NO. 17-16783

IN THE UNITED STATES COURT OF APPEALS
FOR THE NINTH CIRCUIT

HIQ LABS, INC.,

PLAINTIFF-APPELLEE,

v.

LINKEDIN CORPORATION,

DEFENDANT-APPELLANT.

On Appeal from the United States District Court
for the Northern District of California
Case No. 3:17-cv-03301-EMC
The Honorable Edward M. Chen, District Court Judge

BRIEF OF AMICI CURIAE ELECTRONIC FRONTIER FOUNDATION, DUCKDUCKGO, AND INTERNET ARCHIVE IN SUPPORT OF PLAINTIFF-APPELLEE

Jamie Williams
Corynne McSherry
Cindy Cohn
Nathan Cardozo
ELECTRONIC FRONTIER FOUNDATION
815 Eddy Street
San Francisco, CA 94109
Email: jamie@eff.org
Telephone: (415) 436-9333

Counsel for Amici Curiae

DISCLOSURE OF CORPORATE AFFILIATIONS AND OTHER ENTITIES WITH A DIRECT FINANCIAL INTEREST IN LITIGATION

Pursuant to Rule 26.1 of the Federal Rules of Appellate Procedure, Amici Curiae Electronic Frontier Foundation, DuckDuckGo, and Internet Archive each individually state that they do not have a parent corporation and that no publicly held corporation owns 10 percent or more of their stock.

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Cases

Statutes

Legislative Authorities

Other Authorities

STATEMENT OF INTEREST[1]

The Electronic Frontier Foundation ("EFF") is a nonprofit, member-supported civil liberties organization working to protect rights in the digital world. With over 37,000 active donors and dues-paying members, EFF represents the interests of technology users in court cases and broader policy debates surrounding the application of law in the digital age. EFF's interest in this case is in the principled and fair application of computer crime laws generally, and the Computer Fraud and Abuse Act ("CFAA") specifically, to online activities and systems—especially as it impacts Internet users, innovators, and security researchers. EFF has served as counsel or amicus curiae in key cases addressing the CFAA and/or state computer crime statutes, including *United States v. Nosal*, 676 F.3d 854 (9th Cir. 2012) ("*Nosal I*") (en banc) (amicus); *United States v. Nosal*, 844 F.3d 1024 (9th Cir. 2016) ("*Nosal II*") (amicus); *Facebook, Inc. v. Power Ventures, Inc*., 844 F.3d 1058 (9th Cir. 2016) (amicus); *United States v. Valle*, 807 F.3d 508 (2d Cir. 2015) (amicus); and *United States v. Auernheimer*, 748 F.3d 525 (3d Cir. 2014) (co-counsel).

DuckDuckGo is an Internet search engine that distinguishes itself from other search engines by protecting the privacy of its users. In particular, DuckDuckGo

[1] Pursuant to Federal Rule of Appellate Procedure 29(c)(5), no one except for Amici or their counsel authored this brief in whole or in part, or contributed money towards its preparation. Both parties consent to this brief's filing.

does not track user search histories or otherwise profile its users, which enables DuckDuckGo users to search anonymously and avoid the filter bubble of personalized search results. DuckDuckGo was founded in 2008 and is headquartered in Paoli, Pennsylvania.

The Internet Archive is a public nonprofit organization founded in 1996 to build an "Internet library," with the purpose of offering researchers, historians, scholars, artists, and the general public permanent access to historical collections in digital format. Located in San Francisco, California, the Internet Archive receives data donations and collects, records, and digitizes material from a multitude of sources, including libraries, educational institutions, government agencies, and private companies. The Internet Archive then provides free public access to its data—which include text, audio, video, software, and archived Web pages.

INTRODUCTION

This case is about whether private companies can use the Computer Fraud and Abuse Act ("CFAA"), a blunt and outdated 1986 criminal "anti-hacking" statute intended to target malicious computer breaks-ins, to police who gets to access publicly available data on the open Internet, and how. The stakes of this dispute thus go far beyond a skirmish between two commercial services. Open access is a hallmark of today's Internet, and one of the main reasons the Internet has become our "modern public square." *See Packingham v. North Carolina*, 137 S. Ct. 1730, 1732 (2017). The power to limit access to publicly available information on the Internet under color of the law should be dictated by carefully considered rules that balance the various competing policy interests. These rules should not allow the handful of companies that collect massive amounts of user data to reap the benefits of making that information publicly available online—*i.e.*, more Internet traffic and thus more data and more eyes for advertisers—while at the same time limiting use of that public information via the force of criminal law.

The Court should find that the CFAA cannot be used to enforce use restrictions on publicly available information on the open Internet—including restrictions on using automated scripts to access that publicly available information—for three reasons:

First, this Court's holding in *United States v. Nosal*, 676 F.3d 854 (9th Cir. 2012) ("*Nosal I*") (en banc), forecloses LinkedIn's claim that it can rescind a party's authorization to access publicly available information in writing, without placing that information behind a true access barrier, and thereby render their subsequent access of that publicly available information criminal under the CFAA. As *Nosal I* held, the CFAA's "purpose is to punish hacking"—breaking into a private computer system in order to access or alter non-public information—not to create "a sweeping Internet-policing mandate" by punishing those who violate corporate computer use policies. *Id*. at 858, 863. Allowing websites to use the CFAA as a terms of service enforcement mechanism would do precisely what this Court in *Nosal I* sought to avoid: it would "transform the CFAA from an anti-hacking statute into an expansive misappropriation statute" for enforcing computer use policies across the Web. *See id*. at 857. Indeed, LinkedIn's position presents the very "tension" this Court identified in *Facebook, Inc. v. Power Ventures, Inc*., 844 F.3d 1058, 1067 & n.1 (9th Cir. 2016), between *Nosal I's* holding that terms of use cannot be the basis for CFAA liability and purported revocations of authorization based on terms of use violations.

Second, imposing CFAA liability for accessing publicly available information via automated scripts would potentially criminalize all automated "scraping" tools—including societally valuable tools that Internet users,

researchers, and journalists around the world rely on every day. Automated scraping is the process of using Internet "bots" to extract content and data from a website.[2] LinkedIn paints bots as nefarious agents of devious actors, but bots are an essential component of the Internet. The Web crawlers that power tools we all rely on every day, including Google Search and Amici DuckDuckGo and Internet Archive, are Internet bots. News aggregation tools, including Google's Crisis Map, which aggregates critical emergency information, are Internet bots. A "bot" is merely a software application that runs automated tasks over the Internet,[3] and software applications can be used for good as well as bad. In 2016, "good bots" were responsible for 23 percent of global Web traffic.[4] Imposing CFAA liability for automated Internet scraping would chill the development and use of these societally valuable bots.

Third, by criminalizing what is in fact a common and critical online practice, imposing CFAA liability for automated Internet access would fail to provide notice

[2] Web scraping can also be done manually by any Internet user, via copy-and-paste.

[3] Wikipedia, "Internet Bot" (last updated Nov. 10, 2017), https://en.wikipedia.org/wiki/Internet_bot. Bots are not the same as botnets. A botnet, a portmanteau of "robot" and "network," is a network of private computers or devices infected with malicious software and controlled without the owners' knowledge.

[4] *See* Igal Zeifman, Bot Traffic Report 2016, Incapsula (Jan. 24, 2017), https://www.incapsula.com/blog/bot-traffic-report-2016.html.

of what types of automated access are (and are not) criminal, risk arbitrary prosecution, and render the CFAA unconstitutionally vague.

LinkedIn wants to "participate in the open Web" but at the same time abuse the CFAA to avoid "accept[ing] the open trespass norms of the Web." Orin S. Kerr, *Norms of Computer Trespass*, 116 Colum. L. Rev. 1143, 1163 (2016). The Court should not allow it. Instead, it should affirm the district court's finding that hiQ demonstrated a likelihood of success on the merits of its CFAA claim[5] and find that the CFAA cannot be used to enforce restrictions on using automated scripts to access publicly available information on the open Internet.[6]

///

///

///

[5] This amicus brief addresses only the first element of the preliminary injunction standard, that "[a] plaintiff seeking a preliminary injunction must establish that [it] is likely to succeed on the merits[.]" *See Winter v. Natural Res. Def. Council, Inc*., 555 U.S. 7, 20 (2008).

[6] This finding would not prevent LinkedIn from metering access to its website, such as by implementing measures to prevent harmful intrusions, limit abusive behavior, stop malicious hackers and identity thieves, or block DDoS attacks. It would mean only that violations of contractual restrictions on using automated tools to obtain publicly available information cannot give rise to criminal CFAA liability. *See hiQ Labs, Inc. v. LinkedIn Corp*., No. 17-CV-03301-EMC, 2017 WL 3473663, at *8 (N.D. Cal. Aug. 14, 2017) (noting that such a finding "is not to say that a website like LinkedIn cannot employ, e.g., anti-bot measures to prevent, e.g., harmful intrusions or attacks on its server").

ARGUMENT

I. ACCESSING PUBLICLY AVAILABLE INFORMATION ON THE INTERNET CANNOT GIVE RISE TO CFAA LIABILITY.

A. The CFAA Was Meant to Target Computer Break-Ins.

The CFAA makes it a crime to "intentionally access[] a computer without authorization or exceed[] authorized access, and thereby obtain[] . . . information from any protected computer"—which includes any computer connected to the Internet. 18 U.S.C. §§ 1030(a)(2)(C), 1030(e)(2)(B).[7]

Congress intended this language "to address 'computer crime,' which was then principally understood as 'hacking' or trespassing into computer systems or data." *United States v. Valle*, 807 F.3d 508, 525 (2d Cir. 2015) (citing H.R. Rep. No. 98–894, 1984 U.S.C.C.A.N. 3689, 3691–92, 3695–97 (1984); S. Rep. No. 99–432, 1986 U.S.C.C.A.N. 2479, 2480 (1986)). As this Court recognized, Congress sought to "target hackers who accessed computers to steal information or to disrupt

[7] The first incarnation of the computer crime statute—enacted in 1984—was a narrow statute intended to criminalize unauthorized access to computers to obtain national security secrets or personal financial and consumer credit information, or to "hack" into government computers. *See* Comprehensive Crime Control Act of 1984, Pub. L. No. 98-473, 98 Stat. 2190, codified at 18 U.S.C. § 1030(a)(1)–(3). After multiple revisions, "protected computer" now includes not merely computers "used in interstate or foreign commerce or communication," but computers "used in or *affecting* interstate or foreign commerce or communication." 18 U.S.C. § 1030(e)(2)(B) (emphasis added). The practical effect of this seemingly small change allows the CFAA to reach computers as far as the Commerce Clause can extend. Orin S. Kerr, *Vagueness Challenges to the Computer Fraud and Abuse Act*, 94 Minn. L. Rev. 1561, 1570 (2010).

or destroy computer functionality, as well as criminals who possessed the capacity to 'access and control high technology processes vital to our everyday lives[.]'" *LVRC Holdings LLC v. Brekka*, 581 F.3d 1127, 1130–31 (9th Cir. 2009) (quoting H.R. Rep. 98–894, 984 U.S.C.C.A.N. 3689, 3694 (1984)). The statute's legislative history "consistently characterizes the evil to be remedied—computer crime—as 'trespass' into computer systems or data, and correspondingly describes 'authorization' in terms of the portion of the computer's data to which one's access rights extend." *Valle*, 807 F.3d at 525; *see also Cvent, Inc. v. Eventbrite, Inc.*, 739 F. Supp. 2d 927, 932 (E.D. Va. 2010) ("The CFAA is a civil and criminal anti-hacking statute designed to prohibit the use of hacking techniques to gain unauthorized access to electronic data.").[8]

The House Committee Report to the original 1984 computer crime bill explained, "the conduct prohibited is analogous to that of 'breaking and entering'"—and not "using a computer (similar to the use of a gun) in committing the offense." H.R. Rep. 98-894, 1984 U.S.C.C.A.N. 3689, 3706 (1984). As an example of what Congress intended to target, the Report identified an individual who had "stole[n] confidential software" from a previous employer "by tapping into the computer system of [the] previous employer from [a] remote terminal."

[8] The Senate Committee Report to the bill's 1986 amendments "specifically described 'exceeds authorized access' in terms of trespassing into computer systems or files." *Valle*, 807 F.3d at 525 (citing S. Rep. No. 99–432, 1986 U.S.C.C.A.N. 2479, at 2483).

Id. at 3691–92. According to the Report, the individual would have escaped federal prosecution—despite a clear computer break-in—had he not made two of his fifty access calls from across state lines. *Id*. at 3692.[9]

The Report called for a statutory solution to ensure that such computer intrusions would not evade prosecution. It referred to a "recent flurry of electronic trespassing incidents" and described "so-called 'hackers'" who had been able to "access (trespass into) both private and public computer systems, sometimes with potentially serious results" thanks to the "proliferation of computer networking[.]" *Id*. at 3695, 3696. It was this sort of technical, exploitative behavior—breaking into a private computer system for the purpose of accessing or altering non-public information—that Congress sought to outlaw.

B. The CFAA Must Be Interpreted Narrowly to Avoid Transforming the Statute Into an All-Purpose Mechanism For Enforcing Computer Use Policies.

Congress sought to address a narrow problem, not create "a sweeping Internet-policing mandate." *See Nosal I*, 676 F.3d at 858. But Congress passed the CFAA when the Internet was in its infancy, and its attempt to take on computer

[9] The House Committee Report also (incorrectly) characterized the 1983 techno-thriller film *WarGames*—in which a young Matthew Broderick breaks into a U.S. military supercomputer programmed to predict possible outcomes of nuclear war and unwittingly almost starts World War III—as "a realistic representation of the automatic dialing and access capabilities of the personal computer." H.R. Rep. 98-894, 1984 U.S.C.C.A.N. 3689, 3696 (1984).

breaks-ins so early in the Internet's lifecycle resulted in a vague and ill-defined statute. As a result, to avoid perverting the CFAA into a broad policing tool, courts must take care to interpret the statute in light of its core purpose.

Courts should, but sometimes do not, take particular care in interpreting the meaning of "authorized access." While the CFAA defines "exceeds authorized access,"[10] it does not define its most critical term: access "without authorization." At the start of 1986, the total number of networks connected via the Internet was a mere 2,000.[11] Accessing another person's computer was relatively rare. Today, it's possible to check your email—and thus access information on a distant server, *i.e.*, someone else's computer—from a remote campsite, at 30,000 feet above sea level or deep underwater, and even from low Earth orbit.[12] In a world where it is difficult to go a single day, or even sometimes a single waking hour, without "accessing" someone else's computer system, the precise meaning of "without authorization . . . has proven to be elusive.'" *See EF Cultural Travel BV v. Explorica, Inc.*, 274 F.3d 577, 582 n.10 (1st Cir. 2001).

---

[10] To exceed authorized access is "to access a computer with authorization and to use such access to obtain or alter information in the computer that the accessor is not entitled so to obtain or alter[.]" 18 U.S.C. § 1030(e)(6).

[11] Computer History Museum, "Internet History 1962 to 1992," http://www.computerhistory.org/internethistory/1980s/.

[12] Adrienne LaFrance, "The Internet in Space? Slow as Dial-Up," *The Atlantic* (June 11, 2015), https://www.theatlantic.com/technology/archive/2015/06/the-internet-in-space-slow-dial-up-lasers-satellites/395618/.

In interpreting this language, four courts of appeal have strayed from Congress' intent, broadly interpreting "without authorization" and "exceeds authorized access" to include violations of corporate computer use policies.[13] But this Court explicitly rejected their broad reading of the CFAA. *See Nosal I*, 676 F.3d at 862–63 (rejecting *John*, *Citrin*, and *Rodriguez* for failing to "construe ambiguous criminal statutes narrowly so as to avoid 'making criminal law in Congress's stead'") (quotation omitted); *Brekka*, 581 F.3d at 1135 ("[W]e decline to adopt the interpretation of 'without authorization' suggested by *Citrin*.").

Instead, this Court, en banc—along with the Second and Fourth Circuits and various other district courts[14]— narrowly interpreted the CFAA to ensure that the

---

[13] *See, e.g.*, *United States v. John*, 597 F.3d 263, 272–73 (5th Cir. 2010); *Int'l Airport Ctrs. v. Citrin*, 440 F.3d 418, 420 (7th Cir. 2006); *Explorica*, 274 F.3d at 582–84; *United States v. Rodriguez*, 628 F.3d 1258, 1263–64 (11th Cir. 2010).

[14] *See Valle*, 807 F.3d at 527–28; *WEC Carolina Energy Solutions LLC v. Miller*, 687 F.3d 199, 206 (4th Cir. 2012); *Cloudpath Networks, Inc. v. SecureW2 B.V.*, 157 F. Supp. 3d 961, 983 (D. Colo. 2016); *Lane v. Brocq*, 2016 WL 1271051, at *10 (N.D. Ill. March 28, 2016); *Experian Mktg. Solutions, Inc. v. Lehman*, 2015 WL 5714541, at *5 (W.D. Mich. Sept. 29, 2015); *Giles Constr., LLC v. Tooele Inventory Solution, Inc.*, 2015 WL 3755863, at *3 (D. Utah Jun. 16, 2015); *Enhanced Recovery Co., LLC v. Frady*, 2015 WL 1470852, at *6-*7 (M.D. Fla. Mar. 31, 2015); *Cranel Inc. v. Pro Image Consultants Group, LLC*, 57 F. Supp. 3d 838, 845-46 (S.D. Ohio 2014); *Advanced Fluid Systems, Inc. v. Huber*, 28 F. Supp. 3d 306, 329 (M.D. Pa. 2014); *Dresser-Rand Co. v. Jones*, 957 F. Supp. 2d 610, 619 (E.D. Pa. 2013); *Power Equip. Maint., Inc. v. AIRCO Power Servs., Inc.*, 953 F. Supp. 2d 1290, 1295 (S.D. Ga. 2013); *Lewis-Burke Assocs. LLC v. Widder*, 725 F. Supp. 2d 187, 194 (D.D.C. 2010); *Bell Aerospace Servs., Inc. v. U.S. Aero Servs., Inc.*, 690 F. Supp. 2d 1267, 1272 (M.D. Ala. 2010); *Clarity Servs., Inc. v. Barney*, 698 F. Supp. 2d 1309, 1315 (M.D. Fla. 2010); *ReMedPar, Inc. v. AllParts Med., LLC*, 683 F. Supp. 2d 605, 615 (M.D. Tenn. 2010); *Nat'l City Bank, N.A. v.*

statute remained consistent with Congress's intent and to thereby avoid criminalizing common, innocuous online behavior—like checking the score of a baseball game in contravention of an employer's computer use policies. This Court held that an overbroad interpretation of the CFAA would "expand its scope far beyond computer hacking to criminalize any unauthorized *use* of information obtained from a computer" and "make criminals of large groups of people who would have little reason to suspect they are committing a federal crime." *Nosal I,* 676 F.3d at 859 (emphasis added). The Court stated that "[i]f Congress meant to expand the scope of criminal liability to everyone who uses a computer in violation of computer use restrictions—which may well include everyone who uses a computer—we would expect it to use language better suited to that purpose." *Id.* 857. But Congress had a far more narrow purpose: "to punish hacking, the circumvention of technological access barriers[.]" *Id.* at 863.

In *Valle*, the Second Circuit found *Nosal I*'s narrow interpretation of the CFAA to be "consistent with the statute's principal purpose of addressing the problem of hacking, *i.e.*, trespass into computer systems or data." 807 F.3d at 526. And in *WEC Carolina*, the Fourth Circuit put it bluntly: "we are unwilling to

---

Republic Mortg. Home Loans, 2010 WL 959925, at *3 (W.D. Wash. Mar. 12, 2010); *Black & Decker (US), Inc. v. Smith*, 568 F. Supp. 2d 929, 934 (W.D. Tenn. 2008); *Shamrock Foods Co. v. Gast*, 535 F. Supp. 2d 962, 967 (D. Ariz. 2008); *Diamond Power Int'l, Inc. v. Davidson*, 540 F. Supp. 2d 1322, 1343 (N.D. Ga. 2007); *Int'l Ass'n of Machinists & Aerospace Workers v. Werner-Masuda*, 390 F. Supp. 2d 479, 499 (D. Md. 2005).

contravene Congress's intent by transforming a statute meant to target hackers into a vehicle for imputing liability to workers who access computers or information in bad faith, or who disregard a use policy." 687 F.3d at 207.

C. LinkedIn Seeks to Transform the CFAA From an "Anti-Hacking" Statute Into a Tool For Policing Use of Publicly Available Information.

Despite the clear precedent outlined above, LinkedIn nonetheless seeks to persuade this Court to contravene Congress's intent and "transform the CFAA from an anti-hacking statute into an expansive misappropriation statute." *See Nosal I*, 676 F.3d at 857. Specifically, it wants to turn a terms of use violation into a federal computer crime.

LinkedIn's terms of service prohibit the use of technological measures to access publicly available information posted on its website by its uses. After years of permitting hiQ's automated access of publicly available LinkedIn profiles, LinkedIn elected to enforce this contractual term. It choose to do so not by putting the public data hiQ accessed behind a username and password gate, but via a letter purportedly revoking hiQ's authorization to access the still-public data and threatening CFAA legal action. LinkedIn thereafter implemented "targeted IP blocks" designed to block hiQ's automated collection of otherwise entirely public data. LinkedIn Opening Brief, Dkt. 6, pp. 12, 15. These efforts were all designed

to police the use of publicly available information posted by LinkedIn users to the open Internet for all (except hiQ) to see.

LinkedIn recognizes that simply accessing publicly available LinkedIn data via automated scripts in violation of a terms of use policy—a quintessential computer use restriction, just like any other non-technical, policy-based restriction on the "manner" of access—cannot serve as the basis for CFAA liability. *See Nosal I*, at 858, 864. And neither LinkedIn's cease and desist letter followed by targeted IP address blocks, nor this Court's decisions in *Nosal II,* 844 F.3d 1024, and *Power Ventures*, 844 F.3d 1058, change the analysis.

As the district court correctly held, *Nosal II* and *Power Ventures* are distinguishable because neither involved access of public information. *hiQ Labs*, 2017 WL 3473663, at *5. *Nosal II* involved access to a proprietary corporate computer network by an ex-employee whose own credentials to access the non-public information within that propriety network had been expressly revoked upon termination of his employment. *Nosal II*, 844 F.3d at 1035–36. *Power Ventures* involved access to non-public Facebook user data stored within a password-protected computer system—a system constructed by Facebook "to limit and control access to its website" and that requires third-party developers or websites that wish to access Facebook data to do so via Facebook's application programming interfaces (APIs). *Power Ventures*, 844 F.3d at 1063. While *Power*

Ventures involved both a cease and desist letter revoking authorization and IP address blocks, all of the data at issue was private and password-protected, stored behind Facebook's authentication barrier. *Id.* at 1067.

Here, by contrast, the public LinkedIn data hiQ accessed was not protected by any code-based access limitation. And without such barrier to entry, everyone on the Internet was and is "authorized" to access the data. *See, e.g., Pulte Homes, Inc. v. Laborers' Int'l Union of N. Am.*, 648 F.3d 295, 304 (6th Cir. 2011) (the public is presumptively authorized to access "unprotected website"); *Craigslist Inc. v. 3Taps, Inc.*, 964 F. Supp. 2d 1178, 1182 (N.D. Cal. 2013) ("*Craigslist II*") (making information website publicly available gives everyone "authorization" to view it under the CFAA). LinkedIn could have indicated who was and was not authorized to access the data by placing all user information behind a true access barrier—its preexisting username and password gate—to allow authorized users in and keep unwanted individuals out. It instead elected to give users the option of posting their information to the public Internet, and this case involves only the publicly available information of those users who elected to do so.[15]

[15] Placing data behind code-based access barriers is also the only way for LinkedIn to truly protect the privacy of that data. Imposing criminal CFAA liability for automated access of publicly available LinkedIn data will not protect the privacy interests of those users who decide to publish their information publicly, as it will not give LinkedIn any meaningful control over who has access to that data and how they use it. The data will still be freely available on the open Internet for malicious actors and anyone not within the jurisdiction of the United States to

Although LinkedIn set up targeted IP address blocks, IP address blocks are not barriers to access. As Professor Kerr explains, IP addresses change frequently, even without any effort on the part of the Internet users. For example, a person's IP address will change as they move from home to work to a café, and merely turning on and off a modem can cause an IP addresses to change.[16] Privacy protecting tools, like Tor or virtual private networks (or VPNs), also change a user's IP address. As a result, IP address blocks "should be construed as merely speed bumps" (and not access restrictions) because "bypassing an IP block is no more culpable than bending your neck to see around someone who has temporarily

---

access and use however they wish. LinkedIn's contractual use restrictions may provide an illusion of privacy—and deter law-abiding individuals and U.S.-based companies from using automated tools to access that data—but nothing more. LinkedIn's privacy policy in fact acknowledges the inherent lack of privacy in data posted publicly and makes no promises to users about LinkedIn's ability to protect it: "Please do not post or add personal data to your profile that you would not want to be publicly available." The policy also acknowledges that LinkedIn itself "collect[s] public information about you, such as professional-related news and accomplishments . . . and make[s] it available as part of [its] Services[,]" unless a user opts out via adjusting their default privacy settings. *See* LinkedIn, Privacy Policy, §§ 1.1-1.2 (June 7, 2017), https://www.linkedin.com/legal/privacy-policy. As LinkedIn well knows, if it wants to protect the privacy of its users via controlling the access and use of their data, its only option is to put that data behind a true authentication barrier. It has not done so, and this Court should give that decision its due weight.

[16] As this Court has acknowledged, a user whose IP has been blocked "does not receive notice that he has been blocked, [so] he may never realize that the block was imposed and that authorization was revoked. Or, even if he does discover the block, he could conclude that it was triggered by misconduct by someone else who shares the same IP address, such as the user's roommate or co-worker." *Power Ventures*, 844 F.3d at 1068.

blocked your view." Kerr, *Norms of Computer Trespass*, 116 Colum. L. Rev. at 1161, 1168. With "no authentication requirement, the web server welcomes all, and the norm is openness to the world"—including "any one of the billion or so Internet users around the world" or "'bot[s]' . . . running automatically." *Id.* at 1162.

Both *Power Ventures* and *Nosal II* were decided on their "stark" facts—facts that involved access to non-public information stored behind code-based access barriers—and this Court should reject LinkedIn's request to extend their holdings beyond those stark facts to entirely public information. *Power Ventures*, 844 F.3d at 1067, n.1; *Nosal II*, 844 F.3d at 1036.

Indeed, because the only data at issue here is data that was (and that remains) outside LinkedIn's authentication gate, publicly published on the open Internet for all to see, the facts presented here are more akin to the factual scenario this Court in *Power Ventures* specifically declined to address—*i.e.*, a situation in which *Nosal I's* holding that terms of use violations cannot give rise to CFAA liability is in direct "tension" with a supposed revocation of access authorization *based on* a terms of use violation, such as an "automatic boilerplate revocation [of authorization] follow[ing] a violation of a website's terms of use[.]" 844 F.3d at 1067 & n.1. Even the court in *Craigslist Inc. v. 3Taps Inc.*, 942 F. Supp. 2d 962 (N.D. Cal. 2013) ("*Craigslist I*")—the district court case on which LinkedIn relies

to support its claim that the inherent authorization to access a public website can be revoked—acknowledged this tension in its initial order denying a motion to dismiss Craigslist's CFAA claim: "[a]pplying the CFAA to publicly available website information presents uncomfortable possibilities." *Id*. at 969, n.8. That court pointed to a need for clarification by the Ninth Circuit, noting "potential problems with an overly expansive interpretation of the CFAA" but stating that it would assume the CFAA covered restrictions on the use of otherwise of public information "until the Ninth Circuit holds otherwise." *Id.*[17]

This case gives the Court the opportunity to clarify, if further clarification were needed, that obtaining publicly available information, with no true barriers to access, cannot give rise to liability under the CFAA. *See Craigslist I,* 942 F. Supp. 2d at 969, n.8. As in *Nosal I*, this Court should reject LinkedIn's expansive interpretation of the CFAA to ensure that it does not "transform the CFAA from an anti-hacking statute into an expansive misappropriation statute." *Nosal I*, 676 F.3d at 857. Ensuring that the CFAA remains limited to its original purpose is not merely as a matter of principal; it is necessary to ensuring that the statute cannot be

[17] *See also* Mark A Lemley, *Place and Cyberspace*, 91 Cal. L. Rev. 521, 528 (2003) (describing "the judicial application of the [CFAA], which was designed to punish malicious hackers, to make it illegal—indeed, criminal—to seek information from a publicly available website if doing so would violate the terms of [use]" as a serious problem).

used to undermine a hallmark of today's Internet—open access to publicly available information.

II. LINKEDIN'S INTERPRETATION OF THE CFAA WOULD POTENTIALLY CRIMINALIZE A WIDE RANGE OF VALUABLE TOOLS AND SERVICES.

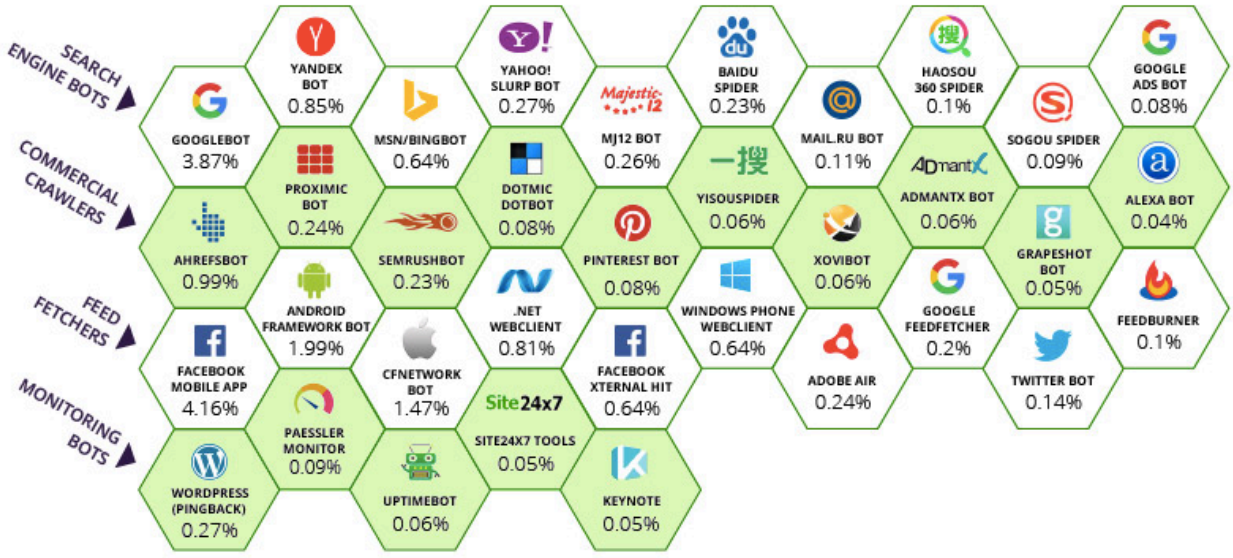
Undermining free and unrestricted access to data published on the public Internet would mean undermining equal access to "the principal sources for knowing current events, checking ads for employment, speaking and listening in the modern public square, and otherwise exploring the vast realms of human thought and knowledge." *Packingham*, 137 S. Ct. at 1737. It would also undermine critical tools for making that information more readily and easily accessible: automated bots. LinkedIn's position, by potentially imposing criminal liability for accessing publicly available information via automated scripts, would create legal uncertainty for all automated bots—including the "good bots" that accounted for 23 percent of global Web traffic in 2016.[18]

Good Internet bots automatically scrape the public Internet to collect, aggregate, and index publicly available information and support various business and operational goals of their owners—from individual Internet users to multinational corporations. Broadly, there are four types of good bots: (i) feed

[18] Igal Zeifman, Bot Traffic Report 2016, Incapsula (Jan. 24, 2017), https://www.incapsula.com/blog/bot-traffic-report-2016.html.

fetcher bots, which "ferry website content to mobile and web applications, which they then display to users"; (ii) search engine bots, also called Web crawlers, which "collect [or scrape] information for search engine algorithms, which is then used to make ranking decisions[,]" and systematically index pages and data; (iii) commercial crawlers, "[s]piders used for authorized data extractions, usually on behalf of digital marketing tools"; and (iv) monitoring bots, which "monitor website availability and the proper functioning of various online features."[19]

Chart 1: The 45 Most Active Good Bots[20]



The 45 most active good bots are responsible for 84.2 percent of all good bot traffic

[19] *Id*.

[20] Chart 1 was included in Imperva Incapsula's Bot Traffic Report of 2016, which was based on a sample of over 16.7 billion bot and human visits collected between August 9, 2016, and November 6, 2016, on 100,000 randomly chosen domains. *See id*.

The 45 most active good bots depicted in Chart 1 already account for 84.2 percent of all good bot traffic on the Internet. LinkedIn's position will potentially criminalize—and therefore undoubtedly chill[21]—many of these automated tools. Prohibitions on automated access are a standard provision in websites' computer use policies, so all Internet bots—other than those operated by established companies that have already been granted express and widespread permission to crawl the Web, like Google's search engine crawler—will be at risk. LinkedIn's position would make it even more difficult for smaller companies to create new automated tools for accessing publicly available information.

In an era of algorithms, machine learning, and artificial intelligence, an order endorsing LinkedIn's overly broad reading of the CFAA would inevitably create an uneven playing field in favor of established players and chill innovation by effectively allowing corporations with the largest datasets to control the use of publicly available information on the Internet. Alterative search engines, like Amicus DuckDuckGo, might never have survived under LinkedIn's proposed rule,

[21] The uncertainty created via some courts' overbroad interpretation of the CFAA has already chilled the work of computer security researchers. *See* Letter from Computer Security Experts to Congress and Members of the Senate and House Committees on the Judiciary (Aug. 1, 2013), https://www.eff.org/document/letter-def-con-cfaa-reform ("Many of our colleagues, and many of us, have directly experienced the chilling effects of the CFAA. Actual litigation or prosecution of security researchers is, to be sure, quite rare. But that's because the mere risk of litigation or a federal prosecution is frequently sufficient to induce a researcher (or their educational or other institution) to abandon or change a useful project. Some of us have jettisoned work due to legal threats or fears.").

as it might have been either blocked from accessing publicly available data across the Web or chilled from even trying thanks to the threat of potential federal criminal prosecution. By the same token, LinkedIn's rule would chill the creation of news or information aggregation tools, including important public safety tools like Google's Crisis Map, which during California's October 2017 wildfires aggregated information about the fire, topology, traffic, and shelter availability and resource needs,[22] or Amicus Internet Archive's Web crawler project, which works to archive as much of the public Web as possible.[23]

LinkedIn's position will also impact journalists, researchers, and watchdog organizations, who (increasingly) rely on automated tools including scrapers to support their work, much of which is protected First Amendment activity. For investigative journalists, scraping is "one of the most powerful techniques for data-savvy journalists who want to get to the story first, or find exclusives that no one

[22] *See* Google, Crisis Map Help: About Google Crisis Map (2017), https://support.google.com/crisismaps ("Crisis Map collects information that's normally scattered across the Web and other resources and makes it easily available through a single map. Find authoritative information as well as crowd-sourced data, all in one place.").

[23] *See* Internet Archive, Heritrix (last updated Feb 27, 2014), https://webarchive.jira.com/wiki/x/8Ao ("Heritrix is the Internet Archive's open-source, extensible, web-scale, archival-quality web crawler" that "seeks to collect and preserve the digital artifacts of our culture for the benefit of future researchers and generations[.]").

else has spotted."[24] ProPublica journalists, for example, investigated Amazon's algorithm for ranking products by price via a "software program that simulated a non-Prime Amazon member" and "scrapped . . . product listing page[s]"; their research uncovered that Amazon's pricing algorithm was hiding the best deals from many of its customers.[25]

Online discrimination researchers also rely on automated access tools for audit testing. One recent study of racial discrimination on Airbnb—which found that distinctively African American names were 16 percent less likely to be accepted relative to identical guests with distinctively white names—"sent inquiries to Airbnb hosts using web browser automation tools" and "collected all data using scrapers[.]"[26] In another study, Carnegie Mellon University researchers looked at discrimination in online ad delivery via "an automated tool that explore[d] how user behaviors, Google's ads, and Ad Settings interact" and found that "setting the gender to female resulted in getting fewer instances of an ad

[24] Leanpub, Scraping for Journalists (2nd edition): About the Book (last updated Sep. 11, 2017), https://leanpub.com/scrapingforjournalists.

[25] Julia Angwin and Surya Mattu, "How We Analyzed Amazon's Shopping Algorithm," *ProPublica* (Sep. 20, 2016), https://www.propublica.org/article/how-we-analyzed-amazons-shopping-algorithm.

[26] Benjamin Edelman, Michael Luca, and Dan Svirsky, *Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment*, 9 American Economic Journal: Applied Economics 1, at 1, 7 (Apr. 2017), available at https://www.aeaweb.org/articles?id=10.1257/app.20160213.

related to high paying jobs than setting it to male."[27] A growing body of evidence shows that proprietary algorithms are causing websites to discriminate among users, including on the basis of race, gender, and other characteristics protected under civil rights laws. Discrimination research has historically proven necessary for ensuring compliance with federal and state anti-discrimination laws,[28] and in today's increasingly data-driven world, in order to uncover whether and how any particular website is treating users differently, researchers need to use a variety of techniques—including automated tools for accessing public information that many websites ban.

Finally, in the academic research community, open access to research and scholarship—which includes "non-restrictively allowing researchers to use automated tools to mine the scholarly literature"—has "ensur[ed] rapid and widespread access to research findings such that all communities have the

[27] Amit Datta, Michael Carl Tschantz, and Anupam Datta, *Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination*, 2015 Proceedings on Privacy Enhancing Technologies 92, at 92 (Apr. 18, 2015), available at https://doi.org/10.1515/popets-2015-0007.

[28] Offline, audit testing has long been recognized as a crucial way to uncover racial discrimination in housing and employment and to vindicate civil rights laws, particularly the Fair Housing Act and Title VII's prohibition on employment discrimination. *Cf. Havens Realty Corp v. Coleman*, 455 U.S. 363, 373 (1982).

opportunity to build upon them and participate in scholarly conversations."[29] And because open access to academic scholarship leads to more media coverage, including via social media, open access allows for broader societal impact.[30]

Imposing potential CFAA liability for automated access will chill the use of these societally valuable research tools. If LinkedIn's position prevails, a website that disagrees with a researcher's purpose or manner of access could render that research criminal by merely updating in terms of use and sending a cease and desist letter. To avoid the threat of criminal prosecution, researchers and journalists will refrain from conducting their socially valuable and constitutionally protected research. In an era of infinite data, a ruling that chills such research will handicap researchers, journalists, and watchdogs and give the handful of corporations with the world's largest datasets the upper hand. A company's choice to prohibit investigative journalism or socially valuable research using information publicly published on the open Internet should not be enforceable as a federal criminal offense under a statute meant to target computer break-ins.

---

[29] Jonathan P. Tennant, *et al*., *The academic, economic and societal impacts of Open Access: an evidence-based review,* F1000Research (2016), https://f1000research.com/articles/5-632/v3.

[30] *Id.*

III. LINKEDIN'S POSITION WOULD RENDER THE CFAA UNCONSTITUTIONALLY VAGUE.

By potentially criminalizing what is in fact a common and critical online practice, LinkedIn's position would not only chill the use of societally beneficial automated access tools, but it would render the CFAA unconstitutionally vague.

Although this is a civil case, the underlying statutory prohibition against accessing a computer "without authorization" is criminal. Constitutional constraints on criminal statutes therefore apply. Due process requires that criminal statutes provide ample notice of what conduct they prohibit. *Connally v. Gen. Constr. Co*., 269 U.S. 385, 390 (1926). Vague laws that do not "provide explicit standards for those who apply them . . . impermissibly delegate[] basic policy matters to policemen, judges, and juries for resolution on an ad hoc and subjective basis[.]" *Grayned v. City of Rockford,* 408 U.S. 104, 108–09 (1972). A criminal statute that fails to provide fair notice of what is criminal—or that threatens arbitrary and discriminatory enforcement—is thus void for vagueness. *Skilling v. United States*, 561 U.S. 358, 412 (2010) (citing *Kolender v. Lawson*, 461 U.S. 352, 357 (1983)).

To avoid fatal vagueness problems, the Rule of Lenity calls for ambiguous criminal statutes like the CFAA to be interpreted narrowly in favor of the defendant—in civil and criminal cases alike. *United States v. Santos*, 553 U.S. 507, 514 (2008). The Rule of Lenity "ensures fair warning by so resolving

ambiguity in a criminal statute as to apply [] only to conduct clearly covered." *United States v. Lanier*, 520 U.S. 259, 266 (1997). The Rule of Lenity "not only ensures that citizens will have fair notice of the criminal laws, but also that Congress will have fair notice of what conduct its laws criminalize. We construe criminal statutes narrowly so that Congress will not unintentionally turn ordinary citizens into criminals." *Nosal I*, 676 F.3d at 863.

The competing interpretations of the CFAA outlined above demonstrate that the statutory language is ambiguous and should, consistent with the Rule of Lenity, be interpreted narrowly. Indeed, vagueness concerns were at the heart of this Court's decisions to adopt a narrow interpretation of the statute in *Nosal I*. The Court recognized that while the CFAA *could* be interpreted to base criminal liability on private computer use policies, allowing "criminal liability to turn on the vagaries of private polices that are lengthy, opaque, subject to change and seldom read" would create "significant notice problems[.]" *Id.* at 860; *see also Valle*, 807 F.3d at 527; *WEC Carolina*, 687 F.3d at 205–06. Specifically, attaching criminal punishment to breaches of vague, boilerplate policies—which companies typically reserve the right to modify at any time—would make it impossible for employees or Internet users to know what conduct is criminally punishable at any given time. *See* Orin S. Kerr, Vagueness Challenges to the Computer Fraud and Abuse Act, 94 Minn. L. Rev. 1561, 1586 (2010) (expansive or uncertain interpretations of

unauthorized access would provide "insufficient notice of what line distinguishes computer use that is allowed from computer use that is prohibited"). It would also allow "private parties to manipulate their computer-use and personnel policies" so as to turn employer-employee or company-consumer relationships—relationships traditionally governed by tort and contract law—"into ones policed by the criminal law." *Nosal I*, 676 F.3d at 860. This would grant employers and website operators the power to unilaterally "transform whole categories of otherwise innocuous behavior into federal crimes simply because a computer is involved." *Id*.

By potentially criminalizing all automated scripts, LinkedIn's interpretation of the CFAA would not only enable websites to pick and choose who did and did not have "authorization" to access to publicly available information on the open Internet, but it would enable prosecutors to pick and choose which types of automated access "are so morally reprehensible that they should be punished as crimes." *See United States v. Kozminski*, 487 U.S. 931, 949 (1988). By giving that inherently legislative power to prosecutors, it would "invit[e] discriminatory and arbitrary enforcement." *See Nosal I*, 676 F.3d at 862. The Constitution, however, "does not leave us at the mercy of *noblesse oblige*" by the government. *United States v. Stevens*, 559 U.S. 460, 480 (2010). Rather, it requires that criminal statutes be clear. LinkedIn's expansive interpretation of the CFAA does meet constitutional standards.

CONCLUSION

The Court should affirm the district court's finding that hiQ demonstrated a likelihood of success on the merits of its CFAA claim, find that the CFAA cannot be used to enforce restrictions on using automated scripts to access publicly available information, and limit the statute to the purpose intended by Congress.

Dated: November 27, 2017

By: /s/ Jamie Williams
Jamie Williams
Corynne McSherry
Cindy Cohn
Nathan Cardozo
ELECTRONIC FRONTIER FOUNDATION
815 Eddy Street
San Francisco, CA 94109
Telephone: (415) 436-9333
jamie@eff.org

Counsel for Amici Curiae

CERTIFICATE OF COMPLIANCE WITH TYPE-VOLUME LIMITATION, TYPEFACE REQUIREMENTS AND TYPE STYLE REQUIREMENTS PURSUANT TO FED. R. APP. P. 32(A)(7)(C)

Pursuant to Fed. R. App. P. 32(a)(7)(C), I certify as follows:

1. This Brief of Amici Curiae Electronic Frontier Foundation, DuckDuckGo, and Internet Archive in Support of Plaintiff-Appellee complies with the type-volume limitation, because this brief contains 6,903 words, excluding the parts of the brief exempted by Fed. R. App. P. 32(a)(7)(B)(iii); and

2. This brief complies with the typeface requirements of Fed. R. App. P. 32(a)(5) and the type style requirements of Fed. R. App. P. 32(a)(6) because this brief has been prepared in a proportionally spaced typeface using Microsoft Word 2011, the word processing system used to prepare the brief, in 14-point font in Times New Roman font.

Dated: November 27, 2017

By: /s/ Jamie Williams
Jamie Williams

Counsel for Amici Curiae

CERTIFICATE OF SERVICE

I hereby certify that I electronically filed the foregoing with the Clerk of the Court for the United States Court of Appeals for the Ninth Circuit by using the appellate CM/ECF system on November 27, 2017.

I certify that all participants in the case are registered CM/ECF users and that service will be accomplished by the appellate CM/ECF system.

Dated: November 27, 2017

By: /s/ Jamie Williams
Jamie Williams

Counsel for Amici Curiae