer Fraud and Deceptive Practices Act, 815 Ill. Comp Stat. §§ 505, *et seq.*
 8. Illinois Consumer Fraud and Deceptive Business Practices Act, 815 Ill. Comp. Stat. §§ 510/2, *et seq.*
 9. New York General Business Law Section 349 *et seq.*
- d) Whether Defendants' collection, scraping, and uses of the protected Non-User Class

Members of protected data constitutes:

1. Common law Negligence;
 2. Unlawful Intrusion upon Seclusion under California laws;
 3. Conversion;
 4. Larceny/Receipt of Stolen Property under Cal. Pen. Code § 496(a) and (c).
- e) Whether as a result of Defendants' collection, scraping, and uses of the protected Non-User Class Members of protected data, Non-User Class Members suffered monetary damages, including but not limited to actual damages, statutory damages, punitive damages, treble damages, or other monetary damages.
- f) Whether as a result of Defendants' collection, scraping, and uses of the protected Non-User Class Members of protected data, Non-User Class Members are entitled to equitable relief, including but not limited to restitution, disgorgement of profits, injunctive and declaratory relief, or other equitable remedies.

Defendants' Collection/Interception Practices of Private Information From ChatGPT User, ChatGPT Plug-In User, ChatGPT Plus User Classes, and Subclasses (including Minors):

- a) Whether Defendants failed to advise the members of Classes and Subclasses the extent to which Defendants intercepted, received, collected Private Information;
- b) Whether Defendants intercepted, received, or collected communications, tracked all activities, chat history, and other Private Information from the Users of Other Platforms Which Integrate ChatGPT without consent of such Users.
- c) Whether Microsoft Defendant intercepted, received, or collected communications, tracked all activities, chat history, and other Private Information of ChatGPT Users, without consent of such Users;
- d) Whether Open AI Defendant aided, abetted, and otherwise conspired with Microsoft Defendant, to allow Defendant Microsoft's interception, receipt, or collection of communications, tracking of all activities, and other Private Information of ChatGPT Users, without consent of such Users;

- 1 e) Whether Defendants' conduct of intercepting, receipt, collection of Private
 2 Information of the members of Classes and Subclasses violated federal and state
 3 privacy laws, anti-wiretapping laws, or other tort laws, including but not limited to:
- 4 1. Electronic Communication Privacy Act, 18 U.S.C. § 2510 *et. seq.*
 - 5 2. Computer Fraud and Abuse Act, 18 U.S.C. § 1030 *et. seq.*
 - 6 3. California Constitution right to privacy;
 - 7 4. California Invasion of Privacy Act, Cal. Pen. Code §§ 630 *et seq.*
 - 8 5. California Unfair Competition Law, Bus. & Prof Code §§ 17200;
 - 9 6. New York General Business Law Section 349 *et seq.*
 - 10 7. Common law Negligence;
 - 11 8. Unlawful Intrusion upon Seclusion under California laws;
 - 12 9. Conversion.
- 13 f) Whether as a result of Defendants' collection, scraping, and uses of the protected
 14 Private Information, ChatGPT User, ChatGPT Plug-In User, ChatGPT Plus User
 15 Class Members and Subclass Members suffered monetary damages, including but
 16 not limited to actual damages, statutory damages, punitive damages, treble damages,
 17 or other monetary damages.
- 18 g) Whether as result of Defendants' interception, collection, receipt, or unauthorized
 19 uses of Private Information, ChatGPT User, ChatGPT Plug-In User, ChatGPT Plus
 20 User Class Members and Subclass Members are entitled to equitable relief,
 21 including but not limited to restitution, disgorgement of profits, injunctive and
 22 declaratory relief, or other equitable remedies.

Additional Questions Affecting Minor Classes

- 24 a) Whether Defendants failed to prevent the uses of ChatGPT and ChatGPT Plug-ins
 25 by minors;
- 26 b) Whether Defendant failed to prevent integration of ChatGPT Plug-ins in the systems
 27 used by minors;
- 28 c) Whether Defendants failed to obtain a verified consent from parents of minors

authorizing the minors' uses of ChatGPT and ChatGPT Plug-ins;

- d) Whether Defendants failed to obtain a verified consent from parents of minors authorizing the tracking, collection, interception, receipt of the Members of Minor Classes and Subclasses' Private Information;
- e) Whether Defendants violated the Members of Minor Classes and Subclasses reasonable expectation of privacy;
- f) Whether Defendants' violation of the Members of Minor Classes and Subclasses reasonable expectation of privacy is highly offensive.

367. **Typicality:** Plaintiffs' claims are typical of the claims of other Class Members in that Plaintiffs and the Class Members sustained damages arising out of Defendants' uniform wrongful conduct and data collecting practices, interception/sharing of the collected data with each other, and use of such data in attempt to train the AI Products, and further develop the Products.

368. **Adequate Representation:** Plaintiffs will fairly and adequately represent and protect the interests of the Members of Classes and Subclasses. Plaintiffs' claims are made in a representative capacity on behalf of the Members of Classes and Subclasses. Plaintiffs have no interests antagonistic to the interests of the other Members of Classes and Subclasses. Plaintiffs have retained competent counsel to prosecute the case on behalf of Plaintiffs and the Class. Plaintiffs and Plaintiffs' counsel are committed to vigorously prosecuting this action on behalf of the Members of Classes and Subclasses.

369. **The declaratory and injunctive relief sought in this case includes, by way of example and without limitation:**

1. Establishment of an independent body of thought leaders (the "AI Council") who shall be responsible for approving uses of the Products before, not after, the Products are deployed for said uses;
2. Implementation of Accountability Protocols that hold Defendants responsible for Product actions and outputs and barred from further commercial deployment absent the Products' ability to follow a code of human-like ethical principles and guidelines and respect for human

- values and rights, and until Plaintiffs and Class Members are fairly compensated for the stolen data on which the Products depend;
3. Implementation of effective cybersecurity safeguards of the Products as determined by the AI Council, including adequate protocols and practices to protect Users' PHI/PII collected through Users' inputting such information within the Products as well as through Defendants' massive web scraping, consistent with the industry standards, applicable regulations, and federal, state, and/or local laws;
4. Implementation of Appropriate Transparency Protocols requiring Defendants to clearly and precisely disclose the data they are collecting, including where and from whom, in clear and conspicuous policy documents that are explicit about how this information is to be stored, handled, protected, and used;
5. Requiring Defendants to allow Product users and everyday internet users to opt out of all data collection and stop the illegal taking of internet data, delete (or compensate for) any ill-gotten data, or the algorithms which were built on the stolen data;
6. Requiring Defendants to add technological safety measures to the Products that will prevent the technology from surpassing human intelligence and harming others;
7. Requiring Defendants to implement, maintain, regularly review and revise as necessary a threat management program designed to appropriately monitor Defendants' information networks for threats, both internal and external, and assess whether monitoring tools are appropriately configured, tested, and updated;
8. Establishment of a monetary fund (the "AI Monetary Fund" or "AIMF") to compensate class members for Defendants' past and ongoing misconduct to be funded by a percentage

of gross revenues from the Products;

9. Appointment of a third-party administrator (the “AIMF Administrator”) to administer the AIMF to members of the class as “data dividends” as fair and just compensation for the stolen data on which the Products depend;

10. Confirmation that Defendants have deleted, destroyed, and purged the PII/PHI of all relevant class members unless Defendants can provide reasonable justification for the retention and use of such information when weighed against the privacy interests of class members; and

11. Requiring all further and just corrective action, consistent with permissible law and pursuant to only those causes of action so permitted.

370. **This case also satisfies Fed. R. Civ. P. 23(b)(3) - Predominance:** There are many questions of law and fact common to the claims of Plaintiffs and Members of Classes and Subclasses, and those questions predominate over any questions that may affect individual Class Members. Common questions and/or issues for Class members include the questions listed above in *Commonality*, and also include, but are not necessarily limited to the following:

- a) Whether Defendants violated the California Invasion of Privacy Act;
- b) Whether Defendants’ unauthorized disclosure of Users’ sensitive information was negligent;
- c) Whether Defendants owed a duty to Plaintiffs’ and Class Members not to disclose their sensitive user information to unauthorized third parties;
- d) Whether Defendants breached their duty to Plaintiffs’ and Class Members not to disclose their sensitive user information to unauthorized third parties;
- e) Whether Defendants represented to Plaintiffs and the Class that they would protect Plaintiffs’ and the Members of Classes and Subclasses Private Information;
- f) Whether Defendants violated Plaintiffs’ and Class Members’ right to privacy;
- g) Whether Plaintiffs and Class members are entitled to actual damages, enhanced

damages, statutory damages, restitution, disgorgement, and other monetary remedies provided by equity and law;

- h) Whether Defendants collected the Personal Information of children;
- i) Whether Defendants had knowledge they were collecting the Personal Information of children;
- j) Whether Defendants obtained parental consent to collect the Personal Information of children;
- k) Whether the collection of Personal Information of children is highly offensive to a reasonable person;
- l) Whether the collection of Personal Information of children without parental consent is sufficiently serious and unwarranted as to constitute an egregious breach of social norms;
- m) Whether Defendants' conduct was unlawful or deceptive;
- n) Whether Defendants were unjustly enriched by their conduct under the laws of California.
- o) Whether Defendants fraudulently concealed their conduct; and
- p) Whether injunctive and declaratory relief and other equitable relief is warranted.

371. **Superiority:** This case is also appropriate for class certification because class proceedings are superior to all other available methods for the fair and efficient adjudication of this controversy as joinder of all parties is impracticable. The damages suffered by individual Members of Classes and Subclasses will likely be relatively small, especially given the burden and expense of individual prosecution of the complex litigation necessitated by Defendants' actions. Thus, it would be virtually impossible for the individual Members of Classes and Subclasses to obtain effective relief from Defendants' misconduct. Even if Class Members could mount such individual litigation, it would still not be preferable to a class action, because individual litigation would increase the delay and expense to all parties due to the complex legal and factual controversies presented in this Complaint. By contrast, a class action presents far fewer management difficulties and provides the benefits of single adjudication, economy of scale, and comprehensive supervision

1 by a single Court. Economies of time, effort, and expense will be enhanced, and uniformity of
2 decisions ensured.

3 372. Likewise, particular issues under Rule 23(c)(4) are appropriate for certification
4 because such claims present only particular, common issues, the resolution of which would advance
5 the disposition of this matter and the parties' interests therein.

6 **CALIFORNIA LAW SHOULD APPLY TO OUT OF STATE PLAINTIFFS' & CLASS**

7 **MEMBERS' CLAIMS**

8 373. Courts "have permitted the application of California law where the plaintiffs' claims
9 were based on alleged misrepresentations [or misconduct] that were disseminated from
10 California." *Ehret v. Uber Technologies, Inc.*, 68 F. Supp. 3d 1121, 1130 (N.D. Cal.
11 2014). "California courts have concluded that state statutory remedies may be invoked by out-of-
12 state parties when they are harmed by wrongful conduct occurring in California." *In re iPhone 4S*
13 *Consumer Litig.*, No. C 12-1127 CW, 2013 U.S. Dist. LEXIS 103058, at *23 (N.D. Cal. July 23,
14 2013) (internal quotation marks and citation omitted).

15 374. Defendant Open AI is headquartered in California; this is where Defendant Open AI's
16 nerve center of its business operations is located. This is where Defendant Open AI has its high-
17 level officers direct, control, coordinate, and manage its activities, including policies, practices,
18 research and development, and other decisions affecting Defendants' Products. This is where the
19 majority of unlawful conduct took place – from development of the AI products, decisions
20 concerning AI Products and training of the AI, web scraping practices, and other major decisions
21 which affected all Class Members. Furthermore, Defendant Microsoft operates in the state of
22 California. Upon information and belief, decisions concerning Defendants' Products were entered
23 into in California.

24 375. Furthermore, Defendant Open AI requires that California law applies to disputes
25 between Defendant Open AI and ChatGPT Users.

26 376. The State of California, therefore, has significant interests to protect all residents and
27 citizens of the United States against a company headquartered and doing business in California, and
28 has a greater interest in the claims of Plaintiffs and the Classes than any other state, and the state

1 most intimately concerned with the claims and outcome of this litigation.

2 377. California has significant interest in regulating the conduct of businesses operating
3 within its borders, and that California has the most significant relationship with Defendants – as
4 Defendant Open AI is headquartered in California, and Defendant Microsoft conducts business (at
5 least as it relates to Defendant Open AI) in California, there is no conflict in applying California
6 law to non-resident consumer claims.

7 378. Application of California law to the Classes’ claims is neither arbitrary nor
8 fundamentally unfair because choice of law principles applicable to this action support the
9 application of California law to the nationwide claims of all Class Members.

10 379. Application of California law to Defendants is consistent with constitutional due
11 process.

12 **COUNT ONE**

13 **VIOLATION OF ELECTRONIC COMMUNICATIONS PRIVACY ACT, 18 U.S.C. § 2510,**

14 **et seq.**

15 **(on behalf of ChatGPT, ChatGPT API User, Microsoft User Classes against Defendants)**

16 380. Plaintiffs repeat the allegations contained in the foregoing paragraphs as if fully set
17 forth herein.

18 381. A violation of the ECPA occurs where any person “intentionally intercepts, endeavors
19 to intercept, or procures any other person to intercept or endeavor to intercept, any ... electronic
20 communication” or “intentionally discloses, or endeavors to disclose, to any person the contents of
21 any ... electronic communication, knowing or having reason to know that the information was
22 obtained through the [unlawful] interception of a[n] ... electronic communication” or “intentionally
23 uses, or endeavors to use, the contents of any ... electronic communication, knowing or having
24 reason to know that the information was obtained through the [unlawful] interception of a[n] ...
25 electronic communication.” 18 U.S.C. §§ 2511(1)(a), (c) - (d).

26 382. In addition, “a person or entity providing an electronic communication service to the
27 public shall not intentionally divulge the contents of any communication . . . while in transmission
28 on that service to any person or entity other than an addressee or intended recipient of such

1 communication or an agent of such addressee or intended recipient.” 18 U.S.C. § 2511(3)(a).

2 383. 18 U.S.C. § 2520(a) provides a private right of action to any person whose wire or
3 electronic communications are intercepted, disclosed, or intentionally used.

4 384. Both Defendants had a tortious and/or criminal intent in (a) obtaining the Private
5 Information, (b) sharing the Private Information with each other; (c) feeding the Private Information
6 into the Products, to train, develop, and commercialize their Products. Their actions were knowing
7 and deliberate, especially since Defendants were well aware that consumers did not want nor allow
8 Defendants to use their Private Information for training of the Products.

9 385. **Electronic Communications.** Electronic communication means any “transfer[s] of
10 signs, signals, writing, . . . data, [and] intelligence of [some] nature transmitted in whole or in part
11 by a wire, radio, electromagnetic, photoelectronic, or photooptical system that affects interstate
12 commerce.” 18 U.S.C. § 2510(12). Here, the following communications qualify as
13 “communications” under the ECPA:

- 14 a) **Communications On ChatGPT:** Plaintiffs’ and Class Members’ communications
15 (including but not limited to chats, comments, replies, searches, keystrokes, signals,
16 mouse clicks, or other data, activity, or intelligence) on ChatGPT intercepted by
17 Defendant Microsoft;
- 18 b) **ChatGPT Intercepted Communications On Platforms Which Integrated**
19 **ChatGPT API:** Plaintiffs’ and Class Members’ communications (including but not
20 limited to chats, comments, replies, searches, keystrokes, signals, mouse clicks, or
21 other data, activity, or intelligence) on various applications, platforms, or websites
22 which integrate ChatGPT API (i.e. Stripe, Snapchat, etc.) intercepted by
23 Defendants;
- 24 c) **Communications on Microsoft Platforms:** Plaintiffs’ and Class Members’
25 communications (including but not limited to chats, comments, replies, searches,
26 keystrokes, mouse clicks, signals, or other data, activity, or intelligence) on
27 Microsoft platforms which integrate ChatGPT API (i.e. Microsoft Teams, Outlook,
28 etc.) intercepted by Defendant Open AI;

386. **Content.** The ECPA defines content, when used with respect to electronic
communications, to “include [] any information concerning the substance, purport, or meaning of
that communication.” 18 U.S.C. § 2510(8).

387. Plaintiffs, and the members of all Classes and Subclasses have an expectation of

1 privacy in their communications, entered keystrokes, chats, comments, replies, searches, signals,
2 and other data, activity, or intelligence, and they exercised a reasonable expectation of privacy
3 concerning the transmission of that content.

4 **388. Interception.** The ECPA defines interception as the “acquisition of the contents of
5 any wire, electronic, or oral communication through the use of any electronic, mechanical, or other
6 device” and “contents . . . include [] any information concerning the substance, purport, or meaning
7 of that communication.” 18 U.S.C. §§ 2510(4), (8).

8 389. Defendants intentionally accessed, and obtained access to the contents of Plaintiffs’,
9 the Classes’, and Subclasses’ protected computers and obtained information concerning the
10 substance, purport, or meaning of communications, thereby, and in doing so, exceeded authority
11 granted by Plaintiffs, the Classes, and Subclasses to access the protected computers.

12 **390. Electronic Communication Service.** The ECPA defines electronic communication
13 service as “any service which provides to users thereof the ability to send or receive wire or
14 electronic communications.” 18 U.S.C. § 2510(15). The following services constitute “electronic
15 communication services:”

- 16 (1) Reddit, Twitter, YouTube, Spotify, TikTok, and other websites which were scraped
17 by Defendants;
- 18 (2) Third Party websites, programs, and applications, which integrate ChatGPT
19 technology;
- 20 (3) Microsoft platforms, programs, applications, and websites, which integrate
21 ChatGPT technology;
- 22 (4) Open AI website and mobile application(s) for ChatGPT.

23 **391. Electronic, Mechanical, or Other Device.** The ECPA defines “electronic,
24 mechanical, or other device” as “any device...which can be used to intercept a[n]...electronic
25 communication[.]” 18 U.S.C. § 2510(5). The following constitute “devices” within the meaning of
26 18 U.S.C. § 2510(5):

- 27 (1) Plaintiffs’ and Classes’, Subclasses’ computing devices (Mac and Windows devices
28 present on computers, mobile phones, tablets, or other devices);

- (2) Plaintiffs' and Classes', Subclasses' browsers;
- (3) Defendants' web-servers, platforms, and applications;
- (4) Third-Party web-servers, platforms, and applications, where ChatGPT API technology was implemented;
- (5) The tracking codes deployed by Defendants to effectuate the sending and acquisition of communications.

I. Interception of Communications Between ChatGPT API Class Members which occurred on Third-Party Websites, Platforms, Applications, Programs which have integrated ChatGPT API. [Microsoft User Class is Excluded]

392. The allegations for violation of 18 U.S.C. § 2510 arising out of Defendants' interception of Plaintiffs', and ChatGPT API Class Members' (collectively referred to as ChatGPT API Class Members) communications which occurred on various applications, platforms, websites which integrate ChatGPT technology (i.e., Stripe, Snapchat, etc.).

393. The transmissions of Plaintiffs', and ChatGPT API Class Members' communications (including but not limited to chats, comments, replies, searches, keystrokes, mouse clicks/movements, signals, browser activity, or other data, activity, or intelligence) on various applications, programs, platforms, and websites which integrate ChatGPT technology (i.e., Stripe, Snapchat, etc.) qualify as "communications" under 18 U.S.C. § 2510(12).

394. By integrating ChatGPT technology on third party platforms, Defendants are in the unique position of having unrestricted, real-time access to the users' every input, move, mouse click, chat, comment, reply, search, keystroke, browser activity, or other data, activity, or intelligence on the third-party platform.

395. As Plaintiffs and ChatGPT API Class Members interact with each other or the third-party entities, Defendants intentionally tap, electrically or otherwise intercept, the lines of internet communications between Plaintiffs and ChatGPT API Class Members, and/or third-party entities.

396. In disregard for Plaintiffs', and ChatGPT API Class Members' privacy rights, Defendants act as a third-party "eavesdropper," redirecting Plaintiffs and ChatGPT API Class Members' electronic communications to Defendants' own servers for appropriation, and training of

1 their Products.

2 397. Defendants' interception of the contents of Plaintiffs' and ChatGPT API Class
3 Members' communications happens contemporaneously with their exchange of such
4 communications, whether such communications are directed to Plaintiffs' and ChatGPT API Class
5 Members' friends, colleagues, or third-party entities. As described above, the ChatGPT API is
6 designed to simultaneously intercept and send a recording of each keystroke, mouse click,
7 movement, writing, or other data, activity, or intelligence to Defendants sufficient to not only
8 identify Plaintiffs and ChatGPT API Class Members also to be able to understand, collect, and use
9 for training Plaintiffs' and ChatGPT API Class Members' communications.

10 398. **Unauthorized Purpose.** Plaintiffs and ChatGPT API Class Members did not
11 authorize Defendants to acquire, access, or intercept the content of their communications on third
12 party platforms, websites, applications. Therefore, such interception and recording of
13 communications invades Plaintiffs', and ChatGPT API Class Members' privacy. Defendants
14 intentionally intercepted the contents of Plaintiffs' and ChatGPT API Class Members' electronic
15 communications for the purpose of committing a tortious act in violation of the Constitution or laws
16 of the United States or of any State – namely, the knowing intrusion into a private place,
17 conversation, or matter that would be highly offensive to a reasonable person.

18 399. **While in Transmission.** Through this calculated scheme of using ChatGPT API to
19 intercept, acquire, transmit, and record Plaintiffs' and ChatGPT API Class Members' electronic
20 communications, Defendants willfully and without valid consent from all parties to the
21 communication, take unauthorized measures to read and understand the contents or meaning of the
22 electronic communications of Plaintiffs, and ChatGPT API Class. The interception and recording
23 of electronic communications occurs while the electronic communications are in transit or passing
24 over any wire, line, or cable, or are being sent from or received at any place.

25 400. In sending and in acquiring the content of Plaintiffs', and ChatGPT API Class
26 Members' communications with third-party platforms, Defendants' purpose was tortious, and
27 designed to violate federal and state legal laws. By intentionally using, or endeavoring to use, the
28 contents of the electronic communications of Plaintiffs, ChatGPT API Class and Subclass Members,

1 while knowing or having reason to know that the information was obtained through the interception
2 of an electronic communication, Defendants violate 18 U.S.C. § 2511(1)(a).

3 401. Plaintiffs, individually, on behalf of the GPT API Class and Subclass Members, seek
4 all monetary and non-monetary relief allowed by law, including actual damages, statutory damages,
5 punitive damages, preliminary and other equitable or declaratory relief, and attorneys' fees and
6 costs.

7 **II. Microsoft's Interception of Communications Between ChatGPT Class Members**

8 402. The allegations for violation of 18 U.S.C. § 2510 arising out of Defendant Microsoft's
9 interception of Plaintiffs, ChatGPT User Class Members' communications which occurred on
10 ChatGPT platform.

11 403. The transmissions of Plaintiffs', ChatGPT User Class Members' communications
12 (including but not limited to chats, comments, replies, searches, keystrokes, mouse
13 clicks/movements, signals, browser activity, or other data, activity, or intelligence) on ChatGPT
14 platform qualify as "communications" under 18 U.S.C. § 2510(12).

15 404. By integrating ChatGPT technology on third party platforms, Defendants are in the
16 unique position of having unrestricted, real-time access to the users' every input, move, mouse click,
17 chat, comment, reply, search, keystroke, browser activity, or other data, activity, or intelligence on
18 the third-party platform.

19 405. As Plaintiffs, ChatGPT User Class Members' interact with each other or the third-
20 party entities, Defendant Open AI intentionally divulges and Defendant Microsoft intentionally
21 taps, electrically or otherwise intercepts the lines of internet communications between Plaintiffs,
22 ChatGPT, and/or third party entities (integrated within ChatGPT through plug-in technologies).

23 406. In disregard for Plaintiffs' and ChatGPT User Class Members' privacy rights,
24 Defendant Microsoft acts as a third-party "eavesdropper," redirecting Plaintiffs' and ChatGPT User
25 Class Members' electronic communications to Defendant Microsoft's own servers for
26 appropriation, and training of their Products.

27 407. Defendant Microsoft's interception of the contents of Plaintiffs', ChatGPT User Class
28 Members' communications happens contemporaneously with their exchange of such

1 communications, whether such communications are directed to Defendant Open AI or third-party
2 entities. As described above, the ChatGPT is designed to simultaneously intercept and send a
3 recording of each keystroke, mouse click, movement, writing, or other data, activity, or intelligence
4 to Defendant Microsoft sufficient to not only identify Plaintiffs, and ChatGPT User Class Members,
5 but also to be able to understand, collect, and use for training Plaintiffs' and ChatGPT User Class
6 Members' communications.

7 **408. Unauthorized Purpose.** Plaintiffs and ChatGPT User Class Members did not
8 authorize Defendant Microsoft to acquire, access, or intercept the content of their communications
9 on third party platforms, websites, applications. Moreover, Plaintiffs and ChatGPT User Class
10 Members did not authorize either Defendant to train their AI Products on private information
11 acquired by Defendants. Therefore, such interception and recording of communications invades
12 Plaintiffs', ChatGPT User Class Members' privacy. Defendant Open AI illegally divulged the
13 content of such communications to Defendant Microsoft. Defendant Microsoft intentionally
14 intercepted the contents of Plaintiffs' and ChatGPT User Class Members' communications for the
15 purpose of committing a tortious act in violation of the Constitution or laws of the United States or
16 of any State – namely, the knowing intrusion into a private place, conversation, or matter that would
17 be highly offensive to a reasonable person.

18 **409. While in Transmission.** Through this calculated scheme of using ChatGPT
19 technology to intercept, acquire, transmit, and record Plaintiffs', and ChatGPT User Class
20 Members' electronic communications, Defendant Microsoft willfully and without any iota of valid
21 consent from all parties to the communication, takes unauthorized measures to read and understand
22 the contents or meaning of the electronic communications of Plaintiffs and ChatGPT User Class
23 Members. The interception and recording of electronic communications occur while the electronic
24 communications are in transit or passing over any wire, line, or cable, or are being sent from or
25 received at any place.

26 **410.** In sending and in acquiring the content of Plaintiffs', and Class Members'
27 communications with third-party platforms, Defendants' purpose was tortious, and designed to
28 violate federal and state laws. By intentionally using, or endeavoring to use, the contents of the

1 electronic communications of Plaintiffs, ChatGPT User Class Members, while knowing or having
 2 reason to know that the information was obtained through the interception of an electronic
 3 communication, Defendant Microsoft violates 18 U.S.C. § 2511(1)(a).

4 411. Plaintiffs, individually, on behalf of the ChatGPT User Class Members, seek all
 5 monetary and non-monetary relief allowed by law, including actual damages, statutory damages,
 6 punitive damages, preliminary and other equitable or declaratory relief, and attorneys' fees and
 7 costs.

8 **III. Defendant Open AI's Interception of Microsoft User Class Members which**
 9 **occurred on Microsoft's Websites, Platforms, Applications, Programs which have**
 10 **integrated ChatGPT.**

11 412. The allegations for violation of 18 U.S.C. § 2510 arising out of Defendant Open AI's
 12 interception of Microsoft User Class Members' (collectively "Microsoft Subclasses")
 13 communications with their friends, family, colleagues, or other individuals or third-party entities,
 14 which occurred on Microsoft platforms (Teams, Bing, Outlook etc.), which integrate ChatGPT API.

15 413. The transmissions of Plaintiffs' and Microsoft Subclasses' communications
 16 (including but not limited to chats, comments, replies, searches, keystrokes, signals, mouse
 17 clicks/movements, signals, browser activity, or other data, activity, or intelligence) on Microsoft's
 18 various applications, programs, platforms, websites which integrate ChatGPT API qualify as
 19 "communications" under 18 U.S.C. § 2510(12).

20 414. By integrating ChatGPT technology within the entire Microsoft suite, Defendant
 21 OpenAI is in the unique position of having unrestricted, real-time access to the users' every input,
 22 move, mouse click, chat, comment, reply, search, keystroke, browser activity, or other data, activity,
 23 or intelligence.

24 415. As Plaintiffs, Microsoft Subclasses interact with each other or the third-party entities,
 25 Defendants intentionally tap, electrically or otherwise intercept, the lines of internet
 26 communications between Plaintiffs, Microsoft Subclasses, and/or third-party entities.

27 416. In disregard for Plaintiffs', Microsoft Subclasses Members' privacy rights, Defendant
 28 OpenAI acts as a third-party "eavesdropper," redirecting Plaintiffs, Microsoft Subclasses Members'

1 electronic communications to Defendants' own servers for appropriation, and training of their
2 Products.

3 417. Defendant Open AI interception of the contents of Plaintiffs', Microsoft Subclasses
4 Members' communications happens contemporaneously with their exchange of such
5 communications, whether such communications are directed to Plaintiffs', Microsoft Subclasses
6 Members' friends, colleagues, or third-party entities. As described above, the ChatGPT API is
7 designed to simultaneously intercept and send a recording of each keystroke, mouse click, signal,
8 movement, writing, or other data, activity, or intelligence to Defendants sufficient to not only
9 identify Plaintiffs, Microsoft Subclasses Members, but also to be able to understand, collect, and
10 use for training Plaintiffs', Microsoft Subclasses Members' communications.

11 418. **Unauthorized Purpose.** Plaintiffs and Microsoft Subclasses did not authorize
12 Defendant Open AI to acquire, access, or intercept the content of their communications which
13 occurred on Microsoft platforms, applications, programs, and websites. Therefore, such interception
14 and recording of communications invades Plaintiffs', Microsoft Subclasses Members' privacy.
15 Defendant Open AI intentionally intercepted (and continues to intercept) the contents of Plaintiffs',
16 Microsoft Subclasses Members' electronic communications for the purpose of committing a tortious
17 act in violation of the Constitution or laws of the United States or of any State – namely, the knowing
18 intrusion into a private place, conversation, or matter that would be highly offensive to a reasonable
19 person.

20 419. **While in Transmission.** Through this calculated scheme of using ChatGPT API to
21 intercept, acquire, transmit, and record Plaintiffs', Microsoft Subclasses Members' electronic
22 communications, Defendant Open AI willfully and without any iota of valid consent from all parties
23 to the communication, implements unauthorized measures to read and understand the contents or
24 meaning of Plaintiffs' and Microsoft Subclasses' communications. The interception and recording
25 of electronic communications occur while the electronic communications are in transit or passing
26 over any wire, line, or cable, or are being sent from or received at any place.

27 420. In sending and in acquiring the content of Plaintiffs', and Class Members'
28 communications with third-party platforms, Defendant Open AI's purpose was tortious, and

designed to violate federal and state laws. By intentionally using, or endeavoring to use, the contents of Plaintiffs' and Microsoft Subclasses' electronic communications, while knowing or having reason to know that the information was obtained through the interception of an electronic communication, Defendant Open AI violated and continues to violate 18 U.S.C. § 2511(1)(a).

421. Plaintiffs, individually, on behalf of the Microsoft Subclasses Members, seek all monetary and non-monetary relief allowed by law, including actual damages, statutory damages, punitive damages, preliminary and other equitable or declaratory relief, and attorneys' fees and costs.

COUNT TWO

VIOLATION OF THE COMPUTER FRAUD AND ABUSE ACT, 18 U.S.C. § 1030

(on behalf of All Plaintiffs against Defendants)

422. Plaintiffs herein repeat, reallege, and fully incorporate all allegations in all preceding paragraphs.

423. Plaintiffs', the Classes', and Subclasses' computer devices (including but not limited to Mac and Windows devices) were, used for interstate communication and commerce and are therefore "protected computers" under 18 U.S.C. § 1030(e)(2)(B).

424. Defendants intentionally accessed Plaintiffs', the Classes and Subclasses Members' protected computers and obtained information thereby, and in doing so exceeded authority granted by Plaintiffs, the Classes, and Subclasses to access the protected computers in violation of 18 U.S.C. § 1030(a)(2)(C). Plaintiffs, the Classes, and Subclasses Members have a civil cause of action for violation of the CFAA under 18 U.S.C. § 1030(g) and have suffered damage or loss.

425. **Chat GPT Plug-In:** Defendants owned and operated their Products and ChatGPT Plug-Ins. Defendants integrated ChatGPT Plug-Ins within various platforms, websites, applications, and programs, and thereby intercepted and obtained Plaintiffs', the Classes', and Subclasses' Private Information, inclusive of keywords, mouse clicks, searches, movements, signals, and other activity and intelligence.

426. **Microsoft GPT Plug-In:** Defendant Microsoft owned and operated its Microsoft platforms, websites, programs, and applications which integrated Defendants' ChatGPT Plug-In.

1 Defendant Open AI intercepted and obtained Plaintiffs', the Classes', and Subclasses' Private
 2 Information, inclusive of keywords, mouse clicks, searches, movements, signals, and other activity
 3 and intelligence. Defendants collected, and transmitted this data to their Products, and used it to
 4 train their Products. Defendants' collected data allows Defendant to determine individual users'
 5 precise locations, unique identifiers, cookies, patterns (including browsing patterns, conversational
 6 patterns), conversational and browsing activities and habits, and a plethora of other Private
 7 Information.

8 427. **ChatGPT:** Defendant Open AI owned and operated its ChatGPT platforms.
 9 Defendant Open AI transmits all data from its ChatGPT platforms to Defendant Microsoft;
 10 Defendant Microsoft thereby intercepted and obtained Plaintiffs', the Classes', and Subclasses'
 11 Private Information, inclusive of keywords, mouse clicks, searches, movements, signals, and other
 12 activity and intelligence. Defendants collected, and transmitted this data to their Products, and used
 13 it to train their Products. Defendants' collected data allows Defendant to determine individual users'
 14 precise locations, unique identifiers, cookies, patterns (including browsing patterns, conversational
 15 patterns), conversational and browsing activities and habits, and a plethora of other Private
 16 Information.

17 428. Defendants accessed, and otherwise transmitted this data without authorized consent
 18 from Plaintiffs, Classes, and Subclasses; or at a minimum, as discussed above, exceed any consent
 19 that was given.

20 429. Defendants were actively involved in implementing the unlawful interception alleged
 21 herein and promoted the use of their Products to U.S. residents and other companies, knowing about
 22 the privacy violations alleged herein. Defendants are also liable for this conduct because it occurred
 23 pursuant to the common enterprise of which they are a part.

24 430. Defendants' conduct caused "loss to 1 or more persons during any 1-year period . . .
 25 aggregating at least \$5,000 in value" under 18 U.S.C. § 1030(c)(4)(A)(i)(I) because the unauthorized
 26 access and collection of Private Information (i) caused a diminution in value of Plaintiffs', Classes',
 27 and Subclasses' Private information, both of which occurred to millions of individuals, easily
 28 aggregating at least \$5,000 in value.

431. For these reasons, and those discussed in this Complaint, Plaintiffs, Classes, and Subclasses are entitled to “maintain a civil action against the violator to obtain compensatory damages and injunctive relief or other equitable relief.” 18 U.S.C. § 1030(g).

COUNT THREE

VIOLATION OF THE CALIFORNIA INVASION OF PRIVACY ACT (“CIPA”), CAL.

PENAL CODE § 631, et seq.

(on behalf of ChatGPT, ChatGPT API User, Microsoft User Classes against Defendants)

432. Plaintiffs repeat the allegations contained in the foregoing paragraphs as if fully set forth herein.

433. Plaintiffs bring this claim individually and on behalf of the members of the proposed Class against Defendants.

434. CIPA § 631(a) imposes liability for “distinct and mutually independent patterns of conduct.” *Tavernetti v. Superior Ct.*, 22 Cal. 3d 187, 192 (1978). Thus, to establish liability under CIPA § 631(a), a plaintiff need only establish that the defendant, “by means of any machine, instrument, contrivance, or in any other manner,” does any of the following:

Intentionally taps, or makes any unauthorized connection, whether physically, electrically, acoustically, inductively or otherwise, with any telegraph or telephone wire, line, cable, or instrument, including the wire, line, cable, or instrument of any internal telephonic communication system,

OR

Willfully and without the consent of all parties to the communication, or in any unauthorized manner, reads or attempts to read or learn the contents or meaning of any message, report, or communication while the same is in transit or passing over any wire, line or cable or is being sent from or received at any place within this state,

OR

Uses, or attempts to use, in any manner, or for any purpose, or to communicate in any way, any information so obtained,

OR

Aids, agrees with, employs, or conspires with any person or persons to unlawfully do, or permit, or cause to be done any of the acts or things mentioned above in this section.

Cal. Penal Code § 631 (Deering 2023).

435. Section 631(a) is not limited to phone lines, but also applies to “new technologies”

such as computers, the Internet, and email. *See Matera v. Google Inc.*, No. 15-CV-04062-LHK, 2016 U.S. Dist. LEXIS 107918, at *61-*63 (N.D. Cal. Aug. 12, 2016) (CIPA applies to “new technologies” and must be construed broadly to effectuate its remedial purpose of protecting privacy); *Bradley v. Google, Inc.*, 2006 WL 3798134, at *5-6 (N.D. Cal. Dec. 22, 2006) (CIPA governs “electronic communications”); *In re Facebook, Inc. Internet Tracking Litigation*, 956 F.3d 589, 598-99 (9th Cir. 2020) (reversing dismissal of CIPA and common law privacy claims based on Facebook’s collection of consumers’ Internet browsing history).

436. Defendants’ ChatGPT platform is a “machine, instrument, contrivance, or ... other manner” used to engage in the prohibited conduct at issue here.

I. Defendants’ Interception of Communications of ChatGPT API Class Members which occurred on Third-Party Websites, Platforms, Applications, Programs which have integrated ChatGPT API. [Microsoft User Subclass is Excluded]

437. The allegations for violation of CIPA § 631(a) arise out of Defendants’ interception of Plaintiffs, ChatGPT API Class Members’ (collectively referred to as Chat-GPT API Class and Subclass) communications which occurred on various applications, platforms, websites which integrate ChatGPT technology (i.e., Stripe, Snapchat, etc.).

438. The transmissions of Plaintiffs’ and ChatGPT API Class Members’ communications (including but not limited to chats, comments, replies, searches, keystrokes, mouse clicks/movements, signals, browser activity, or other data, activity, or intelligence) on various applications, programs, platforms, websites which integrate ChatGPT API (i.e., Stripe, Snapchat, etc.) qualify as “electronic communications” under Cal. Penal Code §629.51(2).

439. By incorporating ChatGPT technology on third party platforms, Defendants are in the unique position of having unrestricted, real-time access to the users’ every input, move, chat, comment, reply, search, keystroke, or other browser activity/communication on the third-party platform.

440. As Plaintiffs and ChatGPT API Class Members interact with the third-party platform, Defendants intentionally tap, electrically or otherwise, the lines of internet communication between Plaintiffs and ChatGPT API Class Members, and/or third-party entities.

1 441. In disregard for Plaintiffs’ and ChatGPT API Class Members’ privacy rights,
2 Defendants act as a third-party “eavesdropper”, redirecting Plaintiffs and Chat-GPT API Members’
3 electronic communications to Defendants’ own servers for appropriation, and training of their
4 Products.

5 442. Defendants’ interception of the contents of Plaintiffs’ and ChatGPT API Class
6 Members’ communications happens contemporaneously with their exchange of such
7 communications, whether such communications are directed to Plaintiffs’ and ChatGPT API Class
8 Members’ friends, colleagues, or third-party entities. As described above, the ChatGPT technology,
9 integrated on various platforms, is designed to simultaneously intercept and send a recording of
10 each keystroke, mouse click, movement, writing, or other data, activity, or intelligence to
11 Defendants sufficient to not only identify Plaintiffs and ChatGPT API Class Members’, but also to
12 be able to understand, collect, and use for training Plaintiffs’ and ChatGPT API Class Members’
13 communications.

14 443. Through this calculated scheme of using ChatGPT technology, integrated on various
15 non-ChatGPT platforms (such as Snapchat, Stripe etc.) to intercept, acquire, transmit, and record
16 Plaintiffs’ and ChatGPT API Class Members’ electronic communications, Defendants willfully and
17 without valid consent from all parties to the communication, take unauthorized measures to read
18 and understand the contents or meaning of the electronic communications of Plaintiffs and ChatGPT
19 API Class. The interception and recording of electronic communications occurs while the electronic
20 communications are in transit or passing over any wire, line, or cable, or are being sent from or
21 received at any place.

22 444. Plaintiffs and ChatGPT API Class Members did not authorize Defendants to acquire
23 the content of their communications for the purposes of training Defendants’ Products.

24 445. Plaintiffs, individually, on behalf of the GPT API Class, also seek all monetary and
25 non-monetary relief allowed by law, including actual damages, statutory damages in accordance
26 with § 637.2(a), punitive damages, preliminary and other equitable or declaratory relief, and
27 attorneys’ fees and costs.
28

II. Microsoft's Interception of ChatGPT User Class Members' Communications on ChatGPT

446. The allegations for violation of CIPA § 631(a) arise out of Defendant Microsoft's interception of Plaintiffs' and ChatGPT User Class Members' communications which occurred on ChatGPT platform.

447. The transmissions of Plaintiffs' and ChatGPT User Class Members' communications (including but not limited to chats, comments, replies, searches, keystrokes, mouse clicks/movements, signals, browser activity, or other data, activity, or intelligence) on ChatGPT qualify as "electronic communications" under Cal. Penal Code §629.51(2).

448. By developing ChatGPT and controlling the extent of training/development of this program, Defendants are in the unique position of having unrestricted, real-time access to the users' every input, move, mouse click, chat, comment, reply, search, keystroke, browser activity, or other data, activity, or intelligence on ChatGPT.

449. As Plaintiffs and ChatGPT User Class Members ask questions, or otherwise interact with Defendant Open AI, Defendant Open AI intentionally aids and abets Defendant Microsoft to intentionally tap and intercept, electrically or otherwise, the lines of internet communications of Plaintiffs' and Chat-GPT User Class Members' searches and communications.

450. In disregard for Plaintiffs' and ChatGPT User Class Members' privacy rights, Defendant Microsoft acts as a third-party "eavesdropper," redirecting Plaintiffs and Chat-GPT User Class Members' electronic communications to Defendant Microsoft's own servers for appropriation, and training of their Products.

451. Defendant Microsoft's interception of the contents of Plaintiffs' and ChatGPT User Class Members' communications happens contemporaneously with their exchange of such communications, whether such communications are directed to Defendant Open AI or third-party entities (for instance, Expedia). As described above, the ChatGPT technology is designed to simultaneously intercept and send a recording of each keystroke, mouse click, movement, writing, or other data, activity, or intelligence to Defendant Microsoft sufficient to not only identify Plaintiffs and Chat-GPT User Members, but also to be able to understand, collect, and use for training

1 Plaintiffs' and Chat-GPT User Class Members' communications.

2 452. Defendant Microsoft intercepted communications including all text entry input as a
3 search within ChatGPT as well as intercepted numerous other forms of a user's navigation and
4 interaction with ChatGPT.

5 453. Through this calculated scheme of using ChatGPT to intercept, acquire, transmit, and
6 record Plaintiffs' and ChatGPT User Class Members' electronic communications, Defendant
7 Microsoft willfully and without any iota of valid consent from all parties to the communication,
8 takes unauthorized measures to read and understand the contents or meaning of the electronic
9 communications of Plaintiffs and Chat-GPT User Class. The interception and recording of
10 electronic communications occur while the electronic communications are in transit or passing over
11 any wire, line, or cable, or are being sent from or received at any place.

12 454. In sending and in acquiring the content of Plaintiffs' and Class Members'
13 communications on ChatGPT, Defendants' purpose was tortious, and designed to violate federal
14 and state laws. By intentionally using, or endeavoring to use, the contents of the electronic
15 communications of Plaintiffs, ChatGPT User Class Members, while knowing or having reason to
16 know that the information was obtained through the interception of an electronic communication,
17 Defendant Microsoft violates CIPA § 631(a).

18 455. Additionally, under the fourth clause of §631(a), Defendant OpenAI aided, agreed
19 with, and conspired with Defendant Microsoft to accomplish the wrongful conduct at issue here.
20 *Graham v. Noom, Inc.*, 533 F. Supp. 3d 823, 831-32 (N.D. Cal. 2021) (while a party to a
21 communication may record the communication without triggering § 631(a) liability, it will be
22 subject to derivative liability where the third party is liable for recording the communications in
23 violation of the first, second or third clauses of § 631(a)); *Revitch v. New Moosejaw, LLC*, No. 18-
24 cv-06827-VC, 2019 WL 5485330, at *2 (N.D. Cal. 2019) (conversation participants may be liable
25 because § 631 "was designed to protect a person placing or receiving a call from a situation where
26 the person on the other end of the line permits an outsider to tap his telephone or listen in on the
27 call.")

28 456. Plaintiffs, individually, on behalf of the GPT ChatGPT User Class Members, seek all

monetary and non-monetary relief allowed by law, including actual damages, statutory damages, punitive damages, preliminary and other equitable or declaratory relief, and attorneys' fees and costs.

III. Defendant Open AI's Interception of Microsoft User Class Members which occurred on Microsoft's Websites, Platforms, Applications, Programs which have integrated ChatGPT.

457. The allegations for violation of CIPA § 631(a) arise out of Defendant Open AI's interception of Microsoft User Class Members' (collectively "Microsoft Subclass") communications with their friends, family, colleagues, or other individuals or third-party entities, which occurred on Microsoft platforms (Teams, Bing, Outlook etc.), which integrate ChatGPT API.

458. The transmissions of Plaintiffs' and Microsoft Subclasses' communications (including but not limited to chats, comments, replies, searches, keystrokes, signals, mouse clicks/movements, , browser activity, or other data, activity, or intelligence) on Microsoft's various applications, programs, platforms, websites which integrate ChatGPT API qualify as "electronic communications" under Cal. Penal Code §629.51(2).

459. By integrating ChatGPT technology within the entire Microsoft suite, Defendant OpenAI is in the unique position of having unrestricted, real-time access to the users' every input, move, mouse click, chat, comment, reply, search, keystroke, browser activity, or other data, activity, or intelligence.

460. As Plaintiffs and Microsoft Subclasses interact with each other or the third-party entities, Defendant OpenAI intentionally taps, electrically or otherwise intercept, the lines of internet communications between Plaintiffs, Microsoft Subclasses, and/or third-party entities.

461. In disregard for Plaintiffs' and Microsoft Subclasses Members' privacy rights, Defendant OpenAI acts as a third-party "eavesdropper," redirecting Plaintiffs and Microsoft Subclasses Members' electronic communications to Defendants' own servers for appropriation, and training of their Products.

462. Defendant Open AI's interception of the contents of Plaintiffs' and Microsoft Subclasses Members' communications happens contemporaneously with their exchange of such

communications on Microsoft platforms, whether such communications are directed to Plaintiffs' and Microsoft Subclasses Members' friends, colleagues, or third-party entities. As described above, the ChatGPT API is designed to simultaneously intercept and send a recording of each keystroke, mouse click, signal, movement, writing, or other data, activity, or intelligence to Defendant Open AI sufficient to not only identify Plaintiffs and Microsoft Subclasses Members, but also to be able to understand, collect, and use for training Plaintiffs' and Microsoft Subclasses Members' communications.

463. Additionally, under the fourth clause of §631(a), Defendant Microsoft aided, agreed with, and conspired with Defendant OpenAI to implement AI technology within its own platforms. The incorporation of such technology shares users' electronic communications with Microsoft platforms with OpenAI in an effort to accomplish the wrongful conduct at issue here. *Graham v. Noom, Inc.*, 533 F. Supp. 3d 823, 831-32 (N.D. Cal. 2021) (while a party to a communication may record the communication without triggering § 631(a) liability, it will be subject to derivative liability where the third party is liable for recording the communications in violation of the first, second or third clauses of § 631(a)); *Revitch v. New Moosejaw, LLC*, No. 18-cv-06827-VC, 2019 WL 5485330, at *2 (N.D. Cal. 2019) (conversation participants may be liable because § 631 "was designed to protect a person placing or receiving a call from a situation where the person on the other end of the line permits an outsider to tap his telephone or listen in on the call.")

464. Plaintiffs, individually, on behalf of the Microsoft Subclasses Members, seek all monetary and non-monetary relief allowed by law, including actual damages, statutory damages, punitive damages, preliminary and other equitable or declaratory relief, and attorneys' fees and costs.

465. Unless enjoined, Defendants will continue to commit the illegal acts alleged here.

466. Plaintiffs and Class Members seek all relief available under Cal. Penal Code § 637.2, including injunctive relief and statutory damages of \$5,000 per violation.

COUNT FOUR

VIOLATION OF CALIFORNIA UNFAIR COMPETITION LAW (Cal. Bus. & Prof. Code

§§ 17200, et seq.)

(on behalf of All Plaintiffs against Defendants)

467. Plaintiffs repeat and reallege the allegations set forth in the preceding paragraphs and incorporate the same as if set forth herein at length.

468. As discussed above, Plaintiffs believe that California law should apply to all Plaintiffs, including out of state residents. Plaintiffs will collectively refer to all classes and subclasses as “Nationwide Classes,” and to subclass of minors as “Minors User Subclass.”

469. California Business & Professions Code, sections 17200, *et seq.* (the “UCL”) prohibits unfair competition and provides, in pertinent part, that “unfair competition shall mean and include unlawful, unfair or fraudulent business practices and unfair, deceptive, untrue or misleading advertising.”

I. Unlawful

470. Defendants engaged in and continue to engage in “unlawful” business acts and practices under the Unfair Competition Law because Defendants took, accessed, intercepted, tracked, collected, or used the Plaintiffs’ and Nationwide Classes’ Private Information, including but not limited to their private conversations, personally identifiable information, financial and medical data, keystrokes, searches, cookies, browser activity and other data, and shared this information with each other, while also using this information to train Defendants’ AI Products. Defendants’ unlawful conduct is as follows:

a) Web-Scraping and Interception of Communications, Private Information and Data:

Defendants scraped nearly the entire internet in order to train their AI Products, and in this process, Defendants accessed, and stole private conversations, personal information, and other private data from websites including Reddit, Twitter, TikTok, Spotify, YouTube, and other websites, without consent of the individuals. Defendants’ illegal web scraping violates privacy laws, and other laws outlined in this complaint. Defendants failed to register as data brokers under California law as required.

b) Defendants’ Intercepted Communications and Accessed, Collected, and Tracked Private Information from Platforms Which Integrated ChatGPT: Defendants

intercepted, tracked, and recorded communications, messages, chats, web activity, user activity, associated cookies, keystrokes and other Private Information through its ChatGPT technology integrated within hundreds of applications (including but not limited to Stripe, Snapchat, Expedia etc.) which were used to train Defendants' Products. Defendants' illegal tracking of such data, which is subsequently used to train Defendants' AI products violates privacy laws, California wiretapping law, and other laws outlined in this complaint.

c) Open AI's Interception of Communications and Accessed, Collected, and Tracked Private Information on Microsoft Platforms: Defendant Microsoft aided Defendant Open AI in intercepting, tracking, and recording communications, messages, chats, web activity, user activity, associated cookies, and other Private Information through its ChatGPT technology integrated within the entire Microsoft suite (Microsoft Teams, Microsoft Outlook, Bing). Defendant's Open AI illegal tracking of such data and Defendant Microsoft's aiding and abetting this conduct violates privacy laws, California wiretapping law, and other laws outlined in this complaint.

d) Microsoft's Interception of Communications and Accessed, Collected, and Tracked Private Information on ChatGPT: Defendant OpenAI aided Defendant Microsoft in intercepting, tracking, and recording communications, messages, chats, web activity, user activity, associated cookies, and other Private Information by sharing access to ChatGPT and sending all communications to Defendant Microsoft and its partners.

471. Defendants' conduct as alleged herein was unfair within the meaning of the UCL. The unfair prong of the UCL prohibits unfair business practices that either offend an established public policy or that are immoral, unethical, oppressive, unscrupulous, or substantially injurious to consumers.

472. Defendants' conduct violates the EPCA, CFAA, CIPA, California Consumer Privacy Act ("CCPA"), Cal. Civ. Code § 1798.100, *et seq.*, and the California Online Privacy Protection Act ("CalOPPA"), Section 5 of the Federal Trade Commission Act ("FTCA"), Cal. Bus. & Prof. Code § 22575, *et seq.*, and other tort claims stated in this lawsuit. The violations of EPCA, CFAA,

1 CIPA, and other tort claims stated in this lawsuit, are incorporated herein by reference.

2 473. Under the CCPA, a business that collects consumers' personal information is
3 required, at or before the point of collection, to provide notice to consumers indicating: (1) "[t]he
4 categories of personal information to be collected and the purposes for which the categories of
5 personal information are collected or used and whether that information is sold or shared"; (2) "the
6 categories of sensitive personal information to be collected and the purposes for which the
7 categories of sensitive personal information are collected or used, and whether that information is
8 sold or shared."; and (3) "[t]he length of time the business intends to retain each category of personal
9 information . . ." Cal. Civ. Code § 1798.100(a).

10 474. "Personal information" is defined by the CCPA as "information that identifies, relates
11 to, describes, is reasonably capable of being associated with, or could reasonably be linked, directly
12 or indirectly, with a particular consumer or household." Cal. Civ. Code § 1798.140(v)(1).

13 475. As alleged, Defendant uses web scraping technology to collect information from
14 webpages across the internet and, in so doing, Defendant gathers and compiles personal information
15 about consumers that is reflected on those webpages.

16 476. Because Defendants conduct web scraping across millions of web pages, without
17 asking the affected consumers their permission to use their content for training, Defendants do not,
18 and cannot provide consumers with the notice required by Cal. Civ. Code § 1798.100(a) at or before
19 the point of collection. Similarly, when Defendants intercept and wiretap users' communications on
20 various platforms which integrate ChatGPT, Microsoft platforms, and ChatGPT platforms, to use
21 these intercepted communications and gathered data to train their Products. Defendants never
22 notified Plaintiffs and affected Nationwide Classes Members of this extensive wiretapping, and
23 more importantly, that this information would be used for commercial purposes and development
24 of Defendants' Products. Therefore, Defendants failed to provide notice to the affected consumers
25 as required by Cal. Civ. Code § 1798.100(a).

26 477. Defendant's failure to provide notice to Plaintiffs and Nationwide Classes Members
27 whose personal information is collected through the process of web scraping and illegal wiretapping
28 is unlawful and violates Cal. Civ. Code § 1798.100(a).

1 478. The CCPA further grants consumers the right to “request that a business that collects
2 a consumer’s personal information disclose to that consumer the categories and specific pieces of
3 personal information the business has collected.” Cal. Civ. Code § 1798.100(b).

4 479. Upon receipt of a verifiable request for disclosure pursuant to Section 1798.110, a
5 business must “disclose any personal information it has collected about a consumer, directly or
6 indirectly, including through or by a service provider or contractor, to the consumer . . .” Cal. Civ.
7 Code § 1798.130 (3)(A).

8 480. Any disclosure must provide the requesting consumer with all of the following: (1)
9 “The categories of personal information it has collected about that consumer”; (2) “The categories
10 of sources from which the personal information is collected”; (3) “The business or commercial
11 purpose for collecting, selling, or sharing personal information” (4) “The categories of third parties
12 to whom the business discloses personal information”; and (5) “The specific pieces of personal
13 information it has collected about that consumer.” Cal. Civ. Code § 1798.110(a).

14 481. Consumers also “have the right to request that a business delete any personal
15 information about the consumer which the business has collected from the consumer.” Cal. Civ.
16 Code § 1798.105(a).

17 482. Pursuant to Cal. Civ. Code §§ 1798.100(b) and 1798.130(a), OpenAI’s privacy policy
18 provides a method by which California residents who have had their data collected may request
19 disclosure of the categories and specific pieces of personal information OpenAI has collected about
20 them.²⁵⁴ Open AI’s privacy policy specifically states that consumers “may have certain statutory
21 rights in relation to their Personal Information,” including the right to “Access your Personal
22 Information.”²⁵⁵

23 483. To exercise their right to access the Personal Information OpenAI has collected about
24 them, consumers are instructed to email their request for disclosure to dsar@openai.com.²⁵⁶

25 484. Under the heading “Additional U.S. State Disclosures,” the privacy policy states that
26

27 ²⁵⁴ *Privacy Policy*, OPENAI, <https://openai.com/policies/privacy-policy> (last updated June 23,
2023).

28 ²⁵⁵ *Id.*

²⁵⁶ *Id.*

1 some users may have “[t]he right to know information about our processing of your Personal
 2 Information, including the specific pieces of Personal Information that we have collected from you
 3 . . .”²⁵⁷ Users are instructed that, “to the extent applicable under local law, [they] can exercise privacy
 4 rights. . . by submitting a request to dsar@openai.com.”²⁵⁸

5 485. Yet OpenAI fails to disclose that once its AI Products have been trained on an
 6 individual’s information, that information has been included into the product and cannot reasonably
 7 be extracted. Whether individuals’ information was collected through web scraping or obtained
 8 through interception from ChatGPT, or other platforms incorporating ChatGPT, this information,
 9 once used to train Products, cannot be extracted. Therefore, Defendants violated and continue to
 10 violate CCPA.

11 486. CalOPPA applies to Defendant OpenAI because it operates a commercial website and
 12 online service that collects personally identifiable information about individual consumers residing
 13 in California. Cal. Bus. & Prof. Code § 22575(a).

14 487. CalOPPA defines personally identifiable information as first and last name; home or
 15 other physical address, including street name and name of a city or town; e-mail address; telephone
 16 number; social security number; any other identifier that permits the physical or online contacting
 17 of a specific individual; information concerning a user that the website or online service collects
 18 online from the user and maintains in personally identifiable form in combination with an identifier
 19 described in this subdivision. Cal. Bus. & Prof. Code § 22577(a).

20 488. OpenAI violates CalOPPA because while its privacy policy instructs consumers
 21 regarding how they can review and request changes to OpenAI’s collection of their data, the
 22 disclosures in this regard are misleading and incomplete in that they do not disclose that data used
 23 to train the Products realistically cannot be deleted from the Products.

24 489. OpenAI also violates CalOPPA by failing to disclose whether other parties may
 25 collect personally identifiable information about an individual consumer’s online activities over
 26 time and across different Web sites when a consumer uses the OpenAI’s website of ChatGPT
 27

28 ²⁵⁷ *Id.*

²⁵⁸ *Id.*

1 service.

2 490. Furthermore, OpenAI also violates CalOPPA by knowingly collecting information
3 from minors under the age of thirteen (“13”) without appropriate measures to ensure parental
4 consent and without ensuring that the full deletion of information about minors is feasible from their
5 products.

6 491. Plaintiffs, individually and on behalf of the Nationwide Classes seek: (i) an injunction
7 requiring OpenAI to revise its privacy policy to include reasonable protections for children and
8 Minors User Subclass, to fully disclose all information required under CalOPPA and COPPA, and
9 to delete all information previously collected in violation of these laws; (ii) an injunction requiring
10 OpenAI to revise its privacy policy to fully disclose all information required under CCPA, and to
11 delete all information previously collected in violation of these laws; (iii) relief under Cal. Bus. &
12 Prof. Code § 17200, et seq., including, but not limited to, restitution to Plaintiffs and other members
13 of the Nationwide Classes of money or property Defendants acquired by means of their unlawful
14 business practices; and, as a result of bringing this action to vindicate and enforce an important right
15 affecting the public interest, (iv) reasonable attorney’s fees (pursuant to Cal. Code of Civ. P. §
16 1021.5).

17 492. Defendants’ unlawful actions in violation of the UCL have caused and are likely to
18 cause substantial injury to consumers that consumers cannot reasonably avoid themselves and that
19 is not outweighed by countervailing benefits to consumers or competition.

20 493. As a direct and proximate result of Defendants’ misconduct, Plaintiffs and Nationwide
21 Classes Members had their private communications containing information related to their sensitive
22 and confidential Private Information intercepted, disclosed, and used by third parties, including but
23 not limited to each Defendant.

24 494. As a result of Defendants’ unlawful conduct, Plaintiffs and Nationwide Classes
25 Members suffered an injury, including violation to their rights of privacy, loss of value and privacy
26 of their Private Information, loss of control over their sensitive personal information, and suffered
27 embarrassment and emotional distress as a result of this unauthorized scraping, interception,
28 sharing, and misuse of information.

II. Unfair

495. Defendants' conduct as alleged herein was unfair within the meaning of the UCL. The unfair prong of the UCL prohibits unfair business practices that either offend an established public policy or that are immoral, unethical, oppressive, unscrupulous or substantially injurious to consumers.

496. Defendants also engaged in business acts or practices deemed "unfair" under the UCL because, as alleged above, Defendants failed to disclose that they scraped information belonging to millions of internet users without the users' consent. Defendants also failed to disclose that they used the stolen information to train their Products, without consent of the internet users. Furthermore, Defendants failed to disclose that they were intercepting, tracking Private Information belonging to millions of ChatGPT users, and the users of other platforms which integrated ChatGPT. Private Information obtained from individual uses of ChatGPT and other platforms which integrate ChatGPT was and is continued to be used to train Defendants' Products, without consent of the users. Defendants also illegally profiled, and targeted Plaintiffs' and Minor User Subclass' members through the collection of Private Information, to profit from such uses.

497. Unfair acts under the UCL have been interpreted using three different tests: (1) whether the public policy which is a predicate to a consumer unfair competition action under the unfair prong of the UCL is tethered to specific constitutional, statutory, or regulatory provisions; (2) whether the gravity of the harm to the consumer caused by the challenged business practice outweighs the utility of the defendant's conduct; and (3) whether the consumer injury is substantial, not outweighed by any countervailing benefits to consumers or competition, and is an injury that consumers themselves could not reasonably have avoided.

498. Defendants' conduct is unfair under each of these tests. As described above, Defendants' conduct violates the policies underlying privacy laws and, with respect to children under the age of thirteen, the mandates of COPPA and CalOPPA. The gravity of the harm of Defendants' illegal scraping, interception and misuse of Private information to train their AI Products, as well as secret tracking, profiling, and targeting of children is significant and there is no corresponding benefit to consumers of such conduct.

1 499. Finally, because Plaintiffs [Initials] and Minor User Subclass Members were minors
2 unable to consent to or understand Defendants' conduct—and because their parents did not consent
3 to this conduct and were misled by their belief that Defendants would follow applicable laws and
4 societal expectations about children's privacy as well as Defendants' statements—they could not
5 have avoided the harm.

6 500. Under the UCL, a business practice that is likely to deceive an ordinary consumer
7 constitutes a deceptive business practice. Defendants' conduct was deceptive in numerous respects.

8 501. Defendants have intentionally and deceptively misled parents and the public about
9 Defendants' intention to use the ChatGPT language model and its free chatbox application to attract
10 children in order to gain access to the Personal Information of such children and to exploit such
11 children's Personal Information for Defendants' financial gain.

12 502. Defendants' misrepresentations and omissions include both implicit and explicit
13 representations.

14 503. Defendants' representations and omissions were material because they were likely to
15 deceive reasonable consumers such as the parents or guardians of Plaintiffs and Class and Subclass
16 Members about the terms under which their children were interacting with the ChatGPT app as well
17 as the fact that Defendant was collecting and profiting from minors' Personal Information without
18 their parents and guardians' knowledge or consent.

19 504. Defendants had a duty to disclose the above-described facts due to the important
20 public interest in securing the privacy of minors' Personal Information and the fact that minors are
21 unable to fully protect their own interests.

22 505. Defendant OpenAI represented, throughout the Class Period, that it would “respect
23 your privacy and [is] strongly committed to keeping secure any information we obtain from you or
24 about you.”

25 506. The expectations of Plaintiffs' parents and guardians included that Defendants would
26 not track their children's online activity, without their consent, in order for Defendants to reap huge
27 profits from building out the fastest growing application ever, and the most advanced AI language
28 models of all time.

1 507. The parents and guardians of Plaintiffs and Minor User Subclass members reasonably
2 expected that Defendants respected children's privacy online, in accordance with societal
3 expectations and public policy as well as state and federal statutes and regulations including
4 COPPA, CalOPPA, and Federal Trade Commission regulations.

5 508. At the same time, Defendants have, at all times throughout the Class Period, been well
6 aware that children, including children under the age of 16 and under the age of 13, access ChatGPT;
7 have actively sought to increase engagement with ChatGPT by children; and have sought to exploit,
8 for commercial purposes and gain, thousands if not millions of minor users of ChatGPT.

9 509. Defendants' knowledge of the widespread use of ChatGPT by children and failure to
10 disclose that they are tracking, profiling, and targeting such children and/or profiting from this
11 behavior, while at the same time representing that OpenAI and ChatGPT comply with law and
12 societal expectation, and does not permit and does not seek to reach children, are likely to and, in
13 fact, did deceive Plaintiffs and Minor User Subclass Members and their parents or guardians.
14 Defendants' conduct therefore constitutes deceptive business practices in violation of Cal. Bus. &
15 Prof. Code §17200.

16 510. Additionally, to the extent that Defendants have represented to Plaintiffs, Minor User
17 Subclass members, and their respective parents and guardians that Defendants can and will disclose
18 to such individuals, upon request, the private information that Defendants have gathered about any
19 such minor user or non-user, and that such information can be deleted, these representations are
20 fraudulent and deceptive because it is functionally impossible for Defendants to "undo" the fact that
21 their LLMs have learned on this private information and incorporated that learning in such a manner
22 that the information cannot be meaningfully segregated, identified, extracted, and deleted.

23 511. Defendants' conduct, as alleged herein, was fraudulent within the meaning of the
24 UCL. Defendants made deceptive misrepresentations and omitted known material facts in
25 connection with the solicitation, interception, disclosure, and use of Plaintiffs' and Minor User
26 Subclass Members' User Data. Defendants actively concealed and continued to assert misleading
27 statements regarding their protection and limitation on the use of the User Data. Meanwhile,
28 Defendants were collecting and sharing Plaintiffs' and Class Members' User Data without their

1 authorization or knowledge in order to profit off of the information, and to deliver advertisements
2 to Plaintiffs and Class Members, among other unlawful purposes.

3 512. Defendants' conduct, as alleged herein, was unlawful within the meaning of the UCL
4 because Defendants violated regulations and laws as discussed herein, including but not limited to
5 HIPAA, Section 5 of the Federal Trade Commission Act ("FTCA"), 15 U.S.C. § 45 and the CIPA.

6 513. Defendants have unlawfully tracked, targeted, and profiled minor Plaintiffs, and
7 Minor User Subclass Members without obtaining parental consent in violation of COPPA,
8 CalOPPA, Federal Trade Commission regulations, and other laws.

9 514. Defendants also engaged in business acts and practices deemed "unlawful" under the
10 UCL as to the Nationwide Classes by unlawfully tracking, targeting, and profiling Plaintiffs' minor
11 children, in violation of the California Constitution.

12 515. Defendants reaped profits from these actions in the form of increased company
13 valuation, investments, improved language model performance, and dominance in the AI field.

14 516. Defendants' unlawful actions in violation of the UCL have caused and are likely to
15 cause substantial injury to consumers that consumers cannot reasonably avoid themselves and that
16 is not outweighed by countervailing benefits to consumers or competition.

17 517. As a direct and proximate result of Defendants' misconduct, Plaintiffs and Nationwide
18 Classes Members had their private communications containing information related to their sensitive
19 and confidential User Data intercepted, disclosed, and used by third parties, including but not limited
20 to each Defendant.

21 518. As a result of Defendants' unlawful conduct, Plaintiffs and Nationwide Classes
22 Members suffered an injury, including violation to their rights of privacy, loss of the privacy of their
23 PHI/PII, loss of control over their sensitive personal information, loss of autonomy over their minor
24 children and their minor children's data, and suffered aggravation, inconvenience, and emotional
25 distress.

26 519. Plaintiffs and Minor User Subclass Members placed trust in Defendants as major and
27 reputable companies that represented they were in compliance with applicable laws and societal
28 interests in safeguarding minors' Personal Information.

1 520. Additionally, Defendants had the sole ability to understand the extent of their
2 collection of Personal Information, and the parents or guardians of Plaintiffs and Minor User
3 Subclass Members could not reasonably have discovered—and were unaware of—Defendants’
4 secret tracking, profiling, and targeting.

5 521. Defendants invaded Plaintiffs’ and Minor User Subclass Members’ privacy without
6 their or their parents and guardians’ consent.

7 522. Because Defendants held themselves out as complying with law and public policy
8 regarding minors’ privacy rights, the parents or guardians of Plaintiffs and California Minor User
9 Subclass Members acted reasonably in relying on Defendants’ misrepresentations and omissions.

10 523. Plaintiffs and Minor User Subclass Members could not have reasonably avoided
11 injury because Defendants’ business acts and practices unreasonably created or took advantage of
12 an obstacle to the free exercise of their decision-making. By withholding the important information
13 that it was collecting and profiting from minors’ Personal Information, Defendants created an
14 asymmetry of information.

15 524. Further, Defendants’ conduct is immoral, unethical, oppressive, unscrupulous and
16 substantially injurious to Plaintiffs, Nationwide Classes Members, and Minor User Subclass, and
17 there are no greater countervailing benefits to consumers or competition.

18 525. Plaintiffs, as well as the Nationwide Classes Members and Minor User Subclass
19 Members, were harmed by Defendants’ violations of Cal. Bus. & Prof. Code §17200. Defendants’
20 practices were a substantial factor and caused injury in fact and actual damages to Plaintiffs and
21 Nationwide Classes and Minor User Subclass Members.

22 526. As a direct and proximate result of Defendants’ deceptive acts and practices,
23 Plaintiffs, Nationwide Classes Members, and Minor User Subclass Members have suffered and will
24 continue to suffer an ascertainable loss of money or property, real or personal, and monetary and
25 non-monetary damages, as described above, including the loss or diminishment in value of their
26 Private Information and the loss of the ability to control the use of their Private Information, which
27 allowed Defendants to profit at the expense of Plaintiffs, Nationwide Classes Members, and Minor
28 User Subclass Members.

1 527. Plaintiffs', Nationwide Classes Members', and Minors' Members' Personal
2 Information has tangible value; it is now in the possession of Defendants, who has used and will
3 continue to use it for financial gain.

4 528. Plaintiffs', Nationwide Classes Members, and Minor User Subclass Members' injury
5 was the direct and proximate result of Defendant's conduct described herein.

6 529. Defendants' retention of Plaintiffs' and Nationwide Minor User Class, and California
7 Minor User Subclass Members' Personal Information presents a continuing risk to them as well as
8 the general public.

9 530. Plaintiffs, individually and on behalf of the Nationwide Classes Members, and Minor
10 User Subclass, seek: (1) an injunction requiring Defendants to permanently delete, destroy or
11 otherwise sequester the Private Information collected without consent, and with respect to minors,
12 without parental consent; (2) compensatory restitution of Plaintiffs', Nationwide Classes Members,
13 and Minor Class Members' money and property lost as a result of Defendants' acts of unfair
14 competition; (3) disgorgement of Defendants' unjust gains; and (4) reasonable attorney's fees
15 (pursuant to Cal. Code of Civ. Proc. § 1021.5).

16 531. Had Plaintiffs, Nationwide Classes Members, and Minor User Subclass Members
17 known Defendants would disclose and misuse their User Data in contravention of Defendants'
18 representations, they would not have used Defendants' Products.

19 529. Defendants' unlawful actions in violation of the UCL have caused and are likely to cause
20 substantial injury to consumers that consumers cannot reasonably avoid themselves and that
21 is not outweighed by countervailing benefits to consumers or competition.

22 532. As a direct and proximate result of Defendants' misconduct, Plaintiffs and Nationwide
23 Classes Members had their private communications containing information related to their sensitive
24 and confidential Private Information intercepted, disclosed, and used by Defendants, to train their
25 Products.

26 533. As a result of Defendants' unlawful conduct, Plaintiffs Nationwide Classes Members,
27 and Minor User Subclass members suffered an injury, including violation to their rights of privacy,
28 loss of the privacy of their Private Information loss of control over their sensitive personal

information, and suffered aggravation, inconvenience, and emotional distress.

COUNT FIVE

VIOLATION OF ILLINOIS’S BIOMETRIC INFORMATION PRIVACY ACT, 740 ILCS

14/1, et seq.

(on behalf of Illinois Plaintiff and Illinois Subclasses against Defendants)

534. The Illinois Plaintiff, individually and on behalf of the Illinois Subclasses, repeats and re-alleges the allegations contained in the foregoing paragraphs as if fully set forth herein.

535. BIPA created statutory duties for Defendants with respect to the collection of biometric identifiers and biometric information of the Illinois Plaintiff and the Illinois Subclasses.

536. Defendants violated BIPA section 15(b)(1) by systematically collecting the Illinois Plaintiff’s and the Illinois Subclasses’ biometric identifiers and biometric identifiers, by taking their photos off of the internet and scanning their facial geometry and related biometric information to train the algorithms on which DALL-E runs, without first informing the Illinois Plaintiff and the Illinois Subclasses.

537. In so doing, Defendants also violated section 15(b)(2) of BIPA by not informing Illinois Plaintiff and the Illinois Subclasses in writing of the purpose for their collection of facial geometry and related biometric information, and by failing to inform them in writing of the length of time Defendants would collect their biometric identifiers and biometric information, including scans of their facial geometry and related biometric information.

538. Defendants violated section 15(b)(3) of BIPA by not receiving a written release executed by Illinois Plaintiff and the Illinois Subclasses, the subjects of the biometric identifiers and biometric information.

539. Section 15(c) of BIPA makes it unlawful for any private entity to among other things, “sell, lease, trade, or otherwise profit from a person’s or a customer’s biometric identifier or biometric information,” which Defendants did by incorporating that information into the Products for commercial gain. Without the facial scans and recording of facial geometry of Illinois Plaintiff and the Illinois Subclass, DALL-E could not exist.

540. BIPA prohibits private entities “in possession of a biometric identifier or biometric

information” from “disclos[ing], redisclos[ing], or otherwise disseminat[ing] a person’s or a customer’s biometric identifier or biometric information unless” any one of four enumerated conditions are met. 740 ILCS 14/15(d)(1)-(4). None of such conditions are met here.

541. Defendants disclose, redisclose and disseminate, and at all relevant times disclosed, redisclosed and disseminated, the Illinois Plaintiff’s and the Illinois Subclasses’s “biometric identifiers,” including but not limited to their face geometry scans, and “biometric information” without the consent of any of them or their “legally authorized representatives.” 740 ILCS 14/15(d)(1). Moreover, the disclosures and redisclosures did not “complete[] a financial transaction requested or authorized by” the Illinois Plaintiff, the Illinois Subclasses or any of their legally authorized representatives. 740 ILCS 14/15(d)(2). Nor are, or at any relevant times were, the disclosures and redisclosures “required by State or federal law or municipal ordinance.” 740 ILCS 14/15(d)(3). Finally, at no point in time were the disclosures ever “required pursuant to a valid warrant or subpoena issued by a court of competent jurisdiction.” 740 ILCS 14/15(d)(4).

542. BIPA mandates that a private entity “in possession of biometric identifiers or biometric information” “develop a written policy, made available to the public, establishing a retention schedule and guidelines for permanently destroying biometric identifiers and biometric information when the initial purpose for collecting or obtaining such identifiers or information has been satisfied or within 3 years of the individual’s last interaction with the private entity, whichever occurs first.” 740 ILCS 14/15(a).

543. But Defendants do not publicly provide any written policy establishing any retention schedule or guidelines for permanently destroying the Illinois Plaintiff’s and the Illinois Subclasses’s “biometric identifiers” and “biometric information.” 740 ILCS 14/15(a).

544. BIPA also commands private entities “in possession of a biometric identifier or biometric information” to: (1) store, transmit, and protect from disclosure all biometric identifiers and biometric information using the reasonable standard of care within the private entity’s industry; and (2) store, transmit, and protect from disclosure all biometric identifiers and biometric information in a manner that is the same as or more protective than the manner in which the private entity stores, transmits and protects other confidential and sensitive information. 740 ILCS 14/15(e).

Based on the facts alleged herein, including Defendants' lack of a public written policy, their failure to inform Users that Defendants obtain such users' "biometric identifiers" and "biometric information," their failure to obtain written consent to collect or otherwise obtain Users' "biometric identifiers" and "biometric information," and their unauthorized dissemination of Users' "biometric identifiers" and "biometric information," Defendants have also violated this provision.

545. Illinois Plaintiff and the Illinois Subclasses have been directly harmed by these violations. They have been deprived of their control over valuable information, and otherwise suffered monetary and non-monetary losses. By depriving them of control over their valuable information, Defendants misappropriated the value of their biometric identifiers and biometric information, and are profiting from this unlawful conduct.

546. Illinois Plaintiff and the Illinois Subclasses seek (i) injunctive and equitable relief requiring Defendants to comply with BIPA; (ii) statutory damages of \$5,000 per intentional or reckless violation of BIPA and statutory damages of \$1,000 per negligent violation of BIPA; and (iii) reasonable attorneys' fees and costs and other litigation expenses as permitted by statute. 740 ILCS 14/20(1)-(4).

COUNT SIX

ILLINOIS CONSUMER FRAUD AND DECEPTIVE BUSINESS PRACTICES ACT

815 ILL. COMP STAT. §§ 505, et seq.

(on behalf of Illinois Plaintiff and Illinois Subclasses against Defendants)

450. The Illinois Plaintiff, individually and on behalf of the Illinois Subclasses, repeats and re-alleges the allegations contained in the foregoing paragraphs as if fully set forth herein.

451. Defendant OpenAI and Defendant Microsoft are "persons" as defined by 815 Ill. Comp. Stat. §§ 505/1(c).

452. The Illinois Plaintiff and Illinois Subclasses Members are "consumers" as defined by 815 Ill. Comp. Stat. §§ 505/1(e).

453. Defendants' conduct as described herein was in the conduct of "trade" or "commerce" as defined by 815 Ill. Comp. Stat. § 505/1(f).

454. Defendants' deceptive, unfair, and unlawful trade acts or practices, in violation of 815

1 Ill. Comp. Stat. § 505/2, include:

- 2 a) Defendants have exploited Non-Users and Users of their Products, by stealing such
3 individuals' data at scale from web crawler caches without permission from the data
4 owners and without any way of segregating out any given Non-Users' or User' data from
5 the datasets used to train OpenAI's LLMs upon request of such individuals—including
6 where such individuals are minors.
- 7 b) Defendants knew that they were collecting and/or profiting from individuals' Personal
8 Information and that the risk of collecting of such Personal Information was highly likely.
9 Defendants' actions in engaging in the above-named deceptive acts and practices were
10 negligent, knowing and willful, and/or wanton and reckless with respect to the rights of
11 the Illinois Plaintiff and members of the Illinois Subclasses;
- 12 c) As described herein, Defendants are misrepresenting that they have and are complying
13 with common law and statutory duties pertaining to the security and privacy of the Illinois
14 Plaintiff's and Illinois Subclass Members' data, including but not limited to duties
15 imposed by the FTC Act, 15 U.S.C. § 45, the Illinois Uniform Deceptive Trade Practices
16 Act, 815 Ill. Comp. Stat. § 510/2(a), and the Illinois Personal Information Act, 815 Ill.
17 Comp. Stat. §§ 530/10(a).
- 18 d) As described herein, Defendants have and are omitting, suppressing, and concealing the
19 material fact that they are stealing and profiting from the mass collection and analysis of
20 the Illinois Plaintiff's and Illinois Subclasses Members' data at scale and without adequate
21 or effective consent; and
- 22 e) Omitting, suppressing, and concealing the material fact that they did not comply with
23 common law and statutory duties pertaining to the security and privacy of Illinois
24 Plaintiff's and the Illinois Subclass Members' data, including but not limited to the fact
25 that they are functionally unable to delete such data once it has been incorporated into
26 their LLMs as training data.

27 455. Defendants' representations and omissions were material because they were likely to
28 deceive reasonable consumers about the terms of use of the Products, as well as the available

1 mechanisms for seeking to exert control over Illinois Plaintiff's and Illinois Subclasses Members'
2 data.

3 456. Defendants intended to mislead the Illinois Plaintiff and Illinois Subclasses Members
4 and induce them to rely on their misrepresentations and omissions.

5 457. The above unfair and deceptive practices and acts by Defendants were immoral,
6 unethical, oppressive, and unscrupulous. These acts caused substantial injury that these consumers
7 could not reasonably avoid; this substantial injury outweighed any benefits to consumers or to
8 competition.

9 458. Defendants acted intentionally, knowingly, and maliciously to violate Illinois's
10 Consumer Fraud Act, and recklessly disregarded Illinois Plaintiff's and Illinois Subclasses
11 Members' rights.

12 459. As a direct and proximate result of Defendants' unfair, unlawful, and deceptive acts
13 and practices, the Illinois Plaintiff and Illinois Subclasses Members have suffered and will continue
14 to suffer injury, ascertainable losses of money or property, and monetary and non-monetary
15 damages, as described herein.

16 460. The Illinois Plaintiff and Illinois Subclasses Members seek all monetary and non-
17 monetary relief allowed by law, including damages, restitution, punitive damages, injunctive relief,
18 and reasonable attorneys' fees and costs.

19 **COUNT SEVEN**

20 **ILLINOIS CONSUMER FRAUD AND DECEPTIVE BUSINESS PRACTICES ACT 815**

21 **ILL. COMP. STAT. §§ 510/2, et seq.**

22 **(on behalf of Illinois Plaintiff and Illinois Subclasses against Defendants)**

23 461. The Illinois Plaintiff, individually and on behalf of the Illinois Subclasses, repeats and
24 re-alleges the allegations contained in the foregoing paragraphs as if fully set forth herein.

25 462. Defendant OpenAI and Defendant Microsoft are "persons" as defined by 815 Ill.
26 Comp. Stat. §§ 510/1(5).

27 463. Defendants engaged in deceptive trade practices in the conduct of its business, in
28 violation of 815 Ill. Comp. Stat. §§ 510/2(a), including:

- a) Representing that goods or services have characteristics that they do not have, 815 Ill. Comp. Stat. § 510/2(a)(5);
- b) Representing that goods or services are of a particular standard, quality, or grade if they are of another, 815 Ill. Comp. Stat. § 510/2(a)(7);
- c) Advertising goods or services with intent not to sell them as advertised, 815 Ill. Comp. Stat. § 510/2(a)(9); and
- d) Engaging in other conduct that creates a likelihood of confusion or misunderstanding, 815 Ill. Comp. Stat. § 510/2(a)(12).

464. Defendants' deceptive acts and practices include those enumerated, *supra*, in paragraph 454.

465. Defendants' representations and omissions were material because they were likely to deceive reasonable consumers about the terms of use of the Products, as well as the available mechanisms for seeking to exert control over the Illinois Plaintiff's and Illinois Subclasses Members' data.

466. Defendants intended to mislead the Illinois Plaintiff and Illinois Subclasses Members and induce them to rely on its misrepresentations and omissions.

467. The above unfair and deceptive practices and acts by Defendants were immoral, unethical, oppressive, and unscrupulous. These acts caused substantial injury that these consumers could not reasonably avoid; this substantial injury outweighed any benefits to consumers or to competition.

468. Defendants acted intentionally, knowingly, and maliciously to violate Illinois's Consumer Fraud Act, and recklessly disregarded the Illinois Plaintiff's and Illinois Subclasses Members' rights.

469. As a direct and proximate result of Defendants' unfair, unlawful, and deceptive acts and practices, the Illinois Plaintiff and the Illinois Subclasses Members have suffered and will continue to suffer injury, ascertainable losses of money or property, and monetary and non-monetary damages, as described herein.

470. The Illinois Plaintiff and Illinois Subclasses Members seek all monetary and non-

monetary relief allowed by law, including damages, restitution, punitive damages, injunctive relief, and reasonable attorneys' fees and costs.

COUNT EIGHT

NEGLIGENCE

(on behalf of All Plaintiffs against Defendants)

547. Plaintiffs herein repeat, reallege, and fully incorporate all allegations in all preceding paragraphs.

548. Defendants owed a duty to Plaintiffs and the Classes to exercise due care in: (a) obtaining data to train their Products; (b) not using individual's private information to train Defendants' AI; (c) ensuring that individuals' private data is not shared with or disclosed to unauthorized parties (including Defendant Microsoft); (d) destroying personal information to which Defendants had no legal right to possess.

549. Defendants' duties to use reasonable care arose from several sources, including those described below. Defendants had a common law duty to prevent foreseeable harm to others, including Plaintiffs and members of the Classes, who were the foreseeable and probable victims of Defendants' unlawful practices. Defendants acknowledge the Products are inherently unpredictable and may even evolve to act against human interests. Nevertheless, Defendants collected and continue to collect Private Information of millions of individuals and permanently feed the data to the Products, to train the Products for Defendants' commercial benefit. Defendants knowingly put Plaintiffs and the Classes in a zone of risk that is incalculable – but unacceptable by any measure of responsible data protection and use.

550. Defendants' conduct as described above constituted an unlawful breach of their duty to exercise due care in collecting, storing, and safeguarding Plaintiffs' and the Classes Members' Private Information by failing to protect this information.

551. Plaintiffs and Classes Members trusted Defendants to act reasonably, as a reasonably prudent manufacturer of AI products, and also trusted Defendants not to use individuals' Private Information to train their AI products. Defendants failed to do so, and breached their duty.

552. Defendants' negligence was, at least, a substantial factor in causing the Plaintiffs and

1 the Classes' Private Information to be improperly accessed, disclosed, used for development and
2 training of a dangerous product, and in causing the Class members' injuries.

3 553. The damages suffered by Plaintiffs and the Classes' members was the direct and
4 reasonably foreseeable result of Defendants' negligent breach of their duties to adequately design,
5 implement, and maintain reasonable practices to (a) avoid web scraping without consent of the
6 users; (b) avoid using Personal Information to train their AI products; and (c) avoid collecting and
7 sharing Users' data with each other.

8 554. Defendants' negligence directly caused significant harm to Plaintiffs and the Classes.

9 **COUNT NINE**

10 **INVASION OF PRIVACY UNDER CALIFORNIA CONSTITUTION**

11 **(on behalf of All Plaintiffs against Defendants)**

12 555. Plaintiffs herein repeat, reallege, and fully incorporate all allegations in all preceding
13 paragraphs.

14 556. Plaintiffs and Nationwide Classes Members had a legally protected privacy interest
15 and reasonable and legitimate expectation of privacy in the Private Information that Defendants
16 acquired illegally, tracked, collected, or otherwise used to train their Products.

17 557. Defendants owed a duty to Plaintiffs and Nationwide Classes Members to (a) not
18 collect via illegal web-scraping the individuals' information; (b) not to train their AI Products on
19 individuals Private Information; and (c) keep the data collected confidential, and not share with
20 Defendant Microsoft.

21 558. Defendants violated Plaintiffs, and Nationwide Classes Members' constitutional right
22 to privacy by tracking, collecting, storing, and misusing their Private Information, in which they
23 had a legally protective privacy interest, and for which they had a reasonable expectation of privacy
24 in a manner that was highly offensive to Plaintiffs and the Nationwide Classes Members. Such
25 violation, and disregard for Plaintiffs' and Nationwide Classes Members' rights was an egregious
26 violation of social norms.

27 559. Defendants knew, or acted with reckless disregard of the fact that a reasonable person
28 in Plaintiffs' and Nationwide Classes Members' position would consider their actions highly

1 offensive.

2 560. As a proximate result of such unauthorized disclosures, Plaintiffs' and Nationwide
3 Classes Members' reasonable expectations of privacy in their Private Information was unduly
4 frustrated and thwarted and caused damages to Plaintiffs and Nationwide Classes Members.

5 561. Plaintiffs seek injunctive relief on behalf of the Nationwide Classes, restitution, as
6 well as any and all other relief that may be available at law or equity. Unless and until enjoined, and
7 restrained by order of this Court, Defendants' wrongful conduct will continue to cause irreparable
8 injury to Plaintiffs and Nationwide Classes Members. Plaintiffs and Nationwide Classes Members
9 have no adequate remedy at law for the injuries in that a judgment for monetary damages will not
10 end the invasion of privacy for Plaintiffs and the Classes.

11 **COUNT TEN**

12 **INTRUSION UPON SECLUSION**

13 **(on behalf of All Plaintiffs against Defendants)**

14 562. Plaintiffs herein repeat, reallege, and fully incorporate all allegations in all preceding
15 paragraphs.

16 563. California adheres to Restatement (Second) of Torts, § 652B with no material
17 variation.

18 564. "One who intentionally intrudes, physically or otherwise, upon the solitude or
19 seclusion of another or his private affairs or concerns, is subject to liability to the other for invasion
20 of his privacy, if the intrusion would be highly offensive to a reasonable person." Restatement
21 (Second) of Torts, § 652B.

22 565. As our digital footprints continue to expand, individuals including Plaintiffs and
23 Nationwide Classes Members, have an increased expectation of privacy in their right to control who
24 has access to their information and how it is used.

25 566. In this context, personal data refers to not only information that can be used to identify
26 an individual but also nuanced data like browsing history, online chats, social media activity,
27 individuals' blogs and posts, keystrokes/online movements and searches, speech/writing patterns,
28 and other habits. The increasing reliance on digital services for everyday activities generates vast

1 amounts of such data, which Defendants collected, stored, and monetized without informed consent.

2 567. The reasonableness of such expectations of privacy is supported by Defendants'
3 unique position to be able to collect, store and track Plaintiffs' and Nationwide Classes Members'
4 data not only from information inserted into the chatbot, but also through a massive scraping of the
5 web. Defendants store, track, collect and disclose every user's information—every click, entry,
6 question, and use. This level of data tracking results in the unauthorized intrusion into sensitive
7 personally identifying data which ChatGPT pulls from users' device or browser, browser history,
8 IP addresses and location, geolocation of the users, searches through ChatGPT, searches through
9 ChatGPT plug-ins, uses of Bing searches – and other data.

10 568. Defendants intentionally intruded on and into Plaintiffs' and Nationwide Classes
11 Members' solitude, seclusion, or private affairs by constructing a system which collects, stores and
12 discloses (1) data inserted by users into the ChatGPT platform; (2) every click, every move, every
13 click, every search, every post, and every comment; (3) information scraped from the massive
14 corpus of the internet; (4) identifying data like IP addresses and location, including geolocation of
15 the users and more.

16 569. These intrusions are highly offensive to a reasonable person. This is evidenced by,
17 *inter alia*, countless consumer surveys, studies, and op-eds decrying tracking of people and children,
18 centuries of common law, state and federal statutes and regulations, legislative commentaries,
19 enforcement actions undertaken by the FTC, industry standards and guidelines, and scholarly
20 literature on consumers' reasonable expectations. Further, the extent of the intrusion cannot be fully
21 known, as the nature of privacy invasion involves sharing Plaintiffs' and Classes Members' personal
22 information with potentially countless third-parties, known and unknown, for undisclosed and
23 potentially unknowable purposes, in perpetuity.

24 570. Plaintiffs and Nationwide Classes Members were harmed by the intrusion into their
25 private affairs as detailed throughout this Complaint.

26 571. Defendants' actions and conduct complained of herein were a substantial factor in
27 causing the harm suffered by Plaintiffs and Nationwide Classes Members.

28 572. As a result of Defendants' actions, Plaintiffs and Classes Members seek injunctive

1 relief, in the form of Defendants' cessation of tracking practices in violation of state law, and
2 destruction of all personal data obtained in violation of state law.

3 573. As a result of Defendants' actions, Plaintiffs and Nationwide Classes Members seek
4 nominal and punitive damages in an amount to be determined at trial. Plaintiffs and Nationwide
5 Classes Members seek punitive damages because Defendants' actions—which were malicious,
6 oppressive, willful—were calculated to injure Plaintiffs and made in conscious disregard of
7 Plaintiffs' rights. Punitive damages are warranted to deter Defendants from engaging in future
8 misconduct.

9 574. Plaintiffs seek restitution for the unjust enrichment obtained by Defendants as a result
10 of the commercialization of Plaintiffs' and Nationwide Classes Members' sensitive data.

11 **COUNT ELEVEN**

12 **LARCENY/RECEIPT OF STOLEN PROPERTY**

13 **Cal. Penal Code § 496(a) and (c)**

14 **(on behalf of All Plaintiffs against Defendants)**

15 575. Plaintiffs herein repeat, reallege, and fully incorporate all allegations in all preceding
16 paragraphs.

17 576. Courts recognize that internet users have a property interest in their personal
18 information and data. *See Calhoun v. Google, LLC*, 526 F. Supp. 3d 605, 635 (N.D. Cal. 2021)
19 (recognizing property interest in personal information and rejecting Google's argument that "the
20 personal information that Google allegedly stole is not property"); *In re Experian Data Breach*
21 *Litigation*, SACV 15-1592 AG (DFMx), 2016 U.S. Dist. LEXIS 184500, at *14 (C.D. Cal. Dec. 29,
22 2016) (loss of value of PII is a viable damages theory); *In re Marriott Int'l Inc. Customer Data Sec.*
23 *Breach Litig.*, 440 F. Supp. 3d 447, 460-61 (D. Md. 2020) ("The growing trend across courts that
24 have considered this issue is to recognize the lost property value of this [personal] information.");
25 *Simona Opris v. Sincera*, No. 21-3072, 2022 U.S. Dist. LEXIS 94192, at *20 (E.D. Pa. May 23,
26 2022) (collecting cases).

27 577. Defendants owned and operated their AI Products and GPT Platforms (ChatGPT,
28 ChatGPT Plug-Ins, ChatGPT API). Defendants illegally obtained vast amounts of private

information to train their AI Products.

**A. Defendants’ Taking of Individual’s Private Information to Train Their AI
Violated Plaintiffs’ Property Interests**

578. Penal Code § 496(a) creates an action against “any” person who (1) receives “any” property that has been stolen or obtained in any manner constituting theft, knowing the property to be stolen or obtained, or (2) conceals, sells, withholds, or aids in concealing or withholding “any” property from the owner, knowing the property to be so stolen or illegally obtained.

579. Under Penal Code § 1.07(a)(38), “person” means “an individual, corporation, or association.” Thus, Defendants are persons under section 496(a).

580. As discussed above, Defendants stole the contents of the internet – everything individuals posted, information about the individuals, personal data, medical information, and other information – all used to create their Products to generate massive profits. At no point did Defendants have individuals consent to take/scrape this information in order to train their AI Products. Defendants meet the grounds for liability under Cal. Penal Code 496(a) because each of them:

- a. Knew that the taken information was stolen or obtained by theft, and with such knowledge;
- b. Concealed, withheld, or aided in concealing or withholding said data from their rightful owners by unlawfully using the data to train their Products;
- c. Defendants moved the data from the internet in order to feed it into their Products for training.

581. Pursuant to California Penal Code 496(c), Plaintiffs, on behalf of themselves and the Nationwide Class, seek actual damages, treble damages, costs of suit, and reasonable attorneys’ fees.

B. Tracking, Collecting, and Sharing Private Information Without Consent

582. As described above, in violation of Cal. Penal Code § 496(a) and (c), Defendants unlawfully collected, used, and exercised dominion and control of Private Information belonging to Plaintiffs and Classes Members.

1 583. Defendants wrongfully took Plaintiffs', ChatGpt User Class', ChatGPT API User
2 Class', and Microsoft User Class' (collectively "User Classes") Private Information to be used to
3 feed into Defendants' AI Products, to train and develop a dangerous technology.

4 584. Plaintiffs and the User Classes Members did not consent to such taking and misuse of
5 their personal data, and Private Information.

6 585. Defendants did not have consent from any state or local government agency allowing
7 them to engage in such taking and misuse of Private Information.

8 586. Defendants' taking of Private Information was intended to deprive the owners of such
9 information from ability to use their Private Information in the way they chose.

10 587. Defendants did so to maximize their profits and become rich at the expense of
11 Plaintiffs and the Classes.

12 588. Defendants collected data allows Defendants and their AI to learn the unique patterns
13 of each individuals, their online activities, habits, and speech/writing patterns.

14 589. Defendants moved Private Information to store and collect it on Defendant
15 Microsoft's servers, and thereafter, feed it to their AI products.

16 590. As a result of Defendants' actions, Plaintiffs and User Classes Members seek
17 injunctive relief, in the form of Defendants' cessation of tracking practices in violation of state law,
18 and destruction of all personal data obtained in violation of state law.

19 591. As a result of Defendants' actions, Plaintiffs, Nationwide Classes, and User Classes
20 seek nominal, actual, treble, and punitive damages in an amount to be determined at trial. Plaintiffs,
21 Nationwide Classes, and User Classes seek treble and punitive damages because Defendants'
22 actions—which were malicious, oppressive, willful—were calculated to injure Plaintiffs and made
23 in conscious disregard of Plaintiffs' rights. Punitive damages are warranted to deter Defendants
24 from engaging in future misconduct.

25 592. Plaintiffs seek restitution for the unjust enrichment obtained by Defendants as a result
26 of the commercialization of Plaintiffs', Nationwide Classes', and User Classes' sensitive data.

COUNT TWELVE**CONVERSION****(on behalf of All Plaintiffs against Defendants)**

471. Plaintiffs herein repeat, reallege, and fully incorporate all allegations in all preceding paragraphs.

593. The Nationwide Classes repeat and incorporate by reference all preceding paragraphs as if fully set forth herein.

594. Property is the right of any person to possess, use, enjoy, or dispose of a thing, including intangible things such as data or communications. Plaintiffs' and Nationwide Classes Members' personal information is their property. *Calhoun v. Google LLC*, 526 F. Supp. 3d 605, 636 (N.D. Cal. 2021).

595. As described in the cause of action for Larceny / Receipt of Stolen Property, Cal. Penal Code § 496(a) and (c), Defendants unlawfully collected, used, and exercised dominion and control over the Nationwide Classes Members' personal and private information without authorization.

596. Defendants wrongfully exercised control over Plaintiffs' and Nationwide Classes' information and have not returned it.

597. Plaintiffs and Nationwide Classes Members have been damaged as a result of Defendants' unlawful conversion of their property.

COUNT THIRTEEN**UNJUST ENRICHMENT****(on behalf of All Plaintiffs against Defendants)**

598. Plaintiffs incorporate, re-allege, and include the foregoing allegations as if fully set forth herein.

599. By virtue of the unlawful, unfair and deceptive conduct alleged herein, Defendants knowingly realized hundreds of millions of dollars in revenue from the use of the Personal Information of Plaintiffs and Nationwide Classes Members for the commercial training of its ChatGPT and other AI language models.

1 600. This Private and Personal Information, the value of the Private and Personal
2 Information, and/or the attendant revenue, were monetary benefits conferred upon Defendants by
3 Plaintiffs and the members of the Nationwide Classes.

4 601. As a result of Defendants' conduct, Plaintiffs and Nationwide Classes Members
5 suffered actual damages in the loss of value of their Private Information and the lost profits from
6 the use of their Private Information.

7 602. It would be inequitable and unjust to permit Defendants to retain the enormous
8 economic benefits (financial and otherwise) it has obtained from and/or at the expense of Plaintiffs
9 and Classes Members.

10 603. Defendants will be unjustly enriched if they are permitted to retain the economic
11 benefits conferred upon them by Plaintiffs and Nationwide Classes Members through Defendants'
12 obtaining the Private Information and the value thereof, and profiting from the unlawful,
13 unauthorized, and impermissible use of the Private Information of Plaintiffs and Nationwide Classes
14 members.

15 604. Plaintiffs and Nationwide Classes members are therefore entitled to recover the
16 amounts realized by Defendants at the expense of Plaintiffs and Nationwide Classes Members.

17 605. Plaintiffs and the Nationwide Classes have no adequate remedy at law.

18 606. Plaintiffs and the members of the Nationwide Classes are entitled to restitution,
19 disgorgement, and/or the imposition of a constructive trust to recover the amount of Defendants'
20 ill-gotten gains, and/or other sums as may be just and equitable.

21 **COUNT FOURTEEN**

22 **FAILURE TO WARN**

23 **(on behalf of All Plaintiffs against Defendants)**

24 607. Plaintiffs incorporate, re-allege, and include the foregoing allegations as if fully set
25 forth herein.

26 608. "[M]anufacturers have a duty to warn consumers about the hazards inherent in their
27 products. . . . The purpose of requiring adequate warnings is to inform consumers about a product's
28 hazards and faults of which they are unaware, so that the consumer may then either refrain from

1 using the product altogether or avoid the danger by careful use.” *Taylor v. Elliott Turbomachinery*
2 *Co., Inc.*, 171 Cal. App. 4th 564, 577 (2009). “A duty to warn or disclose danger arises when an
3 article is or should be known to be dangerous for its intended use, either inherently or because of
4 defects.” *DeLeon v. Commercial Manufacturing and Supply Co.*, 148 Cal. App. 3d 336, 343 (1983).

5 609. Defendants manufactured AI technology and disseminated it to the public, placing
6 this technology into the stream of commerce in a defective and unreasonably dangerous manner
7 such that the foreseeable risks exceeded the benefits associated with the design and/or formulation
8 of the technology.

9 610. Defendant’s AI Products were, and remain, defective due to inadequate warnings and
10 inadequate testing prior to public dissemination.

11 611. Defendants knew, and remain aware of the fact that their technology is novel in the
12 marketplace, and that consumers do not understand the full capabilities of this technology.
13 Defendants likewise know that consumers are not experts in understanding how AI technology
14 works in general.

15 612. Despite the foregoing, Defendants released their AI technology to the public without
16 adequately warning consumers of the danger.

17 613. Defendants disclosed, and continue to disclose private information belonging to
18 Plaintiffs and Nationwide Classes members. Specifically, Defendants monitor, collect, and track the
19 habits, preferences, thoughts, online activity, and geolocation data for their users, including young
20 children.

21 614. Plaintiffs, and all Nationwide Classes Members, could not consent to Defendants’
22 conduct when they were unaware their sensitive information would be collected and used in the first
23 place, and most certainly were unaware that this information would be used to train Defendants’ AI
24 Products.

25 615. As a proximate result of such unauthorized disclosures of private information,
26 Plaintiffs’ and Nationwide Classes Members’ reasonable expectations of privacy in their Private
27 Information was unduly frustrated and thwarted, resulting in damages to the Plaintiffs and
28 Nationwide Classes Members. Defendants scraped personal data and information without consent,

and continue to collect personal data and information, claiming a right to sell this data without notice by integrating this data into their AI Products.

616. Further, Defendants train their AI Products on the data consumers input into the system with no way to remove this data. Once the data has been entered the system, it will remain there indefinitely. Consumers are not adequately informed of this. Because the system is “learning” from consumers, it is being trained on data which is not adequately vetted to ensure accuracy. Nor is there any way to determine how the AI is absorbing and utilizing the inaccurate information. This, in turn, results in the spread of inaccurate misinformation to consumers and the general public.

617. The foregoing results in unreasonable invasions into consumers’ privacy and opens consumers up to having their identities or lives disrupted, as well as leads to the spread of misinformation to consumers.

618. The defective and/or absent warnings to consumers with regards to the foregoing is a substantial factor in bringing about injuries to Plaintiffs and Nationwide Classes Members.

619. Plaintiffs seek injunctive relief on behalf of the Nationwide Classes, restitution, as well as any and all other relief that may be available at law or equity. Unless and until enjoined, and restrained by order of this Court, Defendants’ wrongful conduct will continue to cause irreparable injury to Plaintiffs and Nationwide Classes Members. Plaintiffs and Nationwide Classes Members have no adequate remedy at law for the injuries in that a judgment for monetary damages will not end the invasion of privacy for Plaintiffs and the Classes.

COUNT FIFTEEN

NEW YORK GENERAL BUSINESS LAW

N.Y. GEN. BUS. LAW §§ 349, et seq.

(on Behalf of New York Plaintiff and the New York Subclasses against Defendants)

620. The New York Plaintiff, individually and on behalf of the New York Subclasses, repeats and re-alleges the allegations contained in the foregoing paragraphs as if fully set forth herein.

621. Defendants engaged in deceptive acts or practices in the conduct of its business, trade, and commerce or furnishing of services, in violation of N.Y. Gen. Bus. Law § 349, including:

- 1 a) Defendants have exploited Non-Users and Users of their Products, by stealing
2 such individuals' data at scale from web crawler caches without permission from
3 the data owners and without any way of segregating out any given Non-Users'
4 or User data from the datasets used to train OpenAI's LLMs upon request of
5 such individuals—including where such individuals are minors.
- 6 b) Defendants knew that they were collecting and/or profiting from individuals'
7 Personal Information and that the risk of collecting of such Personal
8 Information was highly likely. Defendants' actions in engaging in the above-
9 named deceptive acts and practices were negligent, knowing and willful,
10 and/or wanton and reckless with respect to the rights of the New York Plaintiff
11 and members of the New York Subclasses;
- 12 c) As described herein, Defendants are misrepresenting that they have and are
13 complying with common law and statutory duties pertaining to the security and
14 privacy of Plaintiff's and Subclass Members' data, including but not limited to
15 duties imposed by the FTC Act, 15 U.S.C. § 45 and N.Y. Gen. Bus. Law §§ 349,
16 *et seq.*
- 17 d) As described herein, Defendants have and are omitting, suppressing, and
18 concealing the material fact that they are stealing and profiting from the mass
19 collection and analysis of New York Plaintiff's and New York Subclasses
20 Members' data at scale and without adequate or effective consent; and
- 21 e) Omitting, suppressing, and concealing the material fact that they did not comply
22 with common law and statutory duties pertaining to the security and privacy of
23 New York Plaintiff's and Subclasses Members' data, including but not limited
24 to the fact that they are functionally unable to delete such data once it has been
25 incorporated into their LLMs as training data.

26 622. Defendants' representations and omissions were material because they were likely to
27 deceive reasonable consumers about the terms of use of their products, as well as the available
28 mechanisms for seeking to exert control over New York Plaintiff's and New York Subclasses

Members' data.

623. Defendants acted intentionally, knowingly, and maliciously to violate New York's General Business Law, and recklessly disregarded the New York Plaintiff's and New York Subclasses Members' rights.

624. As a direct and proximate result of Defendants' deceptive and unlawful acts and practices, the New York Plaintiff and New York Subclasses Members have suffered and will continue to suffer injury, ascertainable losses of money or property, and monetary and non-monetary damages, as described herein.

625. Defendants' deceptive and unlawful acts and practices complained of herein affected the public interest and consumers at large, including millions of New Yorker User Class Members and Non-User Subclass Members.

626. The above deceptive and unlawful practices and acts by Defendants caused substantial injury to the New York Plaintiff and New York Subclasses Members that they could not reasonably avoid.

627. The New York Plaintiff and New York Subclasses Members seek all monetary and non-monetary relief allowed by law, including actual damages or statutory damages of \$50 (whichever is greater), treble damages, injunctive relief, and attorney's fees and costs.

PRAYER FOR RELIEF

WHEREFORE, Plaintiffs on behalf of themselves and the Proposed Class respectfully requests the following relief:

- A. Injunctive relief in the form of a temporary freeze on commercial access to and commercial development of the Products until such time as Defendants can demonstrate completion of some or all of the following to the Court's satisfaction:
 1. Establishment of an independent body of thought leaders (the "AI Council") who shall be responsible for approving uses of the Products before, not after, the Products are deployed for said uses;
 2. Implementation of Accountability Protocols that hold Defendants responsible for Product actions and outputs and barred from further commercial

1 deployment absent the Products' ability to follow a code of human-like ethical
 2 principles and guidelines and respect for human values and rights, and until
 3 Plaintiffs and Class Members are fairly compensated for the stolen data on
 4 which the Products depend;

- 5 3. Implementation of effective cybersecurity safeguards of the Products as
 6 determined by the AI Council, including adequate protocols and practices to
 7 protect Users' PHI/PII collected through Users' inputting such information
 8 within the Products as well as through Defendants' massive web scraping,
 9 consistent with the industry standards, applicable regulations, and federal,
 10 state, and/or local laws;
- 11 4. Implementation of Appropriate Transparency Protocols requiring Defendants
 12 to clearly and precisely disclose the data they are collecting, including where
 13 and from whom, in clear and conspicuous policy documents that are explicit
 14 about how this information is to be stored, handled, protected, and used;
- 15 5. Requiring Defendants to allow Product users and everyday internet users to
 16 opt out of all data collection and stop the illegal taking of internet data, delete
 17 (or compensate for) any ill-gotten data, or the algorithms which were built on
 18 the stolen data;
- 19 6. Requiring Defendants to add technological safety measures to the Products
 20 that will prevent the technology from surpassing human intelligence and
 21 harming others;
- 22 7. Requiring Defendants to implement, maintain, regularly review and revise as
 23 necessary, a threat management program designed to appropriately monitor
 24 Defendants' information networks for threats, both internal and external, and
 25 assess whether monitoring tools are appropriately configured, tested, and
 26 updated;
- 27 8. Establishment of a monetary fund (the "AI Monetary Fund" or "AIMF") to
 28 compensate class members for Defendants' past and ongoing misconduct to

be funded by a percentage of gross revenues from the Products;

9. Appointment of a third-party administrator (the “AIMF Administrator”) to administer the AIMF to members of the class as “data dividends” as fair and just compensation for the stolen data on which the Products depend;

10. Confirmation that Defendants have deleted, destroyed, and purged the PII/PHI of all relevant class members unless Defendants can provide reasonable justification for the retention and use of such information when weighed against the privacy interests of class members; and

11. Requiring all further and just corrective action, consistent with permissible law and pursuant to only those causes of action so permitted.

B. Actual damages for economic and non-economic harm in an amount to be determined at trial;

C. Statutory damages in an amount to be determined at trial;

D. Equitable relief in the form of monetary damages, restitution, and disgorgement;

E. Pre-judgment interest;

F. Post-judgment interest;

G. Reasonable attorneys’ fees and costs of suit incurred by their attorneys, in recognition of the spirit of the consumer protection statutes at issue, which encourage holding businesses to account for unfair business practices;

H. Treble damages allowable under applicable laws;

I. Punitive damages allowable under applicable laws;

J. Exemplary damages allowable under applicable laws;

K. Any and all other such relief as the Court may deem just and proper.

JURY TRIAL DEMANDED

Plaintiffs demand a jury trial on all triable issues.

DATED: June 28, 2023

CLARKSON LAW FIRM, P.C.

/s/ Ryan J. Clarkson

Ryan Clarkson, Esq.
Yana Hart, Esq.
Tracey Cowan, Esq.
Tim Giordano, Esq.
Tiara Avanes, Esq.
Valter Malkhasyan, Esq.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28