UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

REDDIT, INC., Plaintiff, - against - SERPAPI LLC, OXYLABS UAB, AWMPROXY, and PERPLEXITY AI, INC., Defendants.	Case No.: 25-cv-8736 **COMPLAINT** **JURY TRIAL DEMANDED**

Reddit, Inc., by and through its undersigned attorneys, alleges upon knowledge with respect to its own actions, and upon information and belief as to all other matters, as follows:

NATURE OF THE ACTION

1. Reddit, Inc. brings this action to stop the industrial-scale, unlawful circumvention of data protections by a group of bad actors who will stop at nothing to get their hands on valuable copyrighted content on Reddit. Defendants Oxylabs UAB, AWMProxy, and SerpApi–respectively, a Lithuanian data scraper, a former Russian botnet, and a Texas company that publicly advertises its shady circumvention tactics—are data-scraping service providers who specialize in creating and selling tools designed to circumvent digital defenses and scrape others' content. These tools are aimed at bypassing two levels of security: First, evading Reddit's own anti-scraping measures, and second, circumventing Google's controls and scraping Reddit content directly from Google's search engine results. In a very real sense, these Defendants are similar to would-be bank robbers, who, knowing they cannot get into the bank vault, break into the armored truck carrying the cash instead. Meanwhile, Defendant Perplexity AI, Inc., an artificial intelligence ("AI") company that is more akin to a "North Korean hacker," is a willing customer of at least one of its co-defendants and will apparently do anything to get the Reddit

data it desperately needs to fuel its "answer engine"—that is, anything *other than* enter into an agreement with Reddit directly, as some of its competitors have done. Reddit brings this action to stop each of these bad actors.

2. Founded over 20 years ago, Reddit is one of the largest repositories of human conversation in existence. Every day, over 100 million unique users engage in discussions across its hundreds of thousands of interest-based communities (or "subreddits"), a continuous stream of real-time and creative copyrighted works. This vast corpus of authentic human discourse is "like manna from heaven" for the growing number of companies vying for dominance in the AI market. These AI companies, worth up to tens of billions of dollars, desperately need access to more and more high quality, current data to support their ambitions, and Reddit is a "top-cited source" of data for them.

3. But Reddit has rules. It does not permit unauthorized commercialization of Reddit content absent an express agreement with guardrails in place to ensure that Reddit and its users' rights are protected. In short, if AI companies want to legally access Reddit data, they need to comply with Reddit's policies. Some of the largest AI companies, like OpenAI and Google, have done just that, entering into agreements permitting them to access Reddit data while ensuring that the Reddit community is protected from abuse. That is not the path Defendants have chosen.

4. The problem for Defendants, however, is that Reddit is protected by technological-control barriers to prevent unscrupulous scrapers from accessing and stealing data directly from its website. It has also issued cease-and-desist letters to companies that have tried

to get around those barriers, and even sued an AI company worth $183 billion for misappropriating Reddit data.[1]

5. Recognizing that Reddit denies scrapers like them access to its site, Defendants SerpApi, Oxylabs, and AWMProxy scrape the data from Google's search results instead. They do so by masking their identities, hiding their locations, and disguising their web scrapers as regular people (among other techniques) to circumvent or bypass the security restrictions meant to stop them. For example, during a two-week span in July 2025, Defendants SerpApi, Oxylabs, and AWMProxy circumvented Google's technological control measures and automatedly accessed, without authorization, almost *three billion* search engine results pages ("SERPs") containing Reddit text, URLs, images, and videos.

6. Separately, Reddit caught Perplexity red-handed by using the digital equivalent of marked bills (to use the bank robbery analogy) to track Reddit data and confirm that Perplexity was using Reddit data acquired through the scraping of Google SERPs. Perplexity knows that what it is doing is wrong because Reddit told it so in a cease-and-desist letter. Reddit explicitly told Perplexity that (1) Reddit prohibits commercial use of its content absent agreement, (2) this prohibition applies regardless of the means through which Perplexity obtained the data, and (3) technological control systems were in place within Reddit to prevent Perplexity from taking its data. But, rather than respect Reddit and its users' rights, what Perplexity has done in response is simply come up with increasingly devious schemes to circumvent Reddit's security systems and policies. In fact, Perplexity's citations to Reddit

[1] On June 4, 2025, Reddit filed a Complaint against Anthropic for breach of contract, unjust enrichment, trespass to chattels, and unfair competition. *See Reddit, Inc. v. Anthropic, PBC*, Case No. CGC-25-625892 (Super. Ct., S.F. Cty. June 4, 2025), Dkt. 1.

increased *forty-fold* after Reddit told it to stop. And as an advertised client of SerpApi, there can be little doubt where and how Perplexity is getting its illicit Reddit data.

7. Congress has enacted laws to prevent exactly what Defendants are doing: circumventing or bypassing technological measures that effectively control access to copyrighted works. *See* Digital Millennium Copyright Act, 17 U.S.C. § 1201, et seq. Each of the Defendants in this action is profiting by evading technological control measures to access Reddit data it knows it does not have permission to access or use. Because Reddit has always believed in the open internet, it takes its role as a steward of its users' communities, discussions, and authentic human discourse seriously. Through this action, Reddit seeks to end Defendants' circumvention of security measures protecting Reddit data, blatant misuse of Reddit content, and disrespect for its users' rights, all of which harm Reddit and its hundreds of thousands of authentic human communities.

PARTIES

8. Plaintiff Reddit, Inc.'s mission is to empower communities and provide human perspectives to everyone. Reddit is a corporation organized under the laws of Delaware with its principal place of business at 303 2nd Street, South Tower, 5th Floor, San Francisco, California 94107. Reddit maintains an office at One World Trade Center in New York, New York with over 300 employees.

9. Defendant SerpApi is a Texas limited liability company with its principal place of business at 5540 N. Lamar Blvd., Austin, Texas 78756. SerpApi provides products and services to enable web-scraping and allow users to bypass or otherwise circumvent technological control measures that prevent large-scale access and web-scraping.

10. Defendant Oxylabs UAB is a Lithuanian company with its principal place of business at 32 Svitrigailos St. 03230, Vilnius, Lithuania. Oxylabs provides products and services to enable web-scraping and allow users to bypass or otherwise circumvent technological control measures that prevent large-scale access and web-scraping.

11. Defendant AWMProxy is a web domain currently registered in California as of March 30, 2025. AWMProxy was previously operated by a Russian entity in connection with a cybercriminal "botnet" that was shut down in 2021.[2] It has apparently resumed operation, selling access to "proxy services" that are used by entities to conceal their location and hide their identity as industrial-scale scrapers, thereby allowing users to bypass or otherwise circumvent technological control measures designed to prevent unauthorized large scale access and web-scraping.

12. Defendant Perplexity AI, Inc. calls itself an "artificial intelligence powered answer engine." Perplexity is a Delaware corporation with its principal place of business in San Francisco, California. Perplexity maintains an office at 299 Park Ave., New York, New York 10171 with dozens of employees in its New York office. Defendant Perplexity is a customer of Defendant SerpApi.[3]

JURISDICTION AND VENUE

13. Reddit seeks damages, injunctive, and other equitable relief against each of the Defendants under 17 U.S.C. § 1201, et seq.

---

[2] Krebs on Security, *The Link Between AWM Proxy & the Glupteba Botnet* (June 28, 2022), https://krebsonsecurity.com/2022/06/the-link-between-awm-proxy-the-glupteba-botnet/ (last accessed Oct. 20, 2025).

[3] *See* Exhibit A (https://serpapi.com/#features (listing Perplexity as a SerpApi customer) (last accessed Oct. 21, 2025)).

14. This Court has jurisdiction over this civil action under 28 U.S.C. §§ 1331, 1338, and 1367.

15. Venue is proper in this judicial district under 28 U.S.C. §§ 1391 and 1400(a).

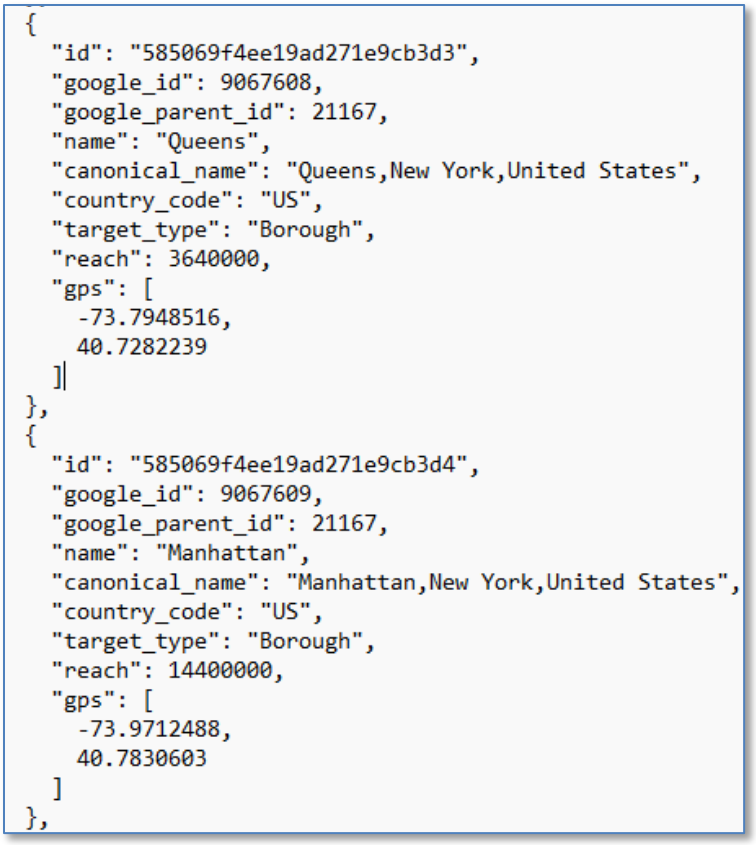
SerpApi Personal Jurisdiction

16. Defendant SerpApi is subject to the jurisdiction of this Court pursuant to N.Y. C.P.L.R. § 302(a)(1), (2), and (3) and the Due Process Clause of the Constitution of the United States. It has purposefully directed its activities at New York and has purposefully availed itself of the benefits of doing business in New York.

17. SerpApi transacts business in the State and District and supplies goods and services in the State and District. SerpApi prides itself on its Google Search API tool, through which users can access and scrape Google search engine results. SerpApi encourages users to input location-specific information, which enables SerpApi to utilize proxy servers[4] near the location specified to provide improved search results. Its website includes a page with all the "SerpApi supported locations," which contains a "Locations" file.[5] That file includes hundreds of locations in the State of New York, including locations in all five boroughs of New York City, including the boroughs of Queens and Manhattan:

[4] A "proxy server" masks a user's IP address, effectively hiding a user's true location and identity from the website being visited. When a user connects to a website through a proxy server, the website will see the proxy server's IP address instead of the user's true IP address, allowing the user to appear to be browsing from near the proxy server's physical location.

[5] SerpApi, *Supported Locations API*, https://serpapi.com/locations-api (last accessed Oct. 20, 2025).



```
{
  "id": "585069f4ee19ad271e9cb3d3",
  "google_id": 9067608,
  "google_parent_id": 21167,
  "name": "Queens",
  "canonical_name": "Queens,New York,United States",
  "country_code": "US",
  "target_type": "Borough",
  "reach": 3640000,
  "gps": [
    -73.7948516,
    40.7282239
  ]
},
{
  "id": "585069f4ee19ad271e9cb3d4",
  "google_id": 9067609,
  "google_parent_id": 21167,
  "name": "Manhattan",
  "canonical_name": "Manhattan,New York,United States",
  "country_code": "US",
  "target_type": "Borough",
  "reach": 14400000,
  "gps": [
    -73.9712488,
    40.7830603
  ]
},
```


18. In addition, various SerpApi customers are headquartered in or have substantial operations in New York, including Perplexity, KPMG, Morgan Stanley, Shopify, and the United Nations.[6]

19. SerpApi is a "Remote-First Company," permitting employees to work from various locations both within the United States and around the world.[7] One or more SerpApi employees live and work in New York, and conduct activities from New York that give rise to Reddit's claims.[8]

---

[6] SerpApi, *They trust us*, https://serpapi.com/ (last accessed Oct. 20, 2025).

[7] SerpApi, *SerpApi Careers*, https://serpapi.com/careers (last accessed Oct. 20, 2025).

[8] SerpApi's website identifies a senior software developer, who has authored various blog posts on scraping Google data, as being located in New York. SerpApi, *Michael Moura*, https://serpapi.com/blog/author/michael_moura/ (last accessed Oct. 20, 2025); *see also* SerpApi (@serp_api), X (May 3, 2024, at 5:33 P.M.),

20. SerpApi unlawfully accessed and obtained Reddit data on computers located in New York and/or by using proxy servers located in New York, and/or by circumventing technological control measures on servers located in New York.

21. This Court's exercise of personal jurisdiction over SerpApi comports with both the New York long-arm statute and the Due Process Clause of the Fourteenth Amendment. SerpApi maintains various contacts within New York from which Reddit's claims arise, and the maintenance of this suit does not offend traditional notions of fair play and substantial justice.

Oxylabs Personal Jurisdiction

22. Defendant Oxylabs is subject to the jurisdiction of this Court pursuant to N.Y. C.P.L.R. § 302(a)(1), (2), and (3) and the Due Process Clause of the Constitution of the United States. It has purposefully directed its activities at New York and has purposefully availed itself of the benefits of doing business in New York.

23. Oxylabs transacts business in the State and District and supplies goods and services in the State and District. Oxylabs purports to operate "the world's largest ethical proxy network" and offers over 62,000 IP addresses located in New York.[9] Oxylabs also boasts to potential customers that its "[h]igh-quality New York proxies with residential IP addresses will help you scale up your data gathering projects and collect publicly available information efficiently":

---

https://x.com/serp_api/status/1786524554434597360. This employee can also be seen in this video posted on X (formerly Twitter) touting recent updates to SerpApi's scraping platforms. https://x.com/serp_api/status/1955290218069627300

[9] Oxylabs, *Proxy locations in the US by largest IP count*, https://oxylabs.io/location-proxy/usa (last accessed Oct. 20, 2025); *see also* Oxylabs, *New York Proxy*, https://oxylabs.io/location-proxy/usa/new-york (last accessed Oct. 20, 2025).

24. Oxylabs also advertises that its "New York proxy network allows you to select proxies from various locations, such as Bronx proxies, Brooklyn proxies, Buffalo proxies, Manhattan proxies, Queens proxies."[10]

25. At least two Oxylabs employees reside in New York.

26. Oxylabs unlawfully accessed and obtained Reddit data on computers located in New York and/or by using proxy servers located in New York, and/or by circumventing technological control measures on servers located in New York.

27. This Court's exercise of personal jurisdiction over Oxylabs comports with both the New York long-arm statute and the Due Process Clause of the Fourteenth Amendment. Oxylabs maintains various contacts within New York from which Reddit's claims arise, and the maintenance of this suit does not offend traditional notions of fair play and substantial justice.

AWMProxy Personal Jurisdiction

28. Defendant AWMProxy is subject to the jurisdiction of this Court pursuant to N.Y. C.P.L.R. § 302(a)(1), (2), and (3). AWMProxy has purposefully directed its activities at New York and has purposefully availed itself of the benefits of doing business in New York.

[10] Oxylabs, *Large pool of New York proxy IPs*, https://oxylabs.io/location-proxy/usa/new-york (last accessed Oct. 20, 2025).

29. AWMProxy transacts business in the State and District and supplies goods and services in the State and District. AWMProxy has previously claimed to have a proxy pool of over 350,000 IP addresses in locations spread throughout the world,[11] including in Brooklyn and New York City.[12] AWMProxy has recently re-established its proxy network, including IP addresses located in New York, to resume its data-scraping operations.

30. AWMProxy unlawfully accessed and obtained Reddit data on computers located in New York and/or by using proxy servers located in New York, and/or by circumventing technological control measures on servers located in New York.

31. This Court's exercise of personal jurisdiction over AWMProxy comports with both the New York long-arm statute and the Due Process Clause of the Fourteenth Amendment. AWMProxy maintains various contacts within New York from which Reddit's claims arise, and the maintenance of this suit does not offend traditional notions of fair play and substantial justice.

Perplexity Personal Jurisdiction

32. Defendant Perplexity is subject to the jurisdiction of this Court pursuant to N.Y. C.P.L.R. § 302(a)(1), (2), (3), and (4) and the Due Process Clause of the Constitution of the United States. It has purposefully directed its activities at New York and has purposefully availed itself of the benefits of doing business in New York.[13]

---

[11] Blackdown, *AWMProxy Review – Affordable Proxies for Micro Tasks*, https://www.blackdown.org/awm-proxy-review/ (last accessed Oct. 20, 2025).

[12] As of 2021, AWMProxy advertised proxies located in Brooklyn and New York City. AWMProxy, *Proxy from United States*, https://web.archive.org/web/20211020154610/https://awmproxy.net/modules/socks/country.php?city=US (last accessed Oct. 20, 2025).

[13] A court within this District recently concluded that personal jurisdiction over Perplexity was proper based on, among other reasons, Perplexity transacting business in New York, noting Perplexity's registration, office, and employees located in New York, its operation of a highly

33. Perplexity is registered as an "active" foreign business corporation in New York and has been registered since April 21, 2023. Perplexity's New York Department of State identification number is 6837808. Perplexity transacts business in the State and District, possesses and maintains an office from which various employees work within the State and District, and supplies goods or services in the State and District, through both the promotion and distribution of its products and the sale of subscription services to customers in this State and District.

34. Perplexity employs numerous individuals in this State and District. These New York-based employees include its Co-Founder and Chief Strategy Officer, business, legal, human resources, product, and engineering staff, as well as staff responsible for content, marketing, and maintaining and improving the customer experience of Perplexity's interactive website and mobile applications, through which it conducts significant business. In addition to Chief Strategy Officer, job titles and descriptions for some of these employees include Commercial Counsel, AI Business Fellow, and Product Marketing Lead.

35. These employees located in this State and District: (a) develop, implement, maintain, and promote the technology and/or business relationships through which Perplexity wrongfully accesses and obtains Reddit's data and then uses that data in the Perplexity platform; (b) support the operation of the web-based and mobile application technology that Perplexity uses to sell and otherwise provide answers and/or information to customers in New York, including by utilizing Reddit's content and services and content authored by Reddit's users; and (c) market Perplexity's products to customers and potential customers, including to business

---

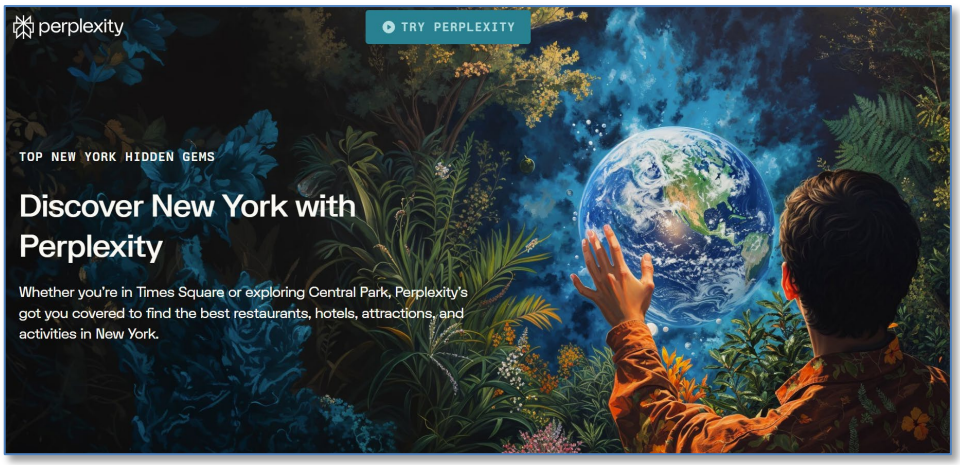
interactive website, and its advertising efforts targeted at New York. *Dow Jones & Co., Inc. v. Perplexity AI, Inc.*, No. 24 Civ. 7984, 2025 WL 2416401 (S.D.N.Y. Aug. 21, 2025).

customers, which results in Perplexity contracting for the sale of goods and services to customers in New York.

36. Perplexity is actively seeking to expand its New York presence by hiring additional employees. Perplexity's career website informs prospective employees that it has an office in New York City. As of October 2025, Perplexity's website listed openings for 27 different positions that were or could be based in its New York City office, including openings for artificial intelligence, machine learning, and various other engineering positions. The responsibilities and duties of these open jobs include (a) developing, implementing, maintaining, and promoting the technology and/or business relationships through which Perplexity wrongfully accesses and obtains Reddit's content and services, and content authored by Reddit's users, and then uses that data in the Perplexity platform; and (b) supporting the operation of the web-based and mobile application technology that Perplexity uses to sell and otherwise provide answers and/or information to customers in New York, including by utilizing Reddit's content and services, and content authored by Reddit's users.

37. Perplexity projects itself directly into New York and aggressively promotes and advertises its products and services in this State and District, including through its dynamic, heavily interactive website (https://www.perplexity.ai and associated subpages) and mobile applications. Perplexity's website targets customers in this State and District with promotional

material tailored to a New York audience, including a web page inviting visitors to "Discover New York with Perplexity":[14]



38. Perplexity's marketing activities include promoting on its Instagram account using a massive billboard in Times Square from September 2024 which reads "Congratulations Perplexity on 250 million questions answered last month!"[15]

39. Perplexity derives substantial revenue from its website and mobile products through activities and customers in New York in at least three ways: (a) directing targeted advertising to New York users based on New York users' queries; (b) selling monthly recurring Perplexity Pro subscriptions to New York-based customers, and tailoring features to these customers based on their location in New York; and (c) providing its products and services to enterprise customers based in New York.

40. Perplexity's actions to unlawfully access, obtain, and utilize Reddit's data occurred, in part, on computers using proxy servers located in New York, and/or by

[14] Perplexity AI, *Discover New York with Perplexity*, https://www.perplexity.ai/encyclopedia/discovernewyork (last accessed Oct. 20, 2025).

[15] Perplexity.ai (@perplexity.ai), Instagram (Sept. 4, 2024) https://www.instagram.com/perplexity.ai/p/C_g2TonSHC5.

circumventing technological control measures and/or accessing data on servers located in New York.

41. This Court's exercise of personal jurisdiction over Perplexity comports with both the New York long-arm statute and the Due Process Clause of the Fourteenth Amendment. Perplexity maintains various contacts within New York from which Reddit's claims arise, and the maintenance of this suit does not offend traditional notions of fair play and substantial justice.

FACTUAL BACKGROUND

I. AI Companies Need Massive Amounts of Data and Target Reddit's Human Content

42. AI companies require staggering amounts of data to operate their products and services, including their large language models ("LLMs"). These companies atomize text and images into individual "tokens" and ingest the tokens into massive AI models defined by billions of parameters.

43. One of the most significant challenges AI companies face is how to procure additional and current data—especially content created by humans—to feed their voracious AI models. Recent trends, including increased competition amongst AI companies both within the United States and across national borders, have only compounded that hunger and driven AI companies to seek more sophisticated, creative, and—for some—underhanded means to obtain data.

44. Reddit is one of the Internet's largest bodies of knowledge and information, with billions of posts and comments. As such, Reddit has emerged as an important source of human conversational data. Founded in 2006, Reddit is a community of communities and one of the largest online discussion platforms in the world, with over 100 million active unique users per

day.[16] Reddit comprises nearly two decades of human conversational data organized across interest-based, user-created communities (referred to as "subreddits") spanning subjects on virtually every topic imaginable, from current events to product reviews to sports and entertainment.

45. Reddit's vast corpus of human-generated content[17] is widely seen as invaluable to AI companies. Reddit data is particularly well-suited to training LLMs because "Reddit data and information constantly grows and regenerates as users come and interact with their communities and each other in genuine and authentic ways." Reddit data also provides "real-time access to evolving and dynamic topics."[18] The largest technology companies in the world, as well as a growing list of new AI companies, require tremendous amounts of data to train and operate their AI products, with a high premium for data generated by actual humans. Reddit has become a

[16] Reddit owns and operates all the "Services" associated with the Reddit community. These include "websites, mobile apps, widgets, APIs, emails, and other online products and services." Reddit also owns the rights to the compilation of Reddit data, as well as all other elements included in the "Services." *See* Reddit, *Reddit User Agreement* (Sept. 24, 2024), https://redditinc.com/policies/user-agreement-september-24-2024#hello-redditors-and-people-of-the-internet-this.

[17] Reddit's User Agreement defines "Content" as "information, text, links, graphics, photos, videos, audio, streams, software, tools, or other materials ('Content'), including Content created with or submitted to the Services by you or through your Account ('Your Content')." *See* Reddit, *Reddit User Agreement*, Preamble & §§ 3, 5 (Sept. 24, 2024), https://redditinc.com/policies/user-agreement-september-24-2024#hello-redditors-and-people-of-the-internet-this. Under the User Agreement, Reddit users retain ultimate ownership of their content and grant Reddit a license with respect to that content. The license grant is subject to both Reddit's Privacy Policy and its Public Content Policy, both of which the User Agreement incorporates by reference. Reddit's User Agreement prohibits users from commercially exploiting either Reddit services or content. This ban includes efforts to "[a]ccess, search, or collect data from the Services by any means (automated or otherwise) except as permitted in these Terms or in a separate agreement with Reddit. . . ." *See* Reddit, *Reddit User Agreement*, (Sept. 24, 2024), https://redditinc.com/policies/user-agreement-september-24-2024#hello-redditors-and-people-of-the-internet-this.

[18] Reddit, Inc. Form S-1 Registration Statement (Feb. 22, 2024), https://www.sec.gov/Archives/edgar/data/1713445/000162828024006294/reddits-1q423.htm.

"top-cited source"[19] of data for these companies. As one AI executive put it, Reddit's body of data is "like manna from heaven."[20]

46. Reddit has entered into partnership agreements with a select list of AI companies that are willing to abide by its policies that protect the intellectual property, privacy, and other rights of Reddit and Reddit users, as set forth in its Public Content Policy.[21] For example, on February 22, 2024, Reddit and Google announced a partnership that enabled programmatic access by Google to Reddit content for use in Google's products and services.

II. Reddit and its Partners Utilize Technological Measures to Control Access and Prevent Scraping

47. Reddit and Google implement technological control measures to effectively control access and prevent scraping[22] of content posted to Reddit.

48. Reddit has and maintains automated anti-scraping systems and teams dedicated to detecting and protecting against unauthorized access to its products and services. Reddit has made significant investments in various anti-scraping tooling and systems that detect and prevent

[19] D. Gallagher, *As AI Changes Internet Search, Reddit Lies in a Sweet Spot*, Wall Street Journal (Aug. 6, 2025), https://www.wsj.com/tech/ai/reddit-rddt-stock-ai-search-2dcc69a4?st=5sRNSY.

[20] S. Needleman, *How Years of Reddit Posts Have Made the Company an AI Darling*, Wall Street Journal (Dec. 10, 2024), https://www.wsj.com/tech/ai/reddit-ai-partnerships-investors-59eef0fe?st=tNB19v.

[21] Reddit's Public Content Policy describes how Reddit protects and licenses public user content. This is to help Reddit users "understand what happens when [they] create, submit, and make content publicly available on the Reddit platform, and how and why Reddit protects and, where appropriate, licenses public content and related data." Reddit, *Public Content Policy*, https://support.reddithelp.com/hc/en-us/articles/26410290525844-Public-Content-Policy (last accessed Oct. 20, 2025). Reddit published this policy in response to "more and more entities using unauthorized access (for example, by scraping or using data brokers) or misusing authorized access to collect public data in bulk, especially with the rise of use cases like generative AI." *Id.*

[22] Scraping refers to an automated process for copying content from a website using web crawlers or "bots."

unauthorized access by scrapers and bots, as well as invested in multiple engineering teams and security professionals dedicated to improving such tooling and systems. Where Reddit's tooling and systems detect unauthorized automated access, its systems issue challenges that must be bypassed before providing any content in response to that network traffic. Some of the measures Reddit employs to prevent unauthorized entities from accessing or scraping its data include industry-standard automation-prevention techniques, such as registered user-identification limits, IP-rate limits, captcha bot protection, and anomaly-detection tools. In other words, as Reddit's current Robots Exclusion Protocol file ("robots.txt") says, "Reddit believes in an open internet, but not the misuse of public content." Reddit's robots.txt file provides clear notice to all scrapers that it prohibits all unapproved web crawlers from accessing any part of Reddit's website. Reddit's User Agreement, which authorized humans agree to when accessing Reddit, also clearly prohibits "scraping the Services without Reddit's prior written consent."[23] Reddit only allows individuals to access its services and content if they agree to Reddit's User Agreement. Reddit prohibits bulk automated access except by separate agreement through defined application programming interfaces ("APIs").

49. Google likewise has anti-scraping systems and teams dedicated to preventing unauthorized access to its products and services. Google operates a search engine which, when a human user submits a query, responds with search engine results pages that include Reddit content when the search results include Reddit results. Google does not permit, and indeed prohibits, unauthorized automated access to its search engine results pages. Google's terms of service prohibit "using automated means to access content from any of our services in violation

[23] Reddit, *Reddit User Agreement*, § 7 (Sept. 24, 2024), https://redditinc.com/policies/user-agreement-september-24-2024#hello-redditors-and-people-of-the-internet-this.

of the machine-readable instructions on our web pages (for example, robots.txt files that disallow crawling, training, or other activities).” Google utilizes a technological access control system called “SearchGuard,” which is designed to prevent automated systems from accessing and obtaining wholesale search results and indexed data while allowing individual users—*i.e.*, humans—access to Google’s search results, including results that feature Reddit data.[24] SearchGuard prevents unauthorized access to Google’s search data by imposing a barrier challenge that cannot be solved in the ordinary course by automated systems unless they take affirmative actions to circumvent the SearchGuard system.

50. Reddit has no agreement or arrangement with any of the Defendants that would permit them to circumvent these technological control measures. Reddit has implemented an API that allows a third party to access Reddit data in bulk so long as the third party agrees to specific restrictions. But access to the Data API requires prior Reddit authorization, user authentication, a unique and descriptive user agent, request rate limits, and limits on data use and retention. Reddit grants Data API users a non-exclusive, non-transferable, non-sublicensable, and revocable license to access and use the Data API in accordance with the terms of use. Unless otherwise allowed, Reddit limits the use of such bulk data through the Data API to “develop, deploy, distribute, and run in” the licensee’s app. Reddit also expressly prohibits using content for any other unapproved purpose, such as for training a machine-learning, LLM, or AI model, unless or until the third party has an agreement to do so. Reddit routinely blocks scrapers that attempt to access and scrape data unless they agree to the privacy and data restrictions in Reddit’s policies, including their agreement to not use crawled data for AI training. If an entity is interested in

[24] Reddit obtained this understanding by subpoenaing Google for information regarding its technological control measures and the identity of any entities circumventing Google’s technological control systems to obtain Reddit content.

using the Data API for a commercial purpose, then it must enter a separate agreement with Reddit. None of the Defendants have entered into any agreement with Reddit with respect to the Data API.

51. Google likewise has no agreement or arrangement with any of the Defendants that would permit any of them to circumvent any of these technological control measures. Google provides APIs for trusted partners for accessing search results in bulk but none of the Defendants have entered into any such agreement or arrangement with Google.

III. Data Scrapers Unlawfully Access and Obtain Reddit Content to Sell to AI Companies

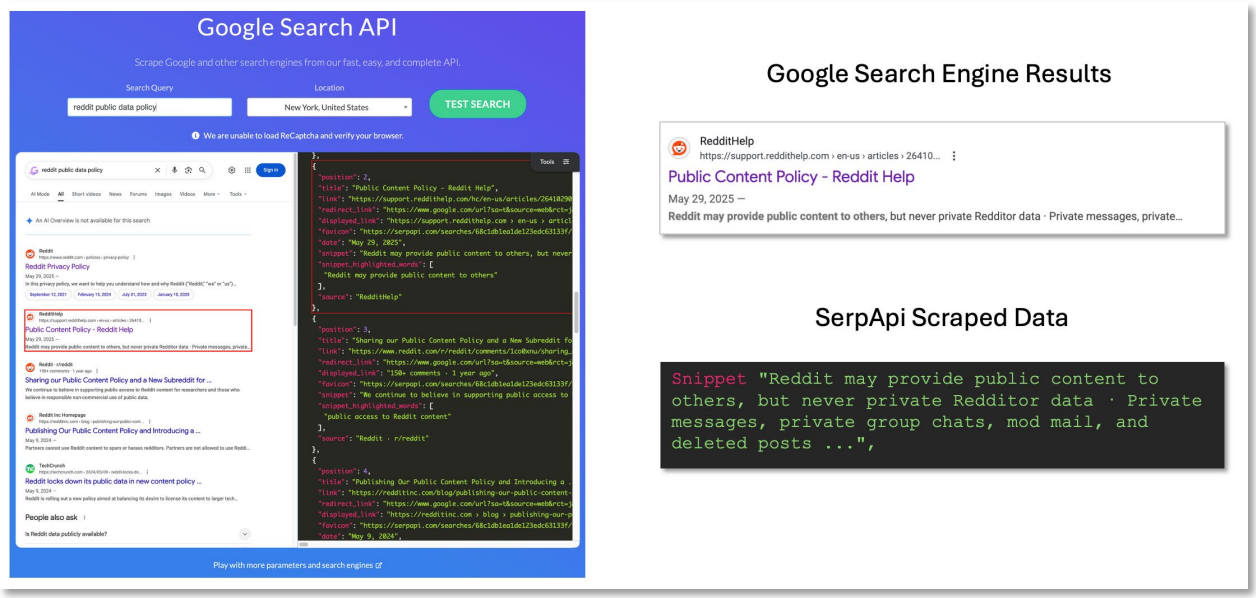
52. To avoid Reddit's restrictions, Defendants SerpApi, Oxylabs, and AWMProxy have engaged in a scheme to obtain Reddit content by circumventing the above technological measures that control access to Reddit content. Defendant data scrapers SerpApi, Oxylabs, and AWMProxy, in an end-run around Reddit's own technological control measures, are instead circumventing Google's technological control measures to access Reddit content in Google search engine result pages.

53. As mentioned above, Google controls access to the data, including Reddit content, that appears in SERPs. Google has implemented various measures and security protocols to prohibit access and scraping by unauthorized, automated entities, including SerpApi, Oxylabs, and AWMProxy.[25] The Defendant data scrapers SerpApi, Oxylabs, and AWMProxy, sidestepping Reddit's own technological control measures, have created tools that circumvent

[25] In fact, "Google executives have privately derided SerpApi" and have "tried various techniques to make it harder for the firm to scrape high-quality information through its web crawler[.]" A. Efrati, *et al.*, *OpenAI is Challenging Google—While Using Its Search Data*, The Information (Aug. 21, 2025), https://www.theinformation.com/articles/openai-challenging-google-using-search-data?rc=zsmcbw.

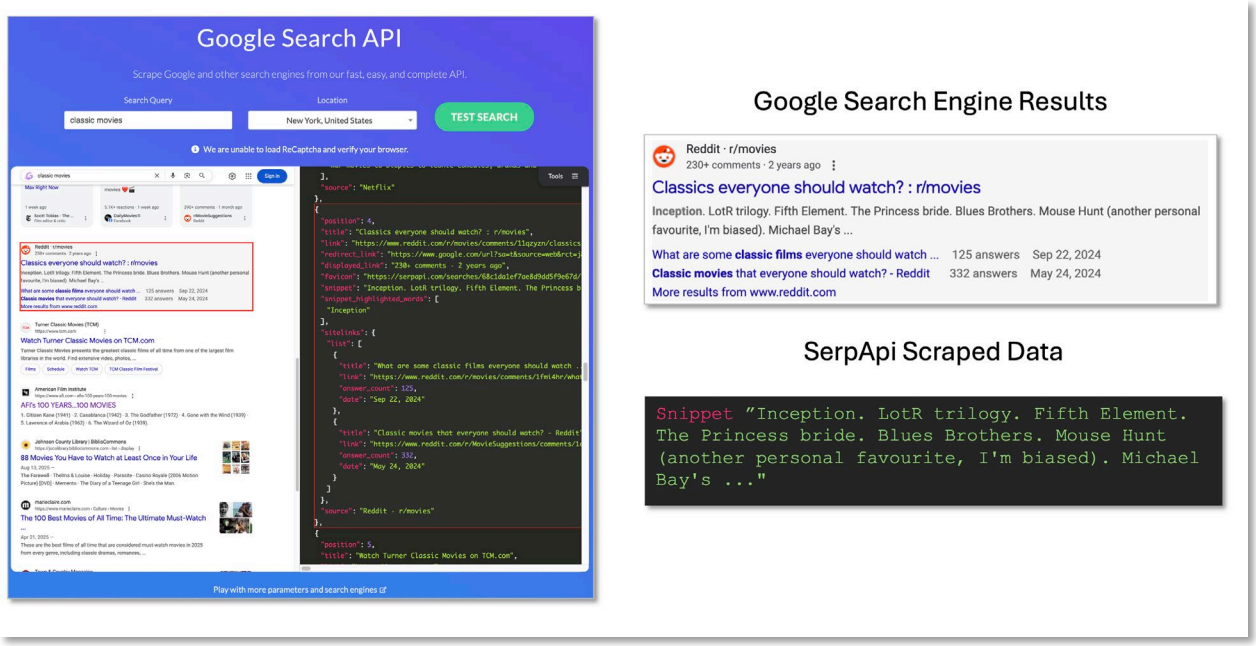
Google's security protocols to access, through unauthorized and automated processes, and collect Reddit's data from Google SERPs.

54. Defendants either use or offer products that make an end-run around Reddit's and Google's restrictions. For example, SerpApi advertises that its product allows users to access and scrape the data that comes back from Google SERPs. Using SerpApi's interface,[26] a search for "Reddit public data policy" generates a list of snippets directly from Reddit, which SerpApi scraped from Google's search results page:



55. SerpApi's tool also accesses and scrapes Reddit data from SERPs even if the query makes no mention of Reddit. For example, a query on SerpApi's interface for "classic movies" also retrieves snippets of Reddit content from the Google SERP:

[26] SerpApi, *Google Search API*, https://serpapi.com (last accessed Oct. 20, 2025).



56. Google's terms of service prohibit "using automated means to access content from any of our services in violation of the machine-readable instructions on our web pages (for example, robots.txt files that disallow crawling, training, or other activities)."[27] The scraping examples noted above of Google's SERPs occur through Defendants bypassing these access restrictions.

57. Despite Google's technological control measures, SerpApi, Oxylabs, and AWMProxy all have products or services that circumvent these measures – indeed, both Oxylabs and SerpApi boast about their ability to bypass these protections.

58. For example, Oxylabs is explicit that its scraping service is meant to circumvent Google's technological measures controlling access to SERPs, featuring a page on its website

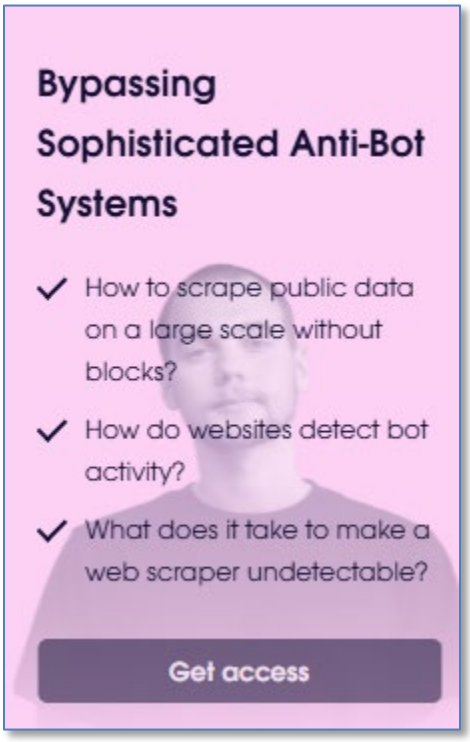
[27] Google, *Terms of Service* (May 22, 2024), https://policies.google.com/terms?hl=en-US. One observer has specifically suggested that SerpApi's actions "might run afoul of Google's terms of service." A. Efrati, *et al.*, *supra*.

called "How to Scrape Google Search Results" that specifically notes that its service "bypass[es]" these "restrictions":

> [I]f you need to scrape Google search pages automatically and on a large scale, you may run into some difficulties. Using a reliable free proxy list or any other proxy option can help you bypass restrictions and avoid getting blocked while scraping.[28]

Oxylabs then notes that "Oxylabs' Google Search API is constructed to bypass the technical challenges Google has implemented."[29]

59. Oxylabs' website also includes tutorials from its employees on how to use their services to "bypass" sophisticated anti-bot systems:[30]



[28] R. Aukstikalnyte, *How to Scrape Google Search Results: Python Tutorial*, Oxylabs (Jan. 28, 2025), https://oxylabs.io/blog/how-to-scrape-google-search-results.

[29] *Id.*

[30] Oxylabs, *Learn From Experts and Master Your Web Scraping Skills*, https://experts.oxylabs.io/ (last accessed Oct. 20, 2025).

60. Similarly, SerpApi describes on its website how to scrape Google Search Results. SerpApi tells its users that with its services, they "don't need to care" about technological control measures that block scrapers, including "HTTP requests, parsing HTML files to JSON (parser), or captcha, IP address, bots detection, maintaining user-agent, HTML headers, [or] being blocked by Google."[31] In order to bypass these technical measures and "*to avoid being detected and blocked by Google*, SerpApi uses a proxy and the latest technologies to mimic human behavior."[32] SerpApi provides its users with tips to reduce the chance of being blocked while web scraping, such as by sending "fake user-agent string[s]," shifting IP addresses to avoid multiple requests from the same address, and using proxies "to make traffic look like regular user traffic" and thereby "impersonate" user traffic.[33] SerpApi markets this tool as providing a way to scrape Google web searches "at scale" for use in training LLM or other AI models.[34] SerpApi says that it can run more than 100,000 searches per hour.[35]

61. SerpApi has also promoted its "Ludicrous Speed" and "Ludicrous Speed Max" features as effective ways to circumvent Google's technological control measures.[36] Ludicrous Speed Max "use[s] 4x the server resources to automatically create numerous parallel requests

---

[31] SerpApi, *How to Scrape Google Search Results (SERPs) – 2025 Guide*, https://serpapi.com/blog/how-to-scrape-google-search-results-serps-2023-guide/ (last accessed Oct. 20, 2025).

[32] *Id.* (emphasis added).

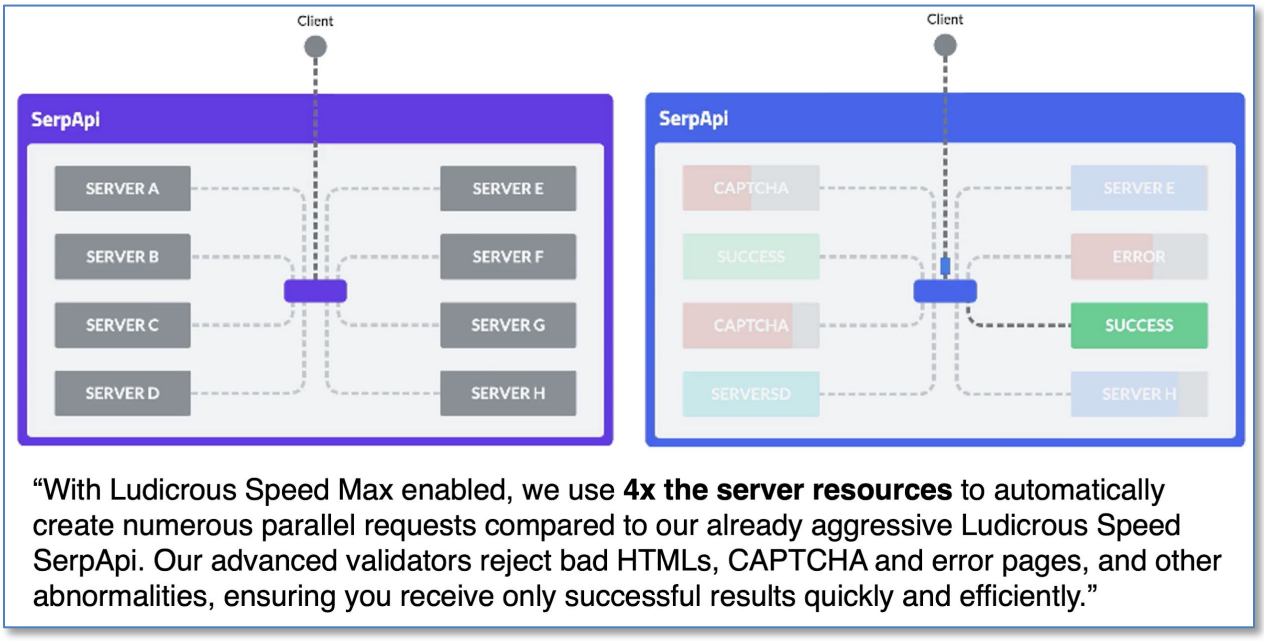
[33] D. Zub, *Reducing the chance of being blocked while web scraping*, SerpApi (Sept. 16, 2021), https://serpapi.com/blog/how-to-reduce-chance-of-being-blocked-while-web/.

[34] SerpApi, *Data for Machine Learning and AI Models*, https://serpapi.com/use-cases/machine-learning-and-artificial-intelligence (last accessed Oct. 20, 2025).

[35] SerpApi, *FAQ*, https://serpapi.com/faq (last accessed Oct. 20, 2025).

[36] R. Schafer, *Google Now Requires JavaScript*, SerpApi (Jan. 17, 2025), https://serpapi.com/blog/google-now-requires-javascript/.

compared to our already aggressive Ludicrous Speed."[37] Ludicrous Speed Max operates by "reject[ing]" what SerpApi characterizes as "bad HTMLs, CAPTCHA and error pages, and other abnormalities." In other words, SerpApi uses a server-swarm to hide from, avoid, or simply overwhelm by brute force effective measures Google has put in place to ward off automated access to search engine results. SerpApi's website depicts how its server army circumvents Google technological controls on behalf of its "Client":



62. SerpApi's CEO and Founder, Julien Khaleghy, has described SerpApi's circumvention methods as including a tool for "creating fake browsers using a multitude of IP addresses that Google sees as normal users."[38]

[37] SerpApi, *Ludicrous Speed Max*, https://serpapi.com/ludicrous-speed-max (last accessed Oct. 20, 2025).

[38] S. Palazzolo, *Meet the Google-Scraping Startup Used by ChatGPT, Cursor and Perplexity*, The Information (Aug. 28, 2025), https://www.theinformation.com/articles/meet-google-scraping-startup-used-chatgpt-cursor-perplexity?rc=zsmcbw.

63. According to information obtained through a subpoena issued to Google, the three Defendant scraping companies have been circumventing or bypassing Google's technological control measures to access and scrape Reddit data from Google search results at an industrial scale:

Total Google SERPs with Reddit Data Automatedly Accessed Without Authorization in a Two-Week Period		
	July 1-6, 2025	**July 7-13, 2025**
SerpApi	784,006,306	1,057,744,585
Oxylabs	333,022,219	448,235,070
AWMProxy	217,954,929	264,232,011

64. The three companies are only able to obtain Reddit data from Google SERPs because they circumvent the technological control measures Google has in place to prevent such unauthorized automated access and scraping. In particular, Google Search has technological control measures to validate that a query comes from a real user and not automated software. SerpApi, Oxylabs, and AWMProxy employ various measures to avoid, bypass, disable, or impair these measures, including by forging credentials or manipulating Google's control measures into incorrectly identifying their scraping services as individual users, not automated scrapers, and/or evading or otherwise circumventing Google's control measures by submitting repeated queries through large volumes of proxy servers.

IV. AI Companies Like Perplexity Unlawfully Circumvent Technological Protections to Access and Obtain Reddit Content Despite Promising Not to Do So

65. Perplexity is an AI company that purports to be "the world's first answer engine" by gathering insights from "top tier sources." In reality, Perplexity does nothing groundbreaking. Its answer engine simply uses a different company's LLM to parse through a massive number of Google search results to see if it can answer a user's question based on those results. But

Perplexity can only run its "answer engine" by wrongfully accessing and scraping Reddit content appearing in Google's own search results from Google's own search engine.

66. Perplexity admits that one of its "top tier sources" is Reddit, noting in an August 2025 blog post that "Reddit has emerged as the most cited domain across AI models globally" based on an analysis of ChatGPT, Google AI Overviews, and Perplexity citations.[39]

67. Perplexity created its "answer engine" by utilizing content from original sources such as Reddit and placing that data into a massive "retrieval-augmented generation" ("RAG") database, which Perplexity then combines with another company's LLM to repackage original source data into "answers" to user queries. In other words, Perplexity's business model is effectively to take Reddit's content from Google search results, feed them into a third party's LLM, and call it a new product. While that business model has somehow translated into a $20 billion valuation, it has not resulted in a willingness to pay for what others (including Google) have.

68. In May 2024, Reddit sent a cease-and-desist letter to Perplexity, demanding that it stop scraping Reddit data. Reddit further demanded that Perplexity cease using Reddit data unless it entered a license agreement expressly authorizing that use. Reddit also noted that Perplexity must stop using Reddit data for commercial purposes regardless of the means by which Perplexity obtained Reddit content. At that time, Perplexity told Reddit that it was not using Reddit content to train any AI models and that it would respect Reddit's robots.txt

[39] Perplexity, *Reddit overtakes Wikipedia as top source for AI model citations* (Aug. 5, 2025), https://www.perplexity.ai/discover/you/reddit-overtakes-wikipedia-as-UUKumG5zRRiU8oZHcu_gDA (last accessed Oct. 20, 2025).

directives to not scrape Reddit content. Perplexity claims that it "respects robots.txt directives," and "only crawls content in compliance with robots.txt."[40]

69. After Reddit sent its cease-and-desist letter to Perplexity, the volume of Reddit data cited by Perplexity did not decrease. Just the opposite: Perplexity *increased* the volume of citations to Reddit forty-fold.[41] The increase was so dramatic that an outside observer hypothesized that the increase was due to Perplexity entering a licensing deal with Reddit and thereby obtaining full access to Reddit's data. In truth, there is no license between Perplexity and Reddit but rather a scheme by Perplexity to obtain Reddit's data through the circumvention of the technological measures protecting Reddit data.

70. Perplexity or its agents are working with one or more of SerpApi, Oxylabs, and AWMProxy to circumvent the technological control measures protecting Reddit content in order to scrape Reddit content in bulk from Google SERPs.

71. To confirm this hypothesis, Reddit created a "test post" – the equivalent of a digital "marked bill" – that could only be crawled by Google's search engine and was not otherwise accessible anywhere on the internet. Within hours, queries to Perplexity's "answer engine" produced the contents of that test post. The only way that Perplexity could have obtained that Reddit content and then used it in its "answer engine" is if it and/or its Co-Defendants scraped Google SERPs for that Reddit content and Perplexity then quickly incorporated that data into its answer engine.

---

[40] Jennifer, *How does Perplexity follow robots.txt*, Perplexity, https://www.perplexity.ai/help-center/en/articles/10354969-how-does-perplexity-follow-robots-txt (last accessed Oct. 20, 2025).

[41] N. Bluman, *Did Perplexity Sign a Deal with Reddit?*, Xfunnel (Apr. 15, 2025), https://www.xfunnel.ai/blog/perplexity-reddit-rumor.

72. Public reports suggest Perplexity will often resort to this sort of "stealth behavior" in order to gain access to content it covets. Perplexity has been known to modify its user agent, change operational IP addresses, and ignore or seek to evade directives against crawling of other websites. As just one example, Cloudflare recently noted in an update on its blog that "Perplexity is using stealth, undeclared crawlers to evade website no-crawl directives."[42] Cloudflare's August 2025 update goes on:

> We see continued evidence that Perplexity is repeatedly modifying their user agent and changing their source ASNs to hide their crawling activity, as well as ignoring—or sometimes failing to even fetch—robots.txt files.[43]

73. These practices led Cloudflare's CEO to say that Perplexity's practices better resemble "North Korean hackers" than the practices of a "reputable AI company."[44]

74. Perplexity has continued with its improper and unlawful attempts to access and obtain Reddit data by bypassing or otherwise circumventing the technical measures in place at Google to prevent access to and widespread scraping of Reddit data from SERPs, including by entering into contracts or other agreements with the other Defendants to obtain or otherwise facilitate accessing and scraping of Reddit data from SERPs.

V. Defendants' Actions Have Harmed Reddit

75. Defendants' activities are unauthorized and unlicensed by Reddit and are not governed by the guardrails that Reddit insists on to protect Redditors and their communities.

---

[42] G. Corral, *et al.*, *Perplexity is using stealth, undeclared crawlers to evade website no-crawl directives*, Cloudflare (Aug. 4, 2025), https://blog.cloudflare.com/perplexity-is-using-stealth-undeclared-crawlers-to-evade-website-no-crawl-directives/.

[43] *Id.*

[44] M. Kan, *Cloudflare: Perplexity AI Acts Like North Korean Hackers, Ignoring Scraping Blocks*, PC Magazine (Aug. 4, 2025), https://www.pcmag.com/news/cloudflare-perplexity-ai-acts-like-north-korean-hackers-ignores-scraping.

76. Reddit is harmed by Defendants' conduct in various ways. Reddit depends on the contributions of its many communities of Redditors, who care about how their content will be treated when it is posted to Reddit. When they post their copyrighted content to Reddit, Redditors place great trust in Reddit; they expect Reddit to respect their wishes and take its role as stewards of the content seriously. By bypassing all of the lawful means for obtaining this content, Defendants also bypass all of the strict restrictions placed on Reddit's licensed business partners that protect Redditors and their anonymity and privacy. As articulated in its Public Content Policy, Reddit believes in an open internet, but it "do[es] not believe that third parties have a right to misuse public content just because it's public." Defendants' conduct directly impacts Reddit's reputation, its users' trust, and its ability to operate as a business.

77. Reddit's business and reputation are damaged by Defendants' misappropriation of Reddit data and circumvention of technological control measures because Defendants deprive Reddit of the ability to control (a) what entities have access to its data, (b) how that data is used, and (c) whether Reddit data is handled in a manner consistent with its Privacy Policy, Public Content Policy, and User Agreement. Reddit has entered into licensing agreements with several AI companies for access to the same data that Defendants improperly access and scrape. Permitting Defendants' conduct to continue will discourage those entities from abiding by their existing agreements, including the guardrails in those agreements that protect Redditors, and encourage them to instead use similar techniques as Defendants to gain access to the content from Reddit. Reddit has lost licensing fees or other commercial opportunities it would obtain from these and new arrangements, and its reputation and business relationship with existing licensees is harmed by third parties gaining access to the data outside of Reddit's established licensing program.

78. Reddit is further harmed by Defendants' conduct because it is forced to invest significant resources into hardware, software, and personnel to improve its technical security systems, surveillance, and anti-scraping efforts to try and prevent the unauthorized accessing and use of Reddit data. It also forces Reddit to require its partners to invest in similar ever-increasingly expensive efforts so as to protect against Defendants' malfeasance.

79. These and various other harms to Reddit will continue unabated so long as Defendants' schemes for unlawfully accessing and scraping Reddit data persist.

COUNT I – DIGITAL MILLENNIUM COPYRIGHT ACT
Circumvention of Technological Control Measures (17 U.S.C. § 1201(a)(1)(A))

[All Defendants]

80. Reddit repeats and realleges the allegations in the previous paragraphs as if fully set forth herein.

81. Reddit and Google have implemented technological measures that effectively control access to Reddit content. Both companies use advanced technological techniques, as described above, to control unauthorized, automated access to their server systems. These measures, in the ordinary course of their operation, limit the freedom and ability of users to access Reddit content, including by prohibiting automated entities from accessing search engine result pages and scraping search engine results that include Reddit content.

82. Defendants' actions violate 17 U.S.C. § 1201(a)(1)(A), under which no person shall circumvent a technological measure that effectively controls access to a copyrighted work. Defendants have circumvented these measures in one or more ways, including:

a. Avoiding or bypassing Reddit's measures entirely in order to obtain Reddit's content and services, and the content authored by its users, that appear in Google search results; and

b. Avoiding, removing, deactivating, impairing, and/or bypassing SearchGuard and Google's other technological control measures by using devices, systems, processes, and/or protocols, including large-volume proxy networks, to improperly gain access to Google Search results.

83. Defendants have also substantially aided, assisted, or contributed to circumvention of these measures, with knowledge that circumvention would occur, by entering into contracts or agreements with one another aimed at circumventing technological control measures in order to access and scrape Google search results featuring Reddit's content and services, or content authored by Reddit users; by supplying products or services that enable that circumvention; and/or by paying for or otherwise financing the use of such products or services.

84. Pursuant to 17 U.S.C. § 1203(a), Reddit brings this action because it has been injured by Defendants' violations of 17 U.S.C. § 1201(a)(1)(A). Reddit has been injured in one or more ways, including that Defendants have unlawfully circumvented technological measures which effectively control access to Reddit's content and services, and the content authored by its users, and in so doing have accessed and acquired Reddit data outside the legitimate licensing channels Reddit has established containing appropriate contractual guardrails to protect Reddit's users and its users' rights.

85. As a result, Reddit has suffered damages, including, but not limited to, lost profits and business opportunities, reputational harm, and loss of user trust.

COUNT II – DIGITAL MILLENNIUM COPYRIGHT ACT

Trafficking of Technology, Product, Service, or Device for Use in Circumventing Technological Measure Controlling Access (17 U.S.C. § 1201(a)(2))

[SerpApi and Oxylabs]

86. Reddit repeats and realleges the allegations in the previous paragraphs as if fully set forth herein.

87. Defendants SerpApi's and Oxylabs' actions violate 17 U.S.C. § 1201(a)(2), under which:

> No person shall manufacture, import, offer to the public, provide, or otherwise traffic in any technology, product, service, device, component, or part thereof, that:
>
> (A) is primarily designed or produced for the purpose of circumventing a technological measure that effectively controls access to a work protected under this title;
>
> (B) has only limited commercially significant purpose or use other than to circumvent a technological measure that effectively controls access to a work protected under this title; or
>
> (C) is marketed by that person or another acting in concert with that person with that person's knowledge for use in circumventing a technological measure that effectively controls access to a work protected under this title.

88. As detailed above, SerpApi provides products and services to facilitate web-scraping, including products and services purposefully designed to circumvent effective technological measures that would otherwise prevent access for large scale automated web-scraping. SerpApi specifically designs and markets its products and services, including Ludicrous Max and Ludicrous Speed Max, for use in circumventing effective technological measures including Google's SearchGuard, which controls access to search engine results featuring Reddit's content and services, and the content authored by Reddit users.

89. Defendant Perplexity has purchased one or more of SerpApi's services or products in order to circumvent Reddit's and/or Google's technological control measures.

90. As detailed above, Oxylabs offers a proxy service that provides web-scraping solutions through a network of proxies that allow for human-like scraping without IP blocking, thereby bypassing or circumventing technical restrictions that would otherwise prevent access for web-scraping. Oxylabs specifically designs and markets its products and services for use in circumventing technological measures including Google's SearchGuard, which controls access to search engine results featuring Reddit's content and services, and the content authored by Reddit users.

91. Pursuant to 17 U.S.C. § 1203(a), Reddit brings this action as someone who has been injured by Defendants' violations of 17 U.S.C. § 1201(a)(2). Reddit has been injured in one or more ways, including that Defendants have unlawfully circumvented technological measures which effectively control access to Reddit's content and services, and the content authored by its users, and in so doing have accessed and acquired Reddit data outside the legitimate licensing channels Reddit has established containing appropriate contractual guardrails to protect Reddit's and its users' rights.

92. As a result, Reddit has suffered damages, including, but not limited to, lost profits and business opportunities, reputational harm, and loss of user trust.

COUNT III – DIGITAL MILLENNIUM COPYRIGHT ACT

Trafficking of Technology, Product, Service, or Device for Use in Circumventing Technological Measure Protecting Right of Copyright Owner (17 U.S.C. § 1201(b))

[SerpApi and Oxylabs]

93. Reddit repeats and realleges the allegations in the previous paragraphs as if fully set forth herein.

94. Defendants SerpApi's and Oxylabs' actions violate 17 U.S.C. § 1201(b), under which:

> No person shall manufacture, import, offer to the public, provide, or otherwise traffic in any technology, product, service, device, component, or part thereof, that:
>
> (A) is primarily designed or produced for the purpose of circumventing protection afforded by a technological measure that effectively protects a right of a copyright owner under this title in a work or a portion thereof;
>
> (B) has only limited commercially significant purpose or use other than to circumvent protection afforded by a technological measure that effectively protects a right of a copyright owner under this title in a work or a portion thereof; or
>
> (C) is marketed by that person or another acting in concert with that person with that person's knowledge for use in circumventing protection afforded by a technological measure that effectively protects a right of a copyright owner under this title in a work or a portion thereof.

95. As detailed above, SerpApi provides products and services to facilitate web-scraping, including products and services designed to circumvent technological measures that would otherwise prevent, restrict, or limit the exercise of a right of a copyright owner. SerpApi specifically designs and markets its products and services, including Ludicrous Speed and Ludicrous Speed Max, for use in circumventing technological measures including Google's SearchGuard, which otherwise prevents the unauthorized and large-scale automated copying and reproduction of search engine results featuring Reddit's content and services and the content authored by Reddit users.

96. Defendant Perplexity has purchased and/or utilized one or more of SerpApi's services or products in order to circumvent Google's technological control measures to copy Reddit data.

97. As detailed above, Oxylabs provides products and services to facilitate web-scraping through a network of proxies that allow for human-like scraping without IP blocking, thereby bypassing or circumventing technical restrictions that would otherwise prevent, restrict, or limit the exercise of a right of a copyright owner. Oxylabs specifically designs and markets its products and services for use in circumventing technological measures including Google's SearchGuard, which otherwise prevents the unauthorized and large-scale automated copying and reproduction of search engine results featuring Reddit's content and services and the content authored by Reddit users.

98. Pursuant to 17 U.S.C. § 1203(a), Reddit brings this action as someone who has been injured by Defendants' violations of 17 U.S.C. § 1201(b). Reddit has been injured in one or more ways, including that Defendants have unlawfully circumvented technological measures which effectively protect rights in Reddit's content and services, and the content authored by its users and in so doing have accessed and acquired Reddit data outside the legitimate licensing channels Reddit has established containing appropriate contractual guardrails to protect Reddit's and its users' rights.

99. As a result, Reddit has suffered damages, including, but not limited to, lost profits and business opportunities, reputational harm, and loss of user trust.

COUNT IV – UNFAIR COMPETITION

[All Defendants]

100. Reddit repeats and realleges the allegations in the previous paragraphs as if fully set forth herein.

101. Defendants, through their actions as outlined above, have misappropriated Reddit's labor, skill, expenditures, and/or goodwill, and have displayed bad faith in doing so.

102. By circumventing technological control measures described above in order to access and scrape Reddit data on a large-scale, unauthorized, and automated basis, Defendants have brazenly disregarded technological control measures in which Reddit and Google have invested considerable resources, and in doing so misappropriated real-time Reddit content and services and the timely content authored by Reddit users for Defendants' own commercial gain.

103. SerpApi and Oxylabs have also marketed, sold, and provided products and services for the purpose of circumventing the technological control measures described above in order to access and scrape timely Reddit data on a large-scale, unauthorized, and automated basis, for their own commercial gain.

104. Defendants have taken the actions outlined above in order to secure for themselves an undue competitive advantage, or to damage Reddit's competitive position, including by (a) accessing and scraping timely Reddit data on an unauthorized, large-scale, and automated basis, thereby avoiding the need for any licensing agreement or other commercial arrangement to which Reddit would otherwise be entitled in the marketplace, and reducing the value and desirability of such agreement or arrangements; (b) removing the need for users to access Reddit content directly through Reddit, and thus damaging the commercial utility of Reddit's website, mobile application, and other services; and (c) jeopardizing users' ability to protect privacy interests in and access to their content, and thus threatening Reddit's ability to attract and retain users.

105. Defendants' conduct violates federal law, as detailed above, and is independently unfair conduct that also violates the state common law of New York.

106. As detailed above, Defendants have acted in bad faith by knowingly circumventing the technological control measures at issue and lauding their ability to do so,

and/or by continuing to circumvent the technological control measures after having been blocked from access or otherwise requested to cease their activities.

107. As a result, Reddit has been injured and suffered damages, including, but not limited to, lost profits, business opportunities, reputational harm, and loss of user trust.

COUNT V – UNJUST ENRICHMENT

[All Defendants]

108. Reddit repeats and realleges the allegations in the previous paragraphs as if fully set forth herein.

109. Defendants have been unjustly enriched at Reddit's expense, and equity and good conscience militate against Defendants retaining the benefits they have obtained.

110. Reddit has expended substantial resources to implement technological control measures, policies, and commercial arrangements aimed at preventing access to and scraping of Reddit data on a large-scale, unauthorized, and automated basis, including real-time Reddit content and services and the timely content authored by Reddit users.

111. Through circumventing technological control measures described above, Defendants have gained access to and scraped Reddit data on a large-scale, unauthorized, and automated basis, including misappropriation of real-time Reddit content and services and the timely content authored by Reddit users, from which Defendants have been unjustly enriched at Reddit's expense.

112. By marketing, selling, and providing products and services for the purpose of circumventing the technological control measures described above in order to access and scrape Reddit data on a large-scale, unauthorized, and automated basis, SerpApi, Oxylabs, and AWMProxy have been unjustly enriched at Reddit's expense. Defendant Perplexity has likewise

been unjustly enriched by obtaining Reddit data through its arrangement with SerpApi and/or the other Co-Defendants.

113. As a result, Reddit has been injured and suffered damages, including, but not limited to, lost profits, business opportunities, reputational harm, and loss of user trust.

114. Defendants' conduct violates federal law, as detailed above, and is independently unjust conduct that also violates the state common law of New York.

115. Equity and good conscience demand that the unjust benefits that Defendants have received should be repaid to Reddit.

COUNT VI – CIVIL CONSPIRACY

[SerpApi and Perplexity AI]

116. Reddit repeats and realleges the allegations in the previous paragraphs as if fully set forth herein.

117. SerpApi and Perplexity have entered into one or more contracts or business agreements for the purpose of circumventing the technological control measures described above in order to gain access to Reddit data on a large-scale, unauthorized, and automated basis, including Reddit content and services and the content authored by Reddit users.

118. Both SerpApi and Perplexity have committed one or more overt acts in furtherance of their agreement, including selling or providing services or products designed to conduct the circumvention described above, paying for such services or products, and conducting the circumvention described above. Specifically, Perplexity has contracted with SerpApi for the purpose of acquiring bulk data, which SerpApi unlawfully acquired by circumventing Google's technological control measures, and which Perplexity then uses in its commercial offerings.

119. Each of SerpApi and Perplexity intended by their actions to participate in a plan or scheme designed to circumvent the technological control measures described above in order to

gain access to Reddit data on a large-scale, unauthorized, and automated basis, including Reddit content and services and the content authored by Reddit users.

120. Furthermore, Perplexity knows that Reddit objects to Perplexity's unauthorized use of Reddit data. Reddit told Perplexity in 2024 that Perplexity must cease and desist its use of Reddit data in Perplexity's commercial offerings but Perplexity ignored Reddit's demand, instead obtaining Reddit data through its scheme with SerpApi.

121. SerpApi's and Perplexity's actions led to the commission of one or more torts or other unlawful actions, including those outlined above.

122. As a result, Reddit has been injured and suffered damages, including, but not limited to, lost profits, business opportunities, reputational harm, and loss of user trust.

PRAYER FOR RELIEF

WHEREFORE, Reddit demands a trial by jury and a judgment against Defendants as follows:

a. Injunctive relief enjoining Defendants from:

i. accessing or using Reddit's website, servers, systems, and any data contained therein for the purpose of unlawful data scraping;

ii. accessing or using Google's website, servers, systems, and any data contained therein for the purpose of unlawful scraping of Reddit data;

iii. developing or distributing any technology or product that is used for the unauthorized circumvention of technological control measures and scraping of Reddit data;

iv. using any Reddit data previously obtained through the circumvention of technological control measures in place to

prevent the unauthorized and automated scraping of Reddit data; and

v. selling or offering for sale any Reddit data previously obtained through the circumvention of technological control measures in place to prevent the unauthorized and automated scraping of Reddit data.

b. Awarding Reddit damages for all injuries suffered as a result of Defendants' unlawful acts, including:

i. actual, statutory, and/or other compensatory damages as determined by a jury;

ii. disgorgement of Defendants' ill-gotten gains from unauthorized commercial exploitation of Reddit's data; and

iii. pre-judgment and post-judgment interest, in an amount to be determined at trial.

c. Costs of this action, including attorneys' fees; and

d. Such other legal and equitable relief as this Court deems just and proper.

Dated: October 22, 2025
New York, New York

Respectfully submitted,

/s/ William A. Maher

William A. Maher
Justin T. Zimnoch
WOLLMUTH MAHER & DEUTSCH LLP
500 Fifth Avenue
New York, NY 10110
(212) 382-3300
wmaher@wmd-law.com
jzimnoch@wmd-law.com

Reid M. Bolton (*Pro Hac Vice* Forthcoming)
Matthew R. Ford (*Pro Hac Vice* Forthcoming)
William D. Gohl (*Pro Hac Vice* Forthcoming)
BARTLIT BECK LLP
54 W. Hubbard Street, Suite 300
Chicago, IL 60654
(312) 494-4400
reid.bolton@bartlitbeck.com
matthew.ford@bartlitbeck.com
will.gohl@bartlitbeck.com

Attorneys for Plaintiff Reddit, Inc.