Colin R. Kass (*pro hac vice*)
PROSKAUER ROSE LLP
1001 Pennsylvania Ave., N.W.
Washington, D.C. 20004
(202) 416-6890
ckass@proskauer.com

David A. Munkittrick (*pro hac vice*)
PROSKAUER ROSE LLP
Eleven Times Square
New York, New York 10036
(212) 969-3000
dmunkittrick@proskauer.com

Attorneys for Defendant Bright Data Ltd.
Additional counsel listed on Signature Page

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA

X CORP., Plaintiff, v. BRIGHT DATA LTD., Defendant.	**PUBLIC VERSION** Case No. 23-cv-03698-WHA Hon. William Alsup

BRIGHT DATA'S COUNTERCLAIM AND ANSWER

COUNTERCLAIM[*]

I. NATURE OF THE ACTION

1. Elon Musk did not pay $44 billion for a bunch of eyeballs. He paid this ungodly sum – more than most people will see in their lifetime – because he had a plan. To transform Twitter, an advertising company, into X, an information monopolist.

2. Upon taking control of the company, X engaged in a multifaceted scheme to extract value from data it does not own. The scheme worked. Between X and its sister company, Musk's $44 billion investment, which seemed so foolish at the time, has now ballooned to over $60 billion in just two short years. It would not have been possible absent X's exclusionary and anticompetitive practices.

3. People think of X as a digital public square, or, in X's own words, a "global platform for public self-expression and conversation in real time."[1] As X would have it, it is a common good for the Common Good. But you can't make money by charging people for the right to speak. The competitive price for tweeting is, in fact, zero.

4. Twitter's founders, nonetheless, focused on building the public square. Build it, and they will come. The money will follow, or so they thought. They were wrong. They focused on selling ads: that, after all, is how newspapers in the days of old made their money. But the advertising market is a competitive one. It is difficult to make a dime. And Twitter did not. It lost money almost every single year of its existence.

5. Elon Musk had a different view. "Twitter was over-reliant on advertising revenue," Musk said, and its "ad-based revenue model is dated."[2] He believed "he could unlock Twitter's true potential by shifting away from an advertising-only model … to other forms of revenue…."[3]

[*] Unless otherwise noted, all emphasis added, internal citations and quotation marks omitted, and capitalizations conformed without brackets.

[1] Twitter Annual Report, at 6 (Feb. 16, 2022).

[2] *Twitter, Inc. v. Musk*, 2022-cv-0613 (Del. Ch. Aug. 4, 2022), Verified Counterclaim, ¶¶ 8, 39.

[3] *Id.* ¶ 39.

We live in the dawn of the information age. And the world is changing, on the cusp of transitioning from the early days of link-based search, like Google, to next generation chatbot-based search, driven by generative AI. As one of the founders of OpenAI, a non-profit organization that developed the technology that powered these models, Musk realized that X was uniquely suited to capitalize on this new world order.

6. Just as internal combustion engines need gas, AI-driven search engines need real-time information. If you are going to be asked, "what happened yesterday," you better know the answer. X has the answers. As Musk said, "what's useful [about X's data] is the fact that it's up to the second."[4] It operates legions of server farms in the background that track every user's movement and collects all the information they generate on the platform. And as the dominant public square platform – with a 95% market share – X has unparalleled access to real-time data about current events and public sentiment, allowing it to capture the cultural zeitgeist of the world like no other. 500 million tweets a day, 180 *billion* tweets a year. That's a lot of data. And it is exactly what artificial intelligence powered search engines need. As Musk tweeted, "[b]ecause it consists of billions of bidirectional interactions per day, Twitter can be thought of as a collective, cybernetic super-intelligence."[5]

7. So, in July 2023, Musk created a sister company, xAI, that would use the data stored in X's servers to power its chatbot. Grok was born.

8. But Musk had a problem. X did not own the data, and most of it – the valuable parts, anyway – were in the public domain. X could not prevent others from making free use of it. And if they could, Musk knew his plan would come crashing down like a house of cards. If others had access to the data, artificial intelligence models would proliferate outside his control. He needed to find a way to yank the information out of the public domain.

[4] *Elon Musk on xAI: We will win*, perma.cc/N8YC-QCP5.

[5] @elonmusk, x.com, perma.cc/C2U7-FSZ4, ***publicly published on X.***

9. So, the same month he founded xAI, Musk tried to lock users' data behind a log-in screen. The public square he had so coveted, and paid $44 billion to buy, would exist no more, instead becoming a walled garden for members only. The law allowed him to do that. But competition did not. Days after X changed its log-in policy, Meta entered the market with Threads, a direct competitive threat to X, the likes of which X had never seen. Musk was forced to retreat. The walls he just erected came tumbling down, and, today, most of the information remains in the public domain, just as it always had been.

10. Musk was undeterred. He did not become the world's richest person by easily accepting defeat. If he could not control the data through lawful means, he would do so through unlawful ones. He turned to contracts, combinations, and conspiracies in restraint of trade. The contracts were X's Terms—unilaterally dictated contracts of adhesion that would, in X's view, bind virtually every person on the planet. These "Terms of Use" would change; they would become "Terms of Abuse." And abuse them, he did.

11. Musk changed X's Terms to give X unfettered rights to feed the data to Grok, with no ability for users to opt out. This gave Grok what it needed, but not what Musk needed. He needed to stop the development of competing A.I. models, particularly OpenAI's ChatGPT. Though OpenAI had been paying X $2 million a year for access to X's data, Musk could not tolerate continued access – he needed exclusive use of this critical data, not non-exclusive use.[6] So X terminated its contract with OpenAI, declaring that it was "[l]awsuit time."[7]

12. But to sue, X needed to shore up its Terms. Prior to Musk, X's Terms were clear: "crawling the services is permissible."[8] As Twitter's former head of Trust and Safety explained: "Scraping was the open secret of Twitter data access. We knew about it. It was fine."[9] Musk

[6] *Twitter Accuses Microsoft of Improperly Using Its Data*, NY Times (May 18, 2023) perma.cc/NK28-YL5L.

[7] @elonmusk, x.com, perma.cc/6W64-FDLB, *also publicly published by X.*

[8] X Terms (v.18 May 18, 2023).

[9] @yoyoel, bsky.app, perma.cc/LKT8-34TR.

would change that. Within one month of acquiring Twitter, Musk tweeted, "Twitter rules will evolve over time,"[10] demonstrating his personal involvement in, and commitment to, modifying the Terms to eliminate competition.

13. In September 2023, X amended its Terms to state that "crawling … the Services in any form, for any purpose … is expressly prohibited."[11] But no law on earth gives X the right to stop the public from searching the public web; and its attempt to secure that right through contract is the definition of a contract in restraint of trade.

14. X's Terms are even more nefarious. They purport to prevent millions of people and businesses not just from accessing or scraping X's platform, but from doing business with X's competitors: companies that provide access to and sell public data on X. According to X, the "Terms prohibit selling any content collected from the platform," prevent any users from "using" the content without express permission, and prevent anyone from "facilitating" any of the foregoing.[12] These provisions, X says, prevent anyone bound by the Terms from purchasing the data from anyone other than X. It creates *de facto* exclusive contracts covering nearly the entire market. That is the definition of exclusionary conduct.

15. But X did not stop there. It knows that, as contracts of adhesion wielded to acquire dominion over information it does not own, the Terms are unenforceable. Indeed, this Court has already voided the Terms' scraping provisions as inconsistent with rights Congress has granted the public. As the Court explained,

> "The upshot is that [X's Terms] would entrench its own private copyright system that rivals, even conflicts with, the actual copyright system enacted by Congress. ***X Corp. would yank into its private domain and hold for sale information open to all, exercising a copyright owner's right to exclude where it has no such right***.
>
> We are not concerned here with an arm's length contract between two sophisticated parties… ***We are instead concerned with a massive regime of adhesive terms***

[10] @elonmusk, x.com, perma.cc/TB26-BEFP, *and this too remains publicly published on X.*

[11] X Terms (v.19 Sept. 29, 2023); *see also* X Terms (v.20 Nov. 15, 2024).

[12] X Terms (v.20 Nov. 15, 2024); SAC ¶ 35.

> ***imposed by X Corp. that stands to fundamentally alter the rights and privileges of the world at large.***"[13]

16. This ruling was devastating to X. Though seemingly limited to a discrete Copyright issue, it nullified X's ability to use contract law to enforce the Terms' anti-scraping provisions. That left only its prohibitions on automated access, or crawling. But those Terms would also be wholly ineffective against data scrapers since X suffers no damage from automated access. Access damages are based on server burdens, not the subsequent use of the data. Because X incurs no real burden in responding to scrapers' server requests – indeed, that is what the Internet was designed for – X was left with no realistic way to enforce its anti-access provisions.

17. With this Court's ruling turning X's world upside down, X had to do something. But X did not do what any law-abiding citizen would do: remove the offending and unenforceable provisions from the contract. Instead, it came up with a plan to scare scrapers into backing down.

18. The *in terrorem* threat it came up with was as simple as it was deceptive. The key was what X euphemistically calls a "liquidated damages" provision.[14] But it is not a "damages" provision at all. It is an illegal penalty. Under this provision, X says anyone who dares to use automated means to access X's platform will have to pay the equivalent of 1.5 cents per tweet viewed or accessed. Its statement that this is a "reasonable estimate" of its actual damages is an intentional lie.[15] At more than ***1400%*** above the price X charges its data subscribers, this fee bears no relation to X's actual costs or damages. It is the definition of a material misrepresentation of fact made with *scienter* and with the expectation that it be relied upon. X knows that as an unenforceable contractual penalty, this provision would not hold up in court. But that did not matter to X, because X does not care about collecting on a judgment; it only wants to wield the threat of financial ruin to eliminate competition in the data market. The use of such provisions for this purpose makes it not just a fraud, but an antitrust violation.

[13] ECF 83, at 20.

[14] X Terms (v.20 Nov. 15, 2024).

[15] *Id.*

19. X's exclusionary conduct did not stop there. It took legal action against Bright Data, claiming that Bright Data's scraping of public websites, such as X, is "illicit," "unlawful," "deceptive," and a panoply of other ills. But X knows this to be untrue. In May of this year, X, through its sister company, xAI, needed to obtain additional third-party data to help train its A.I. models. It sought out Bright Data, requesting use of Bright Data's proxy service and scraping tools to scrape third-party websites, including sites that have Terms of Use similar to X's. X turned to Bright Data precisely because it provides a valuable and lawful public service. So when X runs to Court, claiming that these same services are illicit when used by X's competitors, it does so with unclean hands.

20. Nor was X's attack on data scrapers limited to Term abuse. It used exclusionary and deceptive acts to prevent data scrapers from obtaining alternative sources of information. As the dominant Public Square Platform, with 95% market share, it is currently the only viable source of real-time data on such platforms. But its dominance would be threatened if users defected to other platforms. Data scrapers, who sell data in competition with X, would then have alternative sources of supply. So, X engaged in deceptive practices to throttle web traffic to these other platforms, making it appear that users' connections were just slow or that the other platforms' servers were dysfunctional. In reality, X secretly interfered with users' right to traverse the World Wide Web, trapping them within X's platform, and causing them to abandon their efforts to establish a presence on competing websites. The clear, intended, and foreseeable result was that these platforms and the data scrapers who help them synthesize and distribute the data, all in *direct* competition with X, are deprived of the data needed to do so.

21. X's monopolistic practices are not benign. Since X embarked on its monopolistic scheme, it has limited the public's access to public data, drastically raised the price of its data subscriptions, and hoarded the data for its most valuable use case for X's own, exclusive benefit. Its conduct poses a dangerous threat to the availability of online information, the world economy, consumers, data customers, competing platforms, and data competitors, including Bright Data. If X wants to compete in the data market, it is certainly free to do so. But it cannot obtain a $60

billion information monopoly by eliminating competing data sources. Its antitrust violations must come to an end.

II. THE PARTIES

22. ***Bright Data Ltd.*** ("Bright Data") is incorporated in Israel with its principal place of business at 4 Hamahshev St., Netanya, 4250714, Israel. Bright Data is a revolutionary internet company that has developed best-in-class technologies that provide its customers with the ability to search for publicly-available internet data. Bright Data's services are used by over 20,000 customers worldwide, including X Corp., X's commonly-controlled affiliates, Fortune 500 companies, academic institutions, and small businesses.

23. ***X Corp.*** ("X") is incorporated in Nevada. At times, X maintained a principal place of business at 1355 Market Street, Suite 900, San Francisco, California 94104. On or about September 13, 2024, X sought to escape the clutches of California law by purporting to move its headquarters to Bastrop, Texas. On or about November 15, 2024, X adopted new Terms of Service, which purport to adhere people bound by them to Texas law. X owns and operates the Public Square platform X, formerly known as Twitter.

III. JURISDICTION AND VENUE

24. Bright Data brings this action pursuant to Sections 4 and 16 of the Clayton Act, 15 U.S.C. §§ 14 and 26, for treble damages, injunctive relief, costs of suit, and reasonable attorneys' fees for violations of Sections 1 and 2 of the Sherman Act, 15 U.S.C. §§ 1, 2. The Court has jurisdiction over these claims pursuant to 28 U.S.C. §§ 1331 and 1337(a).

25. Bright Data also asserts state law claims under (i) California's Cartwright Act, Cal. Bus. & Prof. Code § 16720, et seq., (ii) California's Unfair Competition Act, Cal. Bus. & Prof. Code § 17200, et seq., (iii) Nevada's Unfair Business Practices Act, NRS 598A.060; (iv) Texas's Free Enterprise and Antitrust Act, Tex. Bus. Comm. Code 15.05(a); and (v) for tortious interference with customers relations. The Court has diversity jurisdiction over these claims pursuant to 28 U.S.C. § 1332 because there is complete diversity among the parties and the amount in controversy exceeds $75,000. The Court also has supplemental jurisdiction over the state law

claims pursuant to 28 U.S.C. § 1367(a), and Federal Rule of Civil Procedure 13(a), because these claims arise out of the same nucleus of operative facts and form the same controversy as Bright Data's federal claims.

26. The Court has personal jurisdiction over X because it instituted this litigation in this Court and, therefore, has consented to personal jurisdiction or waived any objection to it. The Court also has personal jurisdiction, pursuant to Section 12 of the Clayton Act, 15 U.S.C. § 22, based on X's nationwide contacts. The Court can also assert general personal jurisdiction over X because it has (or for much of the relevant period, had), a principal place of business in California. The Court also has specific personal jurisdiction because X engaged in unlawful conduct in, or expressly aimed such conduct at, California, and caused injury in California.

27. Venue is proper in this District pursuant to 28 U.S.C. § 1391(b), 15 U.S.C. § 22, and Federal Rule of Civil Procedure 13. X resides in, transacts business, and can be found in this District, and a substantial portion of X's unlawful conduct has been carried out in this District.

IV. FACTUAL ALLEGATIONS

A. X's Modest Beginnings to Global Powerhouse.

28. In the summer of 2006, a small San Francisco-based startup unleashed a surprisingly simple social media platform that would go on to transform how the world communicates. Twitter, the brainchild of Jack Dorsey, Biz Stone, Evan Williams, and Noah Glass, began as a side project at Odeo, a podcasting company searching for its next big idea. Initially called "twttr," after the SMS short-code culture of the time, the platform invited users to share their thoughts in bursts of 140 characters. From its modest beginnings, Twitter evolved into a cultural phenomenon and a breaking news-making powerhouse. In Twitter's words:

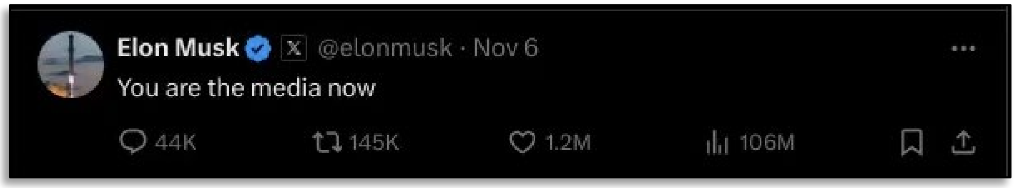
> "Twitter is what's happening in the world and what people are talking about right now. From breaking news and entertainment, to sports, politics, and everyday interests, Twitter shows every side of the story. On Twitter you can join the open conversation and watch highlights, clips, or live-streaming events."[16]

[16] Twitter Annual Report, at 7 (Feb. 20, 2019).

29. Twitter debuted on March 21, 2006, when Jack Dorsey sent the inaugural tweet: "just setting up my twttr."[17] The platform was conceived as a digital space for sharing short updates and thoughts in 140 characters or less. Early adopters were tech enthusiasts who marveled at its simplicity and immediacy. By July 2006, Twitter was launched to the public.

30. The platform's first major breakout moment came during the 2007 South by Southwest (SXSW) festival. Twitter strategically placed large screens in conference venues, streaming live tweets from attendees. The buzz turned heads, and Twitter's user base skyrocketed during the event. The ability to share thoughts in real-time had struck a chord.

31. By the late 2000s, Twitter had cemented itself as the go-to platform for breaking news and real-time updates. Eyewitnesses to major world developments used Twitter to report events as they unfolded, demonstrating its potential as a citizen journalism tool. Indeed, today, tweets are often not just reporting on the news, but are the news themselves. People care about what celebrities or politicians tweet. As X's eventual owner, Elon Musk, stated in his own tweet, "[y]ou are the media now."[18]



Elon Musk @elonmusk · Nov 6

You are the media now

32. Twitter also sparked the hashtag revolution. Introduced in 2007 by user Chris Messina, the hashtag (#) became Twitter's signature feature, revolutionizing the way people organized and discovered content. From #BlackLivesMatter to #MeToo, hashtags amplified social movements and connected global conversations.

33. As Twitter's influence grew, so did its user base. Citizens, celebrities, politicians, brands, and businesses flocked to the platform, recognizing its value as a direct communication channel. In 2013, Twitter went public, debuting on the New York Stock Exchange with a valuation

[17] @jack, twitter.com, perma.cc/KV7G-M638.

[18] @elonmusk, x.com, perma.cc/NG76-ANC9.

of $31 billion.[19] By 2022, the company had surpassed 368 million active users, with over 500-700 million tweets *per day*.[20]

B. Musk Takes Over and Transforms Twitter into a Data Monopolist.

34. In early 2022, Elon Musk, a frequent Twitterist and serial entrepreneur, thought that Twitter management did not understand the business they were in. Management believed Twitter was an advertising company, offering advertisers access to interested eyeballs. But Musk recognized that Twitter was actually in the data business and had ineffectively exploited its information monopoly. Just as Ray Kroc built an empire by recognizing that McDonald's was not in the burger business, but the real estate business, Musk recognized he could build an empire by taking control over the vast amounts of user content generated on the platform.[21]

35. And Musk had a particular end-use in mind. Though most people didn't realize it at the time, one of the most valuable uses for X data is to power artificial intelligence chatbots. As far back as 2015, Musk recognized that internet search engines would transition from being mere URL ranking and listing services to conversation-based oracles of answers. He invested tens of millions of dollars in OpenAI to develop "open source" artificial intelligence technologies.

36. He did not do this for the betterment of mankind. He did this so that he could enlist the help of others to develop the technology behind Large Language Model Artificial Intelligence, without bearing the full costs or risks of doing so. That the resulting technology would be open source did not matter to him because his intended use of the technology – to create a search engine chatbot – required data that he hoped to control. That is, once OpenAI developed the technology, the *engine* powering these artificial intelligence models would be freely available to the public.

[19] *Twitter Goes Public on the New York Stock Exchanges*, CBS News (Nov. 7, 2013), perma.cc/7FTT-5EWX

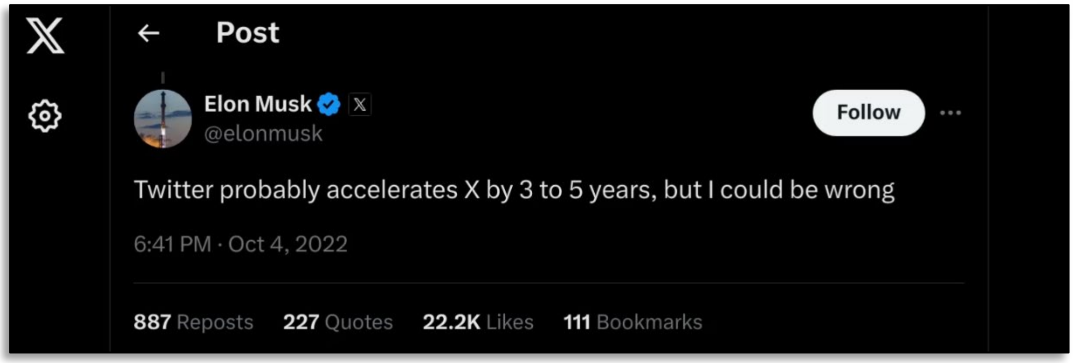
[20] *Twitter, Data, And Elon Musk*, Forbes (Oct. 5, 2022), perma.cc/4ZWQ-6MR2.

[21] *The Founder* (2020) ("You don't seem to realize what business you're in. You're not in the burger business. You're in the real estate business."); *Twitter v. Musk*, Verified Counterclaims, ¶ 39 ("Musk believes that Twitter's ad-based revenue model is dated" and that "he could unlock Twitter's true potential by shifting away from an advertising-only model … to other forms of revenue.").

But the *gas* would be controlled by those with access to the *data* needed to run them. That meant only a handful of companies in the world would be able to develop chatbot-based search engines, such as Microsoft, Google, and Twitter.

37. Musk realized that access to Twitter data was essential because it was the only large-scale source of real-time, conversational data about current events. A conversational search chatbot without access to this data would not work. Given his close connection to OpenAI (as founder, funder, and contributor), Musk knew that, despite OpenAI's vast resources and its access to the data feed from Microsoft's Bing, the world's second largest search engine, not even OpenAI could succeed without Twitter data. So OpenAI shelled out $2 million a year for it.

38. Musk similarly recognized that Twitter data was essential to his own plans to develop his "everything App," which would include A.I. search-engine chatbots. As Musk tweeted, "buying Twitter is an accelerant to creating X, the everything app," and "probably accelerates X by 3 to 5 years."[22]



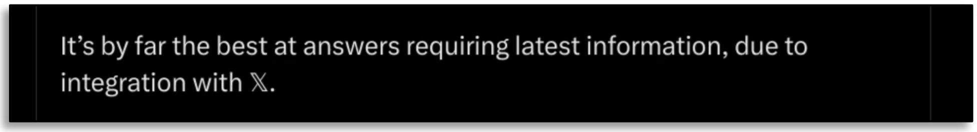
Twitter probably accelerates X by 3 to 5 years, but I could be wrong

6:41 PM · Oct 4, 2022

887 Reposts 227 Quotes 22.2K Likes 111 Bookmarks

[22] @elonmusk, x.com, perma.cc/GBX9-B6AA, perma.cc/XQ2N-W5WW.

39. But Musk was not content with simply paying Twitter for access, as OpenAI had. He needed the power to exclude A.I. competitors and the scrapers who might supply them from gaining access to that data. As Musk recently tweeted, the "integration with X" makes Grok uniquely capable of competing in the A.I. search engine market:[23]



It's by far the best at answers requiring latest information, due to integration with 𝕏.

40. As Musk once said, "It's ok to have all your eggs in one basket, as long as you control what happens to that basket."[24] And he was willing to pay for that control. He offered $44 billion to a money losing platform that only had one year of profit in its 15-year history.

41. The offer was quickly accepted. And the transformation from advertising company to information monopolist began.

42. Prior to Musk's takeover, X derived most of its revenues from advertisers. As its last annual report as a public company stated, it "generated the substantial majority of its revenues from the sale of advertising," with a small percentage derived from "data licensing and other arrangements."[25] Even this was an understatement. In fact, as X recently admitted, "over ninety percent of Twitter's revenue was derived from advertising."[26]

43. Once Musk took control of Twitter, he rebranded the company X, promising to grow other sources of revenue. In his investor presentation, he projected that revenue would quintuple, and committed to reduce reliance on advertising from 90% to just 45%, with new subscriptions and data licensing arrangements making up the remainder.[27]

44. In fact, market realities were far different. Revenue did not quintuple; the bottom fell out. Mr. Musk describes himself as a free speech absolutist, and had political ambitions for

[23] @elonmusk, x.com, perma.cc/4NQ8-3KK3.

[24] @ElonMuskNewsOrg., x.com, perma.cc/VKC4-S74W.

[25] Twitter Annual Report, at 54 (Feb. 16, 2022).

[26] *X Corp. v. World Federation of Advertisers*, 7:24-cv-00114, ECF 62 at 24 (N.D. Tex. 2024).

[27] *Inside Elon Musk's Big Plans for Twitter*, NY Times (Oct. 28, 2022), perma.cc/H3DG-NJYP.

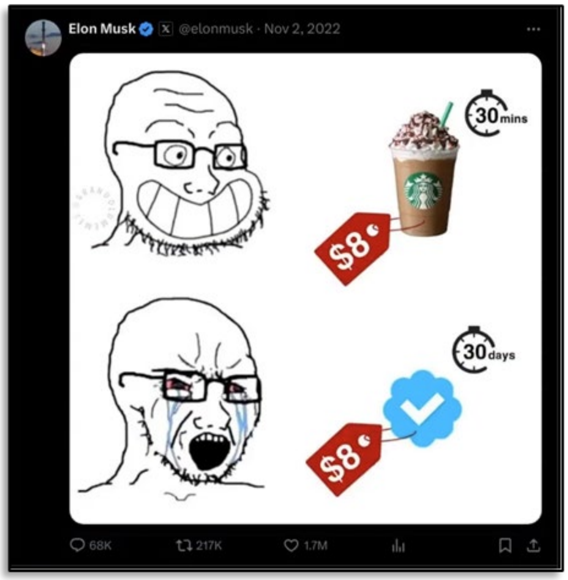
the platform. Acquiring Twitter was not just a business investment, it was a vanity project for the world's richest person. And what he wanted was to eliminate virtually all content moderation efforts. And, as one of his first official acts as CEO of X, he fired most of the staff responsible for policing hate speech.[28] The advertising backlash was immediate. X lost almost half of its advertising revenue within a few months of Musk's takeover, as advertisers fled the platform over concerns about toxic content.

45. X's efforts to make up the loss through user subscriptions fared no better. There is an old saying that people flock to a free meal, but charge them a penny, and you will be eating alone. X experienced this effect – called the "zero price effect" in economics – firsthand.

46. Prior to Musk's takeover, access to X was free to most users. When Musk took control, he tested whether the zero price was really the "competitive" price for user access to social media platforms, and quickly found that it was. X tried to extract cash from its users in many ways, including charging exorbitant fees to place a meaningless blue checkmark next to users' names, and placing certain features behind a subscription paywall.

[28] *Musk's Latest Twitter Cuts: Outsourced Content Moderators*, The Associated Press (Nov. 13, 2022), perma.cc/MH4U-ADW7.

47. Customers were not buying it, as they say. A picture Elon Musk tweeted about his failed strategy speaks volumes:[29]



Elon Musk @elonmusk · Nov 2, 2022

The facial expression of the consumer faced with having to pay $8 for a checkmark says it all: it is far from delight.

48. These and other efforts to extract cash directly from its users were largely unsuccessful, and certainly did not make up for the loss of advertising revenue. According to some published reports, X earned just an additional $5 million in extra revenue per month from selling new subscriptions, in stark contrast to X's multi-billion-dollar advertising slump.[30]

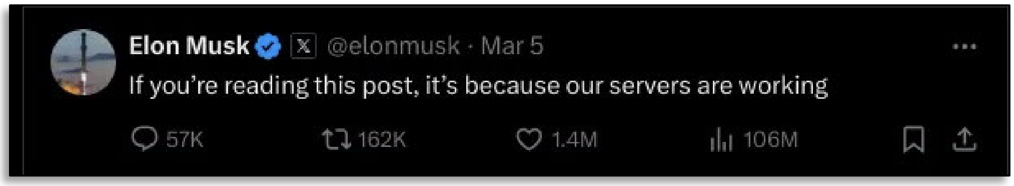
49. The financial distress caused by the advertising loss and anemic response to its subscription efforts was extreme. And so, Musk turned to what he called "drastic" cost cutting. By November 2022, just a few weeks after the acquisition, Musk warned of potential bankruptcy.

[29] @elonmusk, x.com, perma.cc/T6T4-FVQ7.

[30] *X Corp. v. World Federation of Advertisers*, 7:24-cv-00114, ECF 62 at 2 (N.D. Tex. 2024) (claiming Advertisers have withheld "billions of dollars in advertising revenue from Twitter, Inc."); *Elon Musk's X Still Struggles to Grow Subscription Revenue*, TechCrunch (Oct. 15, 2024), perma.cc/4ZFF-ZGJS ("Within the first three months post-relaunch, [X's subscription] service brought in only $11 million in mobile app subscriptions…").

Faced with a $3 billion negative cash flow, he embarked on a three-month plan to save Twitter from bankruptcy by firing 6,000 employees, equating to 80% of the company's staff. [31]

50. X's cost cutting measures caused significant negative effects on the user experience, including platform outages, glitchy system performance, slower bug fixes, and generally degraded platform quality. With tongue-in-cheek, Mr. Musk acknowledged the service disruptions caused by his own cost-cutting:[32]



Elon Musk @elonmusk · Mar 5

If you're reading this post, it's because our servers are working

51. But despite Musk's mismanagement, powerful network effects protected the company from significant user defection, enabling it to continue collecting vast amounts of user-generated data. That was its saving grace.

52. Though Musk's primary motivation for acquiring X was to hoard the data for exclusive use with Grok, back in mid-2023, that day had not yet arrived. Until it did, Musk had investors to appease. X was losing $4 million a day, and he urgently needed to stem the bleeding. With little else to offer, X refocused its efforts on monetizing and monopolizing the data market.

53. Prior to Musk's acquisition, X earned relatively little from data licensing or access fees. Its inability to profit from selling data was not for want of trying. You can't make money from selling something you don't own and that is freely available. And X does not own the user-generated content on its platform. In fact, it tells users that "[y]ou retain your rights to any Content you submit, post or display" on the platform, and "what's yours is yours – you own your content."[33] And those users – all looking for an audience – demanded that X place that content in the public domain. X, in turn, made this information freely available. Not only did X think scraping was

[31] *Elon Musk Says He's Cut About 80% of Twitter's Staff*, CNN Business (Apr. 12, 2023), perma.cc/PVS8-LSS6

[32] @elonmusk, x.com, perma.cc/49ZB-RHTB.

[33] X Terms (v.20 Nov. 15, 2024).

"fine," as Twitter's former head of Trust and Safety explained, but, as Musk would later lament, X had structured its employee compensation system to actively encourage it.[34] It served its goal of increasing engagement with the platform, which in turn made the platform more attractive to advertisers – X's primary focus in the pre-Musk era.

54. That the information was public and freely available, however, does not mean that it lacked value. The data on X is unique, reflecting real-time current events and the cultural zeitgeist of the world as it exists at any point in time. It has many valuable uses. Just look at Grok, the search engine chatbot that X launched through its commonly-controlled affiliate. Trained largely on X's user-generated data, it was recently valued at $60 billion. And that was just one use of the data. Extracting that value was only possible because X found a way to monopolize the data market by excluding all meaningful competitors, including Bright Data and other scrapers.

C. X Has Engaged in a Broad, Multifaceted Scheme to Acquire and Maintain a Monopoly Through Exclusionary and Deceptive Conduct.

55. In seeking to monetize user-generated data, X faced competitive threats from competing platforms and scrapers. To quash those threats, X needed to wrest control over the user-generated content on its platform and prevent emergent platforms from achieving necessary scale. It achieved that goal through a wide-ranging, multifaceted campaign of exclusionary, anticompetitive, and deceptive practices.

1. X's Efforts to Wrest Control Over Data It Does Not Own.

i. X Eschews All Lawful Means of Controlling Access to User Generated Data.

56. X had three options for eliminating data competitors. It chose the illegal one.

57. X could have acquired ownership of its users' content, which would have given X the right of exclusion. That would have been a lawful choice. But it was a non-starter. Consumer backlash would have been extreme if X forced users to relinquish ownership over their own

[34] @yoyoel.com, bluesky.app, perma.cc/29EX-F6RW; *Twitter v. Musk*, Verified Counterclaim, ¶¶ 3, 16 (accusing X of a "scheme" to increase mDAU [*i.e.*, monetizable Daily Active Users] by permitting non-human accounts, and structuring its "executive bonus pool" to achieve that result).

creative works. After years of telling users that "what's yours is yours – you own your content," X could not go back on that promise. X also knew that ownership is a double-edged sword. If X owned the content, it would have no immunity under the Communications Decency Act for false, deceptive, or otherwise unlawful content. This was untenable, especially after X gutted its content-moderation policies, and the amount of false, defamatory, and hateful content on the platform exploded.

58. With the ownership option off the table, X turned to its second option: placing information behind a log-in screen. But that option proved no better. Other social media platforms have done this for much of their content. Facebook, for example, started out as a "members' only club," with access limited to "friends" who logged in with passwords, and actively – not passively – sought out and "friended" those in their social circle. Though Facebook ultimately recognized the benefits of making some of its content publicly available as a means of attracting eyeballs, staying relevant, and avoiding the MySpace death spiral, it continues to password protect much of its content, taking advantage of this industry-standard and legally-enforceable method for restricting access to the non-public content posted on its platform.

59. But X's public nature is its strength, and it is an important structural feature of platforms that successfully compete in the Public Square Platform market. Unlike Facebook, X's platform was never designed for "cliques" of friends to communicate. It is not a members' only club. Its purpose – its raison d'être – is to be an open public square where the world at large can share ideas, and where those ideas would reach the broadest possible audience. As X explained:

> "Our primary product … is a global platform for public self-expression and conversation in real time. We have democratized content creation and distribution so people can consume, create, distribute and discover content about the topics and events they care about most."[35]

Password protection is anathema. Log-in requirements deter users, drive them away, and shrink the number of eyeballs X could claim as its own. Attempts to impose such requirements would

[35] Twitter Annual Report, at 6 (Feb. 16, 2022).

slowly but surely erode the powerful network effects that buttressed the platform's meteoric growth. With the erosion of those barriers to entry, so would go X's economic power.

60. X realized this first-hand. To prevent other chatbots, such as OpenAI's ChatGPT, from competing with Grok, X decided to place more content behind a log-in. In July 2023, the same month xAI was incorporated, X implemented these log-in changes.

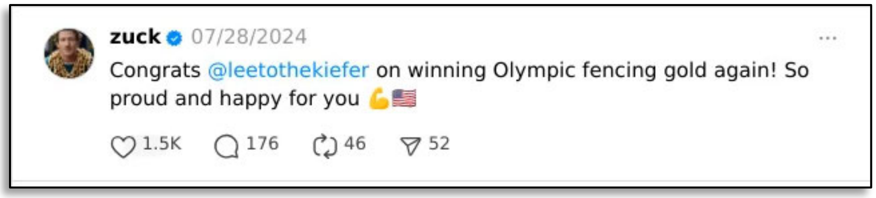
61. In making these changes, Elon Musk deceptively described these changes as an "emergency measure" needed because "we were getting data pillaged so much that it was degrading service for normal users."[36] This was a lie. X's changes were driven by competition concerns. While Musk admitted that he was concerned about A.I. companies scraping X, he falsely claimed that his concern related to server burden and performance, without disclosing that he had just started his own A.I. company and wanted to hoard the data for himself. X's log-in changes also coincided with X's decision to start charging customers for access to information (or quantities of information) that had previously been free through its API. Put simply, X's log-in changes were part-and-parcel of an attempt to exploit its information monopoly, not to reduce X's server costs.

62. The backlash from users, journalists, advertisers, and others was immediate. They depended on X's open access platform, and believed that a smaller, closed-off, walled garden would make the platform less useful and less relevant. But X faced an even greater threat: competition. When X took the "drastic" and "immediate" action of locking up the information behind a log-in, it created a void for competitors to fill. Within weeks, Meta answered the call, introducing Threads.

63. When Threads launched, there was substantial confusion in the market about whether it would require a log-in to access. Though Threads requires a log-in for full viewing access, substantial amounts of Threads posts are available without a log-in, as the following

[36] @elonmusk, x.com, perma.cc/7Y2L-P9YL.

screenshot shows:[37]



zuck 07/28/2024

Congrats @leetothekiefer on winning Olympic fencing gold again! So proud and happy for you 💪🇺🇸

Moreover, even if Threads imposes rate limits, it has a much larger installed base of account holders than X, given Threads' connection to Instagram. And, of course, X knew that Meta could change course at any time. Threads' entry, therefore, posed an existential threat to X, regardless of Threads' log-in practices.

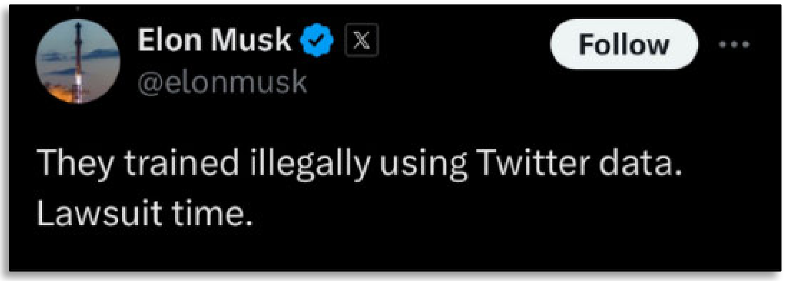
64. X realized that if it had to compete in the password protected market, it would be at a distinct disadvantage. X also knew that if it relaxed its log-in requirements, and Threads imposed one, then X would have a competitive advantage. It was a classic prisoner's dilemma; the kind competitors have faced since the dawn of time. X had no realistic choice. It was forced to quietly rollback the log-in requirements, once again ensuring that substantial amounts of real-time data would be freely available for public use without a log-in.[38]

[37] @zuck, threads.net, perma.cc/WRX7-SMKT.

[38] *Twitter Discreetly Removed Login Requirements to View Tweets*, TechThirsty.com, perma.cc/Y765-CBYG; *Twitter Silently Removes Login Requirement For Viewing Tweets*, Techcrunch, perma.cc/K4LV-7EMP

ii. X Uses Unlawful Exclusionary Contracts of Adhesion to Eliminate Competition.

65. With competition forcing X to maintain its public nature, X turned to the only option it had left: the illegal one – using contracts to wrest control over its users' data and eliminate competition from data competitors. As Musk tweeted, it was time to sue:[39]



Elon Musk @elonmusk

They trained illegally using Twitter data. Lawsuit time.

66. X purports to enter into contracts with every person that interacts with X's platform. According to X, a contract is formed whenever someone creates an account on the platform, or even when they visit the website without logging in. Simply typing x.com or twitter.com into your address bar, X claims, creates a binding contract, the Terms of which X has unilaterally dictated.

67. On July 26, 2023, X filed suit against Bright Data, believing Bright Data to be one of its biggest competitors in the Public Square Data market. This suit came mere days after X adopted its restricted log-in policies (and before it was forced to relax them). But X made a strategic mistake. Though it changed its log-in policy, it hadn't yet changed its Terms.

68. At that time, X's Terms did not purport to prohibit crawling, or most automated access to X's "currently available, published interfaces."[40] To the contrary, the Terms in effect prior to July expressly stated that "crawling the Services is permissible if done in accordance with the provisions of the robots.txt file."[41] And, prior to at least July 2023, the "robots.txt" file expressly permitted automated access or crawling of X's platform with respect to logged-off information. Moreover, even after July 2023 (as well as all dates prior), the Terms permitted

[39] @elonmusk, x.com, perma.cc/6W64-FDLB.

[40] X Terms (v.18 May 18, 2023).

[41] *Id.*

crawling even if not "done in accordance" with the voluntary provisions of X's robots.txt file, since there was no express prohibition of such activities.[42]

69. Recognizing that X's Terms did not prohibit automated access using X's published interfaces, on September 29, 2023, X unilaterally amended its Terms to prohibit "crawling [of] the Services in any form … without prior written consent."[43] But, as to Bright Data at least, that was too little, too late. Bright Data had already terminated its X accounts and served formal notice rejecting the Terms. On May 9, 2024, the Court held that X's post-lawsuit amendment is unenforceable as against Bright Data.[44]

70. The Court then went further and held that ***any*** contractual provisions seeking to strip internet users of their scraping rights are unenforceable against anyone, since *all* internet users have a right – granted by Congress – to scrape public web data. Specifically, in holding X's scraping provisions were preempted by the Copyright Act, the Court held that:

> "Giving social media companies 'free reign to decide … who can collect and use data – data that the companies do not own, that they otherwise make publicly available to viewers, and that the companies themselves collect and use – risks the possible creation of ***information monopolies*** that would disserve the public interest….'
>
> Pursuant to the Terms, X users 'own [their] Content,' [and] retain [their] rights to [it]…. They grant [only] a 'non-exclusive, royalty-free license' to X Corp., [which does not confer the] 'legal right to exclude others'.
>
> Yet that is exactly what X Corp. seeks to do with its claims based on scraping and selling data – to exclude others from using, copying, reproducing, processing, adapting, modifying, publishing, transmitting, displaying, and distributing X users' content.…
>
> The upshot is that [X's contracts] would entrench its own private copyright system that rivals, even conflicts with, the actual copyright system enacted by Congress. ***X would yank into its private domain and hold for sale information open to all, exercising a copyright owner's right to exclude where it has no such right***.

[42] X Terms (v.18 May 18, 2023).

[43] X Terms (v.19 Sept. 18, 2023).

[44] ECF 83, at 26.

> We are not concerned here with an arm's length contract between two sophisticated parties in which one or the other adjusts their rights and privileges under federal copyright law. ***We are instead concerned with a massive regime of adhesive terms imposed by X Corp. that stands to fundamentally alter the rights and privileges of the world at large*** ….
>
> Only by receiving permission and paying X Corp. could Bright Data, its customers, and other X users freely reproduce [information that is otherwise] available for taking and selling as fair use, [and as such, the contractual provisions prohibiting scraping] ***flouts Congress's intent*** [by]' attempt[ing] to protect that which Congress intended to be free from restraint.' …
>
> X Corp. [Terms] would upend the careful balance Congress struck…. In addition to giving itself de facto … ownership in … content that X users designated for public use, X Corp. would give itself de facto … ownership over content that Congress [decided] may not be copyrighted,'… shrink[ing] the public domain, restricting free reproduction … of publicly available … information."[45]

Based on this, the Court voided the two operative provisions of the Terms that purport to prohibit scraping. The first unenforceable provision related to purported limitations on the manner in which internet users may "reproduce, modify, create derivative works, distribute, sell, transfer, publicly display, publicly perform, [or] transmit" information published on the X platform.[46] The second unenforceable provision stated that "scraping the Services in any form, for any purpose without [X's] prior written consent is expressly prohibited."[47] The Court reaffirmed its pre-emption ruling when it denied the scraping portions of X's Motion to Amend the Complaint.[48]

71. Faced with this Court's pre-emption ruling, X scrambled to avoid its consequences. Though X has no right to directly prohibit scraping of publicly-available information, X sought to achieve the same end indirectly by expanding its contractual restrictions on automated access, even in the absence of any actual harm or non-contractual legal right to prevent such access. X's current Terms state that "crawling … the Services in any form, for any purpose without [its] prior written

[45] ECF 83, at 1, 18-19, 20, 26, 24.

[46] X Terms (v.19 Sept. 29, 2023).

[47] *Id.*

[48] ECF 156, at 20.

consent is expressly prohibited."[49] X, therefore, seeks to achieve its goal of preventing automated *scraping* by purporting to prevent automated *access*.[50]

72. In purporting to prevent automated access or scraping, X seeks to use the Terms to *strip* away the rights of the public at large. The Terms do not seek to *grant* selective *rights* to engage in otherwise prohibited automated access but seek to ban otherwise lawful conduct. Standard Internet Protocols do not require specific consent from website operators to use automated means to search, access or scrape the public web. Nor does any federal or state law require that a person obtain such consent. That is, each person has a God Given right to search the public web.

73. As such, X is not *enforcing* pre-existing, independent legal prohibitions, but rather is attempting to create a *new* legal right to restrain trade through contract. Indeed, if X possessed any non-contractual legal right to prevent automated access or scraping, X would not need a contract to ban the practice. So the fact that X relies on its Terms to purportedly ban automated access and scraping, therefore, demonstrates that X is relying on contracts in restraint of trade to achieve its unlawful objectives.

74. In attempting to use the Terms to strip the public of its right to access and scrape public portions of the web, X provides no consideration to data scrapers that do not use an X account. X seeks instead to bind these internet users to the Terms via implied "browser-wrap" consent. But X does not provide such users with *any* rights to use the Services, such as posting information or accessing non-public portions of the site. In fact, X seeks to prohibit such users from engaging in the only type of access they seek, which is automated access for the purpose of scraping publicly available information. Because X seeks to impose *only* obligations on such

[49] X Terms (v.20 Nov. 15, 2024).

[50] Despite the Court's ruling that post-suit amendments are unenforceable against Bright Data and no scraping-related claims can be asserted against Bright Data, X continues to assert that Bright Data is prohibited under the Terms from accessing X's platform through automated means. As such, X is seeking to use one or more versions of the Terms to prevent Bright Data from competing in the relevant data market.

users, rather than granting them *any* rights they seek, X has provided no consideration to support the massive restraints on the public's right to scrape public information published on the X platform.

75. X also seeks to bind data scrapers who do have an account to the Terms via a "click-wrap" account creation process. Scraping of publicly-available information on the X site, however, does not require the use of an account. In fact, X purports to place rate limits on the use of any account, which it claims would prevent account holders from scraping public or non-public portions of the X platform while logged on. Because X can track usage through an account, X can block automated access or scraping regardless of what device or proxy (or how many devices or proxies) are involved in the search.

76. Despite the fact that any logged-off scraping of public information does not require, and cannot be accomplished through the use of an account, X nonetheless purports to condition the benefits of account ownership, such as the right to advertise on the platform or post content for unrelated purposes, on the extraction of a promise not to crawl or scrape the X platform.

77. By requiring registered users to relinquish their right to lawfully access and scrape the public web and to contractually forego their right to compete with X in the monetization or use of data on the platform, X is leveraging its dominant position in the Public Square Platform market. Specifically, X is leveraging its monopoly power over user engagement to obtain or maintain a monopoly in the data market.

78. If X did not have monopoly power in this market, people could simply turn to other competitors, such as BlueSky, to post or view information, without having to relinquish their legal right to compete. But because those are not viable platforms, they have no choice but to accede to X's Terms.

79. Other than protecting X's monopoly, X has no legitimate basis for tying account-holder services to access that does not require or make use of any account. Nor does X have a legitimate basis for limiting internet users' access to the public data published on X's public platform. X does not limit its purported crawling prohibition to situations involving server failures

or injuries. Rather, X seeks to prohibit such conduct even where it involves a single server request, or where such activities are *de minimis*.

80. X's own Complaint demonstrates that the logged-off scraping of information from X's platform constitutes less [redacted] of X's web traffic.[51] X further admits that it maintains [redacted] excess server capacity. X has not alleged or identified any instance in which scraping of public information has resulted in any server failure, server or platform impairment, or degradation of the user experience. Nor has public scraping had such an effect.

81. The lack of such injury is fatal to any breach of contract claim, and thus, would be fatal to any effort to use the Terms to contractually block automated crawling or public scraping. As this Court held in initially dismissing X's access claims without prejudice:

> "Among the elements of a breach-of-contract claim is resulting damage… Under California law, 'a breach of contract without damages is not actionable' …. There is simply no damage alleged that is causally connected with [automated access]. Remember, the only damage that X Corp. plausibly pleaded … is that resulting from scraping and selling. X Corp. has not alleged any damage resulting from access through unauthorized means … without pleading any impairment or deprivation of servers."[52]

Faced with this ruling and the knowledge that X would not, and could not, show actual server impairment or deprivation, X embarked on a two-fold approach to save its contractual prohibitions.

82. *First*, X amended the Complaint with the false allegation that "Bright Data … degrades the user experience … by spreading fake accounts," which X knows to be objectively and subjectively baseless.[53] X knew that Bright Data did not scrape behind a log-in because X modeled its case off a similar case, where the court expressly found that Bright Data only scraped public information.[54] In amending its Complaint, X did not cite any evidence showing that Bright

[51] SAC ¶¶ 39, 80, 87.

[52] ECF 83, at 17.

[53] SAC ¶ 4.

[54] *Meta Platforms, Inc. v. Bright Data Ltd.*, 2024 WL 251406, *4 (N.D. Cal. 2024) (Chen, J.) ("The Court finds on the record before the Court there is no genuine issue of fact whether Bright Data scraped non-public data while logged in—***it did not***.").

Data or its customers use fake accounts. Nor is this the first time in recent years that Mr. Musk and agents under his control made baseless allegations concerning scraping activity on X. In Musk's attempt to reduce the purchase price of Twitter (after he had agreed to it), Musk alleged that Twitter's "claim[] that >95% of daily active users are real, unique humans" appears "false" or materially "misleading."[55] Yet, in this very case, X now admits that, even though the "volume of scraping has only increased" since the acquisition, only "about [redacted] is inauthentic and not attributable to human users."[56] Both statements cannot be true.

83. *Second*, X knew that, even if it survived a motion to dismiss, its victory would be a pyrrhic one because X would have difficulty proving injury at trial, against Bright Data or others. So, X sought to dispense with the need to show any injury whatsoever through contract. On November 15, 2024, X inserted the following provision into its Terms:

> "You further agree that, to the extent permitted by law, if you violate the Terms, or you induce or facilitate others to do so, in addition to all other legal remedies available to us, you will be jointly and severally liable to us for liquidated damages as following for requesting, viewing, or accessing more than 1,000,000 posts (including reply posts, video posts, image posts, and any other posts) in any 24-hour period – $15,000 USD per 1,000,000 posts. You agree that these amounts are (i) a reasonable estimate of our damages; (ii) not a penalty; and (iii) not otherwise limiting our ability to recover from you or other users under any legal or equitable theory…."[57]

84. Though this post-suit amendment is not enforceable against Bright Data, X nonetheless seeks to enforce it against other data scrapers, creating an *in terrorem* threat. Data scrapers now have to weigh vindicating their lawful right to scrape with the risk of having to pay severe or potentially crippling financial penalties if they did not succumb to X's demands to cease scraping.

[55] @elonmusk, x.com, perma.cc/H2XD-ZKSZ; *Twitter, Inc. v. Musk*, Complaint, Ex. 3; *Id.*, Verified Counterclaim, ¶ 203.

[56] SAC ¶¶ 85, 87.

[57] X Terms (v.20 Nov. 15, 2024).

85. X knows that this so-called "liquidated damages" provision is, in fact, a penalty, and not a reflection of its actual damages. As an initial matter, 1,000,000 posts may sound largeto an unsophisticated reader, but this is actually a tiny percentage of daily server requests. Users post approximately 500-700 million tweets per day, with an *exponentially* greater number of posts viewed or accessed each day.[58] As X's Complaint admits, it processes about 400 billion views or "events" per day.[59] X does not incur any incremental costs in responding to server requests to view or access a mere additional million posts, let alone $15,000, or the equivalent of 1.5 cents per viewed post.

86. In the days of yore, certain companies committed fraud by purporting to pass on Xeroxing costs by charging, for example, 10 cents per page when the actual costs were significantly less. X's liquidated damages provision – the equivalent of 1.5 cents for simply viewing a single picture or post – is far more unconscionable. It is also false and deceptive because X knows such fees are a penalty and are not a reasonable estimate of X's damages. Indeed, this penalty is approximately ***1400%*** *higher* than what X charges its API customers for similar types of data.[60]

87. X made this misrepresentation of fact knowingly, with the expectation that it would deter potential scrapers who would rely upon X's representation. Moreover, to the extent X seeks to have courts rely on this fraudulent provision, it is not entitled to immunity because the representation is objectively and subjectively baseless. Although X knew that its liquidated damages provision is not enforceable, and that the fee bears no relation to actual damages, X's goal was not to have courts uphold the fee and to collect on a judgment. X simply wanted a cudgel to wield to deter scrapers from continuing to access publicly available information on its site.

[58] SAC ¶ 39.

[59] *Id.* ¶ 81.

[60] *Twitter's New API Plan Costs Up to $2.5 Million Per Year*, Mashable (Mar. 10, 2023), perma.cc/U56B-W3E5.

Thus, irrespective of the enforceability of the provision, X knew it stood to benefit from the *in terrorem* effect it would have on eliminating competition from data scrapers.

88. The provision plays a critical role in X's efforts to monopolize the relevant markets by eliminating actual or potential competition from internet users who may wish to scrape public information from X's platform, including those using Bright Data's proxy service and scraping tools. The liquidated damages provision also interferes with Bright Data's customer relationships and causes, or threatens to cause, Bright Data significant irreparable injury by depriving Bright Data of revenues from internet scrapers who may wish to use Bright Data's proxy services or scraping tools to scrape public information published on X's platform.

2. X Uses Exclusive Contracts to Ensure that It Is the Sole Source of Public Square Data.

89. In addition to using its Terms to contractually prevent data scrapers, like Bright Data, from exercising their right to access and scrape publicly available data, X sought to eliminate competition by interfering with their customer relationships. X did so by entering into agreements with data scrapers' customers that purport to make X the exclusive source of data published on X's platform.

90. Based on its expansive (but dubious) theory of browser-wrap contract acceptance, X claims to enter into contracts with every person who interacts with its websites, including nearly half of the U.S. adult population. Virtually 100% of all customers interested in purchasing data appearing on the X website would have interacted with the site, meaning that X claims that its Terms contractually bind 100% of customers in the market for such data, prohibiting any of them from purchasing data from data scrapers or companies that provide scraping tools or services, leaving X as the sole and exclusive supplier of such data. Since X has 95% of the Public Square Platform market, X claims it has entered into contracts that cover 95% of the relevant market.

91. X's Terms operate as *de facto* exclusive contracts that restrain trade across the entire market.

92. Different website operators often employ different terms to achieve the same ends. Some website operators are clear in their effort to use their terms as *de facto* exclusive contracts

by purporting to prohibit users from purchasing data appearing on the platform. Others seek to achieve the same ends through their anti-facilitation provisions. X takes the latter approach. X claims that "[i]t is also a violation of these Terms to facilitate or assist others in violating these Terms, including by distributing products or services that enable or encourage violation of these Terms."[61]

93. X takes the position that, through this provision, there is no distinction between the person or entity that does the scraping and the person or entity that receives the fruits of such data scraping. That is, if a customer uses a vendor for datasets or scraping data or services, either the customer is the scraper and the vendor is the facilitator, or it is the reverse. Either way, X takes the position that both entities in this supply arrangement violate the Terms.

94. X's position is not legally sound for several reasons. *First*, the Terms do not expressly state that purchasing or ordering custom datasets constitutes scraping or facilitation of it, and as contracts of adhesion, the Terms cannot be construed in that way. *Second*, as this Court held, customers who purchase scraped data are exercising their lawful right, under the Copyright Act, to purchase *copies* of the data from sellers who are exercising their lawful right, under the Copyright Act, to copy, reproduce, and sell that data.[62] As such, to the extent X interprets its "anti-facilitation" provision to prevent the sale of scraped data, it is pre-empted by the Copyright Act.

95. Despite its legal infirmity, X nonetheless relies on its anti-facilitation provision to contractually prevent customers from purchasing scraped datasets from anyone, making X the sole source of such data. X's use of the anti-facilitation provision does not depend on the legal enforceability of the Terms, or whether X's legal arguments about how its Terms should be interpreted or applied have merit. Certainly, if X's interpretation – that data customers have contractually committed not to purchase scraped data from data scrapers or service providers – is correct and not pre-empted, then the Terms constitute contracts, combinations, or conspiracies in

[61] X Terms (v.20 Nov. 15, 2024).

[62] ECF 83, 156.

restraint of trade. But even if X's Terms could not be enforced in this way, the existence of the Terms creates legal risk sufficient to dissuade customers from purchasing the data from competing sources.

96. Thus, regardless of the Terms' legal enforceability, the existence and use of the Terms *in the marketplace* allows X as a matter of commercial reality to effectively control prices, exclude competition, and exercise substantial monopoly power.

97. Bright Data suffers significant, irreparable antitrust injury by operation of X's purported *de facto* exclusive contracts. But for the anti-facilitation provision in its Terms, there would be substantially greater demand for Bright Data's services. X's provision, therefore, not only insulates X from price competition, but also deprives Bright Data of revenues as a competitor in the relevant data market.

3. X Employs Deceptive Throttling Technologies to Eliminate Users' Links to Competing Social Media Platforms.

98. Recall that X faced two significant competitive threats to its information monopoly: threats from data scrapers and threats from other Public Square Platforms. These are not separate threats, but inter-related ones.

99. If other Public Square Platforms gain sufficient scale, they become viable competitors for user engagement and independent sources of data. In addition, if these other platforms are successful, data scrapers would turn to them to supply their data needs. If data scrapers can get Public Square Data from other platforms, they can sell that data in competition with X. Moreover, just as X, through its commonly-controlled affiliate, used Bright Data to scrape other websites, X's A.I. competitors could also use data scrapers to obtain the information they need for their models.

100. To eliminate competition for user engagement from other Public Square Platforms, and to eliminate competition for data sales from such other platforms and data scrapers, X devised a plan to stymie the growth of competing Public Square Platforms. And stymie them it did.

101. X's rightward swing and elimination of content moderation policies not only drove advertisers away, it alienated many of X's users. Competing platforms, like Meta's Threads and

BlueSky, tried to fill the void by offering an alternative. X also faced some competition from the right from Donald Trump's Truth Social platform. But these competitors were mostly annoyances, unable to overcome the powerful network effects protecting X's monopoly in the Public Square Platform market.

102. But in X's view, any leakage is too much leakage. The internet landscape is littered with once-dominant companies that found themselves relegated to the dustbins of history. Companies like MySpace, WordPerfect, AOL, and others all had their heyday, and were able to exercise monopoly power for substantial periods of time, until one day, other competitors found ways to overcome the powerful network effects that protected them. It can happen in a blink of the eye. Determined not to let history repeat itself, X took a page from those who had avoided this fate – Microsoft, Google, Facebook, Apple, and Amazon – by using exclusionary practices to successfully squash nascent threats.

103. Inter-operability of the Internet is important. The World Wide Web works as a global communication network because web pages are interconnected. Internet users can seamlessly traverse from one page to the next simply by clicking on a link pointing to a web page they wish to visit. X participates in the World Wide Web, and its platform has links to thousands of other websites. For the most part, these links are not selected or chosen by X, but rather by the users who post their content to X and embed links to other websites in their posts.

104. When visitors or users click on a link, X historically complied with World Wide Web protocols by allowing such visitors to leave X's site and traverse to the linked page.

105. In some cases, however, users choose to link to content posted on other Public Square Platforms. Just as one newspaper may scoop another, sometimes another platform may have more interesting, timely, or relevant content. If users look elsewhere for "What's Happening," the public square conversation will also move elsewhere. This poses an existential threat to X's dominance if it happens too often because users who leave X – even for a single post – might be inclined to sign up for an account, spend some time there, and ultimately begin developing a presence on that site.

106. The defections from X are critical because the Public Square Platform market is characterized by strong network effects. Users typically will not devote significant time and energy to creating entirely duplicative internet presences on multiple Public Square Platforms, but will generally choose the one that best fits their needs. As long as other Public Square Platforms remain small, network effects limit the scale of defections. But as they grow, defections become more likely, which makes it all the more important to nip nascent competitors in the bud before they have time to develop and grow.

107. X recognized this and began to secretly throttle other Public Square Platforms. According to published investigative reports, about a year ago, X began to "slow traffic to competing sites," such as BlueSky and Threads, forcing users to wait "more than 60 times longer" for those sites than the "average wait for links to other sites."[63] X did not disclose that it was secretly throttling users' ability to traverse the web, leaving users to believe that the delays were caused by their own slow internet connections, heavy internet usage, or the target sites' own technical problems.

108. X's secret throttling had significant anticompetitive effects. As the same investigative reports noted, the delays X caused can "feel extremely slow to users," and such delays can increase the "probability of a user … abandoning" the target site by 32%.[64] This not only decreases the likelihood of defections by users who are throttled, it dissuades content generators from posting information on competing sites. If it is harder for audiences to access, say BlueSky, than it is to access X, then content generators – who are time and resource constrained and often work on tight deadlines – will choose to develop content on X, rather than BlueSky.

109. X's secret throttling of web traffic has no procompetitive justification. Instead, X engaged in this anticompetitive behavior for the sole purpose of maintaining its monopoly power, both in the Public Square Platform market and the Public Square Data market.

[63] *Twitter Is Still Throttling Competitors' Links—Check for Yourself*, The Markup (Sept. 15, 2023), perma.cc/MC3D-UNSA.

[64] *Id.*

110. But X did not just stop at throttling competing websites. After Musk claimed to want to stop "lazy linking," X also began deprioritizing posts that contained links to other Public Square Platforms or real-time news sites.[65] X's clear purpose was to keep its users on its platform for as long as possible and make it as difficult as possible to leave. X's efforts to prevent defections from its site were deceptive. It neither disclosed to those who post information to the X site with external links that their content would be deprioritized, nor to those who view X posts that X had undertaken methods to entrap them within the X site.

111. X had no legitimate purpose for deceptively burying links to other websites. Its sole purpose in doing so was to eliminate competition from other websites in an effort to prevent user defection and the erosion of network effects protecting its dominant platform.

D. X's Conduct has Allowed it to Acquire, Maintain, and Enhance its Monopoly Power.

1. The Relevant Markets.

112. There are four relevant product markets in which to assess the alleged conduct: (i) the Public Square Platform market; (ii) the Public Square User Engagement market; (iii) the Public Square Data market; and (iv) the Publicly-Available Public Square Data market. The relevant geographic markets for each product market are the United States and the world.

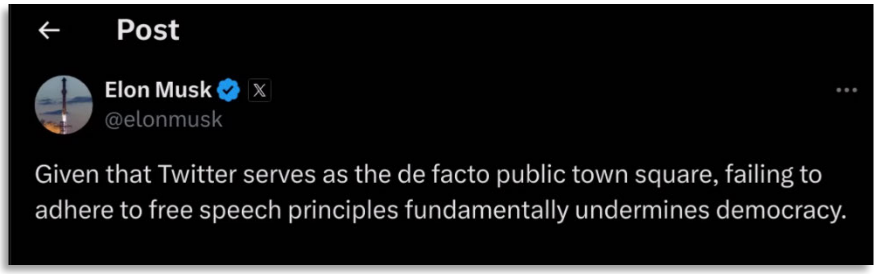
i. The Public Square Platform Market.

113. Public Square Platforms are social media platforms that focus on facilitating public conversation about real-time events, often through a micro-blogging structure that pushes content to broad audiences. As X explains, it provides "a global platform for public self-expression and conversation in real time" that "allows people to consume, create, distribute and discover content and has democratized content creation and distribution."[66] Elon Musk also described X as a "microblogging social media network where users share 280 character messages called 'tweets'"

[65] @elonmusk, x.com, perma.cc/Y8RX-PK54.

[66] Twitter Annual Report, at 6 (Feb. 17, 2021).

that has become the world's "de facto public square" or a "digital town square where matters vital to the future of humanity are debated."[67]



Given that Twitter serves as the de facto public town square, failing to adhere to free speech principles fundamentally undermines democracy.

114. X describes its business as a "multi-sided platform business"[68] that provides a "forum for public conversation on 'what's happening now." Though X describes this market as "multi-sided," it is not a two-sided transaction market, like payment networks, in which the platform sits in the middle of, and facilitates, a one-one transaction between the two sides. Moreover, unlike two-sided transaction markets, there is no privity of contract or any commercial relationship between the participants on each side of the platform. Rather, the Public Square Platform market operates more like a traditional market, with inputs and outputs.

115. The input side of the market consists of users and visitors who create and view the content. Public Square Platforms compete to attract content generated from or supplied by their users. They do so primarily on the basis of quality, not price, providing an attractive, convenient, and useful public forum for engaging in conversation and exchanging thoughts, ideas, opinions, and facts about current events and other matters of interests. Public Square Platforms provide several services to account holders, including the ability to post information and follow other users. But they generally do not pay for these users' content, nor do they charge these users for the privilege of posting or viewing information. For non-registered users, or visitors, Public Square Platforms provide no services. Non-registered users lack the ability to post information and cannot access the portions of the site that have been password protected. Nonetheless, Public Square

[67] @elonmusk, x.com, perma.cc/4S7U-JT7H; *Twitter v. Musk*, Verified Counterclaim, ¶¶ 24, 32.
[68] ECF 117, ¶ 52.

Platforms seek to encourage such interaction – by users and visitors alike – because the platform can monetize their engagement. That is, users and visitors – content creators and eyeballs – are inputs that these platforms seek to monetize in the output side of the market.

116. The input market for user content and engagement is protected by powerful network effects and other barriers to entry. Engagement with a Public Square Platform takes time and resources. While a particular user may use multiple, distinct, complementary social media platforms and other websites for distinct use cases, they typically only meaningfully engage with a single Public Square Platform. That is, a user might engage with LinkedIn – which is ***not*** a Public Square Platform – for job networking purposes, with Facebook – which is also ***not*** a Public Square Platform – to network with friends, and X to participate in public conversations about recent local, national, or world events. But a typical user will not engage with more than one Public Square Platform at a time. Instead, users will choose the Public Square Platform that best suits their personality, preferences, and needs. One of the most important factors, as with any communications network, is the number of other users or visitors that actively engage with the platform. Content providers (*i.e.*, users) are more likely to post content to Public Square Platforms that have large numbers of users or visitors, since it increases the likelihood that the content will go viral or be viewed by the targeted audience. Similarly, content viewers are more likely to devote time on platforms that have large amounts of desired content.

117. Once a platform becomes dominant, it is hard to unseat it from its throne. Due to these network effects, there are significant first mover advantages, which makes it difficult for nascent Public Square Platforms to gain sufficient market share and scale to meaningfully compete against more entrenched, incumbent Public Square Platforms. X has enjoyed the benefits of these barriers to entry for almost two decades.

118. Public Square Platforms that dominate the input side of the market also have structural advantages on the output side of the market, where these platforms make their money by monetizing user and visitor engagement. There are several revenue streams on the output side

of the market, the two most significant of which are advertising revenues and data licensing revenues.

119. Public Square Platforms compete with other Public Square Platforms in the monetization of user and visitor engagement on those platforms. They do so by competing on both price and quality. Only Public Square Platforms are capable of competing for the monetization of user and visitor engagement on these platforms.

120. A hypothetical monopolist of Public Square Platforms in the United States or globally could increase profits through pricing effects without significant loss of user or visitor engagement on the input side of the market. As a practical matter and competitive reality, any such price effect would likely occur on the output side of the market. Such a pricing effect would not adversely affect the input side of the market, and, in fact, may actually increase user engagement, because ads can be annoying.

121. A hypothetical monopolist could profitably raise prices above competitive levels in the output sides of the market because Public Square Platforms possess a unique ability to reach the users and visitors that engage with the site.

122. Account holders, in particular, tend to spend a significant amount of time engaging with the platform. And time spent engaging with the site means time not spent on other websites. These platforms also collect copious amounts of personal information about their users, which they use to target the user with specific ads using complex algorithms. As user engagement increases, so does value of the advertising (because there are fewer off-platform opportunities to reach the user), the effectiveness of the advertising (because greater engagement means more information and more effective targeting), and the amount of advertising that can be displayed (because more time on the site means more opportunities to show ads). Rather than forego the opportunity to advertise on a dominant Public Square Platform, advertisers would pay increased fees to reach the platform's users.

123. Data licensees would also pay increased fees to access this data. As discussed below, Public Square Data is incredibly valuable. Indeed, one of Musk's primary motivations for purchasing X was to gain control of this valuable data, which has no good substitute.[69]

124. Other Social Media Companies are not reasonable substitutes for the products or services offered by Public Square Platforms. Other websites and social media platforms – like Facebook, Instagram, TikTok, LinkedIn – are not Public Square Platforms and provide *complementary* services, not substitutable services, to their respective users. These social media companies cater to different audiences for different purposes. LinkedIn, for example, is focused largely on facilitating job networking. Facebook is focused largely on connecting with friends and other acquaintances. Instagram and TikTok are focused more on sharing personal videos and photos, or using such media to hawk wares. None of these platforms are designed to provide a public forum to discuss current events in real-time. Because the user content available on these social media platforms differs significantly from the content available on Public Square Platforms, and the nature of the user engagement differs, these social media platforms do not compete in the same output markets as Public Square Platforms.

ii. The Public Square User Engagement Market.

125. A relevant market exists for Public Square User Engagement. This market is an input market, which X is seeking to leverage for the purpose and with the effect of obtaining a monopoly, or a dangerous probability of obtaining a monopoly, in the Public Square Data markets, described below.

126. Public Square Platforms compete for user engagement, which they then seek to monetize. These platforms compete on quality and features of their service, and seek to maximize both the number of users and visitors on the platform, as well as the time spent on the platform.

127. Public Square Platforms serve a specific function of providing a digital forum for real-time public discourse about current events, usually through a micro-blogging structure. As

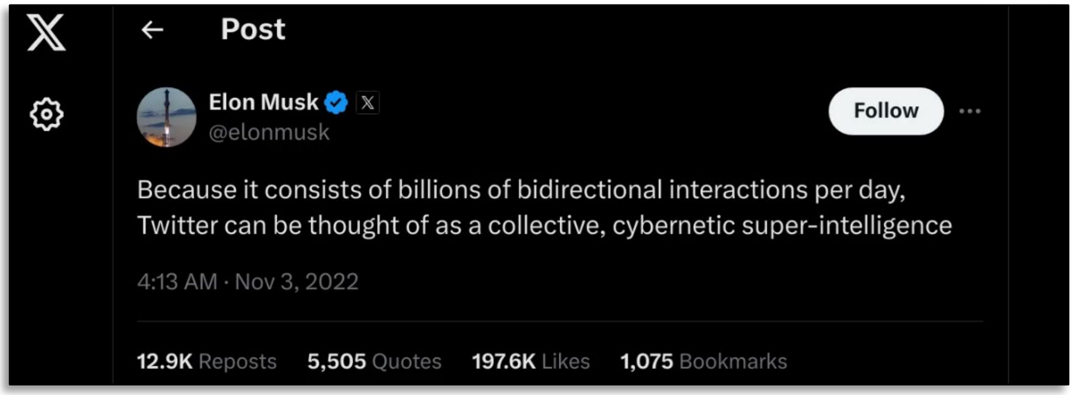
[69] *Uncovering Elon's Data Empire*, 53 STETSON L. REV. 405, 407 (2024).

described above, there are no reasonable substitutes for the services Public Square Platforms provide for users. And a hypothetical monopolist or monopsonist of the Public Square User Engagement market could reduce quality or features without losing so many users as to render the reductions unprofitable.

iii. The Public Square Data Market.

128. A relevant market exists for Public Square Data. Public Square Data is unique. It is public, global, real-time, and multi-participant or conversational.

129. Public Square platforms offer users a unique ability to engage with each other in real-time *conversations*. As Musk explained, "what's useful [about X's data] is the fact that it is up to the second."[70] It is a multi-lateral conversation, not a one-sided monologue. The bi-lateral, or multi-lateral, nature of this information makes it particularly valuable for certain end-use data applications, such as chat-bots. Chat bots are designed to mimic real conversations, so they need to be trained on real conversations, the type that occurs on Public Square Platforms, but not other social media websites. As Musk noted,



Elon Musk
@elonmusk

Because it consists of billions of bidirectional interactions per day, Twitter can be thought of as a collective, cybernetic super-intelligence

4:13 AM · Nov 3, 2022

12.9K Reposts 5,505 Quotes 197.6K Likes 1,075 Bookmarks

130. Other types of data from other sources are not reasonably substitutable, and a hypothetical monopolist of Public Square Data in the United States or globally could profitably raise prices above competitive levels.

[70] *Elon Musk on xAI: We will win*, perma.cc/N8YC-QCP5; @elonmusk, x.com, perma.cc/FQW2-BCDC.

131. Internet search engines, like Google or Bing, are not effective substitutes for Public Square Data because, while they index and have access to large amounts of Internet data, the data is not originally generated on those platforms. Indeed, the Public Square Data available on Google is largely derived from X. When one searches for Musk's posts on Google, for example, it does not show content that is originally created on Google, but links back to X.

132. Similarly, the content on other types of social media platforms is not substitutable for Public Square Data. Whereas Facebook is a one-sided memoir, the equivalent of a diary or a Christmas letter, with some ability to like or comment, X is designed as a multi-lateral conversation about current events and issues. Social media platforms that focus on longform, personal posts generated for member-only viewing such as Facebook have led to different informational content than the information on Public Square Platforms like X. Want to know what people are saying about themselves? Go to Facebook. Want to know what people are saying about events around the world? Go to X. Nor are messaging applications, such as Meta's WhatsApp or Slack, reasonably substitutable. The conversations occurring through those Apps are largely private, and often not retained by the platform's servers, and so cannot be sold as data in competition with Public Square Data.

133. Data scrapers compete in the Public Square Data market. Some scrapers, such as Bright Data and scrapers that use Bright Data's services, only compete in portions of this market because they only have access to publicly-available data. But because Public Square Platforms drive user engagement by being as open and public as possible, there remains a significant amount of Public Square Data that is publicly available and freely accessible by data scrapers. Competition from data scrapers place some competitive pressure on the prices that Public Square Platforms can charge, and thus, are properly considered part of the market, even if none of them are capable on their own of driving prices down to competitive levels.

134. A hypothetical monopolist of Public Square Data could profitably raise prices above competitive levels. Indeed, X's own conduct, as an actual monopolist, demonstrates that a hypothetical monopolist of this data could profitably raise prices. X has granted itself, or an

affiliate under its common control, the exclusive right to use data generated on the X platform for the purpose of training artificial intelligence chatbots. The value of that business went from zero to over $60 billion in about one year, primarily reflecting the value of this data. Moreover, since X embarked on its plan to monopolize the Public Square Data market, the prices it has charged for this data have skyrocketed, including the elimination of free access to the data for some data tiers, and 100% price increases for other data tiers. The result is that X now charges significantly more for access to users' posts than other news and data organizations charge for their data.

iv. The Publicly-Available Public Square Data Market.

135. A separate relevant market exists for Publicly-Available Public Square Data. This market is a submarket of the Public Square Data market. For reasons just discussed, this data differs significantly from data available from other data sources, such as news reporting agencies, other social media platforms, and other websites.

136. Because this information is freely available, competition in this market would drive the price down to the reasonable costs of collecting this information through scraping services, such as those offered by Bright Data. But a hypothetical monopolist of such data in the United States or globally could profitably raise prices above competitive levels without losing significant business. X has been able to profitably raise the price of Publicly-Available Public Square Data above competitive levels since Musk acquired the platform.

2. X Possesses Market Power, Monopoly Power, and a Dangerous Probability of Obtaining Monopoly Power in Each of the Alleged Markets.

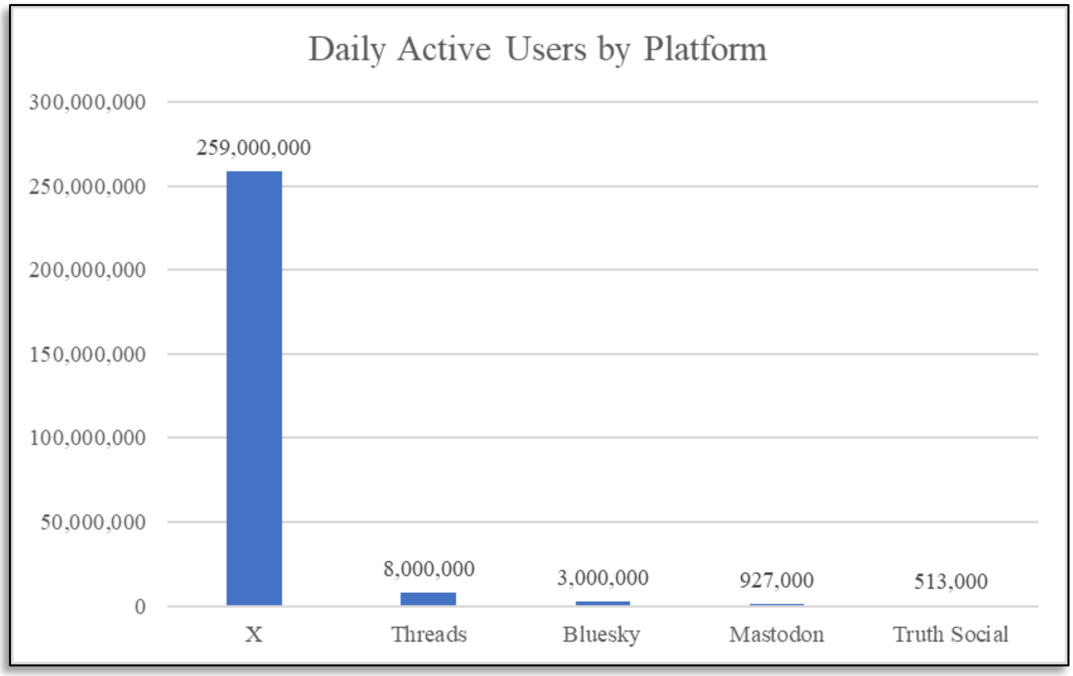
137. X is the dominant provider in each of the alleged markets, and possesses market power, monopoly or monoposony power, and a dangerous probability of obtaining monopoly or monoposony power. Indeed, X describes itself as "one of the largest public, real-time information platforms in the world."[71] X has over 250 million active daily users, who post more than 500

[71] X, Compliance Firehose API, perma.cc/M96W-5FMX.

million post per day.[72] According to some estimates, over 49% of Americans are active users of the platform.

138. X's share of each of these markets exceeds 95%, with an Herfindahl–Hirschman index (HHI) exceeding 9,100.[73]

139. In the Public Square Platform market and Public Square User Engagement market, X faces competition primarily from Mastodon, Bluesky, Threads, and Truth Social. None of these competitors, individually or collectively, have been able to garner significant market share, or overcome the strong network effects protecting X's monopoly or monopsony, as the following chart shows:



[72] @elonmusk, x.com, perma.cc/TU5T-4AB3 (touting "Record usage of the X platform!," and noting that, of October 4, 2024, X had 363.6B daily user-seconds, 547.9M monthly active users, and 8.9B daily video views).

[73] X: perma.cc/7LQ3-R3ZC; Threads: perma.cc/G45Q-NC7X; Bluesky: perma.cc/7GF4-8Y6U; Mastodon: perma.cc/GCX8-C9NV; Truth Social: perma.cc/Z7M7-QVU6.

140. X also has approximately 95% of both the Public Square Data market and the Publicly-Available Public Square markets. The same Public Square Platform competitors compete in each of the data markets and have data sales comparable to or below their respective shares in the Public Square Platform Market. Though data scrapers also compete in the Public Square data markets, their market shares are small relative to the Public Square Platform competitors.

141. X has a demonstrated history of controlling prices and successfully exploiting its monopoly power to raise prices above competitive levels in each of the alleged relevant markets. Prior to Musk's acquisition, Twitter allowed the public to access 2 million tweets per month through its API for free. In March 2023, just months after Musk's acquisition, X eliminated that tier. The closest equivalent now costs $60,000 per year for half the data. Prices for other tiers also increased. X's profits from data licensing have increased as a result of its exclusionary conduct, and the resulting price increases it was able to impose.

E. X's Conduct Has Caused and Threatens to Cause Significant and Irreparable Antitrust Injury.

142. X's conduct has caused and threatens to cause significant harm to competition. The alleged conduct substantially forecloses, eliminates, or reduces competition from X's competitors in the Public Square Platform and the two Public Square Data markets. It does this by preventing other Public Square Platform competitors from achieving sufficient scale, or obtaining the necessary data, to meaningfully compete in either market, and by preventing data scraping companies, including Bright Data, from competing in the Public Square Data markets.

143. The conduct enables X to charge supra-competitive prices for Public Square Data, thereby harming customers in all three alleged markets.

144. The conduct further injures customers by interfering with their ability to freely browse the World Wide Web.

145. Bright Data suffered actual, threatened, and irreparable injury to the extent the Terms apply to Bright Data and purport to prevent it from accessing or scraping public data hosted on the X platform, or from selling services that could be used for such purposes. Bright Data also

suffers actual, threatened, and irreparable injury because the conduct interferes with Bright Data's actual and prospective customer relationships, and its ability to sell data that it has or would lawfully scrape but for X's alleged unlawful conduct.

146. All of Bright Data's injuries constitute antitrust injuries. Bright Data is a competitor in the Public Square Data and Publicly-Available Data Square markets and has been competitively foreclosed, or is threatened with competitive foreclosure, from that market. As a scraper, Bright Data also constrains X's ability to exercise market power or monopoly power in the Public Square Platform market, and thus, its injuries are inextricably intertwined with X's ability to acquire or maintain monopoly power in that market. Bright Data is also a consumer of Public Square Platform services, Public Square Data, and Publicly-Available Public Square Data, and has standing as a consumer or customer in those markets.

V. COUNTS

A. Count I: Section 1 of the Sherman Act, 15 U.S.C. § 1.

147. Bright Data incorporates all other paragraphs of this Counterclaim as if realleged herein.

148. X's conduct violates Section 1 of the Sherman Act, 15 U.S.C. § 1.

149. The markets for Public Square Platforms, Public Square User Engagement, Public Square Data, and Publicly-Available Public Square Data in the United States or globally are each relevant antitrust markets.

150. X possesses market power in each of the alleged relevant markets because it has a dominant share of each alleged relevant market, and possesses the power to control prices or exclude competition from such markets.

151. X entered into a contract, combination, or conspiracy with all persons subject to the Terms, including data scrapers, customers of data scraping services, and customers of scraped data.

152. The Terms unreasonably restrain trade in the alleged relevant markets by purporting to contractually prohibit data scrapers, customers of data scraping services, and customers of

scraped services from exercising their otherwise lawful right to access, scrape, copy, distribute, purchase, or use publicly available data published on the X platform.

153. The unreasonable restraints of trade include:

- Adopting and enforcing procedurally and substantively unconscionable Terms of Adhesion with the purpose and effect of eliminating competition in the relevant market.
- Adopting and enforcing Terms that purport to (i) prohibit crawling or automated access; (ii) prohibit scraping; (iii) prohibit facilitation of the foregoing; and (iv) impose liquidated damages.
- Adopting and enforcing Terms that purport to require users to relinquish their right to engage in otherwise lawful scraping or access as a condition of using X, including for unrelated purposes.
- Adopting and enforcing Terms that operate as *de facto* exclusive contracts with a substantial portion of the market that purports to prevent customers from purchasing data published on the X site, including any publicly-available data, from any other source.

154. There is no procompetitive justification for X’s conduct. I anticompetitive effects of X’s conduct outweigh any procompetitive justifications for such conduct, and/or there are less restrictive means for achieving any procompetitive justifications for X’s conduct.

155. X’s conduct harmed competition in each of the alleged relevant markets. X’s conduct allowed X to foreclose competitors from the alleged relevant markets and to raise prices or reduce quality in all of the alleged relevant markets.

156. X’s conduct occurred in and affected interstate commerce.

157. Bright Data suffered, and is threatened with, ongoing, substantial, direct, and foreseeable antitrust injury directly, proximately, and foreseeably caused by X’s unlawful conduct.

B. Count II: Monopolization Under Section 2 of the Sherman Act, 15 U.S.C. § 2.

158. Bright Data incorporates all other paragraphs of this Counterclaim as if realleged herein.

159. X’s conduct constitutes unlawful monopolization under Section 2 of the Sherman Act, 15 U.S.C. § 2.

160. The markets for Public Square Platforms, Public Square User Engagement, Public Square Data, and Publicly-Available Public Square Data in the United States or globally are each relevant antitrust markets.

161. X possesses monopoly power in each of the alleged relevant markets because it has a dominant share of each alleged relevant market exceeding 95% and possesses the power to control prices or exclude competition from such markets.

162. X willfully acquired, maintained, or enhanced monopoly power in each of the alleged relevant markets through exclusionary conduct. X's exclusionary conduct includes:

- Adopting and enforcing procedurally and substantively unconscionable Terms of Adhesion with the purpose and effect of eliminating competition in the relevant market.
- Adopting and enforcing Terms that purport to (i) prohibit crawling or automated access; (ii) prohibit scraping; (iii) prohibit facilitation of the foregoing; and (iv) impose liquidated damages.
- Adopting and enforcing Terms that purport to require users to relinquish their right to engage in otherwise lawful scraping or access as a condition of using X, including for unrelated purposes.
- Adopting and enforcing Terms that operate as *de facto* exclusive contracts with a substantial portion of the market that purports to prevent customers from purchasing data published on the X site, including any publicly-available data, from any other source.
- Throttling web traffic or burying posts with links to competing Public Social Media Platforms and other websites.
- Leveraging its monopoly power in the Public Square Platform and Public Square User Engagement markets to obtain, maintain or enhance a monopoly in the Public Square Data and Publicly-Available Public Square Data markets.

163. There is no procompetitive justification for X's conduct. The anticompetitive effects of X's conduct outweigh any procompetitive justifications for such conduct, and/or there are less restrictive means for achieving any procompetitive justifications for X's conduct.

164. X's conduct harmed competition in each of the alleged relevant markets. X's conduct allowed X to foreclose competitors from the alleged relevant markets and to raise prices or reduce quality in all of the alleged relevant markets.

165. X's conduct occurred in and affected interstate commerce.

166. Bright Data suffered, and is threatened with, ongoing, substantial, direct, and foreseeable antitrust injury directly, proximately, and foreseeably caused by X's unlawful conduct.

C. Count III: Attempted Monopolization Under Section 2 of the Sherman Act, 15 U.S.C. § 2.

167. Bright Data incorporates all other paragraphs of this Counterclaim as if realleged herein.

168. X's conduct constitutes unlawful monopolistic conduct under Section 2 of the Sherman Act, 15 U.S.C. § 2.

169. The markets for Public Square Platforms, Public Square User Engagement, Public Square Data, and Publicly-Available Public Square Data in the United States or globally are each relevant antitrust markets.

170. X possesses a dangerous probability of obtaining monopoly power in each of the alleged relevant markets because it has a substantial share of each alleged relevant market, and possesses or has a dangerous probability of obtaining the power to control prices or exclude competition from such markets.

171. X acted with specific intent to acquire, maintain, or enhance monopoly power in each of the alleged relevant markets through exclusionary conduct. X's exclusionary conduct includes:

- Adopting and enforcing procedurally and substantively unconscionable Terms of Adhesion with the purpose and effect of eliminating competition in the relevant market.
- Adopting and enforcing Terms that purport to (i) prohibit crawling or automated access; (ii) prohibit scraping; (iii) prohibit facilitation of the foregoing; and (iv) impose liquidated damages.
- Adopting and enforcing Terms that purport to require users to relinquish their right to engage in otherwise lawful scraping or access as a condition of using X, including for unrelated purposes.
- Adopting and enforcing Terms that operate as *de facto* exclusive contracts by purporting to prevent customers from purchasing data published on the X site, including any publicly-available data, from any other source.

- Throttling web traffic or burying posts with links to competing Public Social Media Platforms and other websites.
- Leveraging its monopoly power in the Public Square Platform and Public Square User Engagement markets to achieve a dangerous probability of obtaining a monopoly in the Public Square Data and Publicly-Available Public Square Data markets.

172. There is no procompetitive justification for X's conduct. The anticompetitive effects of X's conduct outweigh any procompetitive justifications for such conduct, and/or there are less restrictive means for achieving any procompetitive justifications for X's conduct.

173. X's conduct harmed competition in each of the alleged relevant markets. X's conduct allowed X to foreclose competitors from the alleged relevant markets and to raise prices or reduce quality in all of the alleged relevant markets.

174. X's conduct occurred in and affected interstate commerce.

175. Bright Data suffered, and is threatened with, ongoing, substantial, direct, and foreseeable antitrust injury directly, proximately, and foreseeably caused by X's unlawful conduct.

D. Count IV: California Cartwright Act, Cal. Bus. & Prof. Code, § 16720, et seq.

176. Bright Data incorporates all other paragraphs of this Counterclaim as if realleged herein.

177. X's conduct violates California's Cartwright Act, Cal. Bus. & Prof. Code, § 16720, et seq.

178. The markets for Public Square Platforms, Public Square User Engagement, Public Square Data, and Publicly-Available Public Square Data in the United States or globally are each relevant antitrust markets.

179. X possesses market power in each of the alleged relevant markets because it has a dominant share of each alleged relevant market and possesses the power to control prices or exclude competition from such markets.

180. X entered into a contract, combination, or conspiracy with all persons subject to the Terms, including data scrapers, customers of data scraping services, and customers of scraped data.

181. The Terms unreasonably restrain trade in the alleged relevant markets by purporting to contractually prohibit data scrapers, customers of data scraping services, and customers of scraped services from exercising their otherwise lawful right to access, scrape, copy, distribute, purchase, or use publicly available data published on the X platform.

182. The unreasonable restraints of trade include:

- Adopting and enforcing procedurally and substantively unconscionable Terms of Adhesion with the purpose and effect of eliminating competition in the relevant market.
- Adopting and enforcing Terms that purport to (i) prohibit crawling or automated access; (ii) prohibit scraping; (iii) prohibit facilitation of the foregoing; and (iv) impose liquidated damages.
- Adopting and enforcing Terms that purport to require users to relinquish their right to engage in otherwise lawful scraping or access as a condition of using X, including for unrelated purposes.
- Adopting and enforcing Terms that operate as *de facto* exclusive contracts with a substantial portion of the market that purports to prevent customers from purchasing data published on the X site, including any publicly-available data, from any other source.

183. There is no procompetitive justification for X's conduct. The anticompetitive effects of X's conduct outweigh any procompetitive justifications for such conduct, and/or there are less restrictive means for achieving any procompetitive justifications for X's conduct.

184. X's conduct harmed competition in each of the alleged relevant markets. X's conduct allowed X to foreclose competitors from the alleged relevant markets and to raise prices or reduce quality in all of the alleged relevant markets.

185. X's conduct occurred in or substantially affected California commerce.

186. Bright Data suffered, and is threatened with, ongoing, substantial, direct, and foreseeable antitrust injury directly, proximately, and foreseeably caused by X's unlawful conduct.

E. Count V: California Unfair Competition Act, Cal. Bus. & Prof. Code § 17200, et seq.

187. Bright Data incorporates all other paragraphs of this Counterclaim as if realleged herein.

188. X engaged in unlawful, unfair, or fraudulent business acts and practices in violation of California's Unfair Competition Law ("UCL), Cal. Bus. & Prof Code §17200, et seq.

189. X violates the unlawful prong of the UCL because its conduct violates the laws alleged in the other Counts of this Counterclaim. X's conduct also violates Section 5 of the FTC Act, 15 U.S.C. § 5, because it constitutes an unfair method of competition and deceptive practices.

190. X violates the unfair competition prong of the UCL by:

- Adopting and enforcing procedurally and substantively unconscionable Terms of Adhesion with the purpose and effect of eliminating competition in the relevant market.
- Adopting and enforcing Terms that purport to (i) prohibit crawling or automated access; (ii) prohibit scraping; (iii) prohibit facilitation of the foregoing; and (iv) impose liquidated damages.
- Adopting and enforcing Terms that purport to require users to relinquish their right to engage in otherwise lawful scraping or access as a condition of using X, including for unrelated purposes.
- Adopting and enforcing Terms that operate as *de facto* exclusive contracts by purporting to prevent customers from purchasing data published on the X site, including any publicly-available data, from any other source.
- Throttling web traffic or burying posts with links to competing Public Social Media Platforms and other websites.
- Leveraging its monopoly power in the Public Square Platform and Public Square User Engagement markets to obtain, maintain or enhance (or to achieve a dangerous probability of obtaining) a monopoly in the Public Square Data and Publicly-Available Public Square Data markets.

191. There is no procompetitive justification for X's conduct. The anticompetitive effects of X's conduct outweigh any procompetitive justifications for such conduct, and/or there are less restrictive means for achieving any procompetitive justifications for X's conduct.

192. X's conduct harmed competition in each of the alleged relevant markets. X's conduct allowed X to foreclose competitors from the alleged relevant markets and to raise prices or reduce quality in all of the alleged relevant markets.

193. X violates the unfair, fraudulent, and deceptive practices prong of the UCL by: (i) falsely representing that the Liquidated Damages provision of the Terms is not a penalty, but a

"reasonable estimate" of X's actual damages caused by automated access or scraping; and (ii) secretly throttling web traffic or burying posts with links to competing Public Social Media Platforms and other websites. Such statements constitute material misrepresentations of fact justifiably relied upon and made with scienter. Such conduct also constitutes a deceptive practice.

194. X's conduct occurred in or substantially affected California commerce.

195. Bright Data suffered, and is threatened with, ongoing, substantial, direct, and foreseeable injury directly, proximately, and foreseeably caused by X's unlawful conduct.

F. Count VI: Nevada Unfair Trade Practice Act (UTPA) – NRS 598A.060.

196. Bright Data incorporates all other paragraphs of this Counterclaim as if realleged herein.

197. X engaged in unfair trade practices that violate NRS 598A.060(1)(e) of Nevada state law.

198. The markets for Public Square Platforms, Public Square User Engagement, Public Square Data, and Publicly-Available Public Square Data in the United States or globally are each relevant antitrust markets.

199. X possesses or has a dangerous probability of possessing monopoly power in each of the alleged relevant markets because it has a dominant share of each alleged relevant market exceeding 95% and possesses the power to control prices or exclude competition from such markets. X also has market power in each such market.

200. X entered into a contract, combination, or conspiracy with all persons subject to the Terms, including data scrapers, customers of data scraping services, and customers of scraped data that substantially restrained trade and commerce.

201. X also attempted to, and did willfully, acquire or maintain monopoly power in the Public Square Platforms, Public Square User Engagement, Public Square Data, and Publicly-Available Public Square Data markets. X's exclusionary conduct and unreasonable restraints of trade include:

- Adopting and enforcing procedurally and substantively unconscionable Terms of Adhesion with the purpose and effect of eliminating competition in the relevant market.

- Adopting and enforcing Terms that purport to (i) prohibit crawling or automated access; (ii) prohibit scraping; (iii) prohibit facilitation of the foregoing; and (iv) impose liquidated damages.

- Adopting and enforcing Terms that purport to require users to relinquish their right to engage in otherwise lawful scraping or access as a condition of using X, including for unrelated purposes.

- Adopting and enforcing Terms that operate as *de facto* exclusive contracts by purporting to prevent customers from purchasing data published on the X site, including any publicly-available data, from any other source.

- Throttling web traffic or burying posts with links to competing Public Social Media Platforms and other websites.

- Leveraging its monopoly power in the Public Square Platform and Public Square User Engagement Markets to obtain, maintain or enhance (or to achieve a dangerous probability of obtaining) a monopoly in the Public Square Data and Publicly-Available Public Square Data markets.

202. There is no procompetitive justification for X's conduct. The anticompetitive effects of X's conduct outweigh any procompetitive justifications for such conduct, and/or there are less restrictive means for achieving any procompetitive justifications for X's conduct.

203. X's conduct harmed competition in each of the alleged relevant markets. X's conduct allowed X to foreclose competitors from the alleged relevant markets and to raise prices or reduce quality in all of the alleged relevant markets.

204. X's conduct occurred in or substantially affected Nevada commerce.

205. Bright Data suffered, and is threatened with, ongoing, substantial, direct, and foreseeable antitrust injury directly, proximately, and foreseeably caused by X's unlawful conduct.

G. Count VII: Texas Free Enterprise & Antitrust Act – Tex. Bus. Comm. Code 15.05(a), (b).

206. Bright Data incorporates all other paragraphs of this Counterclaim as if realleged herein.

207. X's conduct violates Texas's Business and Commerce Code sections 15.05(a) and (b).

208. The markets for Public Square Platforms, Public Square User Engagement, Public Square Data, and Publicly-Available Public Square Data in the United States or globally are each relevant antitrust markets.

209. X possesses or has a dangerous probability of possessing monopoly power in each of the alleged relevant markets because it has a dominant share of each alleged relevant market exceeding 95%, and possesses the power to control prices or exclude competition from such markets. X also has market power in each such market.

210. X entered into a contract, combination, or conspiracy with all persons subject to the Terms, including data scrapers, customers of data scraping services, and customers of scraped data that substantially restrained trade and commerce.

211. X also attempted to, and did willfully, acquire or maintain monopoly power in the Public Square Platforms, Public Square User Engagement, Public Square Data, and Publicly-Available Public Square Data markets. X's exclusionary conduct and unreasonable restraints of trade include:

- Adopting and enforcing procedurally and substantively unconscionable Terms of Adhesion with the purpose and effect of eliminating competition in the relevant market.

- Adopting and enforcing Terms that purport to (i) prohibit crawling or automated access; (ii) prohibit scraping; (iii) prohibit facilitation of the foregoing; and (iv) impose liquidated damages.

- Adopting and enforcing Terms that purport to require users to relinquish their right to engage in otherwise lawful scraping or access as a condition of using X, including for unrelated purposes.

- Adopting and enforcing Terms that operate as *de facto* exclusive contracts by purporting to prevent customers from purchasing data published on the X site, including any publicly-available data, from any other source.

- Throttling web traffic or burying posts with links to competing Public Social Media Platforms and other websites.

- Leveraging its monopoly power in the Public Square Platform and Public Square User Engagement markets to obtain, maintain or enhance (or to achieve a dangerous probability of obtaining) a monopoly in the Public Square Data and Publicly-Available Public Square Data markets.

212. There is no procompetitive justification for X's conduct. The anticompetitive effects of X's conduct outweigh any procompetitive justifications for such conduct, and/or there are less restrictive means for achieving any procompetitive justifications for X's conduct.

213. X's conduct harmed competition in each of the alleged relevant markets. X's conduct allowed X to foreclose competitors from the alleged relevant markets and to raise prices or reduce quality in all of the alleged relevant markets.

214. X's conduct occurred in or substantially affected Texas commerce.

215. Bright Data suffered, and is threatened with, ongoing, substantial, direct, and foreseeable antitrust injury directly, proximately, and foreseeably caused by X's unlawful conduct.

H. Count VIII: Tortious Interference With Prospective Customer Relationships.

216. Bright Data incorporates all other paragraphs of this Counterclaim as if realleged herein.

217. X has tortiously interfered with Bright Data's current and prospective customer relationships in violation of the common law of California and the several States.

218. As a leading internet data company offering a suite of proxy services, scraping tools, and datasets, Bright Data has existing and prospective economic relationships with customers interested in accessing, scraping, or purchasing publicly-available data hosted on the X platform. In the absence of the alleged conduct, such relationships would likely result in significant future economic benefit to Bright Data.

219. X is aware of Bright Data's business, the way its services operate, and the demand for Bright Data's services from customers who seek to access, scrape, purchase, or use publicly-available data hosted on the X platform. X, therefore, either knows about Bright Data's customer relationships or is aware of circumstances suggesting the existence of such current and potential relationships.

220. X has intentionally engaged in acts designed to disrupt these economic relationships. This includes adopting and enforcing unconscionable, unenforceable, and anticompetitive Terms that, among other things, falsely and baselessly threaten customers with extreme penalties for using Bright Data's services. X's wrongful acts also include forcing account holders to relinquish their right to access, scrape, or purchase publicly available data on the X platform either for no consideration, in the case of visitors purportedly bound by browser-wrap, or as a condition of opening an account for unrelated purposes.

221. X's conduct does not constitute mere competition and no legitimate business reason excuses or justifies the wrongful conduct.

222. X's conduct has actually disrupted or is likely to disrupt potential customer relationships.

223. Bright Data suffered, and is threatened with, ongoing, substantial, direct, and foreseeable injury directly, proximately, and foreseeably caused by X's unlawful conduct.

VI. JURY DEMAND

224. Bright Data requests, pursuant to Rule 38(b) of the Federal Rules of Civil Procedure, a trial by jury for all claims and issues so triable.

VII. PRAYER FOR RELIEF

WHEREFORE, Bright Data prays that the Court award the following relief:

1. Declare and adjudge X's conduct unlawful as set forth in the each of the above Counts;

2. Enjoin X from continuing to engage in the unlawful conduct alleged in this Counterclaim;

3. Award damages, treble damages, and punitive damages to the extent permitted by each of the above Counts;

4. Award reasonable attorneys' fees and costs of suit to the extent permitted by each of the above Counts; and

5. Award such other relief as the Court deems just and proper.

ANSWER

Pursuant to Fed. R. Civ. P. 8(b)(3), Bright Data, Ltd. "generally denies" all allegations of the Second Amended Complaint (ECF No. 117-4), including in its footnotes, headings, and sub-headings, "except those specifically admitted," as set forth below.*

INTRODUCTION

1. Defendant Bright Data Ltd. ("Bright Data") has built an illicit data-scraping business on the backs of innovative technology companies like X Corp., which operates the social media platform formerly known as Twitter and now known as X. Bright Data scrapes and sells millions of records from the X platform. Bright Data also sells tools enabling others to unlawfully scrape data from X without (in its own words) being "flagged or blocked" or "detect[ed]." To facilitate the scraping it knows X is trying to stop, Bright Data sells tools enabling its customers to "[a]void IP blocks and bans," to "[b]ypass geo-restrictions and CAPTCHAs," and to "overcome anti-bot detection" measures that X employs to safeguard its platform.

Answer: Denied. This allegation is entirely false. Bright Data is a leading internet-based company that offers a suite of technologies and services that help Fortune 500 companies, academic institutions, and small businesses to retrieve and synthesize vast amounts of public information. Far from being "illicit," Bright Data's tools only permit users to scrape information that website operators, including X, do not own and have chosen to make publicly available. As this Court has already held, federal law gives Bright Data and the public the unfettered legal right to scrape such information, and has pre-empted all state law to the contrary.

X itself believes that Bright Data's industry-leading proxy service and scraping tools are lawful and provide highly valuable services that drive technological innovation. In May 2024, X, through its commonly-controlled affiliate, reached out to Bright Data to request use of Bright Data's proxy service and scraping tools to scrape third-party websites, including sites with Terms of Use similar to X's.

X sought out Bright Data precisely because its tools and services – including the ones referenced

* For any response in which Bright Data asserts that it does not know or is not aware of a particular alleged fact, Bright Data specifically means and avers that it "lacks knowledge or information sufficient to form a belief about the truth of an allegation," and so denies the allegation under Fed. R. Civ. P. 8(b)(5).

in this Paragraph – enhance the user experience by making the (clearly ***not*** illicit) search for public information more efficient. Through the use of these tools, Bright Data has scraped some publicly-available data on X, but such data represents an infinitesimal portion of the data published on the platform. Nor is it true that, for much of the relevant period, X sought to "stop" automated access or scraping on its platform. In fact, X actively *encouraged* it. Since Mr. Musk's take-over of the platform, X has shifted gears, as it seeks – not to "safeguard" the platform – but to improperly acquire and monetize an information monopoly over data it does not own.

To the extent not admitted, Bright Data denies the allegations of this Paragraph.

2. Bright Data purports to be the world's largest provider of illicit data scraping. It markets itself as the "leading proxy provider" allowing users around the globe to circumvent technological measures that platforms like X deploy to prevent scraping. It is a substantial source of all data scraping that takes place on X's platform. And it drives massive amounts of scraping activity even as it goes to great lengths to conceal its involvement and prevent X from tracking its activities.

Answer: Bright Data is a leading proxy provider, with over 5,000 patent claims issued by the United States government protecting its innovative technology. Such technology is far from "illicit." Bright Data's proxy service has many uses, unrelated to scraping, though some of its services can also be used for scraping publicly-available information consistent with all applicable laws and regulations. Bright Data's technologies do not allow users to scrape information behind a log-in screen, nor do they circumvent technological measures that restrict access to such non-public data. Nor is it true that Bright Data is a "substantial source of all data scraping that takes place on X's platform" or that Bright Data "drives massive amounts of scraping activity." While Bright Data does not know how much scraping takes place on X, Bright Data's scraping is an infinitesimal portion of X's own estimates of the amount of such scraping. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

3. X implements stringent technological measures to stop scraping by Bright Data, Bright Data's customers, and others. Those measures restrict much of the content on X's platform to logged-in users, shielding that content from the public writ large. Data scrapers can only access the content they want through logged-in accounts, so scrapers use fake, automatically created accounts to obtain it. Even if Bright Data does not itself scrape such non-public data, it knowingly enables its customers to do so and markets circumvention tools to customers for that purpose.

Answer: Denied. X's platform is designed to be an open digital "public square," and so, X makes a substantial amount of the information on its platform publicly available without a log-in. Many data customers "want" this data, and so, data scrapers also "want" to scrape such publicly-available data. They do not need fake accounts (or any accounts at all) and do not need to "log in" to access or scrape such information. And when using Bright Data's scraping tools, they cannot, and may not, do so. As such, it does not appear that this allegation is properly directed to Bright Data and its customers, as opposed to the hypothetical "others." But if it is directed to Bright Data, to the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

4. Unlawful data scraping requires X to spend [redacted] buying excess server capacity to absorb the unwanted scraping requests with which Bright Data bombards its servers. Bright Data's circumvention further degrades the user experience on X's platform by spreading fake accounts and promoting spam. It threatens X user privacy by allowing real-time automated surveillance of user activity. And it deprives X of the revenue it would obtain if developers paid for legitimate access through X's Application Programming Interfaces ("APIs").

Answer: Denied. Bright Data does not know how X makes decisions about server expenditures, but it does not spend "[redacted]" to "absorb" Bright Data's requests. It doesn't even spend an extra penny. Nor do Bright Data's services "degrade[] the user experience" on X. To the contrary, Bright Data's process is reliant on ensuring that the host website functions exactly as intended, with no degradation, delay, or performance issues. Bright Data does not use fake accounts (or any accounts) when scraping websites and does not "spread[] fake accounts" or "promot[e] spam." Nor does Bright Data threaten user privacy; Bright Data only scrapes public information. Indeed, X threatens its users' privacy to an exponentially greater degree by selling access to information that users deem private, and by preventing users from opting out of the use of their private information for many purposes, including X's A.I. Chatbots. Bright Data also has not deprived X of revenue. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

5. Bright Data's business depends on circumventing the technological measures X employs to protect its platform and users. That circumvention violates X Corp.'s Terms of Service, by which Bright Data is bound, and induces Bright Data's customers to breach their own similar agreements with X Corp. Bright Data even proudly advertises that it "avoids" and "bypasses" the technological measures X has deployed to protect and control access to data on the X platform – in

direct contravention of the Digital Millennium Copyright Act ("DMCA"), Computer Fraud and Abuse Act ("CFAA"), and California's Computer Data Access and Fraud Act ("CDAFA").

Answer: Denied. Bright Data allows its customers to access publicly-available information, and does not circumvent technological measures that restrict such access. Bright Data's services do not violate X's Terms, which do not apply to Bright Data. Nor does Bright Data induce its customers to violate X's Terms, to the extent such Terms remain enforceable after this Court voided the Terms' scraping-related provisions. Despite X's attempt to create a contrary impression through its misquoting of Bright Data's product descriptions, Bright Data's services comply with all applicable laws and regulations. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

6. Bright Data is a sophisticated actor that is aware its activities violate X Corp.'s Terms of Service, Privacy Policy, and the Rules and Policies (together, the "Terms"), because the company and its executives are (or were) registered X account holders who agreed to abide by those Terms and thus have full knowledge of them.

Answer: Bright Data offers services that comply with all applicable laws and regulations. As X impliedly notes, Bright Data is no longer an X account holder, and is not bound by the Terms. But Bright Data does not deny that – as of the filing of this lawsuit – it was aware of X's Terms. Without further investigation, Bright Data does not know when it first learned of the provisions in X's Terms but denies that the mere fact that it is a business, that it had an account it used for limited advertising purposes (and not for scraping), or that certain of its employees may have had their own personal accounts, establishes actual, implied, or imputed knowledge of the Terms. In any event, as noted above, Bright Data's services do not violate X's Terms. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

7. X Corp. brings this action for injunctive relief to halt Bright Data's unauthorized use of X Corp.'s platform and other unlawful conduct and for damages caused by Bright Data's breach.

Answer: This paragraph is definitional and does not require a response. To the extent it does, it is denied.

THE PARTIES

8. Plaintiff X Corp. is a privately held corporation duly organized and existing under the laws of the State of Nevada with its principal place of business at 1355 Market Street, Suite 900, San

Francisco, California, 94103. X Corp. owns and operates the social media platform X, formerly known as Twitter.

Answer: Bright Data admits that X operates X. Bright Data does not know how X is structured. But the allegations in this Paragraph appear to be incorrect or outdated. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

9. On information and belief, Defendant Bright Data was incorporated in Israel in 2008 as Zon Networks Ltd. and changed its name to Bright Data Ltd. in 2021. Bright Data has its principal place of business at 4 Hamahshev St., Netanya 4250714, in Israel. Bright Data has at times maintained an office at L415 Mission Street, 37th Floor, in San Francisco, California.

Answer: Bright Data admits that it is incorporated and has its principal place of business in Israel. Bright Data does not have an office in California, nor did it maintain its own office in California during the relevant period (for a short period, it purchased two WeWork passes for temporary access to a shared office workspace in California). To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

10. Defendant Bright Data operates brightdata.com, where it sells data scraped from numerous websites and social media platforms, including X, along with tools and services to scrape data from X and other platforms.

Answer: Bright Data operates Brightdata.com. The website describes Bright Data's offerings of products and services, including its standardized datasets of public internet information. One such dataset includes public information hosted on X. Sales of such data have been *de minimis*. Bright Data also operates a general proxy network, which customers can use for approved purposes and use cases. In addition, Bright Data offers scraping tools that can be used to scrape publicly-available information, including public information on X. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

JURISDICTION AND VENUE

11. This Court has jurisdiction over this action under 28 U.S.C. § 1332 because complete diversity exists, and the amount in controversy exceeds $75,000. Plaintiff X Corp. is incorporated in Nevada with its principal place of business in California. Defendant Bright Data is incorporated in Israel with its principal place of business in Israel.

Answer: Bright Data does not contest the Court's subject matter jurisdiction over this action at this time. Bright Data admits it is incorporated in Israel with its principal place of business in Israel, and that X is incorporated in Nevada. Bright Data does not know if X has a principal place of business in California, but X does business everywhere and interacts with Bright Data's communications network outside of California. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

12. This Court has personal jurisdiction over Defendant because Defendant has consented to X Corp.'s Terms, which require all disputes related to the Terms be brought in the federal or state courts located in San Francisco, California. As part of its agreement to those Terms, Defendant also consented to personal jurisdiction in California.

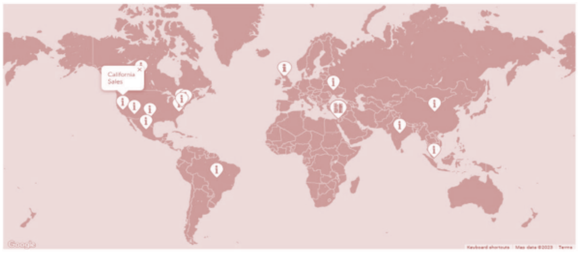
Answer: Denied.

13. Additionally, this Court has personal jurisdiction over Defendant because Defendant knowingly directed prohibited conduct to California and California residents. Defendant offers its data sets and scraping tools for sale in California and to California residents, and has targeted X Corp., which has its principal place of business in California, as well as X Corp.'s users located in California.

Answer: Bright Data sells its services globally. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

14. Defendant markets and sells its products to California residents and businesses via a sales office in California, according to its website:

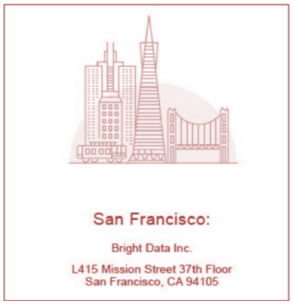
Figure 1: Screenshot of Bright Data's website on November 14, 2023



Answer: Denied. Bright Data does not have a sales office in California. Bright Data sells its services globally and does not make or market products based on customer location. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

15. As recently as October 19, 2022, Defendant encouraged customers to contact Bright Data at its California sales office, as shown in Figure 2.

Figure 2: Screenshot from Bright Data's "Contact Us" page on October 19, 2022



San Francisco:

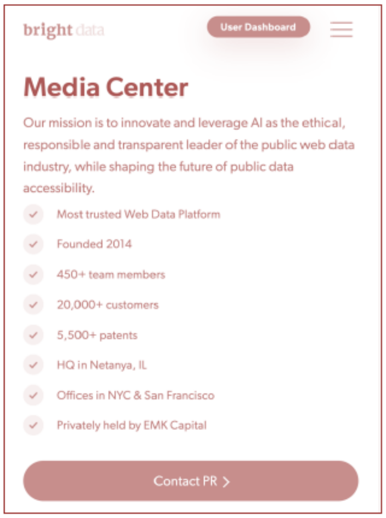
Bright Data Inc.

L415 Mission Street 37th Floor
San Francisco, CA 94105

<u>Answer:</u> Bright Data does not have a sales office in California. Bright Data admits that it had an outdated reference to a WeWork address on its website in 2022. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

16. Members of Defendant's business development and sales team are also located in California. For example, Defendant's Chief Revenue Officer, who oversees Bright Data's sales operations, is based in the San Francisco Bay Area. Defendant's Global Head of Presales is also based in California, along with numerous other Bright Data employees. To this day, Defendant advertises its San Francisco presence to clients, as shown in Figure 3.

Figure 3: Screenshot from Bright Data's "Media Center" page on August 11, 2024



Media Center

Our mission is to innovate and leverage AI as the ethical, responsible and transparent leader of the public web data industry, while shaping the future of public data accessibility.

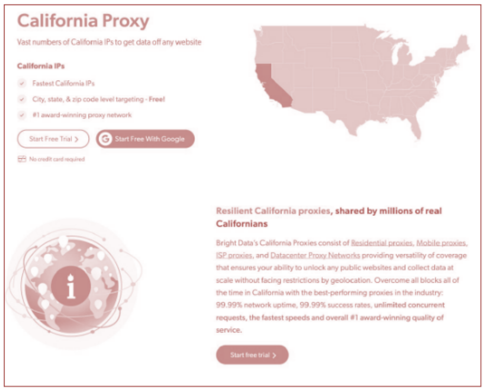
- ✓ Most trusted Web Data Platform
- ✓ Founded 2014
- ✓ 450+ team members
- ✓ 20,000+ customers
- ✓ 5,500+ patents
- ✓ HQ in Netanya, IL
- ✓ Offices in NYC & San Francisco
- ✓ Privately held by EMK Capital

Bright Data, *Media Center*, https://perma.cc/Q9YW-V3C4.

<u>Answer:</u> Bright Data admits that two of its employees have a home in California. Those individuals perform their duties on a global basis, with no distinction based on the location of their home residences. Bright Data denies that "numerous other Bright Data employees" are based in California. As to Figure 3, the excerpted page of Bright Data's website is outdated and references a WeWork facility to which Bright Data no longer has access. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

17. Defendant also targets its products at the California market. For example, Defendant's interactive website, through which California residents can purchase Defendant's scraping tools and scraped data sets, offers a "California Proxy" product that promises "[v]ast numbers of California IPs to get data off any website."

Figure 4: Screenshot from Bright Data's website on August 11, 2024



California Proxy

Vast numbers of California IPs to get data off any website

California IPs

- ✓ Fastest California IPs
- ✓ City, state, & zip code level targeting - Free!
- ✓ #1 award-winning proxy network

Resilient California proxies, **shared by millions of real Californians**

Bright Data's California Proxies consist of Residential proxies, Mobile proxies, ISP proxies, and Datacenter Proxy Networks providing versatility of coverage that ensures your ability to unlock any public websites and collect data at scale without facing restrictions by geolocation. Overcome all blocks all of the time in California with the best-performing proxies in the industry: 99.99% network uptime, 99.99% success rates, unlimited concurrent requests, the fastest speeds and overall #1 award-winning quality of service.

These proxy IP addresses are designed to evade usage restrictions and anti-scraping technology, such as those implemented by X. In fact, Defendant advertises that its California proxies allow users to "[o]vercome all blocks all of the time in California." Bright Data, *California Proxy*, https://perma.cc/L7GE-CXUP.

<u>Answer:</u> Denied.

18. On information and belief, Defendant has sold its scraping tools, scraped data sets, and IP proxies to X users, including X users in California, and has scraped data from X Corp.'s servers in California.

<u>Answer:</u> Bright Data does not know which, if any, of its customers are X users. Nor does Bright Data know where X locates its servers. Bright Data denies that it targets X users in California or elsewhere. As noted above, Bright Data's network is indifferent to the location of both customer and website operator. Moreover, Bright Data does not make a state-level proxy selection when it or its customers scrape X. As such, Bright Data never directs search traffic to X's California servers. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

19. Venue is proper in this district under 28 U.S.C. § 1391, because a substantial part of the events or omissions giving rise to the claims occurred in this judicial district. During all relevant times, Defendant repeatedly, knowingly, and intentionally targeted its wrongful acts at X Corp., which has its principal place of business in this district. Defendant also, on information and belief, sold its scraping tools, scraped data sets, and IP proxies to residents of this district, including through Defendant's sales office located in this district and employees located in this district.

Answer: Denied.

20. Pursuant to Civil L.R. 3-2(d), this case may be assigned to either the San Francisco or Oakland division because X Corp. is located in San Francisco County.

Answer: Denied.

FACTUAL ALLEGATIONS

A. X Corp.'s Platform and Terms of Service

21. Plaintiff X Corp. owns and operates the social media platform X, accessible through twitter.com, X.com, and various mobile and online applications.

Answer: Admitted.

22. X serves multiple audiences, including users and developers. For its users, X contributes to the public conversation. X serves its users content on a variety of interests and topics and allows its users to engage with and post their own content. For its developers, X provides automated and programmatic access to the content on X so that businesses, researchers, and developers have real-time access to the global conversation subject to safeguards that protect the platform's integrity and the X's user experience. The following paragraphs address each in turn.

Answer: X operates the X platform, which hosts public, user-generated content concerning various topics. Only account holders can post content. Visitors, unregistered users, and members of the public can view most of this content. X has made data available through its APIs, which, at times, were available for free to the public. Bright Data denies that X's purported contractual restraints on "programmatic access" were designed to "protect the platform's integrity and the … user experience," as such restraints were designed only to protect X's information monopoly. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

1. *X's Services for Users & Terms of Service*

23. The X user platform has hundreds of millions of active users worldwide. More than 23 million X accounts have been registered from California.

Answer: Bright Data admits that X has millions of users worldwide. Bright Data does not know the precise number of X's active users or the number of X accounts registered from California. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

24. To create a forum for a global conversation about "what's happening," X allows its registered users to post and share content, including written comments, images and videos, known as "posts," and to share, like, and comment on other users' posts.

Answer: Bright Data admits that X purports to provide a "forum for a global conversation about what's happening." Bright Data further admits that X allows its registered users to post and share content, including comments, images, and videos, and to interact with others. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

25. To post content on X or to re-post, like, or otherwise interact with posts by others, users must register for an account and log in to that account. As of July 2023, X has also strictly limited the access that individuals (or bots) have when not logged into a registered account. To gain full access to the platform, an individual must be signed into an account.

Answer: Bright Data admits that an account is needed to post, re-post, like, comment, and "gain full access" to content on X. However, an account is *not* needed to access publicly available information on the X website – namely, information X's users choose to make public, and X chooses not to place behind a log-in screen. As X itself has explained, this is information that is "visible to anyone, whether or not they have a X account." X makes substantial amounts of information available for unrestricted view and use to visitors who are logged out, including virtually all information that X's registered users have placed in the public domain. X does not attempt to restrict access to such information. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

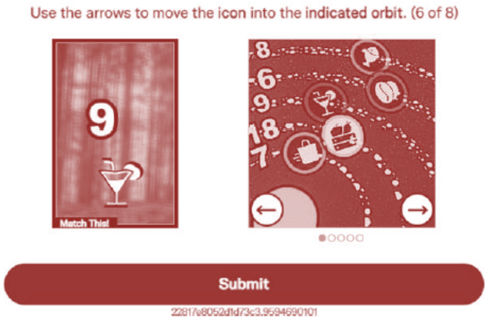
26. If an unregistered user attempts to access the platform through an app, that user is prompted to create an account (which requires agreement to the Terms) and cannot use the app until they do so. If an unregistered user visits the X homepage at x.com or twitter.com, the user is invited to create an account or sign in.

Answer: Bright Data admits that X makes substantial amounts of information available on its website for unrestricted view and use to visitors who are logged out, including virtually all information that X's registered users have placed in the public domain. Bright Data does not know the requirements to use X's app nor does it know which visitors X invites to create an account or to sign in. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

27. To register for an account, users must provide their name, phone number or email address, and date of birth. To prevent automated services from creating numerous fake accounts, X imposes a "CAPTCHA" process to ensure that a human (rather than an automated process) is creating the account. CAPTCHA stands for "Completely Automated Public Turing test to tell Computers and Humans Apart." X then verifies registrants through email or phone confirmation.

Answer: Bright Data does not know what procedures and information X requires to register for an account. Bright Data does not use an X account when engaging in or offering scraping services. Bright Data notes that X allows multiple accounts to be associated with a single phone number and does not require that accounts correspond to the account holder's real name (*i.e.,* a user can use a pseudonym or a display name that is not his or her legal name). To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

28. A picture of X's CAPTCHA is below. To finish the registration, the prospective user must supply the information requested by the CAPTCHA:



Use the arrows to move the icon into the indicated orbit. (6 of 8)

Answer: Bright Data does not know what procedures and information X requires to register for an account because Bright Data does not use an X account when engaging in or offering scraping services. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

29. The registration process also requires express agreement to X Corp.'s Terms. The Terms state that a user may not "access, tamper with, or use non-public areas of the Services, our computer systems, or the technical delivery systems of our providers" or "breach or circumvent any security or authorization measures."

Answer: Bright Data does not know what is required during X's account registration process. Bright Data does not use an X account in engaging in or offering scraping services. Bright Data denies that it is a user or that it is otherwise bound by X's Terms. In any event, Bright Data has not accessed or scraped non-public information on X's platform; nor has it breached or circumvented any security or authorization measures while accessing or using the Services (as that term is used in the Terms). Indeed, Bright Data denies that it uses the Services at all. Bright Data notes that X omits substantive language from the sentence of the Terms it purports to quote. What the Terms actually prohibit, to whom they apply, and whether they are enforceable, are each legal conclusions that may depend on the facts, including facts not alleged here. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

30. X Corp.'s Terms also state a user may not "access or search or attempt to access or search the Services by any means (automated or otherwise) other than through our currently available, published interfaces that are provided by us (and only pursuant to the applicable terms and conditions), unless you have been specifically allowed to do so in a separate agreement."

Answer: Bright Data denies that it is a user or that it is otherwise bound by X's Terms. In any event, Bright Data only accesses or searches X through published interfaces that X makes publicly available. Bright Data denies that it uses the Services at all. Bright Data notes that X omits substantive language from the sentence of the Terms it purports to quote. What the Terms actually prohibit, to whom they apply, and whether they are enforceable, are each legal conclusions that may depend on the facts, including facts not alleged here. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

31. In addition, X Corp.'s Terms specifically state that "crawling or scraping the Services in any form, for any purpose without our prior written consent is expressly prohibited."

Answer: Denied. X's allegation constitutes fraud on the Court, or at least negligence, because X ***fails to disclose*** that the quoted provision does not appear in the operative Terms (*i.e.*, those that existed at the time the original complaint was filed). Instead, this provision first appears in the September 29, 2023 version of the Terms, which post-dated the filing of this case.

The Court specifically ruled on May 9, 2024 that theories of breach based on post-suit amendments to the Terms are not viable, and declined to further analyze such claims. ECF 83 at 25-26 ("X Corp. has not cited any case, and our research has revealed none, where a party was permitted unilaterally to amend a contract midway through litigation concerning that contract."); *see also* ECF 156 at 3 ("[T]he order did not engage X Corp.'s contract breach theories based on changes X Corp. made to X's Terms after filing suit; X Corp. did not cite caselaw or mount argument for why the claims emanating from the revised contract should be in this suit."). As such, X was under an obligation, at a minimum, to disclose to the Court that it was relying on terms that post-dated the original Complaint.

The operative Terms do not prohibit crawling. Specifically, the May 18, 2023 Terms state that "crawling the Services ***is permissible*** if done in accordance with the provisions of the robots.txt file." With respect to scraping, the Court has already voided those provisions as being pre-empted by the Copyright Act.

In any event, Bright Data denies that it is bound by X's Terms, or that it uses the Services at all. Moreover, what the post-suit Terms actually prohibit, to whom they apply, and whether they are enforceable, are each legal conclusions that may depend on the facts, including facts not alleged here.

To the extent not expressly admitted, Bright Data denies the allegations in this Paragraph.

32. Under the Terms, users may not "forge any TCP/IP packet header or any part of the header information in any email or posting, or in any way use the Services to send altered, deceptive or false source-identifying information."

<u>Answer:</u> Bright Data does "not forge any TCP/IP packet header or any part of the header information in any email or posting," nor does it "use the Services" at all. Bright Data denies that it is a user or that it is otherwise bound by X's Terms. In any event, what the Terms actually prohibit, to whom they apply, and whether they are enforceable, are each legal conclusions that may depend on the facts, including facts not alleged here. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

33. Users are also prohibited under the Terms from any conduct that would "interfere with, or disrupt, (or attempt to do so), the access of any user, host or network, including … overloading, flooding, spamming … or by scripting the creation of Content in such a manner as to interfere with or create an undue burden on the Services."

<u>Answer:</u> Bright Data does not do any of the things identified in this Paragraph. Bright Data has not "interfere[d] with or create[d] an undue burden on" X's systems. Nor has Bright Data "interfere[d] with, or disrupt[ed], (or attempt[ed] to do so), the access of any user, host or network" to X's systems. Bright Data also denies that it is bound by the Terms, or that it uses the Services at all. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

34. The Terms also incorporate by reference X Corp.'s Platform Manipulation and Spam Policy (the "Policy"), which prohibits "coordinated harmful activity that encourages or promotes

behavior which violates [X Corp.'s] Rules." The Policy also prohibits "leveraging X's open source code to circumvent remediations or platform defenses."

Answer: Denied. The operative Terms (*i.e.,* those in effect on the date this suit was filed) do not incorporate by reference X's Platform Manipulation and Spam Policy. In any event, the quoted language is irrelevant, as there is no allegation that Bright Data engages in "coordinated" activity, or that it "leverage[es] X's open source code." To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

35. The Terms prohibit selling any content collected from the platform. Users may not "reproduce, modify, create derivative works, distribute, sell, transfer, publicly display, publicly perform, transmit, or otherwise use the Services or Content on the Services" unless otherwise authorized by the Terms or a developer agreement.

Answer: Denied. X improperly crops the actual quote from the Terms. The Terms expressly permit all such activities so long as the information is obtained through the "interface and the instructions [X] provides." Nor do the Terms prohibit selling "any content collected from the platform." In any event, this Court has already voided the Terms to the extent they purport to contractually prohibit the copying, distribution, or sale of information on X's platform. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

36. The Privacy Policy allows X users to choose their privacy settings from a menu of options giving them the ability to withdraw their consent for data sharing or to choose what content is publicly shared. X users may choose to limit access to their content so that only specified individuals – not all X users – may view their content. They may adjust their privacy settings whenever they wish.

Answer: Bright Data admits that X's users can generally choose what content to make public, and that X can, in turn, choose what information to place behind a log-in screen. But, in reality, X has repeatedly prevented users from protecting their information for commercial purposes. For example, it recently went so far as to *affirmatively* permit blocked accounts to view a user's posts, creating an unsafe environment for its users. Similarly, under X's most recent November 15, 2024 Terms, X eliminated any ability for users to opt out of consenting to the use of public and private information by X's A.I. Chatbot. Put simply, X does not give users the "ability to withdraw their consent for data sharing."

X's Terms further make clear that X does not grant users the right to limit distribution of their content, public or private. The Terms, for example, inform users that "by submitting, posting, or displaying Content" on X, X may "use, copy, reproduce, … publish, transmit, display, and distribute such Content," and does not limit its ability to do so based on any privacy setting. X's Privacy Policy also tells users that once information enters the public domain, the public is free to use such information without restriction, explaining that businesses that receive data through X's APIs "are not affiliated with X, and their offerings may not reflect updates you make on X."

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

37. X also gives its users the option to delete their posts or their interactions with other users' posts.

Answer: Bright Data does not know what options X gives users to delete their content. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

38. X also allows users subject to the jurisdiction of, for example, the California Consumer Privacy Act or the European Union's General Data Protection Regulation, to exercise their privacy rights under that act and regulation, including by making certain deletion requests.

Answer: Bright Data does not know what specific privacy protections X affords to users, or how those might change based on where a user lives. But as this allegation shows, X does not care about protecting user's privacy beyond the bare minimum required by law. If it did, it would afford the same protections to, say, customers in Oregon that it affords to customers in California. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

2. *X's Services for Developers & Developer Agreement*

39. The data available on X – including the way X Corp. has organized it on its platform – has tremendous value to many parties. The data includes both user-generated content (like user posts and profiles) and non-user-generated content (like follower lists and other information about the relationships between users). The latter type of data typically reflects insufficient originality to warrant copyright protection. But just as important as the data available on the X platform is how the platform organizes it. There are more than 500 million posts on X per day. Determining which posts to show to which users at which times, and on which parts of the website or app, is a significant part of the platform's value. Beyond individual user posts, the *aggregate* data set across X's user base – amalgamating different posts or user interactions to discern macro trends – also delivers unique value to advertisers and developers. X users may have a copyright interest in the individual content they post to X (though they give X a broad license to that content), but they have no copyright interest in much of the most valuable data available on X – including the non-user generated content, the organization of that content, and the aggregate data across X's platform.

Answer: Bright Data admits that there are millions of posts on X per day and that many parties might find publicly available data posted on X to be of interest for a variety of purposes. Bright Data denies that how the platform organizes the data it hosts is "just as important as the data" itself. Bright Data further denies that non-user generated content, the organization of that content, and the aggregate data across X's platform is "the most valuable data available on X." Bright Data does not know if such data is valuable at all. In any event, Bright Data does not copy X's organization of content. Nor does Bright Data infringe X's (or anyone else's) intellectual property rights, and indeed, no such claim has been asserted. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

40. Businesses, researchers, and others seek the data available on X for a variety of purposes, including measuring user sentiment toward various products or events, gauging market reactions to current events, more effectively tailoring advertisements, and more.

Answer: Bright Data admits that businesses, researchers, and others might find public content on X to be of interest for a variety of purposes. X inappropriately monopolizes such data by attempting to enforce unlawful contractual provisions prohibiting the scraping of user public data. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

41. The data is primarily valuable when it can be analyzed at scale to discern trends, gauge market reactions, or refine ad campaigns. Data scrapers are not primarily interested in individual posts in which the user may hold a copyright. That is because any individual piece of content in isolation – like a user-generated video – has limited value to businesses looking to discern broader trends in the data (though, of course, users can and do enjoy individual posts). For that reason, scrapers like Bright Data typically target everything, siphoning off data from X at massive scale.

Answer: Bright Data has developed best-in-class technologies to provide solutions and access to public data, retrieving and synthesizing information that, though theoretically accessible by anyone with internet access, is unusable simply because of its sheer volume. But Bright Data denies that it targets "everything" on X. When Bright Data scrapes a website domain, it focuses on specific pieces or types of public information, and, through its own creative efforts, Bright Data creates a product offering. In some cases, customers find Bright Data's resulting datasets to be of value, and in other cases, Bright Data has been unsuccessful in finding customers for its offerings. In the case of X, Bright Data's sales of X-related datasets have been *de minimis*. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

42. X has copyrights to its website and app, which Bright Data accesses when it scrapes data in circumvention of X's technological safeguards. By contrast, this complaint disclaims any

exclusive copyright in X's users' posts, and X does not seek to enforce any copyright its users retain in their own creative works. Indeed, X acknowledges that its users retain the right to sell or license their posts to others, including Bright Data, just as they retain the right to exclude others from exploiting those posts. But Bright Data may not scrape those posts from X's platform – without consent from X or its users – and sell them as part of the massive data packages it markets.

Answer: Bright Data admits that X does not have any copyright interest in user-generated content. Bright Data does not know if X possesses any copyrights in its website or app, or whether such rights are enforceable. Regardless, X does not assert any copyright claim, or any infringement of such rights. And X's assertion – that "Bright Data may not scrape [user posts] from X's platform" – is just a legal contention, which this Court has already rejected when it dismissed X's scraping claims with prejudice. ECF 156 at 21. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

43. X has historically made APIs available to developers to allow them to programmatically interact with X in a controlled, secure environment. This is known as X's Developer Platform.

Answer: X has made information available through its API, often for free, to efficiently disseminate such information to a broad audience. The API was just an additional outlet by which X efficiently distributes the same information that it distributed through the web (X also employs other efficient methods of distribution, such as its website and mobile applications). Bright Data does not know X's historic practices with respect to APIs. But the recent limits X placed on its Developer Platform were not established to create a "controlled, secure environment" for the benefit of its users, but rather to exploit its information monopoly. To the extent not expressly admitted, Bright Data denies the allegations of this paragraph.

44. X launched the current iteration of its Developer Platform in February 2023. It has four tiers of access: Free, Basic, Pro, and Enterprise.

Answer: Bright Data does not know when X launched the current iteration of its API. Bright Data admits that, in an attempt to monetize user data that it doesn't own, X now sells API products. Bright Data admits that X's API offerings include "Free," Basic, Pro, and Enterprise categories of access. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

45. The Free tier does not allow developers to obtain data from the X platform. This tier allows developers only to post to X and test the API.

Answer: For much of the relevant period, X made significant amounts of public information available for free through its API Developer Program. As part of X's effort to exploit its information monopoly, X shut off free access through some of its APIs. Bright Data does not know what rights X gives developers under its "Free" tier. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

46. The Basic and Pro tiers are for hobbyists, researchers, and smaller businesses testing out the API. These tiers cost $100 per month (for Basic) or $5,000 per month (for Pro). Both levels impose rate limits that limit the amount of data that can be obtained over a set amount of time (e.g., approximately 900 posts per 15 minutes for Pro), and likewise impose monthly caps on data that can be obtained (e.g., one million posts per month for Pro).

Answer: Bright Data does not know how much X charges, what rate limits X imposes, or who uses the various tiers of X's API, but denies that the Basic or Pro API products are sufficient for the needs of researchers. In response to a questionnaire fielded by the Coalition for Independent Technology Research, public interest researchers listed over 250 projects that would be jeopardized by ending free and low-cost API access, including research into the spread of harmful content, (dis)information flows, crisis informatics, news consumption, public health, elections, and political behavior. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

47. The Enterprise tier is for business and scaled commercial projects that need access to data on X's platform at great scale and in real-time. This tier is ideal for large businesses tracking consumer sentiment, financial companies analyzing market trends, or advertisers. The usage and rate limits are far higher. The entry level fees start at [redacted] and rise depending on the usage and rate limits, but more importantly, the use case and industry potentially served.

Answer: Bright Data denies that X's Enterprise tier is ideal for large businesses but admits that the Enterprise tier represents X's attempt to extract maximum monopolistic prices from commercial customers. Since Elon Musk acquired Twitter, X has consistently raised prices or added additional fees, but Bright Data does not know the specifics of what fees X charges. Bright Data also does not know how such prices change based on volume. In any event, Bright Data denies that the rate limits for the Enterprise Tier are "far higher," or that they are high at all. Indeed, X historically provided researchers with low-cost access to its Decahose, a real-time sample of 10% of all tweets. As of March 2023, that equated to roughly a billion tweets per month; the most expensive Enterprise tier would cut that by 80% at about 400 times the price. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

48. Because of the large amount of available data, X requires that developers seeking access to the Enterprise tier submit their proposed use cases so that X can ensure those uses are consistent with a healthy platform and do not undermine the interests of X's users. X has rejected several requests for access to Enterprise tier – forgoing the associated revenue – where the proposed use cases would have been bad for X users, including by infringing on X users' privacy.

Answer: Bright Data does not know the requirements to access X's Enterprise API offering but denies that any such requirements are rooted in a desire for a "healthy platform" or user interests. Instead, X offers its API products for sale in order to monopolize and monetize user data that it does not own. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

49. No API tier allows developers unrestricted access to all data on X. For example, X does not allow any API access originating from IPs in certain high-risk jurisdictions, such as those that might promote terrorism. Instead, X's API blacklists IPs from those geographic locations, with the goal of preventing malign actors from gaining unfettered access to X's data. X also does not allow API access to any governments, to reduce the risk of government surveillance of X's users.

Answer: Bright Data does not know what restrictions, if any, X places on its API tiers, or its motivations for each such restriction. Many of its restrictions, however, were implemented since the Musk takeover, and such restrictions were designed to exploit X's information monopoly, not to prevent malign actors from misusing the data. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

50. Not all X user content, including some X user content which is publicly accessible on the user side of the X platform, is available through the API. This includes, for example, certain information related to user's specific geographical locations and other information detailed below.

Answer: Bright Data does not know what information is available through X's API, whether the API and the website contain the same or different public information, or the degree of overlap between the two methods of access. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

51. X Corp. requires each developer using its API – at any tier – to agree to X's Developer Agreement (last updated Nov. 14, 2023), https://perma.cc/2YCC-E6E7. That agreement imposes the following restrictions:

a. X developers must obtain X user consent before sharing an individual user's content to promote a product or service, and before storing or sharing non-public or confidential X user information.

b. X developers may not circumvent user blocking or account protections.

c. X developers must delete from their databases any content that is deleted on X, whether deleted by a user or deleted by X for violations of the Terms or applicable laws – for example, revenge-porn content. As part of this, developers must delete X user content after an X user deactivates or deletes their X account.

d. X developers must modify any data modified on X, including when content is made private or deleted.

e. X developers must not collect X user geodata on a standalone basis, barring them from engaging in user tracking, activity heat maps, or similar activities.

f. X developers must not use the X data to create spam, X bot accounts, or automate processes on the X user side such as bulk X user following.

g. X developers must not use X data to infer certain protected characteristics of X users which X does not share with developers even if an X user publicly posts this content on X's user platform. These protected characteristics include: health (including pregnancy), negative financial status or condition, political affiliations or belief, racial or ethnic origin, religious or philosophical affiliation or beliefs, sex life or sexual orientation, trade union membership, and whether the user has actually or is alleged to have committed a crime.

h. X developers must comply with X user requests which X forwards to them under applicable privacy laws, including the California Consumer Privacy Act and General Data Protection Regulation.

i. X developers must not attempt to match X content, usernames, or accounts with a person, household, device, browser, or other off-X identifier without the user's express consent, unless the information was provided by the user or is otherwise publicly available (i.e., for public figures).

j. X developers must not use acquired data for tracking or targeting sensitive groups, such as political activists or dissidents, performing background checks or personal vetting, credit or insurance risk analysis, individual profiling or psychographic segmentation, or the development of facial recognition software.

Answer: Bright Data does not know what terms govern subscribers to X's Developer Program. Bright Data denies, however, that the cited document contains all of the alleged restrictions in this Paragraph. In any event, Bright Data has a roughly comparable Acceptable Use Policy. To the extent X's policies differ from Bright Data's, such differences do not relate to the use or misuse of data but are driven primarily by X's desire to maintain or exploit its information monopoly for commercial purposes. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

3. *Balancing User, Developer, and Advertiser Interests*

52. To balance the complex, interrelated nature of its multi-sided platform business and simultaneously provide a forum for global public conversation on "what's happening," X must be able to credibly enforce its Terms on its users, developers, and advertisers. For example, businesses (including advertisers) may want greater access to data on X for the purpose of targeting specific consumers, but some of X's users may be sensitive to such practice. X Corp.'s Terms restrict such targeting to protect its users' interests. To take another example, some businesses may want unrestricted API access so they can send solicitations to users by direct message or offer to sell increased "follower counts," but X bars such practices because they degrade the quality of X for almost all users. And finally, some governments may seek unrestricted API access to influence or monitor public opinion – again, a practice that disserves X's users.

Answer: Denied. X does not use restrictions on access or scraping to manage the "multi-sided" nature of its platform; it uses the Terms to wrest control over, and extract monopoly fees, for data it does not own. To achieve this goal, X must create "credibl[e]" threats of legal liability, regardless of whether such Terms are legally enforceable.

Bright Data denies that X adopted or uses its purported restrictions on automated access or scraping to balance advertisers' "want[s]" and users' "sensitiv[ities]." Indeed, X's allegations are inconsistent with X's Terms and Privacy Policy, which clearly contemplate that X will "share[] information" with advertisers and other third-parties to show users more "relevant" ads, help them discover "affiliates, third-party apps, and services," and to train X's own machine learning and artificial intelligence models. As such, X affirmatively targets ads to users despite "user sensitiv[ities]."

X's allegation is further belied by the fact that X sells access to the very same data through its API service that it now claims users believe to be too "sensitive" to permit such data to be used for commercial purposes. For similar reasons, Bright Data further denies X's allegations about follower counts, as X is perfectly happy – and obviously does not believe it "degrades the user experience" – to provide this publicly available information to businesses *for a fee* through its API offerings.

Bright Data also denies that X bars automated access or scraping to prevent governments from influencing or monitoring public opinion, or that X believes that such practices disserve X's users. Bright Data does not know what API access, if any, is sought by what government entities, or the reasons for such requests. However, Bright Data denies that X is interested in preventing the influencing of elections. Indeed, Musk himself has turned X into a political megaphone, and has used the platform to influence the last presidential election. As *Fortune* reported, Mr. Musk "has turned his personal feed into a non-stop pro-Trump megaphone, sharing MAGA talking points to his roughly 203 million followers on X." X, obviously, did not believe that its own monitoring and use of the platform for such purposes "disserve[d] X's users."

Bright Data denies that X uses the Terms to try to balance user sensitivities. Far from using the Terms to protect users, X has demonstrated disregard for such user sensitivity when it amended its Terms on November 15, 2024 to strip users' rights to prevent their content from being used for the benefit of X's artificial intelligence business. As recent reporting shows, X has used this information to publish artificially generated images of real users' images and likenesses, without express permission.

If X really cared about users' sensitivities, safety, and well-being, it would not have gutted its content moderation policies. Nor are X's anti-scraping and automated access provisions related to any such interests, and thus, are not a means for "balancing" such interests. When X abandoned its content moderation policies, it knowingly and intentionally allowed advertisers' products to be displayed alongside false, defamatory, and hateful speech that had proliferated on the platform. When advertisers fled the platform in droves, X revised its anti-scraping and automated-access provisions to support its effort to monetize and exploit its information monopoly, not to "balance" its business interests with users' or the public's interests.

Indeed, X's Terms do not benefit users, but actually injure users (to the extent the Terms are enforceable) by forcing them to relinquish their legal right to freely search the web.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

53. X Corp. can police these problems only if it can credibly enforce its policies, including specifically on programmatic or API access to X. For example, over a recent six-month period, X has enforced its platform manipulation and spam policies by:

a. Suspending 169,396 accounts and removing 15,275 instances of conduct for violations of X's policies against sharing improper impersonation;

b. Suspending 2,563 accounts and removing 62,537 instances of conduct for violations of X's policies against sharing personally identifiable information; and

c. Suspending 119,508 accounts and removing 571,902 instances of conduct for violations of X's policies on illegal and regulated goods.

Answer: Bright Data does not know how many accounts or "instances of conduct" X suspended or "remov[ed]" during the unspecified "six-month" period referenced in this Paragraph, or why X took such actions. But none of the conduct alleged in this Paragraph has anything to do with Bright Data or scraping. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

54. X also maintains an active civic integrity policy.[1] This policy prohibits users from posting false and misleading information about how to participate in a civic process, including elections. For example, users may not post information intended to mislead or intimidate voters to dissuade them from participating in elections.

Answer: X's Civic Integrity Policy has no bearing on this case. Putting aside the fact that the policy was not incorporated by reference into the operative version of the Terms, the policy relates to posting false information to X for purposes such as interfering with elections. It does not relate to data access or data scraping, which do not involve "posting" information of any kind, and thus has nothing to do with this case. Nor is there any allegation of any breach of X's Civic Integrity Policy by Bright Data or users of its services. To the extent X's Civic Integrity Policy is relevant to this case, Bright Data denies that X prohibits users from posting misleading information about elections. Indeed, as a New York Times exposé uncovered – based on public information that X tried to bury and was only capable of being compiled with Bright Data's assistance – X affirmatively reinstated the accounts of individuals involved in election fraud, insurrection, and treason against the United States government in connection with the 2020 election. *See Musk Lifted Bans for Thousands on Twitter*, NY Times (Dec. 22, 2022), perma.cc/7ZSG-KLFD.

55. X's civic integrity policy also bars users from posting from accounts with deceptive identities. In February 2021, X disclosed it removed 373 accounts and related instances of content attributed to state-linked information operations originating from Iran, Armenia, and Russia. X again disclosed in December 2021 that it removed 3,465 accounts connected to state-linked information operations from six distinct jurisdictions: Mexico, the People's Republic of China (PRC), Russia, Tanzania, Uganda, and Venezuela. Every account and piece of content associated with these operations was permanently removed from the X platform.

Answer: X's "Civic Integrity Policy" has no bearing on this case. Putting aside the fact that

[1] X Help Center, *Civic integrity policy* (Aug. 2023), https://help.twitter.com/en/rules-and-policies/election-integrity-policy.

the policy was not incorporated by reference into the operative version of the Terms, the policy relates to posting false information to X for purposes such as interfering with elections. It does not relate to data access or data scraping, and thus, has nothing to do with the facts at issue. Nor is there any allegation of any breach of X's Civic Integrity Policy by Bright Data or users of its services. Bright Data does not know how many, if any, foreign accounts X removed under its Civic Integrity Policy or the circumstances that led to any such removal. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

56. X further maintains policies against the use of deceptive marketing or misrepresentative business practices,[2] as well as the advertising of certain high-risk financial products and certain content related to cryptocurrencies.[3]

Answer: Bright Data does not know what X's unspecified policies may or may not prohibit. But the allegation is entirely unconnected to this case, as there is no allegation of any breach of such policies by Bright Data, nor do such policies have any bearing on X's access or (now-dismissed) scraping claims. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

57. The availability of unrestricted third-party programmatic access (including scraping data) weakens X's ability to effectively prevent and deter market manipulation, scams, and fraud, increasing the risk of harm to consumers.[4] This is because the most effective way for X to police such abuses is by having a business relationship with and approving the use cases of those third parties given access to the X API. The goal of data scrapers, by contrast, is to conceal their conduct from X and sell the data they scrape without restriction.

Answer: Denied. The availability of third-party programmatic access to publicly available data hosted on X weakens X's ability to monopolize the market for such data. It does not weaken X's ability to effectively prevent and deter market manipulation, scams, and fraud, nor does it increase the

[2] X Business, *Deceptive & Fraudulent Content Policy*, https://perma.cc/XAU6-S648.

[3] X Business, *Financial products and services*, https://business.x.com/en/help/ads-policies/ads-content-policies/financial-services.html.

[4] *See, e.g.*, Press Release, SEC, *SEC Charges Eight Social Media Influencers in $100 Million Stock Manipulation Scheme Promoted on Discord and Twitter* (Dec. 14, 2022) https://www.sec.gov/news/press-release/2022-221; Press Release, U.S. Att'y's Off., N.D. Cal., *Scottish Citizen Indicted For Twitter-Based Stock Manipulation Scheme* (Nov. 5, 2015), https://perma.cc/8NPH-GN3W.

risk of harm to consumers. Any purported rationale for restricting third-party programmatic access to publicly available data for those purposes is pretextual. If X wanted to police the use of its data, it would protect the information with a password, which is "the most effective way for X to police" such conduct, as it may create legally-enforceable rights. Bright Data does not know the goals of other scrapers, but Bright Data denies that its goal is to conceal its conduct. As relevant here, Bright Data's goal is to use industry-standard internet protocols to lawfully search for and use public information. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

B. X's Battle Against Unapproved Programmatic Access Including Data Scraping

58. Many malign actors neither want to pay for API access to X nor want to comply with X's Developer Agreement. Indeed, many want programmatic access to X for purposes that X's API disallows – such as spamming X users with solicitations or disinformation or scraping vast swaths of data without regard to user privacy. Because these actors cannot acquire X's data through legitimate channels, they turn to services like Bright Data to get the job done.

Answer: Denied. That a person does not want to pay X's monopolistic prices for public data that X does not own does not make him or her a "malign actor." People that want access to public web data, including public data on X, can use Bright Data's tools for approved use cases consistent with all applicable laws and regulations, as wells as Bright Data's policies, including its Acceptable Use Policy, its Privacy Policy, and its Data Protection Addendum. Bright Data does not permit its services to facilitate spamming, disinformation, or violation of third-party privacy rights. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

59. One form of programmatic access – data scraping – is particularly harmful to X. Scraping is the process of using automated means to collect content or data from a website or app. The process involves using an automated process to make a request to a website's server or app, downloading the results, and parsing them to extract the desired data. Data scrapers typically send large volumes – in the millions or even billions – of these requests, taxing the capacity of servers and diminishing the experience for legitimate users. Data scraping is also an end-run around X Corp.'s restrictions on API access that protect the user experience on X.

Answer: Bright Data denies that programmatic access harms X. According to X's own Complaint, less than one-half of one percent of server requests on the X platform are attributable to *any form* of automated access or scraping of publicly available data. And Bright Data's own scraping would be just an infinitesimal fraction of that. As such, scraping of publicly available data is not taxing on the capacity of X's servers nor does it diminish the experience of any of X's users. Any purported concerns about "protect[ing] the user experience" are just a pretextual excuse for X's anticompetitive behavior and monopolization. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

60. X Corp. utilizes a variety of technological measures to detect and prevent automated systems – colloquially known as "bots" – from scraping data from its platform:

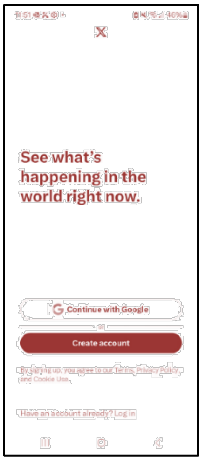
Answer: Bright Data does not know what measures X uses to obstruct the public from scraping public information on X. But such measures interfere with the public's lawful right to search for and use such information. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

1. User Login Requirements

61. Since July 2023, X does not make most of its content available to the public without logging in through a registered account and agreeing to the Terms.

Answer: Prior to July 2023, X made most of its content available to the public without a log-in. As X itself has explained, this is information that is "visible to anyone, whether or not they have a X account." For a few weeks, X made a self-described "*temporary*" change, placing a greater portion of users' content behind a log-in. Faced with user backlash and competitive threats from Threads and others, X relaxed its policies, and within a few weeks, partially reverted to making substantial amounts of information available to the public without a log-in. Bright Data does not know precisely how much information X places behind a log-in versus how much is public, but Bright Data only scrapes the latter. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

62. X's apps cannot be used without logging in. Whenever the app is opened without being logged in, the app prompts the user to login or create an account and cannot be used further until the user does so.



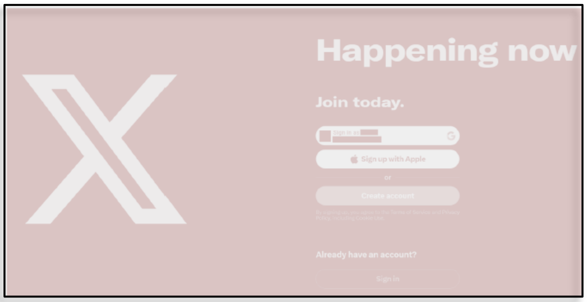
See what's happening in the world right now.

By signing up, you agree to our Terms, Privacy Policy, and Cookie Use.

Have an account already? Log in

Answer: Bright Data does not know the requirements to use X's App, nor does it know which App users X invites to create an account or to log in. But X's App is only one way that X makes information available through the internet. X also distributes the information through other methods, such as X's website, which does not require an account or log-in to access. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

63. The homepages of X's website, x.com and twitter.com, also cannot be used without logging in. If a user is not logged in, the homepage prompts the user either to login or create an account and displays no X content until the user does so.



Happening now

Join today.

By signing up, you agree to the Terms of Service and Privacy Policy, including Cookie Use.

Already have an account?

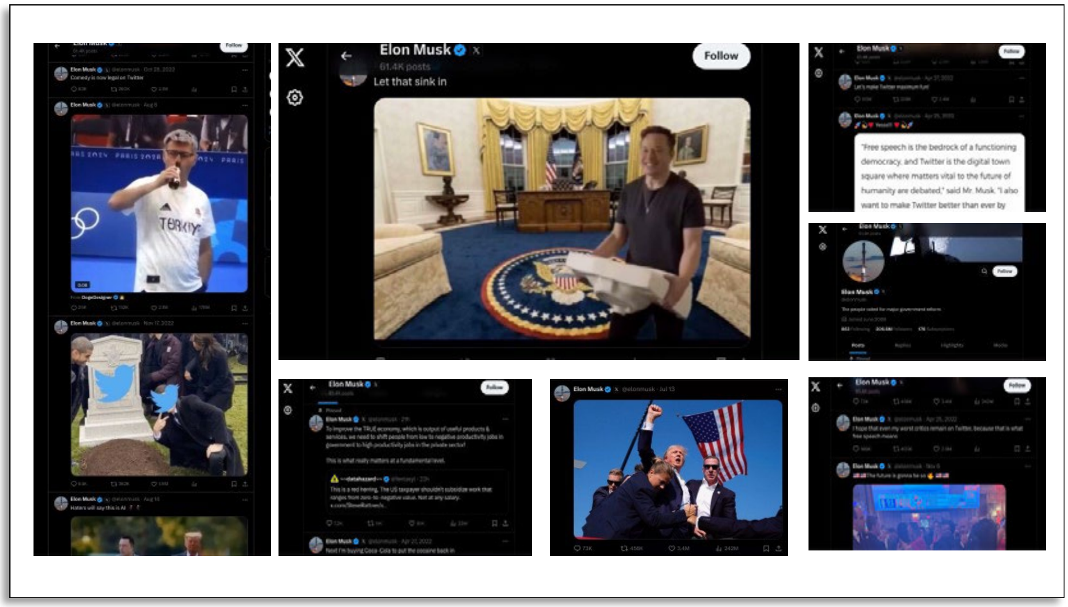
Answer: This allegation is misleading because the homepage is not the only way, and not the most common way, of accessing X.com or Twitter.com. Most, if not all, of the public information that X posts is freely available without having to circuitously navigate through X's homepage. As to X's homepage, Bright Data does not know what log-in or account creation prompts X uses, what content may be available without a log-in, or how that may have changed during the relevant period. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

64. X makes a limited amount of content available to individuals who are not logged in. Examples include users who navigate to X's website from another source, such as a search engine or an X post that is embedded within another website. X also allows Google to "crawl" certain posts on its website to index certain posts so that they are shown in Google's search results. Not all posts on X are accessible from another source.

Answer: X makes substantial, not "limited," amounts of information publicly available on its platform for unrestricted view and use to visitors who are logged out. X does not attempt to restrict access to such information. As X itself has explained, this is information that is "visible to anyone, whether or not they have a X account." X further allows search engines, such as Google (but not competing A.I. search engines, such as OpenAI's ChatGPT), to scrape this public information and publish or distribute to the world at large. Bright Data denies that X only allows Google to "crawl" and "index" X's website. Rather, Google uses automated means to copy, save, reproduce, analyze, and publish all or most of the information that X has made public. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

65. If non-logged-in individuals access a post on X in this manner, they can view that specific post and certain limited information about the post (e.g., number of likes, replies, and reposts). But their visibility is limited. The individual cannot see any replies to the post or the identities of who liked or reposted that content – thus disabling a key platform feature.

Answer: Bright Data does not know when specific information was placed behind a log-in. But prior to July 2023, most public information on X was available without a log-in. Today, a substantial amount of information is still available without a log-in, including post content, the number of followers, number of current followers, and in many cases the tweets that are being replied to. The following are screenshots of a small amount of public information available today on a *single* public page on X without a log-in:[5]



Comedy is now legal on Twitter

Haters will say this is AI

Elon Musk

61.4K posts

Let that sink in

Follow

Let's make Twitter maximum fun!

"Free speech is the bedrock of a functioning democracy, and Twitter is the digital town square where matters vital to the future of humanity are debated," said Mr. Musk. "I also want to make Twitter better than ever by

To improve the TRUE economy, which is output of useful products & services, we need to shift people from low to negative productivity jobs in government to high productivity jobs in the private sector!

This is what really matters at a fundamental level.

This is a red herring. The US taxpayer shouldn't subsidize work that ranges from zero-to-negative value. Not at any salary.

Next I'm buying Coca-Cola to put the cocaine back in

I hope that even my worst critics remain on Twitter, because that is what free speech means

The future is gonna be so 🔥

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

66. A non-logged-in individual accessing X in this manner also has a limited ability to access other content on X. The individual can access the profile of the person who posted the content but can see only a curated list of other posts from that individual (not all posts) and cannot see that user's list of followed accounts or followers. The non-logged-in individual can also click on other linked posts, which show those posts subject to the same restrictions above.

Answer: As the screenshot included above shows, users' "curated" public pages have a wealth of information. Indeed, because it is "curated," it may even have more value and significance than random posts. For example, the picture Musk posted of himself taking even the kitchen sink from the White House on the day Donald Trump won the election says something. These curated pages also include links to other curated pages and information, as X admits in this Paragraph. Other public pages

[5] @elonmusk, x.com, perma.cc/P8FL-LDGH (selected screenshots).

have individual posts and links to individual posts. X does not restrict access to this information. While X does include certain prompts to log-in pages for certain areas on these pages, such prompts are not designed to discourage scraping, but to encourage visitors to become account holders that can be monetized. But X also encourages visitors to view and use the public information on its site without a log-in. Thus, even today, X makes substantial amounts of user-generated content publicly available without a log-in. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

67. When users are logged out, X circumscribes their ability to continue accessing different profiles or linked posts. After several clicks on different profiles or posts, X prompts the user who is not logged in to either login or create an account. At that point, the user may no longer click on additional profiles or posts until a period of time has passed. Thus, the ability to access content when not logged in is both rate limited and qualitatively restricted.

Answer: As just noted, X makes substantial amounts of information available for unrestricted view and use to logged-off visitors. Bright Data does not know what rate limiters X has chosen to employ, whether such limiters are triggered by the number of "clicks" of the visitor's mouse, or whether there is any "qualitative" component to the operation of X's secret rate limits. X's Terms do not disclose the existence of any rate limits, to what they apply, or how they operate. Nor do X's Terms require any agreement to such secret, undisclosed, and ostensibly ever-changing rate limits.

X's rate limiters are not designed to be, nor are they, effective measures to restrict access to the public domain information X has chosen to make available without a log-in. As an initial matter, rate limiters are not "access restriction" technologies, but mere volume "throttling" technologies. They do not prevent anyone from accessing anything.

Moreover, X does not limit or attempt to limit the amount of data or content that any particular logged-out visitor can access since X's rate limits are not designed to apply to distinct persons or entities who visit the site, but only to specific accounts, devices, or IP addresses. X could have assigned each visitor a unique identifier (even without creating an account) that would have allowed X to throttle the amount or types of information that each visitor could view, but X chose not to do so. Instead, X intentionally chose to make information available without a log-in so as to maximize the distribution and use of the information on its platform, in the hopes that doing so would make the platform more attractive to monetizable users and increase the network effects that insulate X from competition. Indeed, X actively encourages scraping of its site for this purpose. To the extent X imposes rate limits, it does so to encourage visitors to create accounts and become monetizable users, not to reduce server load, prevent server failure, or improve the "user experience."

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

68. As set forth above, X enforces these rate limits by IP address and through other anomaly-detection tools. To access X's website, a user must provide their IP address, supply identifying markers associated with their web browser, and also make requests about how they would like to interact with the X platform. X's technological measures analyze this information to determine whether the request comes from a legitimate human user or an unauthorized automated process.

Answer: Denied. As set forth above, X does not require visitors to agree to limit the amount of information they view or use. As such, even if X has chosen to throttle requests by secretly placing limits on the amount of information its servers will deliver to a particular IP address in a given period of time, it is not "enforc[ing]" any limits. Similarly, Bright Data denies that X "enforces" any rate limits through what it calls "other anomaly-detection tools." Aside from the fact that visitors have not agreed to any limits on their access to public domain data, "anomaly-detection tools" are not enforcement mechanisms, but are at most investigative or research aids, which do not themselves restrict or limit access. In any event, X does not specify what these so-called tools are, and Bright Data does not know what they are, making it impossible for Bright Data to have intentionally circumvented them.

X is also wrong that, to access X's website, visitors are required to "provide their IP address," to "supply identifying markers associated with their web browser," or to disclose information about "how they would like to interact" with X's public domain web pages. Instead, X only receives the information that is included in any valid search of any website by any Internet-connected device in accordance with standard Internet Protocols. Neither standard Internet Protocols nor X's Terms require that a visitor disclose "their" IP address or prevent the use of proxies. Even without requiring the creation of an account, X could have required visitors to supply personally-identifiable information in order to limit the volume or types of information that can be viewed, as many websites have done. X has chosen not to do so.

To the extent X seeks to use standard Internet Protocol information as part of its throttling efforts, such use is not consistent with that information's intended purpose. Bright Data, of course, does not know how, if at all, X chooses to internally "analyze" Internet Protocol information. But Internet Protocols were not designed to distinguish between humans typing at keyboards or moving mouses on particular web pages, as opposed to computers running scripts typed by humans. X's purported effort to make inferences based on such information is, thus, neither a reasonable nor effective measure for differentiating between what X pejoratively calls "unauthorized automated process[es]" from "legitimate human" endeavors.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

69. The only feasible means of data scraping the limited content that is accessible to non-logged-in users is by circumventing X's IP rate limits and anomaly-detection tools.

Answer: Denied. As set forth above, Bright Data denies that X enforces any rate or access limits through any choice X has made to throttle requests by secret limits on the amount of information its servers will deliver to a particular IP address in a given period of time or through any other so-called "anomaly-detection tools." Bright Data does not know what X means by "the only feasible means of data scraping." Bright Data denies that only "limited" content is accessible to non-logged-in users. As set forth above, X makes substantial amounts of information available for unrestricted view and use to visitors who are logged out, including for most of the relevant period, all or most information that X's registered users have placed in the public domain. X does not attempt to restrict access to such information. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

2. Account Creation Restrictions

70. To ensure that malicious actors cannot automatically create numerous accounts for impermissible purposes, including data scraping, X Corp. requires potential registrants to pass a CAPTCHA by inputting the required information; to enter a valid phone number or email address; and to supply a verification code X sends to that email or phone number.

Answer: Bright Data does not know what procedures and information X requires to register for an account because Bright Data does not use an account in engaging in or offering scraping services. Bright Data notes, however, that X allows multiple accounts to be associated with a single phone number. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

3. Rate Limits

71. X further limits user access to X content even once logged in, in an effort to fight data scraping through logged-in accounts. In July 2023, X imposed limits for the number of posts that a single account can view in a day: 500 for new unverified accounts, 1,000 for unverified accounts in regular use, and 10,000 for X premium subscribers.

Answer: Denied. Bright Data does not know what specific throttling limits X imposes on various subscribers, as Bright Data does not scrape behind a log-in and does not use any account to view information published on X's website. But by imposing a tiered structure for different types of accounts, X seeks – not to "fight data scraping" – but to increase the number of monetizable active daily users, by providing upgrade incentives to users. Indeed, X faced significant financial difficulties resulting from Mr. Musk's highly-leveraged private take-over of Twitter (now X). Upon Mr. Musk's taking control of the platform, X sought ways to reduce reliance on advertisers (who Mr. Musk antagonized through the elimination of ethical content moderation policies) and to increase the number of paying subscribers. The tiered structure referenced in this paragraph, which X admits it instituted only in July 2023, was part and parcel of this strategy. With respect to the purported limits imposed on X's so-called "premium subscribers," Bright Data does not know X's true motivation but notes that Mr. Musk has expressly stated that his purpose was to stymie "outrageous" competition from A.I. competitors. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

72. These reasonable limits for human use make mass data scraping impossible. Accordingly, the only way to data scrape X is to circumvent these rate limits by creating large numbers of bot accounts. And to create large amounts of bot accounts, the data scraper must circumvent X's account-creation restrictions.

Answer: Denied.

4. Anomaly Detection Tools

73. X also employs anomaly detection tools to detect attempted use of many accounts for data scraping or other impermissible purposes, including at the account-creation stage. These tools analyze the IP address being used, certain identifying information about the web browser being used, and characteristics of the type of use (such as patterns that indicate automated access).

Answer: Bright Data does not know what tools X uses to detect the use of fake accounts, or how those tools operate, because Bright Data does not use fake accounts or permit its services to be used with fake accounts to scrape websites. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

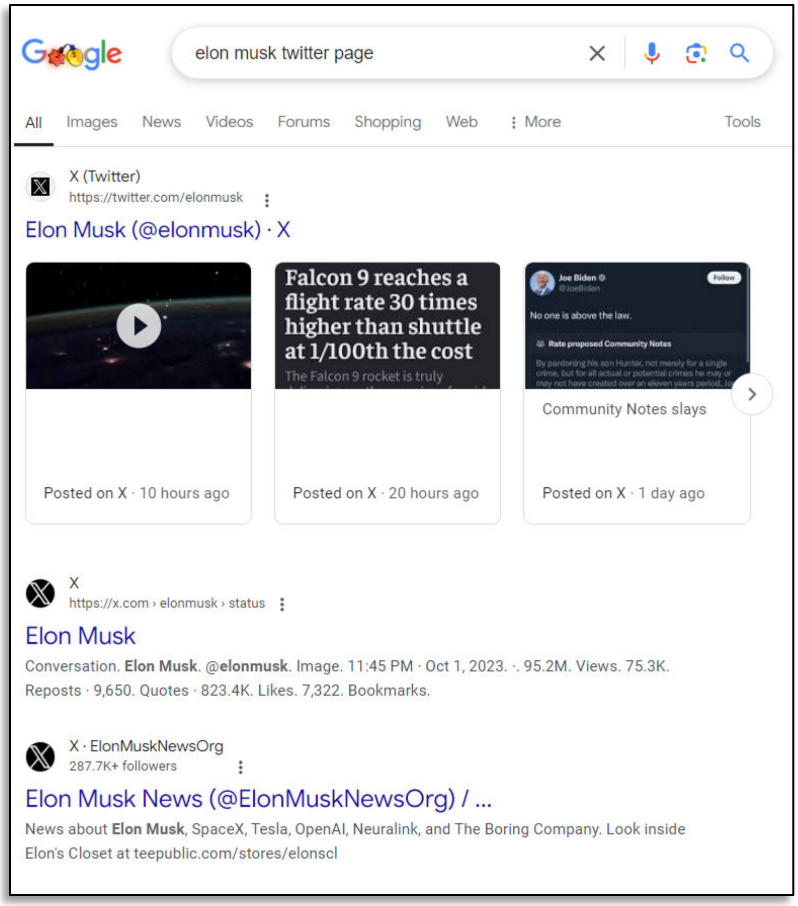
74. To access X's service, users must provide the information relied upon by these anomaly detection tools, including the user's IP address, identifying information for the web browser, and the actions the user wishes to perform on X.

Answer: Visitors do not need to provide any information to access, search, copy, and use the publicly available information on X's site beyond the information that is inherently included in connection with any internet communication under standard Internet Protocols. As to X's "users," who are by definition logged-in, Bright Data does not know what additional information, if any, they are required to provide, since Bright Data does not engage in logged-in scraping. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

5. Robots.txt

75. In July 2023, X Corp. modified its robots.txt instructions to prohibit all forms of automated access to X's website except for Google's web crawler, which is used to index websites for placement on Google's search engine. X Corp.'s robots.txt instructions explicitly forbid any scraping of its website by others, like Bright Data or Bright Data's customers.

Answer: Bright Data denies that X's robots.txt instructions forbid automated access of X's platform. X's robots.txt instructions are not code that X runs or that is incorporated into the operational aspects of X's platform. The robots.txt file simply provides suggested instructions to programmers who may wish to refer to it. It is entirely voluntary, and nothing requires that the robots.txt file be incorporated into the instructions used by internet crawlers. Prior to July 2023, X's robots.txt instructions *actively encouraged* automated access to public portions of X's platform. Bright Data does not know when X changed its robots.txt instructions, but even today, such instructions do not prohibit scraping. Bright Data admits that X permits widespread automated access and scraping by Google and its Googlebot, but denies that Google merely "index[es]" websites. Rather, Google actively scrapes public information published on X's platform, and widely republishes or distributes such information to its own users in response to any applicable search inquiry. The following is just a small sample of the information that Google scrapes from X in virtually real-time:



X (Twitter)
https://twitter.com/elonmusk

Elon Musk (@elonmusk) · X

Posted on X · 10 hours ago

Falcon 9 reaches a flight rate 30 times higher than shuttle at 1/100th the cost
The Falcon 9 rocket is truly

Posted on X · 20 hours ago

Joe Biden
@JoeBiden
Follow
No one is above the law.
Rate proposed Community Notes
By pardoning his son Hunter, not merely for a single crime, but for all actual or potential crimes he may or may not have created over an eleven years period, Jo

Community Notes slays

Posted on X · 1 day ago

X
https://x.com › elonmusk › status

Elon Musk

Conversation. **Elon Musk**. @**elonmusk**. Image. 11:45 PM · Oct 1, 2023. ·. 95.2M. Views. 75.3K. Reposts · 9,650. Quotes · 823.4K. Likes. 7,322. Bookmarks.

X · ElonMuskNewsOrg
287.7K+ followers

Elon Musk News (@ElonMuskNewsOrg) / ...

News about **Elon Musk**, SpaceX, Tesla, OpenAI, Neuralink, and The Boring Company. Look inside Elon's Closet at teepublic.com/stores/elonscl

X's decision to allow automated access and scraping by Google demonstrates that X's robots.txt instructions are designed to attract active daily users that X may monetize, and that this motivation overrides any pretextual arguments about server impairment or the potential adverse impact of scraping on user privacy. Indeed, since at least July 2023, X has eschewed using the specific robots.txt syntax that is expressly designed to manage server load or protect sensitive information.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

76. Robots.txt is a file placed on a website that instructs automated bots which content (if any) they are allowed to access. Robots.txt is an industry-standard tool used by website operators since 1999.

Answer: As just noted, robots.txt is a file that some website operators place on their sites. The file provides voluntary suggestions about how to facilitate automated access or scraping of various portions of a website without server impairment. While some search engines comply with robots.txt files when implemented for proper purposes (*e.g.*, directing bots to public, rather than private, portions of a website), it is not an "industry-standard."

This means that not everyone complies with robots.txt files, especially when a website operator, such as X, uses their robots.txt instructions for improper or anticompetitive purposes. An example illustrates the point: The Internet Archive is a non-profit organization founded in 1996 dedicated to maintaining a free and open internet, and to preserving archival copies of it for future generations of scholars and researchers. This organization operates the Wayback Machine, which appears to be the only place where archived versions of X's own robots.txt files are publicly available, including the pre-2023 versions referenced in this Complaint. In 2017, the Internet Archive announced that it would no longer refer to robots.txt instructions when creating archival copies of webpages, noting that website operators were increasingly using robots.txt files *not* to protect sensitive information or manage server load, but instead to "remove entire … domain[s] from view."[6] Such uses are not an intended or permissible use of the robots.txt instructions, which were not meant to allocate legal rights between website operators and website visitors. Specifically, robots.txt instructions are not a vehicle to strip website visitors of their rights to search the web, nor were they designed to magically create property rights for website operators that did not otherwise exist.

Putting aside X's own improper use of robots.txt instructions, as this example shows, robots.txt is neither a standard, compulsory, nor effective technological measure to restrict access to publicly available content, and many "legitimate operators of automated bots," as X calls them, have chosen not to refer to such instructions. Indeed, because robots.txt instructions were designed from inception to be voluntary, they were never intended to be an effective technological measure to restrict access to content. Rather, if a website operator wishes to restrict information from automated scraping or search engines, the operator needs to password protect the information, which is the *only* industry-standard and potentially legally-enforceable technological measure for preventing automated access or scraping.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

[6] *Robots.txt Meant For Search Engines Don't Work Well for Web Archives* (Apr. 17, 2017), perma.cc/FTV7-JRMT.

77. Bright Data, as one of the largest (if not the largest) providers of data-scraping tools, is familiar with the existence and meaning of instructions in a website's robot.txt file.

Answer: Bright Data is a leading web data company that offers a suite of technologies and services, including scraping services. Bright Data does not know the size of all the other entities that scrape X, but doubts that it is "one of the largest" of such scrapers, as its sales of datasets and scraping tools to customers who scrape X is near *de minimis*. Nonetheless, Bright Data admits that it is familiar with the concept of robots.txt instructions, though not necessarily the particular wording of each website operator's version of them. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

78. All legitimate operators of automated bots (such as search engines like Google) comply with the instructions in a robots.txt file. Bright Data, by contrast, willfully ignores them. Bright Data claims it will force compliance with robots.txt directives for its "immediate access" mode in its "Residential Proxy Network" because that is "respectful and compliant." But it also makes clear that it offers its customers "full . . . access" – that is, not constrained by robots.txt directives – so long as the customer provides Bright Data with basic contact information. *See* Bright Data, *Residential Proxy Network Policy*, https://perma.cc/B945-FX6U. This process is a wink-and-a-nod that Bright Data will ignore robots.txt directives in facilitating full-scale access to data on the platforms it targets.

Answer: As set forth above, many "legitimate operators of automated bots" do not follow the suggestions included in website operators' robots.txt files, especially where – as here – the website operator abuses the file or employs it for improper or anticompetitive purposes.

Indeed, X's own commonly-controlled affiliate, xAI, does not appear to follow robots.txt, and does not provide website operators the information they need to include instructions in their robots.txt file to specifically give direction to xAI's bots.

Regardless of whether other bot operators comply with robots.txt instructions, Bright Data denies that it willfully ignores robots.txt instructions. Indeed, X's own allegations in this Paragraph disprove the point. As X notes, Bright Data customers who seek immediate off-the-shelf access to Bright Data's smart residential proxy network can use Bright Data's standard service, which contains logic that blocks end-points consistent with robots.txt instructions. Customers that do not wish to use Bright Data's standard service because they have unique or bespoke needs are required to go through a rigorous KYC process to validate their particular use case. Bright Data denies that this is a "wink-and-a-nod" process (whatever that is supposed to mean). To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

* * *

79. Because of the combined technological measures X erects to protect its platform from automated access, the only way to data scrape X at the scale Bright Data promises is by: (1) circumventing X's account-creation restrictions like CAPTCHA to automatically open legions of fake accounts; (2) circumventing X's use of IP address monitoring to permit these fake accounts to bombard X's servers at rates otherwise barred by X's limits; and (3) circumventing X's other anomaly-detection tools by having an automated process submit requests to appear as if it were a regular human user.

Answer: Denied. X has chosen not to use password protection – the only industry-standard technological and potentially legally-enforceable method that effectively restricts access to website content – across its platform, and instead has chosen to make vast amounts of user-generated content publicly available. No fake bots or circumvention of any effective technological measure to restrict or limit website access is required to view or copy such information through automated means. Nor has Bright Data opened any accounts ("fake" or otherwise) or intentionally circumvented any so-called "anomaly-detection tools" to access or scrape information on X. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

C. The Harm X Suffers from Bright Data-Enabled Automated Access

80. X has expended vast sums of money developing a service that Bright Data appropriates for illicit ends. X's service is far more sophisticated than one that simply displays user posts. On average, 250 million registered users login to X daily, and 550 million unregistered users visit monthly. On average, there are about 500 million posts per day. Displaying, organizing, and managing all this content requires both money and expertise. For example, X's Recommendation Algorithm must decide hundreds of millions of times per day which of billions of posts to display to its users, and that must happen all in a matter of milliseconds every time. That process involves multiple complex steps to filter posts down to a smaller set of candidate posts, which are then ranked by a continuously trained neural network.[7]

Answer: Bright Data does not appropriate X's service. Instead, Bright Data scrapes publicly available information that X does not own and offers its customers tools and services that can be used for such purposes in accordance with all applicable laws. Bright Data denies that these are "illicit ends."

[7] *See* Twitter, *Twitter's Recommendation Algorithm (Mar. 31, 2023), https:// blog.x.com/ engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm.*

Bright Data does not know how "sophisticated" X's service is, nor does it know how much money or expertise is required in connection with the service. Bright Data also lacks knowledge of the workings of X's Recommendation Algorithm.

In any event, Bright Data's activities on the X platform are *de minimis*. Bright Data does not know whether the statistics in this Paragraph and Paragraphs 39 and 87 are accurate. But even taken at face value, they show that *any form* of unauthorized access or scraping of publicly available data – by all scrapers worldwide – constitutes less than one-half of one percent of server requests on the X platform. And Bright Data's own scraping would be an infinitesimal fraction of that.

X does not incur incremental expense in handling such requests, but rather the funds it expends in operating its platform are ordinary course expenses as X tries to monetize user engagement. For example, while Bright Data does not know how much money X spends on "displaying, organizing, and managing … content" or providing users recommendations, such expenses relate to X's own efforts to increase user engagement with, and the addictiveness of, its platform. These expenses are thus entirely unaffected by any automated access to X's platform by data scrapers.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

81. The infrastructure needed to power that Recommendation Algorithm (and everything else on X) is massive, spanning two large datacenters owned and operated by X and supported by other cloud providers. That infrastructure processed approximately 400 billion events generating petabyte-scale data every day (as of October 22, 2021) and has cost billions of dollars.[8]

Answer: Bright Data does not know the workings or infrastructure of X's Recommendation Algorithm. Bright Data similarly lacks knowledge of X's datacenters, cloud capacity, and the universe of events they processed. In any event, these statistics underscore the fact that Bright Data's activities on the X platform are *de minimis*. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

[8] See Lu Zhang & Chukwudiuto Malife, Processing Billions of Events in Real Time at Twitter, X Engineering (Oct. 22, 2021), https:// blog.x.com/ engineering/ en_us/ topics/ infrastructure/ 2021/ processing-billions-of-events-in-real-time-at-twitter-.

82. Data scraping massively taxes X's infrastructure and requires X to spend [redacted] per year to ensure that its service remains responsive to its actual users.

Answer: Bright Data denies that data scraping in the aggregate – much less Bright Data's own scraping – "massively taxes" X's infrastructure. Indeed, Bright Data's *de minimis* activities on the X platform do not "tax" X's infrastructure at all, as the massive infrastructure X just described is certainly more than capable of handling such requests. In fact, X likely doesn't spend any incremental scraping-related dollars to "ensure that its service remains responsive" to users. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

83. [redacted] And each individual request from a scraper is at least as taxing as an ordinary user request. Most ordinary users are served with a relatively small number of popular posts that X has cached, and the caching makes them relatively efficient for X to serve. When scrapers seek to request *all* posts, by contrast, it forces X to bypass its cache and serve posts from more-expensive storage. This process is far more burdensome for X than caching.

Answer: Denied. If the statistics in Paragraphs 39, 80, and 87 are accurate (which Bright Data does not know and cannot admit to), all global data scrapers of the public portions of X's website collectively constitute far less than one-half of one percent of the server requests submitted by normal users. As noted, such *de minimis* levels do not "tax" or "burden" X at all, as X's "massive" infrastructure (as described in paragraph 81) is more than capable of handling such requests.

X's attempt to mislead the Court by pointing to the number of posts that a single "ordinary user" is served with speaks only to whether such so-called ordinary user's own device(s) are "taxed," not whether X's servers are taxed.

Bright Data denies that scraping "forces X to bypass its cache" or "serve posts from more expensive storage." Indeed, X presumably makes its caching decisions based on the volume and frequency of requests, deciding when it makes sense to spend incremental dollars on caching infrastructure for particular end-points on its platform. Such decisions are independent of any scraping activity. Indeed, the only way data scrapers could increase the storage cost of responding to data requests is if they caused X to move more data into its caching infrastructure, which is the opposite of X's allegations.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

84. Data scrapers compound the harm because they do not use X's APIs, which are optimized for high-volume automated requests targeting the precise information a developer seeks (such as the content of relevant posts). Instead, data scrapers access the X platform as a normal user would, which involves X serving *more* information than the data scraper is actually seeking (such as similar recommended posts or additional users to follow), thus imposing more burden on X's servers than would occur if those data scrapers used the API tailored for the information they are seeking.

Answer: Denied. X's platform is optimized for handling the number and frequency of server requests it receives for each end-point and each type of access, whether through the web, a mobile platform, an API, or otherwise. If requests are shifted from one type of access to another, X may or may not have to increase its expenditures for the type of access that increased, just as it may or may not have to decrease its expenditures for the type of access that decreased. Bright Data does not know how a particular shift of any particular request to a similar request made through a different medium (*e.g.,* X's APIs) would impact X's costs, but denies that such impact would be material, given the infinitesimal volume of Bright Data's (or all scrapers') scraping activity relative to the overall usage of X's platform.

As for the reasons X prefers that data seekers use X's API Developer Platform, it is not to reduce costs or manage server load, but to meter the data and charge monopolistic prices for it. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

85. Further, since the filing of this suit in July 2023, the volume of data scraping on X has only *increased* despite further attempts by X to stop it. As Bright Data touts on its website, data scraping (euphemistically called the "alternative data market") is growing at a 46.5% annual rate, due in no small part to Bright Data.

Answer: Bright Data denies that *its* volume of scraping of X's site is greater today than it was prior to July 2023. X's misquoting of statements about scraping in general is misleading, as those quotes have nothing to do with X.

Bright Data does not know if the volume of data scraping by other data scrapers on X has increased since July of 2023. But if the amount of scraping increased, it is not "despite … attempts by X to stop it," but *because* X has jacked up prices for the data to monopolistic levels, forcing customers to find alternatives. It should be no surprise that when X chose to exit the market for the provision of free public square data, customers would look for alternatives.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

86. X monitors platform usage to identify uses that are likely inauthentic and not attributable to human users. The statistics identified below reflect data gleaned from that monitoring and represent actual burdens on X's server capacity, at a minimum.

Answer: Bright Data does not know how X identifies uses it believes to be "inauthentic" or "not attributable to human users." To the extent X purports to rely on Internet Protocol information, Bright Data notes that Internet Protocols were not designed to distinguish between humans typing at keyboards or moving mouses on particular web pages, as opposed to computers running scripts typed by humans. X's purported effort to make inferences based on any such information is, thus, neither a reasonable nor effective measure for differentiating between what X pejoratively calls "inauthentic" users from "human" ones. Bright Data does not know whether the statistics in the paragraph that follows are accurate, but denies that they represent meaningful burdens on X's server capacity. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

87. On average, about [redacted] is inauthentic and not attributable to human users. Data scrapers prefer to access X through the web, while most of X's traffic from authentic human users comes via iOS or Android. So the numbers below refer to web traffic. On a normal day, X generally [redacted] X's monitoring shows that specific types of end points are more likely to be targeted by data scrapers, including up to [redacted] of certain types of access being by data scrapers. Several are identified below. Notably, each of the types of access listed below are available *only* to logged in users. This demonstrates that most data scraping occurs through logged-in accounts that have agreed to X Corp.'s Terms (only to blatantly violate them), not through logged-out scraping of public data:

a. [redacted] of web requests to view [redacted] are anomalous or inauthentic.
b. [redacted] of the web requests to look up [redacted] are anomalous
c. or inauthentic.
d. [redacted] of the web requests to look up [redacted] are anomalous
e. or inauthentic.
f. [redacted] of web requests to look up [redacted] are anomalous or inauthentic.
g. [redacted] of web [redacted] are anomalous or inauthentic.

Answer: Bright Data does not know whether the statistics in this Paragraph are accurate, or how X generated them. But Bright Data denies that they would represent meaningful burdens on X's server capacity. In fact, the statistics show the exact opposite. Coupled with the statistics in paragraph

39, 80, and 81 (the accuracy of which Bright Data does not know and cannot admit to), this Paragraph shows that public scraping constitutes less than one-half of one percent of server requests.

As for the end-point-specific statistics, X has manipulated the reported statistics to yield the misleading results it wanted. By excluding the ways that "most" humans access the platform from a reported statistic, X has artificially inflated the statistic to show a higher percentage of automated searches than exists in reality. X also fails to allege the volume of such requests, so the percentage is meaningless as a measure of server burden.

In any event, if it is true that such end-points are "available *only* to logged in users," then it would not be information that Bright Data scraped. Bright Data, however, does not know whether X's allegation that such end-points are available only with a log-in is truthful.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

88. X has also encountered and investigated specific instances of massive data scraping, which include: In September 2023, [redacted] of all requests for [redacted] were coming from data scrapers. In October 2023, data scrapers were submitting [redacted]. In December 2023, [redacted] of all "[redacted]" traffic originated from data scrapers.

__Answer:__ Bright Data does not know what instances of data scraping, much less "massive" data scraping, X believes it has encountered or investigated. Nor does Bright Data know how X purports to identify what traffic or requests it believes originate from data scrapers. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

89. Because X allocates specific server capacity to [redacted], these requests lead to [redacted] for the [redacted] scraping targets. Even when specific servers are overloaded, X's [redacted] attempts to ensure that the system overall is less likely to fail entirely, but even isolated failures lead to [redacted]

__Answer:__ Bright Data does not know how X allocates server capacity but denies that any scraping by Bright Data has led to [redacted]," or that X has spent any money on any day to handle X's requests.

Bright Data does not know if X has ever spent a penny to handle other scraping requests. But, if it did, X concedes that data scraping – by all data scrapers – has not led to any significant server failure. To the extent X has alleged otherwise, Bright Data denies such allegations. Bright Data also denies that [redacted].” That argument is just lawyer sophistry. For example, when a person takes a screen shot on his or her computer, it uses [redacted] from any other program.

[redacted]

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

90. To avoid the catastrophic failure of X's platform servers and systems due to impermissible data scraping, X obtains an average projected additional [redacted] of over-provisioned headroom for its server load capacity that it needs to accommodate the anomalous bot activity facilitated by data scrapers like Bright Data. This additional server load capacity imposes additional costs ranging from [redacted] each month.

Answer: Bright Data does not know how X makes its server capacity decisions or how much, if any, is specifically designed to "accommodate … anomalous bot activity," as opposed to other activity, such as just to handle peak times like the World Cup. As Elon Musk tweeted, "World Cup traffic hit almost 20,000 tweets per second today! Great work by Twitter team managing record usage."[9] Obviously, expenditures to accommodate increased traffic due to breaking world events have nothing to do with data scraping or automated access, but likely drive virtually all of X's peak load and/or excess server capacity decisions and expenditures. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

91. If X Corp. did not purchase this additional server load capacity, the X user platform would risk failure and significant degradation of the X user experience. Because X has a reputational and business imperative to be available 24/7 as the platform known for "what's happening" moment to moment, it cannot afford those risks and must pay for additional capacity.

Answer: Denied. As just noted, X purchases server capacity for reasons unrelated to Bright

[9] @elonmusk, x.com, perma.cc/RY7C-VKBR.

Data or data scraping.

92. X also employs a dedicated team of operational engineers to remedy anomalous access on X, which is constantly evolving as X must understand, assess, and stymie new methods of automated, anomalous, and inauthentic access of its platform. Over the last year, this team responded to and remedied at least [redacted] of known data scraping where the demands on X's server capacity jumped extreme amounts. Of course, this team cannot respond to all scraping, much of which goes unremediated – not because it does not tax X's server capacity but because it is so prevalent.

<u>Answer:</u> Bright Data does not know what X's employees do. Nor does it know what these so-called, [redacted] of data scraping are, or what X means when it says server "capacity jumped extreme amounts." So, to the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

93. Each of these scraping incidents could occur only through the coordinated use of proxy networks with massive amounts of rotating IP addresses (which Bright Data provides) and the creation of numerous fake accounts automatically through CAPTCHA circumvention (which Bright Data likewise enables).

<u>Answer:</u> Denied.

94. In attempting to stop this activity, X must balance the harms of data scraping with the risk of false positives blocking legitimate access. Data scrapers (like Bright Data or the companies that it enables) succeed by making their activity as indistinguishable from normal user activity as possible – including by use of IP proxies that route traffic through ordinary consumers' devices. Thus, X suffers harm even in attempting to impose technological measures to stop data scrapers. [redacted]

<u>Answer:</u> At least as to the type of scraping that Bright Data and its customers engage in, X does not need to balance the harms of data scraping with the risk of false positives. It could use the industry-standard and potentially legally-enforceable method of putting information behind a log-in. Bright Data does not know why X's employees [redacted], but denies that X or its employees did so through the exercise of reasonable care and competence in the exercise of their job duties. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

95. Beyond the direct costs posed by the data scraping activity, the continued presence of bots, fake profiles, and data scrapers impacts X's relationship with its users, who participate in X's services on a presumption of trust that X can enforce its policies relating to privacy and user content

authenticity, among others. The presence of bots creates risks to privacy and security, subjects users to spam and other unwanted content, and degrades the overall user experience in a way that risks users leaving X. That risks costing X content, data, and other benefits of its large and robust platform.

Answer: Denied. X's only concern (that is not pretextual) as to public data scraping on its platform relates to X's ability to acquire, maintain, and exploit its information monopoly. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

D. Defendant and Its Customers Have Agreed to X Corp.'s Terms of Service

96. Defendant has expressly agreed to X Corp.'s Terms and is therefore bound by those Terms.

Answer: Denied.

97. Initially, by using the X platform, Defendant, which is well aware of the Terms, agrees to be bound by them. The Terms specifically state:

> These Terms of Service ("Terms") govern your access to and use of our services, including our various websites, SMS, APIs, email notifications, applications, buttons, widgets, ads, commerce services, and our other covered services (https://help.x.com/rules-and-policies/x-services-and-corporate-affiliates) that link to these Terms (collectively, the "Services"), and any information, text, links, graphics, photos, audio, videos, or other materials or arrangements of materials uploaded, downloaded or appearing on the Services (collectively referred to as "Content"). By using the Services you agree to be bound by these Terms.

Answer: Denied. Bright Data does not "use" X, rather, it exercises its lawful right to search the public internet. The Terms only bind account holders. X's website makes clear that only "by signing up, you agree to the Terms of Service." Bright Data has not "signed up." Nor do the Terms purport to bind unregistered visitors. The Terms only state that "by using the Services you agree to be bound by these Terms." It is well established that the term "use," in the context of internet activity, refers to activities related to or performed through a registered account. Bright Data does not have an X account and does not scrape or access X through an account. Indeed, the Terms themselves draw distinctions between "access" and "use," and do not bind those who merely "access" the site. Rather, they only govern account holders when they access the site using the account. To the extent Bright Data is deemed to "use" the X platform, however, Bright Data denies that it is bound by the Terms. Aside from the fact that the Terms are unenforceable in whole or in part, they do not constitute a binding agreement with

Bright Data because they are not supported by any consideration and Bright Data affirmatively manifested express rejection of X's Terms. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

98. In addition to agreeing to the Terms by using X services, Defendant, which has maintained a registered account on X (@bright_data) since at least February 2016, expressly accepted and agreed to the Terms when registering its account. Bright Data's X account frequently posts content promoting the company's products and services.

Answer: Denied. Bright Data does not have any active accounts on X. Bright Data terminated any account it had or may have had by September 25, 2023, through formal notice to X. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

99. Defendant's top executives are also registered X users and expressly agreed to X Corp.'s Terms when registering their accounts, further demonstrating that Bright Data had knowledge of the Terms:

a. Bright Data's CEO, Or Lenchner, has maintained a registered account on X (@orlench) since at least December 2012 and regularly posts from that account.
b. Bright Data's CMO, Yanay Sela, has maintained a registered X account (@yanay_sela) since at least December 2014.
c. Bright Data's Managing Director for North America, Omri Orgad, has maintained a registered X account (@omri_orgad) since at least November 2011.
d. Bright Data's Vice President of Product, Erez Naveh, has maintained a registered X account (@nerez) since at least August 2009.
e. Bright Data's Global Communications Manager, Zachary Keyser, has maintained a registered X account (@KeyserZachary) since at least December 2019.
f. Bright Data's Founder, Ofer Vilenski, has maintained a registered X account (@vilenski) since at least November 2008.

Answer: Bright Data does not know which, if any, of its employees have registered as X users or have had agreements with X. Bright Data denies that it became aware of the Terms by virtue of Bright Data or any Bright Data employee having registered for an X account. Bright Data further notes that it has never used any such account for the purpose of the activities challenged herein. In any event, X cannot – and seemingly does not seek to – bring a claim against Bright Data based on any alleged breach of alleged individual account contracts. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

100. On information and belief, several other employees and agents of Defendant involved in Defendant's data-scraping activities are also X account holders, including, by way of example, Artem Shibakov, a Bright Data software engineer who has maintained a registered X account (@ashibakow)

since at least February 2013. These account holders have also expressly accepted and agreed to X Corp.'s Terms.

Answer: Bright Data does not know which, if any, of its employees or agents have or have had X accounts. Nor has Bright Data used any such account for the purpose of the activities challenged herein. In any event, X cannot – and seemingly does not seek to – bring a claim against Bright Data based on any alleged breach of alleged individual account contracts. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

101. Defendant is additionally subject to the Terms as an advertiser on X. Beginning on March 7, 2016, Defendant (then known as Luminati Networks) purchased advertising on the X platform. Defendant purchased additional advertising on X from 2019 to 2021. As stated in X Corp.'s Ad Policies, to which Defendant expressly agreed, all advertisers are bound by the platform's Terms and Rules.

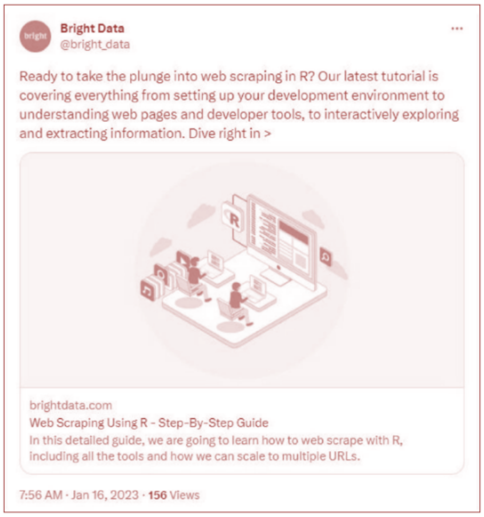
Answer: Bright Data has not advertised on X since at least the time it terminated its X account(s) and rejected the Terms. Bright Data admits that at some point before this lawsuit began, Bright Data placed one or more advertisements on the X platform, which X approved and profited from. When X accepted Bright Data's money and approved its ads (including those cited in this Complaint), X did not claim that Bright Data's services were illegal, illicit, unlawful, deceptive, or otherwise wrongful. In fact, by approving such ads, X granted Bright Data express and implied consent for the advertised activities. In any event, Bright Data denies that it is bound by the Terms for any activities relating to Bright Data's scraping or its sale of proxy services or scraping tools to third parties by virtue of the fact that it placed advertisements on X at some distant point in the past. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

102. Defendant and its executives have repeatedly used these X accounts to discuss and promote their data-scraping products and services, including but not limited to the following posts:

a. On January 1, 2023, Defendant posted a video on X entitled "How to Scrape UNSCRAPABLE data!" which demonstrated how to use Defendant's tools and services for unauthorized data scraping.

b. On January 16, 2023, Defendant encouraged users in a post on X to "take the plunge into web scraping" using a "step-by-step guide" to Defendant's tools and services.

Figure 5: Screenshot of Bright Data's X post on July 11, 2023



Bright Data
@bright_data

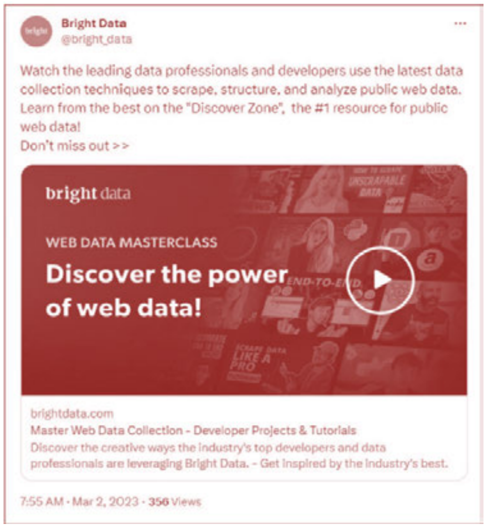
Ready to take the plunge into web scraping in R? Our latest tutorial is covering everything from setting up your development environment to understanding web pages and developer tools, to interactively exploring and extracting information. Dive right in >

brightdata.com
Web Scraping Using R - Step-By-Step Guide
In this detailed guide, we are going to learn how to web scrape with R, including all the tools and how we can scale to multiple URLs.

7:56 AM · Jan 16, 2023 · 156 Views

c. On March 2, 2023, Defendant posted a video on X to a "masterclass" that showed "the latest data collection techniques to scrape, structure, and analyze public web data" using Defendant's tools and services.

Figure 6: Screenshot of Bright Data's X post on July 11, 2023



Bright Data
@bright_data

Watch the leading data professionals and developers use the latest data collection techniques to scrape, structure, and analyze public web data. Learn from the best on the "Discover Zone", the #1 resource for public web data!
Don't miss out >>

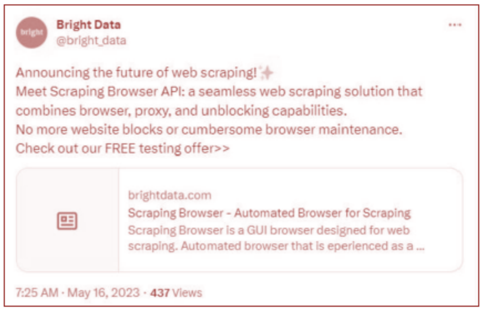
brightdata.com
Master Web Data Collection - Developer Projects & Tutorials
Discover the creative ways the industry's top developers and data professionals are leveraging Bright Data. - Get inspired by the industry's best.

7:55 AM · Mar 2, 2023 · 356 Views

d. On March 23, 2023, Defendant posted and promoted its "Web Unblocker" and its ability to "bypass[] multiple anti-bot solutions" in a post on X.

e. On May 16, 2023, Defendant promoted its "Scraping Browser API: a seamless web scraping solution that combines browser, proxy, and unblocking capabilities" with a link to a "FREE testing offer" in a post on X.

Figure 7: Screenshot of Bright Data's X post on July 11, 2023



Bright Data
@bright_data

Announcing the future of web scraping! ✨
Meet Scraping Browser API: a seamless web scraping solution that combines browser, proxy, and unblocking capabilities.
No more website blocks or cumbersome browser maintenance.
Check out our FREE testing offer>>

brightdata.com
Scraping Browser - Automated Browser for Scraping
Scraping Browser is a GUI browser designed for web scraping. Automated browser that is eperienced as a ...

7:25 AM · May 16, 2023 · **437** Views

Answer: Bright Data admits that prior to its termination of its account(s) and rejection of the Terms, Bright Data may have posted a handful of tweets discussing Bright Data's products, including the posts on January 1, January 16, March 2, March 23, and May 16, 2023. As those posts and related links make clear, Bright Data's services can only be used to scrape "public web data." To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

103. Bright Data's customers are similarly bound by X Corp.'s Terms by use of the X platform, and whenever they sign up for accounts that explicitly require agreement to the Terms.

Answer: Bright Data denies that its customers are bound by X's Terms by virtue of access or scraping publicly-available data. Bright Data does not know which, if any, of its customers have signed up for X accounts or have agreements with X. What the Terms actually prohibit, to whom they apply, and whether they are enforceable, are each legal conclusions that may depend on the facts, including facts that are not alleged here. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

104. Bright Data is also aware that its customers are bound by X Corp.'s Terms. It is a sophisticated entity that knows those customers are bound by the same Terms to which Bright Data agreed. Further, Bright Data is aware of and complicit in customers' use of Bright Data's tools for the purpose of willfully violating X Corp.'s Terms that prevent data scraping. Bright Data knows that the most commercially valuable uses of data on the X platform require scraping large amounts of data that can only be done through logged-in accounts. Bright Data makes its tools available for the purpose of circumventing X's technological measures designed to prevent that activity.

Answer: Denied. None of Bright Data's customers become bound to X's Terms by virtue of using Bright Data's tools to scrape public information on X. Bright Data does not know if any of its customers have otherwise agreed to X's Terms for reasons unrelated to their use of Bright Data's services. Nor has Bright Data become "complicit" in any unlawful activity. Using Bright Data's tools to search, access, or scrape public information is not unlawful, and Bright Data's services cannot be used to scrape behind a log-in. Nor does Bright Data make available any tools to circumvent any log-in requirement or any other technological measures restricting or protecting non-public information. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

E. Defendant's Unauthorized Scraping and Sale of Scraping Tools

105. Defendant, per its own admissions, has engaged in widespread scraping of X Corp.'s data, circumventing X Corp.'s technical barriers and violating the Terms to which it agreed. Defendant has also facilitated the scraping of data from X and induced X users to violate the Terms.

Answer: Denied.

106. X Corp. has not granted Defendant permission to scrape data from the X platform.

Answer: Denied. While permission is not required to scrape public information, X has affirmatively granted Bright Data and its customers permission to scrape data from the X platform when it accepted substantial amounts of money from Bright Data and approved Bright Data's advertisements of its services. In knowingly and intentionally publishing Bright Data's advertisements to X's own users, X actively encouraged its users to use Bright Data's services for the advertised purposes. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

107. X Corp permits paying developers to lawfully access certain categories of X data, subject to contractual usage limits and other restrictions designed to protect the X platform and user experience, as detailed above. Rather than lawfully acquire X data through authorized means, Bright Data elected to scrape the data (and enable others to do so), using proxies and other illicit methods to shield its identity and scraping activities.

Answer: Denied. One way of accessing certain categories of data on X is by paying X's supracompetitive and monopolistic prices. But it is not the only lawful way, at least as it relates to publicly-available information, which is the only type of information Bright Data scrapes. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

108. Bright Data markets itself as the leading provider of tools that enable unlawful data scraping of X's platform. Bright Data states that it is the "leading residential proxy provider" in the

market, controlling "the best proxy network in the world" and "offer[ing] the best web scraping proxies." *See* Bright Data, *The Top 10 Residential Proxies of 2024*, https://perma.cc/T4MU-MMGQ.

Answer: Denied. Bright Data is a leading web data company that offers a suite of technologies and services, including proxy services that can be used to scrape public internet information. Bright Data, however, does not permit scraping of non-public information, and does not "enable unlawful data scraping." Nor is Bright Data a "leading provider of tools that enable … scraping of X's platform." Bright Data does not know how many people scrape X using their own, or other non-Bright Data, tools. But scraping on X by Bright Data (or its customers) is *de minimis*. As such, it is unlikely that Bright Data is a "leading provider" of scraping tools used to scrape X. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

109. Investors have described Bright Data (which used to call itself "Luminati") as "the world's leading enterprise IP proxy network" and "the only mass-scale residential IP proxy network in the world." *See* EMK Capital, *EMK Acquires Luminati – The World's Largest IP Proxy Network Which Brings Transparency to the Internet* (Aug. 10, 2017), https://perma.cc/X4TZ-T9FB.

Answer: Bright Data offers both an enterprise IP proxy network, as well as a residential IP proxy network. The former is not at issue in this case. Nor does Bright Data know what its share of the enterprise IP proxy network is, or whether it is in fact the "leading" provider. Bright Data believes that it is a leading provider of residential proxy network services, supported by thousands of patent claims issued by the United States Patent and Trademark Office. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

110. Investors also tout Bright Data as the only product on the market that can evade the technical safeguards of the websites targeted for scraping: "Unlike Bright Data, other IP proxy networks and their customers are prone to being blocked, slowed or spoofed by websites they visit." *See* EMK Capital, *Bright Data*, https://perma.cc/9JM8-MVUS.

Answer: Bright Data offers best-in-class proxy services that can be used for many purposes, including scraping public information. Bright Data's proxy services and scraping tools cannot be used to evade technical measures that restrict access to non-public information. Bright Data does not know how effective the proxy services or scraping tools of other providers are. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

111. A 2020 research report described Bright Data as the "[m]arket leader" in proxy network services; another report issued the prior year stated that Bright Data "is the world's largest proxy

service." *See* Robert Cavin, *Description – Global Internet Protocol Proxy Network Market, Forecast to 2025*, Frost & Sullivan (2020), https://perma.cc/CDY6-YJ2K; Business Wire, *Frost & Sullivan Names Luminati the 2019 Global Market Leader in the Enterprise IP Proxy Networks Market* (July 23, 2019), https://www.businesswire.com/news/home/20190723005394/en/Frost-Sullivan-Names-Luminati-the-2019-Global-Market-Leader-in-the-Enterprise-IP-Proxy-Networks-Market.

Answer: Bright Data does not know the size of other proxy service providers. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

112. Defendant has not publicly disclosed how it evades X Corp.'s technical safeguards against scraping. Rather, Defendant takes extraordinary steps to conceal its activity from X. Defendant's entire business is built on bypassing the technological measures that X puts in place to stop data scraping – a business that only works if it can evade detection by X.

Answer: Denied. Bright Data admits that ***X*** has engaged in efforts to obstruct the public's right of access to public information. Just as X does not disclose its proprietary information about its obstructionist efforts, Bright Data does not disclose all of its proprietary information about its technology. But Bright Data has published substantial information about how its products and services work.

Bright Data denies that its "entire business is built on bypassing technological measures that X puts in place to stop data scraping." Scraping of X's website by Bright Data (and its customers) is *de minimis*. Moreover, X does not engage in technological measures to stop scraping of public information, since it does not put such information behind a log-in, and Bright Data does not evade any technical measures protecting non-public information.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

113. Because Bright Data conceals its misconduct, X cannot pinpoint all the instances where Bright Data's customers have used Bright Data's scraping tools. However, because Bright Data is the market leader in providing such tools, because of the increasing level of data scraping that occurs on X, and because of Bright Data's own marketing statements targeting X, it is virtually certain that many major data scrapers of the X platform use Bright Data's tools.

Answer: Bright Data does not engage in any misconduct. Nor does Bright Data know what activity X is able to "pinpoint." But because scraping of the X platform by Bright Data and its customers is *de minimis*, it is unlikely that the major data scrapers on the X platform use Bright Data's services. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

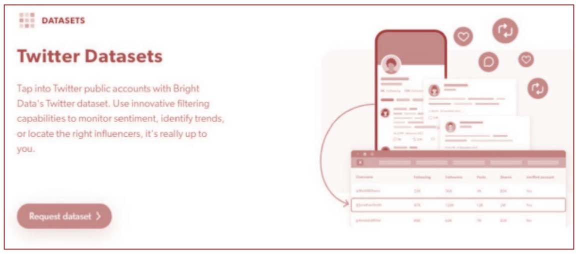
114. Indeed, Defendant's website makes clear that the company itself engages in prohibited scraping of X on an industrial scale and brazenly advertises that Defendant sells tools and services that encourage and enable others to engage in prohibited scraping.

<u>Answer:</u> Denied.

1. Defendant Scrapes and Sells X Corp. Data

115. Defendant offers X Corp.'s data for sale on its website.

Figure 8: Screenshot of Bright Data's website on July 10, 2023



DATASETS

Twitter Datasets

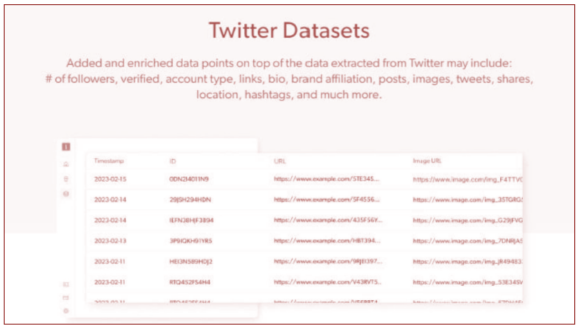
Tap into Twitter public accounts with Bright Data's Twitter dataset. Use innovative filtering capabilities to monitor sentiment, identify trends, or locate the right influencers, it's really up to you.

Request dataset >

<u>Answer:</u> Bright Data admits that it offers to sell a dataset consisting of public information scraped from X's website. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

116. According to Defendant's website, the X Corp. data sets offered for sale by Defendant include "millions of pages and tens of millions of data points." Specifically, these data sets include the following user information: "# of followers, verified, account type, links, bio, brand affiliation, posts, images, tweets, shares, location, hashtags, and much more."

Figure 9: Screenshot of Bright Data's website on July 10, 2023



Twitter Datasets

Added and enriched data points on top of the data extracted from Twitter may include:
# of followers, verified, account type, links, bio, brand affiliation, posts, images, tweets, shares, location, hashtags, and much more.

Timestamp	ID	URL	Image URL
2023-02-15	0DN2I4011N9	https://www.example.com/STE345...	https://www.image.com/img_F4TTV
2023-02-14	29JSH294HDN	https://www.example.com/SF4556...	https://www.image.com/img_35TGRG
2023-02-14	IEFN3BHJF3894	https://www.example.com/435F56Y...	https://www.image.com/img_G29JFVG
2023-02-13	3P9IQKH91YR5	https://www.example.com/HBT394...	https://www.image.com/img_7DNRJAS
2023-02-11	HEI3N589HDJ2	https://www.example.com/9RJEI397...	https://www.image.com/img_JR49483
2023-02-11	RTQ452F54H4	https://www.example.com/V43RVT5...	https://www.image.com/img_53E345V

<u>Answer:</u> Bright Data admits that its website describes the Twitter dataset, as it existed in July 2023, as shown in the screenshot above. The dataset consists only of information that was publicly available at the time of collection. Bright Data does not know, without further investigation, how many records are contained in any already created dataset, and thus, does not know if such datasets contain "millions of pages" or "tens of millions of data points." To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

117. Defendant could have only obtained this data by engaging in prohibited scraping of X's platform.

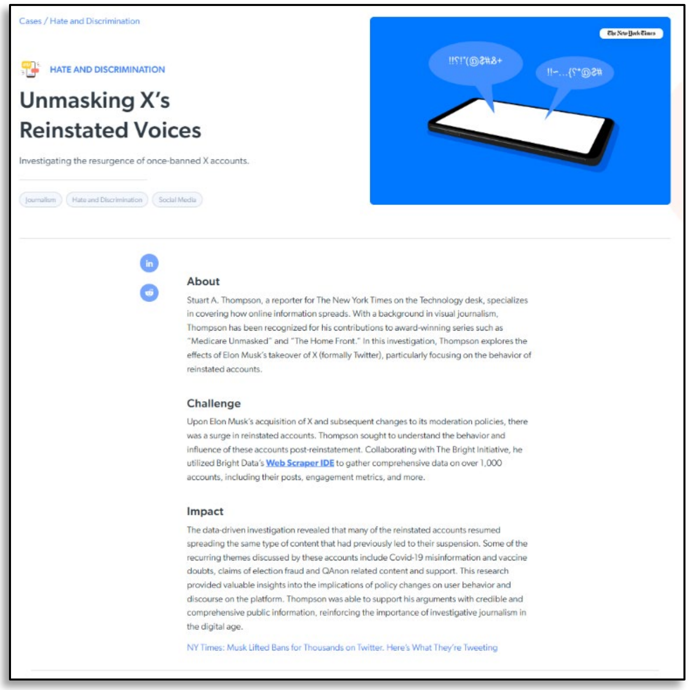
Answer: Denied.

118. Defendant offers this unlawfully obtained data for sale starting at $.01 per record, but also offers customized packages of X Corp.'s data. Defendant also offers several options for delivery of X Corp.'s data, and even offers its customers the opportunity to update its data sets with additional data scraped from X at regular intervals.

Answer: Bright Data denies offering "unlawfully obtained" data for sale. Bright Data does offer for sale publicly available information that it has compiled or may compile at a specific customer's request at competitive rates. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

119. Defendant does not transform the content it scrapes from X, nor does it offer that content for social or educational purposes. Rather, as Defendant's website makes clear, its purpose is commercial: it packages up X content and data and sells it to the highest bidder. Indeed, the "Twitter Datasets" Bright Data sells merely duplicate the data that X's own API tiers already make available, without X's privacy and other safeguards. In doing so, Bright Data does not obtain consent from X's users before scraping and selling their data for commercial purposes. On information and belief, X's users broadly remain unaware that Bright Data is doing so.

Answer: Denied. Bright Data is committed to providing access to public information for many social and educational purposes, and often does so free of charge. For example, when Mr. Musk decided to transform X into a haven for hate speech, *The New York Times* sought to investigate. It turned to Bright Data's services, free of charge, to help shed light on this threat to democracy and human safety.[10]



Cases / Hate and Discrimination

HATE AND DISCRIMINATION

Unmasking X's Reinstated Voices

Investigating the resurgence of once-banned X accounts.

About

Stuart A. Thompson, a reporter for The New York Times on the Technology desk, specializes in covering how online information spreads. With a background in visual journalism, Thompson has been recognized for his contributions to award-winning series such as "Medicare Unmasked" and "The Home Front." In this investigation, Thompson explores the effects of Elon Musk's takeover of X (formally Twitter), particularly focusing on the behavior of reinstated accounts.

Challenge

Upon Elon Musk's acquisition of X and subsequent changes to its moderation policies, there was a surge in reinstated accounts. Thompson sought to understand the behavior and influence of these accounts post-reinstatement. Collaborating with The Bright Initiative, he utilized Bright Data's Web Scraper IDE to gather comprehensive data on over 1,000 accounts, including their posts, engagement metrics, and more.

Impact

The data-driven investigation revealed that many of the reinstated accounts resumed spreading the same type of content that had previously led to their suspension. Some of the recurring themes discussed by these accounts include Covid-19 misinformation and vaccine doubts, claims of election fraud and QAnon related content and support. This research provided valuable insights into the implications of policy changes on user behavior and discourse on the platform. Thompson was able to support his arguments with credible and comprehensive public information, reinforcing the importance of investigative journalism in the digital age.

NY Times: Musk Lifted Bans for Thousands on Twitter. Here's What They're Tweeting

Bright Data also has a long history of helping researchers address societal ills and assisting people in danger from oppressive governments, war, and natural disasters.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

120. Defendant claims dubiously that its datasets sourced from X contain only publicly accessible data – that is, data available to users not logged into X. On information and belief, that claim is inaccurate for the reasons stated above. But even if it were true, there is no way Bright Data could have obtained this amount of data without circumventing the technological measures X put in place to prevent scraping of data available to users that are not logged into X. Specifically, Bright Data would have needed at the very least to circumvent X's IP-based rate limits through proxy networks, and to circumvent X's anomaly detection measures. Without doing so, Bright Data could not have obtained the high volume of data it sells, because of the stringent limits X places on the amount of data that can be accessed before being required to login or create an account.

Answer: Bright Data denies that X is truly "dubious" of Bright Data's representations. The

[10] *Unmasking X's Reinstated Voices*, The Bright Data Initiative, perma.cc/9WA5-3SH8.

information X cites in its Complaint makes clear that Bright Data only scrapes public information, and any reasonable due diligence by X would have confirmed this fact. As to the remaining allegations about X's obstruction technologies, Bright Data refers to its answers to the Paragraphs above. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

2. Bright Data Sells Automated Tools to Scrape X Corp.'s Data

121. Defendant also offers for sale on its website automation software that allows users to scrape data directly from the X platform in violation of X Corp.'s Terms.

Answer: Denied. Bright Data offers tools that allow its customers to scrape publicly-available information. Bright Data denies that such activities violate X's Terms. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

122. Defendant is aware that its customers use its tools to scrape data from X through logged-in accounts. On information and belief, Bright Data tools have the capability to determine whether its tools are being used to login to X's platform and could stop such use if Bright Data wanted. However, Bright Data knowingly allows logged-in scraping to occur so long as Bright Data users comply with a sham "Know Your Customer" process. Defendant's website states: "If you don't want to purchase a Twitter dataset, you can start scraping Twitter data using our Twitter scraping tool." Bright Data, *Twitter Datasets*, https://perma.cc/L72J-WGVL.

Answer: Denied. Bright Data is not aware of any of its customers scraping data from the X platform through logged-in accounts. To the extent Bright Data has the ability to distinguish between logged-on and logged-off activity (which it does for users of its Web Unlocker and scraping tools), Bright Data employs technological measures to prohibit logged-on scraping. Because the connection between X and Bright Data's users are encrypted (not by Bright Data, but according to standard Internet Protocols), Bright Data does not have the ability to monitor the communications, if any, between X and those who visit or use X when using Bright Data's proxy network. For customers that use Bright Data's smart residential proxy network service, Bright Data employs technological measures to follow the host website's voluntary robots.txt files. For other customers of Bright Data's residential proxy network (who are not using Bright Data's Web Unlocker or scraping tools), Bright Data requires such customers to satisfy Bright Data's rigorous KYC procedures, in which each use case is reviewed and must be specifically approved. Logged-on scraping is not an approved use case and is prohibited by Bright Data's Acceptable Use Policy. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

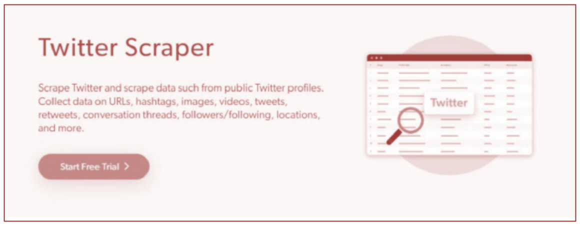
123. Defendant's Web Scraper tool allows individuals to evade detection utilizing a proxy network in order "to remain anonymous, avoid IP blocking, access geo-restricted content, and improve scraping speed." Bright Data, *Twitter Scraper*, https://perma.cc/4EFH-PKVT. The tool also includes an "unblocking solution" that is designed to evade anti-scraping measures like those employed by X Corp. *Id*. Defendant specifically advertises that its Web Scraper tool can be used to "[e]asily scrape data from any geo-location while avoiding CAPTCHAs and blocks." Bright Data, *Web Scraper APIs*, https://perma.cc/MHA9-Q3GG.

<u>Answer:</u> As just noted, Bright Data's scraping tools do not permit logged-in scraping. As with any VPN, Bright Data's tools allow logged-off visitors to websites to remain anonymous, to visit websites that oppressive governments may censor, and to improve their scraping efficiency. As X does not use Captchas to block access to content, Bright Data is, without further investigation, unaware of any instances where Bright Data has solved or avoided Captcha's employed by X. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

124. In addition to its Web Scraper tool, Defendant sells at least four additional tools designed to scrape information specifically from the X Platform: Twitter Scraper, Twitter Profile Scraper, Twitter Image Scraper, and Twitter Followers Scraper.

a. Defendant offers a Twitter Scraper to automatically scrape data from the X platform, including "URLs, hashtags, images, videos, tweets, retweets, conversation threads, followers/following, locations, and more." Bright Data, *Twitter Scraper API*, https://perma.cc/49MC-3HUL. Bright Data advertises its Twitter Scraper's ability to circumvent X's defenses: "Maintain full control, flexibility, and scale without worrying about infrastructure, proxy servers, or getting blocked." *Id*.

Figure 10: Screenshot of Bright Data's Website on July 10, 2023

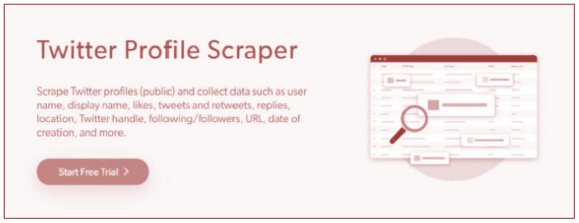


Twitter Scraper

Scrape Twitter and scrape data such from public Twitter profiles. Collect data on URLs, hashtags, images, videos, tweets, retweets, conversation threads, followers/following, locations, and more.

b. Defendant offers a Twitter Profile Scraper to automatically "collect data such as user name, display name, likes, tweets and retweets, replies, location, Twitter handle, following/followers, URL, date of creation, and more." *See also* Bright Data, *Twitter Profile Scraper API*, https://perma.cc/3TAQ-VGVC.

Figure 11: Screenshot of Bright Data's website on July 10, 2023

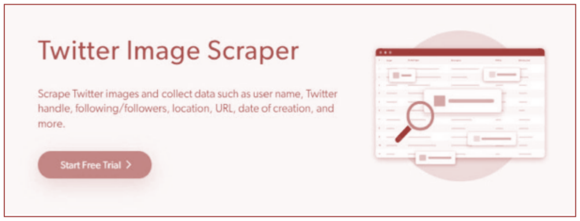


Twitter Profile Scraper

Scrape Twitter profiles (public) and collect data such as user name, display name, likes, tweets and retweets, replies, location, Twitter handle, following/followers, URL, date of creation, and more.

c. Defendant also offers a Twitter Image Scraper to automatically "collect data such as user name, Twitter handle, following/followers, location, URL, date of creation, and more." *See also* Bright Data, *Twitter Images Scraper API*, https://perma.cc/L8SX-W5WC.

Figure 12: Screenshot of Bright Data's website on July 10, 2023

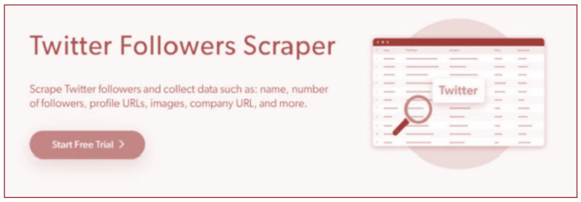


Twitter Image Scraper

Scrape Twitter images and collect data such as user name, Twitter handle, following/followers, location, URL, date of creation, and more.

d. Defendant has also offered a Twitter Followers Scraper to automatically collect data such as "name, number of followers, profile URLs, images, company URL, and more." *See also* Bright Data, *Twitter Followers Scraper*, https://perma.cc/92LX-4PVK.

Figure 13: Screenshot of Bright Data's website on July 10, 2023



Twitter Followers Scraper

Scrape Twitter followers and collect data such as: name, number of followers, profile URLs, images, company URL, and more.

__**Answer:**__ The Twitter Scrapers are simply implementations of the Web Scraper and share the same restrictions and limitations. To facilitate and streamline the search design process, Bright Data also offers customers its Twitter Scraper API service, which delivers scraped public data to customers. Sales of such services have been *de minimis*. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

125. For each of these products, Defendant claims it “[u]tilizes proprietary technology to unlock sites” and allows customers to “collect as much data as you need quickly and completely.”

Answer: The statement X purports to cite refers to the amount of data, not the type of data, that can be collected. Bright Data does not allow customers to collect non-public data. As to the amount, this statement simply refers to the fact that Bright Data’s services can be tailored for the amount of data to efficiently complete the task. Virtually all data companies say something similar. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

126. In addition to these X-specific scraping tools, Bright Data offers an automated “Scraping Browser” that simplifies the act of scraping data from the X platform. Bright Data markets this product for scraping X Corp.’s data.

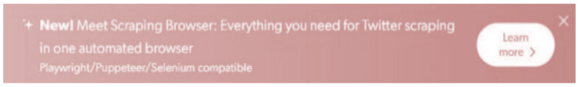


Figure 14: Screenshot of Bright Data’s website on July 10, 2023

New! Meet Scraping Browser: Everything you need for Twitter scraping in one automated browser

Playwright/Puppeteer/Selenium compatible

Answer: Bright Data’s Web Scraping browser is simply an add-on to Bright Data’s Web Unlocker product. As noted, the Web Unlocker product does not permit scraping behind a log-in. Bright Data admits that the Scraping Browser is designed to reduce the need for each customer to engage in complex coding of their own, and thus, “simplifies” the act of lawful scraping of public information. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

127. Defendant advertises this “Scraping Browser” as containing “all website unlocking operations under the hood, including: CAPTCHA solving, browser fingerprinting, automatic retries, selecting headers, cookies, & Javascript rendering, and more.” Defendant also claims its Scraping Browser “automatically learns to bypass bot-detection systems as they adapt, saving you the hassle and cost.”

Answer: As noted, the Scraping Browser is an add-on to the Unlocker product, and thus, contains many of the features and all of the scraping restrictions of that product. The Web Unlocker product uses lawful means to increase the odds of a successful connection despite a website operator’s efforts to obstruct users’ lawful right to search for public information. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

128. The Scraping Browser allows Defendant's customers to "appear as a real user browser to bot-detection system[s]," such as those used by X Corp.

Answer: Bright Data does not know how X's bot-detection systems, if any, work, and thus, does not know how Bright Data's Web Scraper may appear to such systems. But the Web Scraper only uses lawful means to increase the odds of a successful connection despite a website operator's efforts to obstruct visitors' lawful right to search for public information. It does not include any false, misleading, or deceptive statements in its operation, and does not fail to disclose any information that it has a legal duty to disclose. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

3. Bright Data Sells Proxy Services to Facilitate Data-Scraping

129. Bright Data touts that it has more than 72 million "residential IPs" available for use on a rotating basis to support "unlimited concurrent sessions" by data scrapers – that is, massive coordinated scraping. Bright Data, *Residential Proxies*, https://perma.cc/PM7C-75UZ. Bright Data appears to "source" those IPs in two ways. First, it persuades individuals to download EarnApp by paying them incentives; the app then allows Bright Data to route data scrapers' requests through those individuals' IPs to hide the true source of the scraping requests to X's servers. Second, Bright Data bundles its "Bright SDK" with other apps, and when individuals install those other apps, Bright Data can again route data scrapers' requests through those individuals' IPs to hide the true source of the traffic.

Answer: Bright Data admits that its residential proxy network has access to millions of residential IP addresses, though the specific number available at any time varies. Bright Data admits that it leases use of these residential IP addresses from persons or entities that have control over them, and that two of several ways it obtains such leases is through the EarnApp program and the Bright SDK program. As with any VPN or proxy service, a server request initiated through such a VPN or proxy service will reflect the IP addresses of the servers or devices that the VPN or Proxy Service leases, and not the IP address associated with the human initiating the request. The information transmitted with such requests conforms to and complies with all standard Internet Protocols. Bright Data does not include any false, misleading, or deceptive statements in its operation, and does not fail to disclose any information that it has a legal duty to disclose. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

130. Defendant also facilitates violations of X Corp.'s Terms by offering proxy services specifically designed to evade anti-scraping measures, including X Corp.'s CAPTCHAs and its user ID and IP rate limits. These tools allow unregistered users to impersonate registered X users to bypass X Corp.'s digital fence and gates.

Answer: Denied. Bright Data's services do not violate, or facilitate the violation, of, X's Terms. Bright Data's proxy service does not include any Captcha solving capabilities. To the extent Bright Data's Web Unlocker product (which is distinct from its proxy service) involves Captcha solving, the product is limited to the scraping of public information. Moreover, X has not erected any "digital fence" or "gate" around public information on X and Bright Data does not evade any technological measure that restricts access to non-public information. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

131. These proxy services imitate requests from legitimate users to conceal the true requestor's IP address and location. Defendant advertises that these proxies will "avoid[] IP bans and CAPTCHAs" and allow users to "[g]ather vast amounts of public web data with total anonymity."

Answer: Denied. Bright Data and its customers are legitimate internet users, and when scraping X, are legitimate visitors to the X platform. Bright Data has addressed how its technologies operate above, and incorporates its answers here, as applicable. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

132. Bright Data's Proxy Services are intentionally deceptive. Bright Data does not allow data scrapers merely to use one proxy IP – like, for example, a user in another country simulating a U.S. person to access content available in the United States. Rather, Bright Data provides millions of proxies that can be used concurrently and on a rapidly rotating basis. These rotating proxies allow a single data scraper to spread millions of requests across many different IPs and fake accounts. The deceptive proxy services also impede X's efforts to ban requests by individual IPs because, as soon as one IP is banned, Bright Data's software automatically routes the request through a series of new IPs. Although Bright Data's proxy servers make this nearly impossible to track, on information and belief X has blocked numerous accounts for scraping and other violations of its Terms that Bright Data's proxy servers then allowed to regain access to the platform through a different IP. *See* Bright Data, *Proxy Solutions*, https://perma.cc/CL2J-3L3T ("When using Bright Data's award-winning rotating proxies, you can scrape data from any website in the world with a 99.99% success rate. These rotating proxies will constantly replace your IP address, ensuring that you won't get flagged or blocked.").

Answer: Denied. Bright Data's services are not deceptive. Rather, information transmitted with any server request using Bright Data's proxy network or scraping tool conforms to, and complies with, all standard Internet Protocols. Bright Data does not include any false, misleading, or deceptive statements in such requests, and does not fail to disclose any information that it has a legal duty to disclose. With respect to the use of proxies, X appears to concede in this Paragraph that "a user in another country simulating a U.S. person to access content available in the United States" is a permissible use of a proxy service. There is no law, rule, or other practice or custom that suggests using more than one proxy, or using rotating proxies, is deceptive or otherwise impermissible. Indeed, there are many legitimate uses for rotating proxies, including web scraping, data aggregation, ad verification, testing and quality assurance, accessing geo-restricted content, fraud prevention, and academic research.

Bright Data is not aware of any of its customers that have had their accounts blocked on X, and subsequently turned to Bright Data's services to engage in the same conduct. Regardless, any such customer, if one existed, would be limited to lawfully scraping public data.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

133. As Bright Data explains on its website, a "rotating proxy is a proxy server that rotates your IP each time you connect through the proxy server, using proxies from a massive pool of millions of IPs." *See* Bright Data, *Rotating Proxies*, https://perma.cc/S264-G97X. This allows "limitless concurrent connections" that can be used to scrape massive amounts of data. *See* Bright Data, *Residential Proxies*, https://perma.cc/PM7C-75UZ.

Answer: The statement – that a "rotating proxy is a proxy server that rotates your IP each time you connect through the proxy server" – is just definitional. It describes what a rotating proxy is, as distinguished from a static proxy. The definition is consistent with industry custom and understanding of these familiar and frequently employed types of rotating proxy services. The number of proxies that a particular rotating proxy service can draw upon depends on the particular service.

The second sentence – that Bright Data's residential proxy network allows "limitless concurrent connections" – is misleading because it is not a complete quote. The sentence simply means that the service does not impose arbitrary limits on the use of the service, and that multiple searches can be run at the same time. That is, it is no different than having two tabs on your web-browser open at the same time and running in the background. This has nothing to do with the concept of "rotating proxies," which are sequential – not "concurrent" – in nature.

To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

134. There is no legitimate purpose for such tools other than to circumvent X's technological measures. Ordinary proxy services, like a VPN, might be used to maintain a user's privacy in countries with oppressive regimes. That is not what Bright Data markets and sells.

Answer: Bright Data has already addressed these allegations above, and incorporates its answers here, as applicable. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

135. The identity masking that Bright Data's proxy services facilitate further harm X and its users by impeding X from fulfilling its obligations to give users control over their data, including under statutes like General Data Protection Regulation and the California Consumer Privacy Act that require companies to delete consumer data upon request. X cannot ensure that its users' data is deleted once it has been anonymously scraped. By contrast, if the data is obtained through X's API, X has a business relationship with the developer and can demand the deletion of data through established contractual processes.

Answer: The allegations in this Paragraph make no sense. The use of rotating proxies has nothing to do with users' control over their data. To the extent X asserts it is a better steward of privacy than Bright Data, the allegation is denied for the reasons discussed above. To the extent not admitted, Bright Data denies the allegations of this Paragraph.

136. Bright Data does not in its Acceptable Use Policy provide analogous privacy features and control over X user data scraped from the X platform:

a. Bright Data does not require itself or its customers to delete or make private any data scraped from X after an X user has deleted or privatized it;
b. Bright Data does not prohibit tracking of individuals based on the characteristics which X protects;
c. Bright Data does not prohibit matching X usernames to users' legal identities or personally identifying information;
d. Bright Data does not prohibit users from using geodata to track X users, including the creation of heat maps or user location profiles, even in circumstances in which that information is potentially highly sensitive;
e. Bright Data does not restrict scraping from high-risk jurisdictions or take any steps to ensure that X's data is not used by malign actors;
f. Bright Data does not prohibit its scraped data from being used (including by foreign governments) to assist in election interference, voter suppression, or the tracking and targeting of sensitive groups, including activists and political dissidents; and
g. Bright Data does not prohibit its scraped data being used for individual profiling, psychographic segmentation, background and credit checks, or the development of facial recognition.

Answer: Bright Data's Acceptable Use Policy, its Data Protection Addendum, and its other terms and policies contain roughly analogous, if not superior, privacy protections as compared to X's terms, policies, and agreements. To the extent X's policies differ, such differences bear no relation to the use or misuse of data and/or are driven by X's desire to maintain or exploit its information monopoly for commercial purposes— not to protect users from data misuse. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

FIRST CAUSE OF ACTION
(Breach of Contract)

137. X Corp. realleges and incorporates all preceding paragraphs herein.

Answer: Per ECF 156, the Court has dismissed this cause of action with prejudice to the extent it asserts any scraping-based claim, so no answer to the allegations relating to that portion of this Count is necessary. To the extent an answer is required to the allegations in this Count, Bright Data realleges and incorporates its answers to all preceding Paragraphs herein.

138. Use of the X platform and use of X Corp.'s services are governed by X Corp.'s Terms.

Answer: Denied.

139. X users, including Defendant, accept the Terms as a condition of using the platform.

Answer: Denied.

140. Moreover, by virtue of having X accounts, Defendant has expressly accepted and agreed to X Corp.'s Terms.

Answer: Denied.

141. The Terms are enforceable and binding on Defendant.

Answer: Denied.

142. Defendant has repeatedly violated the Terms, including by (i) accessing the X platform through automated means without authorization from X Corp.; (ii) scraping data from the X platform without authorization; (iii) selling tools that enable others, including X users, to access the X platform by automated means and to scrape data; (iv) selling proxy services that enable others, including X users, to access the X platform by automated means and evade X Corp.'s anti-automation and anti-scraping tools; and (v) selling data that Defendant scraped from the X platform.

Answer: Denied.

143. Defendant has breached and continues to breach the Terms by scraping data from X Corp.'s platform without prior consent from X Corp. X Corp. has never authorized Defendant to access its platform through automated means and has never given Defendant consent to scrape data.

Answer: Denied.

144. Despite being bound by the Terms, Defendant has repeatedly accessed the X Corp. platform through automated means and scraped data in violation of the Terms.

Answer: Denied.

145. Defendant has breached, and continues to breach, X Corp.'s Terms by accessing the platform through unauthorized means and scraping data from the platform.

Answer: Denied.

146. Defendant has breached, and continues to breach, X Corp.'s Terms by selling tools that allow other X users to access the platform by automated means and scrape data, and by selling proxy services that allow the same.

Answer: Denied.

147. Defendant has breached, and continues to breach, X Corp.'s Terms by selling data that Defendent has scraped from X Corp.'s platform.

Answer: Denied.

148. Defendant's conduct – both accessing X Corp.'s platform in volumes and manners that violate the Terms as well as selling data scraped from X Corp.'s platform – has damaged X Corp. and caused and continues to cause irreparable harm and injury to X Corp.

Answer: Denied.

149. X Corp. is entitled to compensatory damages, injunctive relief, declaratory relief, and/or other equitable relief.

Answer: Denied.

SECOND CAUSE OF ACTION
(Tortious Interference with Contract)

150. X Corp. realleges and incorporates all preceding paragraphs herein.

Answer: Per ECF 156, the Court has dismissed this cause of action with prejudice to the extent it asserts any scraping-based claim, so no answer to the allegations relating to that portion of this Count is necessary. To the extent an answer is required to the allegations in this Count, Bright Data realleges and incorporates its answers to all preceding Paragraphs herein.

151. All X users must agree to abide by the Terms, which constitute a valid and enforceable agreement between X Corp. and each user.

Answer: Denied.

152. As a user of X Corp.'s platform, as well as a present or former X account holder, Defendant is aware of the Terms and that they govern all users who choose to interact with the X platform. Defendant is also aware of the Terms because several of its executives and employees are present or former X account holders.

Answer: Denied.

153. Nevertheless, Defendant has marketed and sold its scraping tools to X users and account holders, including X users and account holders residing in California, through its interactive website accessible in California and elsewhere, through its sales office and employees in California and elsewhere, and by using the X platform to market its scraping services to other X users and account holders.

Answer: Bright Data admits that it sold scraping services to one or more California residents. Bright Data does not know whether such customers have X accounts or are users of X services, but denies that they are bound by X's Terms just by virtue of their use of Bright Data's services to scrape public information on X. Bright Data's website is a universally-accessible website that does not distinguish between California customers or websites and non-California customers or websites. Bright Data does not have a California sales office or any sales employees responsible specifically for California. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

154. Defendant has also sold proxy services and tools to facilitate automated access and scraping of the X platform by X users and account holders, including by locally offering a "Superior California Proxy" with "[v]ast numbers of California IPs to get data off any website."

Answer: Bright Data offers proxies that are geo-located to all areas of the globe. The use of a proxy has nothing to do with where the customer or host website is located. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

155. By offering services and tools designed to provide automated access to the X platform, and to scrape data from the platform, Defendant induced a breach or disruption of the Terms by X users.

Answer: Denied.

156. On information and belief, those who purchased Defendant's scraping services and tools used them to access X through unauthorized, automated means and to scrape data from the X platform, in violation of the Terms.

Answer: Denied.

157. Defendant's conduct has damaged X Corp. and caused and continues to cause irreparable harm and injury to X Corp.

Answer: Denied.

158. X Corp. is entitled to compensatory damages, injunctive relief, declaratory relief, and/or other equitable relief.

Answer: Denied.

THIRD CAUSE OF ACTION
(Unjust Enrichment, in the alternative)

159. X Corp. realleges and incorporates all preceding paragraphs herein.

Answer: Per ECF 156, the Court dismissed this cause of action with prejudice, and so no answer to this Count is necessary. To the extent an answer is required to the allegations in this Count, Bright Data realleges and incorporates its answers to all preceding Paragraphs herein.

160. If Defendant's acts are found not to be in breach of contract, then Defendant's acts as alleged herein constitute unjust enrichment at X Corp.'s expense.

Answer: Denied.

161. Defendant used X Corp.'s service, platform, and computer network without authorization to scrape data from the X platform.

Answer: Denied.

162. Defendant receives benefits in the form of profits from its unauthorized scraping of data from the X platform.

Answer: Denied.

163. Defendant's retention of the profits derived from its unauthorized scraping of data would be unjust.

Answer: Denied.

164. Defendants' conduct has damaged X Corp., including but not limited to hampering the user experience for authentic X users and customers, in addition to the time and money spent investigating and mitigating Defendants' unlawful conduct.

Answer: Denied.

165. X Corp. seeks actual damages from Defendants' unlawful activities, an accounting, and disgorgement of Defendants' profits in an amount to be determined at trial, compensatory damages, injunctive relief, declaratory relief, and/or other equitable relief.

Answer: Denied.

FOURTH CAUSE OF ACTION
(Trespass to Chattels)

166. X realleges and incorporates all preceding paragraphs herein.

Answer: Bright Data realleges and incorporates its answers to all preceding Paragraphs herein.

167. The X platform and all underlying technological infrastructure are the personal property of X Corp.

Answer: Denied.

168. Defendant intentionally entered into, and made use of, X Corp.'s technological infrastructure, including its software and servers located in California, to obtain information for its own economic benefit.

Answer: Denied.

169. Defendant knowingly exceeded the permission granted by X Corp. to access its personal property, including its technological infrastructure and servers.

Answer: Denied.

170. Defendant's acts have diminished the server capacity that X Corp. can devote to its legitimate users, thereby injuring X Corp. by requiring X Corp. to purchase additional service capacity and diminishing the condition and quality of X Corp.'s service to its legitimate users.

Answer: Denied.

171. Through its acts, Defendant also caused other persons, including X users and account holders based in California and elsewhere, to knowingly exceed the permission granted by X Corp. to access its personal property, further injuring X Corp.

Answer: Denied.

172. X Corp. has never consented to Defendant's conduct.

Answer: Denied.

173. Defendant's conduct constitutes trespass to X Corp.'s chattels.

Answer: Denied.

174. Defendant's acts have caused injury to X Corp. and if continued, expanded, and/or replicated unchecked by others, will cause damage in the form of impaired condition, quality, and value of its servers, technology infrastructure, services, and reputation.

Answer: Denied.

FIFTH CAUSE OF ACTION

(Unlawful, Unfair, or Fraudulent Business Practices (Cal. Bus. & Prof. Code § 17200 et seq.))

175. X Corp. realleges and incorporates all preceding paragraphs herein.

Answer: Per ECF 156, the Court has dismissed the "unfair business acts" portion of this cause of action with prejudice, and so no answer to allegations relating to that portion of this Count is necessary. To the extent an answer is required to the allegations in this Count, Bright Data realleges and incorporates its answers to all preceding Paragraphs herein.

176. Defendant's actions described above constitute unlawful, unfair, and fraudulent acts or practices in the conduct of a business, in violation of California's Business and Professions Code Section 17200, et seq.

Answer: Denied.

177. Defendant's actions violate the Unfair Competition Law's ("UCL") "unlawful" prong because they constitute trespass and tortious interference with business relationships in violation of the law, and violations of the DMCA, CFAA, and CDAFA.

Answer: Denied.

178. Defendant's actions violate the UCL's "unfair" prong because they are unethical, oppressive, unscrupulous, or substantially injurious to consumers and offend established public policy or are immoral. For instance, Defendant circumvents anti-scraping measures designed to protect consumers' privacy rights, including those under the California Consumer Privacy Act, the European Union's General Data Protection Regulation, and other acts with which X Corp. expends substantial resources to comply. The availability of unrestricted third-party access to scraped data also harms X's ability to accurately and effectively prevent and deter market manipulation, scams, and fraud.

Answer: Denied.

179. Defendant's actions violate the UCL's "fraudulent" prong because Defendant and its customers have engaged in widespread scraping of data that is accessible only to X users, developers, or advertisers who are logged into registered, password-protected accounts, and have evaded X Corp.'s

blocking measures by, among other things, creating new fake accounts to engage in the same scraping activity for which a prior account was blocked.

Answer: Denied.

180. Scraping data, as well as circumventing X Corp.'s ability to police its own platform, has caused substantial injury to X Corp., in the form of costs to investigate, remediate, and prevent Defendant's wrongful conduct, among other injuries.

Answer: Denied.

181. As a result of Defendant's various acts and omissions, X Corp. has suffered and continues to suffer irreparable harm for which there is no adequate remedy at law, and which will continue unless Defendant's actions are enjoined.

Answer: Denied.

SIXTH CAUSE OF ACTION
(Misappropriation)

182. X Corp. realleges and incorporates all preceding paragraphs herein.

Answer: Per ECF 156, the Court has dismissed this cause of action with prejudice, and so no answer to this Count is necessary. To the extent an answer is required to the allegations in this Count, Bright Data realleges and incorporates its answers to all preceding Paragraphs herein.

183. X Corp. has invested substantial time, labor, skill, and financial resources into the creation and maintenance of X, its computer systems, and servers, including system and server capacity, as well as the aggregated data at scale. Defendant has not invested any of its own time nor resources to the development of the X platform.

Answer: Bright Data does not know what, if any, time, labor, skill, or financial resources X has invested into X. Bright Data admits that it did not develop or invest in X's platform. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

184. Defendant used automated means – in violation of X Corp.'s Terms – to wrongfully access the X platform, systems and servers, including systems and servers located in California, and obtain aggregated data at scale from the X platform.

Answer: Denied.

185. Defendant appropriated this aggregated data at scale at little or no cost to Defendant, free-riding on X Corp.'s substantial investment of time, effort, and expense to aggregate this data at scale.

Answer: Denied.

186. As a result of Defendant's misappropriation, X Corp. has been forced to expend additional time, labor, skill and financial resources to investigate and mitigate Defendant's wrongful conduct. Defendant has been able to exploit and profit from X Corp.'s substantial investments in the X platform and the creation of its aggregated data at scale.

Answer: Denied.

187. X Corp. has been and will continue to be damaged as a result of Defendant's misappropriation.

Answer: Denied.

188. X Corp. has suffered and will continue to suffer irreparable injury, and its remedy at law is not itself adequate to compensate it for injuries inflicted by Defendant.

Answer: Denied.

SEVENTH CAUSE OF ACTION

(Violation of the DMCA, 17 U.S.C. § 1201(a)(1)(A) and (a)(2))

189. X Corp. realleges and incorporates all preceding paragraphs herein.

Answer: Bright Data realleges and incorporates its answers to all preceding Paragraphs herein.

190. X Corp. owns valid copyrights in its websites, including twitter.com and X.com, and mobile and online applications.

Answer: Bright Data does not know what copyrights X purports to own or whether they are valid. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

191. X Corp. employs technological measures to control access to its websites and applications. These measures include CAPTCHA, login requirements, rate limits, robots.txt restrictions, and anomaly detection tools.

Answer: Denied.

192. These measures require the application of information, or a process or a treatment to gain access to X Corp.'s websites and applications. They require inputting information (CAPTCHA), providing a valid username and password (logins), identifying the IP and/or account of the individual requesting access (rate limits), and use of the service that reflects human use as opposed to automated access (anomaly detection tools).

Answer: Denied.

193. Defendants' tools – including Proxy Solutions, Scraping Browser, Web Unlocker, and Scraper API – circumvent these technological measures by avoiding, bypassing, and impairing each of these measures. That is the primary design and only commercially significant purpose and use of these tools. Defendants' tools automatically crack CAPTCHA prompts; they enable mass use of accounts to bypass rate limits on any individual account; they provide millions of IP addresses on a rotating basis

to evade IP-based rate limits; and they mimic human behavior through automated means to avoid anomaly detection, too.

Answer: Denied.

194. Defendant has used and continues to use automated systems to engage in widespread scraping of data on the X platform, repeatedly bypassing, avoiding, disabling, deactivating, or impairing the technological measures controlling access to X Corp.'s copyrighted websites and applications in violation of 17 U.S.C. § 1201(a)(1)(A).

Answer: Denied.

195. Defendant also offers its tools to the public. These tools are primarily designed to circumvent X Corp.'s technological measures, and do not have more than limited commercially significant purposes other than circumventing X Corp.'s technological measures. Defendant openly markets these tools as being available to circumvent technological measures, including X Corp.'s specifically. This violates 17 U.S.C. § 1201(a)(2)(A).

Answer: Denied.

196. Defendant's acts constituting DMCA violations have been and continue to be performed without the authorization or consent of X Corp. and in violation of its Terms.

Answer: Denied.

197. Defendant has violated Section 1201 of the DMCA willfully.

Answer: Denied.

198. Defendant's conduct has caused damage to X Corp., in the form of costs to investigate, remediate, and prevent Defendant's wrongful conduct, among other injuries, and has unjustly enriched Defendant.

Answer: Denied.

199. X Corp. has suffered and will continue to suffer irreparable injury, and its remedy at law is not itself adequate to compensate it for injuries inflicted by Defendant.

Answer: Denied.

EIGHTH CAUSE OF ACTION

(Violation of the Computer Fraud & Abuse Act
(CFAA) 18 U.S.C. § 1030(a)(2)(C) and (a)(4))

200. X Corp. realleges and incorporates all preceding paragraphs herein.

Answer: Bright Data realleges and incorporates its answers to all preceding Paragraphs herein.

201. The CFAA provides that "[w]hoever . . . intentionally accesses a computer without authorization or exceeds authorized access, and thereby obtains . . . information from any protected computer," is subject both to criminal and civil liability. 18 U.S.C. § 1030(a)(2)(C).

<u>Answer:</u> This Paragraph does not contain any factual allegations, and as such, no response is required. Nonetheless, Bright Data denies that the statute referenced was violated.

202. Defendant has repeatedly and intentionally accessed X Corp.'s servers without authorization and in violation of X Corp.'s Terms and has continued to do so even after X Corp. filed this lawsuit.

<u>Answer:</u> Denied.

203. X Corp.'s servers are a "computer" within the meaning of the CFAA, which defines that term to include "any data storage facility or communications facility directly related to or operating in conjunction with [a computer]." *Id.* § 1030(e)(1).

<u>Answer:</u> This Paragraph states a legal conclusion to which no response is required. To the extent a response is required, Bright Data does not know whether X's servers are covered by the definition of a "computer" under the CFAA.

204. X Corp.'s servers also constitute a "protected computer" within the meaning of the CFAA, because they are connected to the Internet and are used in and affect interstate and foreign commerce and communications. *Id.* § 1030(e)(2)(B).

<u>Answer:</u> Denied.

205. Defendant has circumvented X Corp.'s technological barriers and access restrictions, including CAPTCHAs, login requirements, rate limits, robots.txt restrictions, and anomaly detection tools to engage in widespread scraping of non-public data on X Corp.'s servers that is accessible only to X users, developers, or advertisers who are logged into registered, password-protected accounts.

<u>Answer:</u> Denied.

206. Defendant has also sold and advertised tools – including Proxy Solutions, Scraping Browser, Web Unlocker, and Scraper API – that circumvent X Corp.'s technological barriers and access restrictions and facilitate the scraping of non-public data on the X platform, thereby directing, encouraging, or inducing others to intentionally access X Corp.'s servers without authorization in violation of the CFAA.

<u>Answer:</u> Denied.

207. The CFAA also prohibits "knowingly and with intent to defraud, access[ing] a protected computer without authorization, or exceed[ing] authorized access, and by means of such conduct further[ing] the intended fraud and obtain[ing] anything of value." *Id.* § 1030(a)(4).

<u>Answer:</u> This Paragraph does not contain any factual allegations, and as such, no response is required. To the extent a response is required, Bright Data denies that it has violated this provision.

208. By engaging in misrepresentations, Defendant directly or indirectly accesses X Corp.'s platform and computer network to evade anti-scraping measures and scrape data that is accessible only to X users, developers, or advertisers who are logged into registered, password-protected accounts.

Answer: Denied.

209. Those misrepresentations include using "straw man" accounts and millions of rotating and deceptive IP addresses that mask the true requester. For instance, Defendant's tools circumvent X Corp.'s rate limits and usage restrictions by flooding X Corp.'s servers with requests all originating from the same entity but routed through millions of different IP addresses and/or many different user accounts to conceal that the same entity is making the request.

Answer: Denied.

210. Defendant has caused X Corp. substantial damages and losses, including, without limitation, harm caused by the increased burden on and interruption of service to X Corp.'s website, data and/or underlying databases, amounts expended attempting to prevent the Defendant's unauthorized scraping, and other losses in an amount well over $5,000 aggregated over a one-year period.

Answer: Denied.

211. X Corp. has suffered and will continue to suffer irreparable injury, and its remedy at law is not itself adequate to compensate it for injuries inflicted by Defendant.

Answer: Denied.

NINTH CAUSE OF ACTION
(Violations of the California Comprehensive Computer Data Access and Fraud Act (CDAFA) California Penal Code § 502)

212. The CDAFA provides for criminal and civil liability against "any person" who, among other specified misconduct, (a) "[k]nowingly accesses and without permission takes, copies, or makes use of any data from a computer, computer system, or computer network, or takes or copies any supporting documentation, whether existing or residing internal or external to a computer, computer system, or computer network"; (b) "[k]nowingly and without permission uses or causes to be used computer services"; (c) "[k]nowingly and without permission provides or assists in providing a means of accessing a computer, computer system, or computer network in violation of this section"; or (d) "[k]nowingly and without permission accesses or causes to be accessed any computer, computer system, or computer network." Cal. Penal Code § 502(c).

Answer: This Paragraph does not contain any factual allegations, and as such, no response is required. Nonetheless, Bright Data denies that the statute referenced was violated.

213. X Corp.'s servers constitute and are made up of one or more "computer networks," "computer systems," and "computer programs" or "software," and provide "computer services" to X users, developers, and advertisers.

Answer: Bright Data does not know the constitution of X's servers. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

214. A substantial portion of these networks and systems is located in the State of California.

Answer: Bright Data does not know where X's networks and systems are located. To the extent not expressly admitted, Bright Data denies the allegations of this Paragraph.

215. As set forth above, Defendant has repeatedly and knowingly accessed X Corp.'s servers without X Corp.'s permission and in violation of its Terms and anti-scraping measures, including in California. Among other things, Defendant has circumvented, and caused others to circumvent, X Corp.'s technical and code-based barriers that restrict scraping of non-public data on X's platform.

Answer: Denied.

216. Defendant has caused X Corp. substantial damages and losses, including, without limitation, harm caused by the increased burden on and interruption of service to X Corp.'s website, data and/or underlying databases, amounts expended attempting to prevent the Defendant's unauthorized scraping, and other losses and damage.

Answer: Denied.

217. X Corp. has suffered and will continue to suffer irreparable injury, and its remedy at law is not itself adequate to compensate it for injuries inflicted by Defendant.

Answer: Denied.

PRAYER FOR RELIEF

WHEREFORE, Plaintiff prays for relief, as follows:

1. Preliminary and permanent injunctive relief enjoining Bright Data, its agents, officers, employees and successors from:

a. accessing or using X Corp.'s website, servers, systems, and any data contained therein for purposes of unlawful data scaping;

b. developing or distributing any technology or product that is used, or could be used, for the unauthorized scraping of data from X;

c. facilitating the scraping of data by other users;

d. selling or offering for sale any data previously obtained from X;

e. utilizing any proxies to access X's website, servers, systems, and any data contained therein; and

f. selling or offering for sale any proxies that can be used to access X's website, servers, systems, and any data contained therein.

2. That Defendant be required to identify the location of any and all data obtained from the X platform and to destroy any and all such data;

3. That Defendant be required to identify any and all recipients of data obtained from the X platform;

4. Compensatory, statutory, and punitive damages, as permitted by law and in such amounts to be proven at trial;

5. Reasonable costs, including reasonable attorneys' fees;

6. Pre- and post-judgment interest, as permitted by law;

7. An accounting of Defendant's profits from its scraping activities and disgorgement those profits; and

8. Any other remedy to which Plaintiff X Corp., Inc. may be justly entitled.

Answer: This Paragraph does not contain any factual allegations, and as such, no response is required. To the extent a response is required, Bright Data denies that it is liable and denies that X is entitled to any relief.

AFFIRMATIVE DEFENSES

1. Bright Data asserts the following affirmative defenses. In doing so, Bright Data does not assume any burden of proof, persuasion, or production with respect to any issue where the applicable law places the burden upon plaintiffs.

A. First Affirmative Defense (Lack of Personal Jurisdiction)

2. Bright Data is not subject to the personal jurisdiction of this Court for any of X's state law claims.

3. Bright Data is not found in California. Bright Data is an Israeli company headquartered in Netanya, Israel. It does not have headquarters or a principal place of business in California. Bright Data does not have an office in California.

4. Bright Data does not have sufficient minimum contacts with California to support the exercise of specific personal jurisdiction for any of the asserted state law claims.

5. Bright Data operates a universally accessible global, internet-based communications network. Based in Israel, it allows users from around the world to anonymously search the web, wherever in cyberspace those websites may be located. Its network operates the same way in every zip code, state, and country in the world.

6. Bright Data did not change, alter, or tailor its services for California. It offers proxy services for every jurisdiction. As such, Bright Data did not go beyond the baseline connection that Bright Data's internet presence creates with every jurisdiction through its universally accessible platform.

7. Bright Data did not engage in tortious or unlawful conduct in, or aimed at, California. None of Bright Data's conduct in California bears a direct nexus to X's claims.

8. Bright Data did not knowingly induce any California resident to breach any contract between such resident and X.

9. Bright Data did not cause injury in California. X is not at home or domiciled in California; it is not incorporated in California; and it does not have its principal place of business in California. X has servers located outside of California, and a substantial portion of the search requests it challenges did not touch a California server, and thus, any injury that occurred (including any access-related injury) for such server requests occurred outside of California. To the extent any injury occurred in California, the location of such injury is random, fortuitous, and/or attributed to X's own server-location decisions, and not Bright Data's conduct.

10. The Court lacks supplemental personal jurisdiction because X's federal claims fail to state a claim for which relief can be granted, and its state law claims substantially predominate over X's federal law claims.

11. Bright Data did not consent to personal jurisdiction and did not otherwise waive any defense based on personal jurisdiction.

B. Second Affirmative Defense (Dormant Commerce Clause)

12. X's California state law claims are barred, in whole or in part, by the Dormant Commerce Clause.

13. Bright Data operates a global proxy network and offers its services to customers throughout the world. Bright Data's services allow customers to search the internet from anywhere and to visit websites regardless of where the website operator places its services or chooses to domicile itself. Any state law that seeks to regulate such services places an undue burden on Interstate and Foreign Commerce, in violation of the Dormant Commerce Clause.

14. In particular, the use of California state law to prohibit Bright Data from accessing public information freely available on the internet, or from offering scraping tools or proxy networking services that enable customers to engage in public internet search, places an excessive burden on interstate commerce. The ability of customers to access public information is an important interest, and restrictions on such access constitute a substantial burden on internet commerce.

15. Such restrictions also pose a substantial threat of creating information monopolies in favor of website operators that are domiciled in California, at the expense of out-of-state users and service providers. It is not workable, practical, or financially economical to create different scraping or access tools based on the location of the website operator or the location of its servers. X does not publish the specific locations or server paths it uses to handle server requests. As such, any California law prohibiting automated access or scraping of public information in California is tantamount to a worldwide ban. For this reason, applying California law to prohibit the alleged conduct would frustrate Bright Data's, its customers', and others' ability to operate nationally or internationally.

16. Applying California law to prevent the sale or use of proxy services and scraping tools used to access publicly-available information outside the state of California also unduly impinges on the policy choices of Congress, other states, and other nations who may place greater value on the free flow of information than California.

17. The burdens on interstate commerce are excessive in relation to any local benefits of applying state law to prevent users from using automated means to access publicly-available information. As an initial matter, all of X's state law claims, other than the California CDAFA,

are common law claims not directed specifically to the issue of access or scraping. As such, the legislature has not made any determination that there are any local benefits of prohibiting the automated access or scraping of public information, that such conduct causes any harm in California, or that such harms are sufficient to overcome the burdens on interstate commerce. Moreover, there are less restrictive alternatives that can protect any legitimate state interest since X has the ability to protect any interests it has, including interests in user privacy or server failure, by simply putting information behind a log-in screen. X could then monitor, control, or limit such access simply by tracking account usage.

C. Third Affirmative Defense (Copyright Pre-Emption)

18. Each of X's state law claims are expressly and impliedly pre-empted by the Copyright Act, 17 U.S.C. § 101, et seq., to the extent they relate to, depend upon, or seek damages resulting from any scraping activities.

19. The Court has already dismissed X's claims to the extent they relate to scraping (ECF 83 and 156). To the extent Bright Data is required to assert affirmative defenses relating to any state law claims that relate to, depend upon, or seek damages resulting from any scraping activities, such claims are pre-empted by the Copyright Act, 17 U.S.C. § 101, et seq.

20. X's claims are pre-empted because they seek to assert legal or equitable rights that are equivalent to one or more of the exclusive rights that come within the subject matter of copyright as specified by sections 102, 103, and 106 of the Copyright Act. Enforcement of X's state law claims would also substantially conflict with the purposes and operation of the Copyright Act.

21. Each of X's state law claims are related to scraping, which is the copying and distribution of information appearing on a website. Bright Data does not use, and is not aware of its customers using, its services for the purpose of scraping information in a way not within the general scope of copyright.

22. The information that Bright Data scrapes and (to Bright Data's knowledge, the information its customers scrape) concerns information that X does not have a legally protectible

interest in, including any ownership or copyright rights. With respect to user generated content, X is merely a publisher or distributor of such content and does not own the content or have an exclusive license to it.

23. For data that X's platform generates, such as usage statistics, such information falls within the general scope of copyright because X's website is a broader copyrightable work based on the arrangement and look of the site and because the site and each individual web page contains both user-generated copyrightable content as well as facts and figures.

24. To the extent X's state law claims are related to access, such claims are also expressly and impliedly pre-empted, since any such access is incidental to scraping. Moreover, any access incidental to scraping is not a prohibited activity. X has consented to automated access to the logged-off portions of its website in its Terms, which only purport to prohibit scraping. X has also consented to automated access to the logged-off portions of its website by virtue of its connection of its servers to the public internet.

D. Fourth Affirmative Defense (CUTSA Pre-Emption or Supersession)

25. Each of X's state law tort claims are expressly and impliedly pre-empted by the California Uniform Trade Secrets Act, Cal. Civil Code ("CUSTA") § 3426, et seq., to the extent they relate to, depend upon, or seek damages resulting from any scraping activities.

26. The Court has already dismissed X's claims to the extent they relate to scraping (ECF 83 and 156). To the extent Bright Data is required to assert affirmative defenses relating to any state law statutory or tort claims that relate to, depend upon, or seek damages resulting from any scraping activities, such claims are pre-empted by CUTSA.

27. To the extent X seeks to assert state law claims to prohibit scraping, or seeks damages based on injuries that flow from the act of scraping, such alleged conduct would be covered under misappropriation of trade secrets claims.

28. CUTSA provides the exclusive civil remedy for misappropriation of trade secrets, except to the extent to which the claim falls within its savings clause. The savings clause does not apply to X's statutory or tort law claims.

E. Fifth Affirmative Defense (Lack of Standing)

29. X lacks standing to assert any claims relating to the rights or interests belonging to third-parties, including any intellectual property or privacy rights of persons who post content on X.

F. Sixth Affirmative Defense (Lack of Copyright Interests)

30. To the extent X's claims, including the DMCA claim under 17 U.S.C. § 1201(a)(1)(A) and (a)(2), require a showing that X has a valid and enforceable copyright in the content accessed, such claims fail because X does not possess such rights.

31. X has affirmatively disclaimed any ownership or legally protectable copyright interest in the information accessed or scraped.

32. Bright Data does not scrape any information or content that is subject to X's copyrights in the broader arrangement and look of its website or application.

G. Seventh Affirmative Defense (Lack of Effective Technical Measures)

33. To the extent any of X's access claims are premised on circumvention of technological measures for information not placed behind a log-in, such claims fail because X has not reasonably employed technological measures to prevent such access.

34. Technological measures do not include restrictions placed in the Terms, which may or may not have been "accepted."

35. Anomaly detection tools are not "technological measures" to prevent access, but rather are investigative tools to determine whether there has been access with certain characteristics.

36. Captchas are not effective technological measures because X only uses them in connection with account creation, and therefore, this measure does not apply to logged-off access.

37. Rate limiters are not an effective technological measure because such limits do not apply to persons or entities conducting the search, but only to accounts, devices, or IP addresses. Such limiters also do not, by design, prevent access to the underlying content or information. X has not set any rate limiter to zero.

38. Robots.txt is not a technological measure to prevent access to any website.

39. To the extent that any of the above are deemed to be technological measures, Bright Data has not circumvented them.

H. Eighth Affirmative Defense (Consent)

40. X's claims are barred to the extent that such claims depend on X's lack of authorization or consent to the conduct alleged in the Complaint.

41. X consented to receiving requests for information when it connected its servers to the public internet. Bright Data's conduct falls within the range of conduct associated with ordinary internet usage. Bright Data's conduct was not intended to cause any disruption in X's services or degradation in user experience. And indeed, Bright Data's conduct did not cause any such disruption or degradation.

42. X also consented to automated access through its Terms. The relevant Terms expressly permit "access or search or attempt[s] to access or search the Services by any means (automated or otherwise) … through [its] currently available, published interfaces that are provided by [X]." Bright Data's access, search, or its attempts to access or search X's platform involved only currently available, published interfaces provided by X.

43. X also consented to Bright Data's activities when it accepted money for, and approved, Bright Data's advertisements, and through the placement of such ads, actively encouraged visitors or users to use Bright Data's services.

I. Ninth Affirmative Defense (Termination and Rejection)

44. X cannot assert breach of contract claims for conduct after September 25, 2023.

45. On September 25, 2023, Bright Data affirmatively rejected X's Terms and informed X that it terminated any and all accounts. Specifically, Bright Data informed X that it has "exercised [its] right to terminate the account and any agreement between Bright Data and X. …. [A]ll Twitter or X accounts that are or were owned, controlled, or managed by Bright Data have now been terminated. Bright Data is, therefore, no longer "using" X's services and is now not subject to X's User Agreement. …. To be clear, as of now at least, ***you are on notice that Bright***

Data affirmatively rejects and does not consent to any agreement to your user terms in any context for any reason. Nor will there be any future contract – or "meeting of the minds" – between the parties relating to the activities at issue absent a written agreement physically hand-signed by both parties."

46. On October 12, 2023, Bright Data again informed X that "any future or ongoing claims made by you are foreclosed as we have no contractual relationship from here on out."

47. On February 13, 2024, Bright Data again reiterated that "any claims made by you after that date are foreclosed, as Bright Data has no current, existing, or ongoing contractual relationship with X."

48. On November 13, 2024, Bright Data again repeated that, "this letter shall serve as further formal notice of express rejection of the Terms, any future amendments to the Terms, and any current or future contractual relationship between Bright Data and X."

J. Tenth Affirmative Defense (Inapplicability of Post-Suit Amendments to the Terms)

49. X cannot rely on any post-suit amendments to its Terms.

50. Amendments to X's terms do not purport to apply retro-actively.

51. Prospective enforcement of unilaterally-drafted contracts of adhesion after filing of the initial Complaint in this case, on July 26, 2023, violates public policy.

K. Eleventh Affirmative Defense (X's Termination of the Pre-Suit Terms)

52. X has affirmatively terminated the pre-suit Terms of Service, and therefore, cannot enforce them.

53. The Terms of Service in effect at the time of the lawsuit and Bright Data's account termination and rejection specified that the Terms cease to "govern" any relationship with anyone once X revises its Terms. X revised its Terms on or about September 29, 2023 and November 15, 2024. As such, the May 18, 2023 version of the Terms (and any prior versions of the Terms) do not apply to conduct after September 29, 2023.

L. Twelfth Affirmative Defense (Lack of Consideration)

54. X's Terms are unenforceable because X has not provided any consideration to persons who search the web, including X.com or Twitter.com.

55. X's Terms affirmatively disclaim any consideration being given to persons who do not have an account and exclusively search X's platform in ways that X does not permit. By prohibiting the specific use case that such persons are engaged in, X has not provided any actual performance to such persons. Nor has X made any promise to such persons.

56. Any consideration that X claims supports its Terms was also for X's own benefit and was not bargained for consideration.

M. Thirteenth Affirmative Defense (Unconscionability)

57. X's claims are barred due to the doctrine of unconscionability.

58. X's Terms are procedurally unconscionable. X's Terms are contracts of adhesion, and involve terms unilaterally imposed by X and not as a result of arms-length or good faith bargaining.

59. X's Terms are also procedurally unconscionable to the extent they purport to apply to persons who do not have an account or who have terminated their accounts. X's account creation procedures purport to say that by signing-up for an account, the account creator is agreeing to the Terms. This creates the reasonable expectation that signing-up for and using the account is a condition precedent to the application of the Terms. When opening an account, users reasonably believe that the Terms will only govern their use of the account itself. The average user would not suspect that by agreeing to the Terms they are relinquishing rights to activity unrelated to the use of their account.

60. X's Terms are substantively unconscionable to the extent they purport to, or are deemed to, prevent (i) scraping or automated access to X's website; (ii) scraping or automated access to any public portion of X's website, including any portion of the website that is not placed behind a log-in screen; or (iii) any conduct that does not involve the use of an account.

61. X's Terms unconscionably seek to assert legal rights X does not possess. No law gives X a right to block access or scraping of publicly available information. The public has a right to search for and use information that is publicly available via the Internet. X's Terms unconscionably strip the public of its right to search the Internet.

62. X's Terms are substantively unconscionable to the extent they seek to control logged-off activity because X has not provided any consideration for the restraints it seeks to impose.

63. X's Terms are substantively unconscionable because they constitute an unreasonable restraint of trade.

64. X's September 29, 2023 Terms added a survivability clause, stating that "For the avoidance of doubt, these Terms survive the deactivation or termination of your account." ECF 61-1 at Ex. 4 § 4. This clause was only added after Bright Data terminated its account and rejected the Terms. Enforcement of a survival clause, added only after a party has terminated an agreement, is unconscionable.

65. X's November 15, 2024 Terms added a clause for liquidated damages. To the extent that this provision is deemed to apply to Bright Data, or its customers, the provision is unconscionable. The provision bears no relation to X's actual damages, but rather was intended to be, and operates as, a penalty. The provision also unreasonably restrains trade and interferes with Bright Data's customer relationships.

N. Fourteenth Affirmative Defense (Unreasonable Restraint of Trade)

66. X's Complaint is barred, in whole or in part, as X unreasonably restrained trade.

67. Bright Data incorporates all facts and allegations of its Counterclaim as if fully restated herein.

68. According to X's Terms, users "retain ownership and rights to any of [their] Content [they] post or share." ECF 62-1, Exhibit 4. Through its adhesive Terms, X claims a "broad, royalty-free license to make [users'] Content available to the rest of the world and to let others do the same." *Id.* Although X does not specify that this "license" is exclusive, X

impermissibly uses its Terms to bar users from "crawling or scraping the Services in any form." ECF 62-1, Exhibit 4 § 4.

69. In place of giving others unfettered access to the *public* portions of its website, X decided to monetize its API and began selling API products (including the X API v.2) allowing consumers to "Tap into millions of Posts," "look up X users to analyze networks," "curate and manage lists of accounts," and "identify geographic trends." *See* developer.twitter.com/en/products/twitter-api.

70. Effectively, X's Terms work to give X a complete monopoly over the public data on its site that it simultaneously disavows any ownership interest over. In doing so, X blocks public search and thereby unreasonably restrains trade.

O. Fifteenth Affirmative Defense (Statute of Limitations)

71. The Complaint is barred, in whole or in part, by the applicable statutes of limitations.

P. Sixteenth Affirmative Defense (Laches)

72. X's claims are barred by the equitable doctrine of laches.

73. At all relevant times, X knew that Bright Data offered proxy network services and scraping tools that could, among their many uses, be used for purposes of scraping X.

74. X knew about, encouraged, and profited from Bright Data's conduct.

75. X knew about, encouraged, and profited from automated access and scraping to increase the number of monetizable Active Daily Users.

76. X knew about, encouraged, and profited from such use by accepting advertising, and approving ads, for such services.

77. X actively sought to encourage the widespread distribution of content posted on X, including by other internet companies that engaged in automated access and/or scraping of the X platform.

78. X changed its log-in practices in July 2023, and sued Bright Data a few days later. X knew of Bright Data's activities prior to the log-in changes in July 2023, and did not provide Bright Data notice of its position that such conduct violated the Terms as they existed at that time.

79. It would be inequitable to permit X to recover for, or prohibit, conduct that it encouraged and profited from.

Q. Seventeenth Affirmative Defense (Unclean Hands)

80. X's claims are barred by the doctrine of unclean hands.

81. X, acting for itself and through its affiliates, also engaged in extensive scraping activities, including scraping websites that have Terms similar to X's.

82. X and its affiliates have used Bright Data to scrape such websites. For example, in 2024, xAI, which is partially owned by X, is an alter ego of X, and/or is under common control with X, used Bright Data's services to scrape third-party websites.

83. To the extent that X's claims are premised on the impropriety of Bright Data's proxy network or its scraping tools, X and its affiliates engaged in the same conduct when using Bright Data's services. Thus, if such conduct constitutes unlawful conduct, X engaged in the same unlawful conduct.

84. X's, or its affiliates', use of Bright Data's services to access or scrape websites bears a direct relation to the subject matter of this lawsuit.

85. Allowing X to assert that Bright Data's conduct is lawful when X uses Bright Data's services, but unlawful when Bright Data or any other person does so is inequitable and would undermine the integrity of the judicial process.

R. Eighteenth Affirmative Defense (Choice of Law)

86. X's California state law claims are barred, in whole or in part, because California law does not apply to some or all of the alleged conduct.

87. Bright Data did not consent to the application of California law with respect to the conduct alleged.

88. X has disclaimed or waived application of California law with respect to any conduct that occurred after November 15, 2024. In X's November 15, 2024, version of the Terms, X states that, "The laws of the State of Texas, excluding its choice of law provisions, will govern these Terms and any dispute that arises between you and us." As such, as of November 15, 2024, X's contracts with users that may use Bright Data's proxy network or scraping tools are, at least as of this date, no longer governed by California law.

89. Moreover, as of at least September 2024, X is no longer a California citizen and is not found in California. As such, it has not suffered any injuries in California after that date.

90. Regardless of the effective date, California law does not apply to any scraping activity or access to X's website that occurs outside of California. X operates a global Public Square Platform, with operations, subsidiaries, and servers located in various states and countries throughout the world. Automated scraping requests that do not touch, or cause failure, of a California server owned or operated by X are not injuries that occur in California, and such access is not regulated by California law.

S. Nineteenth Affirmative Defense (Failure to State a Claim)

91. Bright Data incorporates all facts set forth in its Answer to the Complaint and each of the paragraphs of its other Affirmative Defenses, as if re-alleged and fully set forth herein.

92. Each of X's claims fail to state a claim upon which relief could be granted.

PRAYER FOR RELIEF

WHEREFORE, Bright Data respectfully requests that judgment be entered against X, and that Bright Data be awarded costs of suit and such other relief this Court deems equitable and just, as provided by law.

JURY TRIAL DEMANDED

Bright Data demands a trial by jury, pursuant to Rule 38(b) of the Federal Rules of Civil Procedure, of all issues so triable.

Dated: December 17, 2024

Respectfully submitted,

/s/ Colin R. Kass
Colin R. Kass*
Erica T. Jones*
Proskauer Rose LLP
1001 Pennsylvania Ave., N.W.
Washington, D.C. 20004
(202) 416-6890
ckass@proskauer.com
ejones@proskauer.com

David A. Munkittrick*
Reut N. Samuels*
Timothy E. Burroughs*
Michael R. Clifford Beckwith*
Peter C. Angelica*
Proskauer Rose LLP
Eleven Times Square
New York, New York 10036
(212) 969-3000
dmunkittrick@proskauer.com
rsamuels@proskauer.com
tburroughs@proskauer.com
mbeckwith@proskauer.com
pangelica@proskauer.com

Robert C. Goodman (Bar No. 111554)
Lauren Kramer Sujeeth (Bar No. 259821)
Rogers Joseph O'Donnell, PC
311 California Street, 10th Floor
San Francisco, CA 94104
(415) 956-2828
rgoodman@rjo.com
lsujeeth@rjo.com

Sehreen Ladak (Bar No. 307895)
Briana Seyarto Flores (Bar No. 342002)
Proskauer Rose LLP
2029 Century Park East, Suite 2400
Los Angeles, CA 90067-3010
(310) 284-5652
sladak@proskauer.com
bseyartoflores@proskauer.com

Attorneys for Defendant Bright Data Ltd.
*Admitted pro hac vice