Portfolio Media. Inc. | 111 West 19th Street, 5th Floor | New York, NY 10011 | www.law360.com
Phone: +1 646 783 7100 | Fax: +1 646 783 7161 | customerservice@law360.com

LinkedIn Data Scraping Case Shows 9th Circ. Shift On CFAA

By **Joseph Mornin and Bethany Lobo** (October 24, 2019, 5:36 PM EDT)

Joseph Mornin

Bethany Lobo

When may a company legally scrape data from another company's website? Does it matter whether the website is open to the public or only to logged-in users?

This is a contested area of law under the Computer Fraud and Abuse Act, 18 U.S.C. Section 1030, the federal statute that imposes liability for hacking and other forms of unauthorized access.

The U.S. Court of Appeals for the Ninth Circuit weighed in on these questions in its recent opinion in hiQ Labs Inc. v. LinkedIn Corp., holding that scraping publicly accessible data likely does not constitute a CFAA violation premised on accessing a computer "without authorization."[1]

Scraping and the CFAA: The Pre-hiQ Landscape

"Scraping" means automatically accessing and extracting information from a website, which can be done for an array of purposes, both legitimate and malicious. Google Inc., for example, crawls the public web and scrapes data to be included in its search results and its database of cached pages. Likewise, the Internet Archive scrapes transient pages to make permanent archival copies.

Other types of actors also engage in scraping — for instance, automated bots scrape websites in search of email addresses and other identifiers to add to marketing lists, and drive-by hackers scrape a variety of websites for signs of known security vulnerabilities.

Scraping case law often considers the notion that the internet has both public spaces open to all visitors and gated spaces open only to authenticated users, the latter of which are protected by a username and password or other technological barrier. Historically, scraping case law has offered website owners some degree of protection as to both types of spaces, recognizing that owners often take steps to protect even public spaces from scraping — for instance, by adopting terms-of-use provisions that prohibit automated access or by blocking requests from IP addresses known to be associated with bot traffic.

The CFAA is a key statute available to website owners resisting unauthorized scraping of their sites. The CFAA provision typically invoked in scraping cases is Title 18 of the U.S. Code Section 1030(a)(2)(C), which imposes liability on "[w]hoever ... intentionally accesses a computer without authorization or exceeds authorized

access, and thereby obtains ... information from any protected computer."

The Ninth Circuit has played an important role in construing this subsection. Prior to hiQ, two of its central decisions were Facebook Inc. v. Power Ventures Inc. and United States v. Nosal.[2] Both treated the owner's permission (or lack thereof) to access its data as a key factor governing whether the access violated the CFAA.

In its 2016 Power Ventures decision, the Ninth Circuit considered Facebook's assertion that Power Ventures was violating the CFAA by scraping its site "without authorization." Power Ventures' service aggregated users' social media accounts at their request. A user would provide Power Ventures with her login credentials for social media sites like Facebook, which would then access those sites on the user's behalf, scrape data and display it to the user in a single feed.[3]

Power Ventures ran a promotional campaign for its service on Facebook, which caused material to be posted to Facebook on users' behalf. Facebook then sent a cease-and-desist letter instructing Power Ventures to stop accessing Facebook's platform.[4] Power Ventures refused, continuing to scrape data from Facebook's platform despite Facebook's technical efforts to block it, which prompted Facebook to bring suit.[5]

In considering Facebook's CFAA "without authorization" claim, the Ninth Circuit drew two rules from existing CFAA case law. First, a defendant can be liable under the CFAA by accessing a computer without permission or when permission has been revoked. Second, violating a website's terms of use, without more, is not a CFAA violation.[6]

Applying these rules, the court concluded that Facebook's cease-and-desist letter revoked Power Ventures' access to Facebook's computers. Thus, after Power Ventures received the letter (but not before), its access was "without authorization" in violation of the CFAA.[7]

The Ninth Circuit also decided Nosal in 2016. There, the court held that a former employee violated the CFAA's "without authorization" provision by using current employees' credentials to access confidential information on company computers. Importantly, the company had revoked the former employee's own access credentials prior to his usage of current employees' credentials, manifesting the former employee's lack of authorization to continue to access the computers.[8]

Both Power Ventures and Nosal found CFAA violations resulting from continued access to data after the owner had made plain that the access was unauthorized.[9] Both cases considered access to gated (i.e., not publicly viewable) data, and neither case expressly addressed whether (and if so, when) the CFAA permits scraping of public data. On the other hand, neither case expressly limited its holding to data that was not publicly viewable. Rather, Nosal expressly held that "circumvent[ion of] ... a technological access barrier," like a "password requirement," was not necessary to establish a CFAA violation.[10]

Scraping Public Data: hiQ

hiQ scraped LinkedIn profiles and offered analytics to employers, such as "Keeper" (which identified employees at risk of being poached) and "Skill Mapper" (which summarized employees' skills in the aggregate).[11] Unlike Power Ventures and Nosal, hiQ accessed public information only, it did not log in to users' accounts on their behalf or otherwise access nonpublic material.[12]

LinkedIn sent a letter demanding that hiQ cease its scraping activity.[13] In response, hiQ filed a declaratory relief action, seeking (among other forms of relief) a preliminary injunction and a declaration that hiQ's

conduct did not violate the CFAA.[14] The district court granted hiQ's request for a preliminary injunction, and LinkedIn appealed to the Ninth Circuit.[15]

LinkedIn's appeal required interpretation of the same CFAA "without authorization" prohibition at issue in Power Ventures and Nosal. Based on review of the CFAA's legislative history, the court concluded that the CFAA is focused on prohibiting hacking, which the court viewed as analogous to breaking and entering in physical space.[16]

The court thus reasoned that the CFAA's "prohibition on unauthorized access is properly understood to apply only to private information — information delineated as private through use of a permission requirement of some sort" — and not to public information.[17] The court held that both Power Ventures and Nosal were distinguishable because the scrapers in those cases were gathering data for which a username and/or password was required, but "the data hiQ was scraping was available to anyone with a web browser."[18]

"It is likely," the panel concluded, "that when a computer network generally permits public access to its data, a user's accessing that publicly available data will not constitute access without authorization under the CFAA." On that basis, it affirmed.[19] This ultimate conclusion suggests that the linchpin of a CFAA "without authorization" analysis is whether the data is publicly viewable or not, rather than whether the website owner has authorized a particular actor's access to the data.

However, the court did not definitively resolve this issue on the merits, because the legal question at issue on this appeal from entry of a preliminary injunction was whether hiQ had shown a "likelihood of success on the merits."[20]

On Oct. 11, LinkedIn filed a petition for rehearing and rehearing en banc, arguing, inter alia, that the hiQ decision conflicts with the Ninth Circuit's previous CFAA decisions — particularly given the court's seeming shift away from lack of website owner authorization as the linchpin of a CFAA violation.

Specifically, LinkedIn argued that Power Ventures deems the computer owner's permission necessary for CFAA authorization and Nosal specifies that a defendant need not circumvent a technological access barrier for their access to be deemed "without authorization," and the hiQ decision was inconsistent with both principles. LinkedIn also argued that the panel's decision is inconsistent with the CFAA's plain language and decisions from other circuits. The court's ruling on the petition is pending.

Additional Considerations

The panel's hiQ decision spotlights certain other evolving or unresolved issues in scraping jurisprudence, including the following:

- Typically, companies whose websites have been scraped invoke contract theories against scrapers — for instance, by alleging that the scraper has breached the company's terms of service. However, in hiQ, it was the scraper which sought to weaponize a contract-related principle against the company. Specifically, hiQ alleged that LinkedIn's attempts to preclude hiQ's scraping were tortiously interfering with hiQ's own contracts, since hiQ's ability to fulfill its contractual commitments was contingent on its ability to utilize scraped data from LinkedIn's site. The court held that hiQ had shown a likelihood of success on the merits of this claim, in addition to its CFAA claim.[21] It will be interesting to see whether other scrapers follow hiQ's lead in asserting tortious interference with contract claims against companies that attempt to prevent their scraping activities.

- While the court resolved the CFAA and tortious interference with contract claims in hiQ's favor (at least on a preliminary injunction standard), the court noted that "entities that view themselves as victims of data scraping are not without resort, even if the CFAA does not apply: state law trespass to chattels claims may still be available. And other causes of action, such as copyright infringement, misappropriation, unjust enrichment, conversion, breach of contract, or breach of privacy, may also lie."[22] The court's decision thus cannot be read as a clear holding that scraping publicly viewable data is likely lawful (as opposed to likely not a CFAA "without authorization" violation).

- The panel's decision to uphold entry of the preliminary injunction may have been influenced by its seeming conclusion that neither LinkedIn nor its users had a strong property interest in preventing scrapers' access to the relevant data. The panel asserted that LinkedIn did not have a "protected property interest in the data contributed by its users, as the users retain ownership over their profiles."[23] In turn, "the users quite evidently intend the[ir profile data] to be accessed by others, including for commercial purposes — for example, by employers seeking to hire individuals with certain credentials."[24] It is thus possible that the panel's analysis may have differed if it had concluded that the scraped data belonged to LinkedIn itself.

Conclusion

Partly given its procedural posture, hiQ stops short of establishing a per se rule that scraping of public data does not violate the CFAA's "without authorization" provision. However, hiQ certainly supports the view that scraping such data is not a CFAA violation. The decision reflects a potential shift in the Ninth Circuit's CFAA jurisprudence, from focus on whether the computer owner has authorized a particular actor to access the data to whether the data is publicly accessible in the first instance.

Unless the decision is superseded on rehearing before the Ninth Circuit or on appeal to the U.S. Supreme Court, it may prompt companies opposed to scraping of their sites to weigh the benefits of protecting their content from public access to maximize their ability to invoke the CFAA against scrapers, on one hand, against the potential commercial downsides of making their content inaccessible to the general public, on the other.

Joseph D. Mornin is an associate and Bethany C. Lobo is a partner at Cooley LLP.

[1] hiQ Labs, Inc. v. LinkedIn Corp., 938 F.3d 985 (9th Cir. 2019). On October 11, 2019, LinkedIn filed a petition for rehearing and rehearing en banc, which remains pending as of this writing.

[2] Facebook, Inc. v. Power Ventures, Inc., 844 F.3d 1058 (9th Cir. 2016); United States v. Nosal, 844 F.3d 1024 (9th Cir. 2016) (typically cited as "Nosal II").

[3] Id. at 1062-63.

[4] Id. at 1063.

[5] Id.

[6] Id. at 1067. Note, however, that other federal circuits have held that a violation of a computer use restriction does constitute a CFAA violation. See, e.g., EF Cultural Travel BV v. Zefer Corp., 318 F.3d 58, 62 (1st Cir. 2003) (noting that that a "lack of authorization [for CFAA purposes] could be established by an explicit statement on the website restricting access," and affirming a preliminary injunction entered against a scraper); United States v. John, 597 F.3d 263, 272 (5th Cir. 2010) (following EF Cultural Travel).

[7] 844 F.3d at 1068.

[8] 844 F.3d at 1038.

[9] As discussed below, there is some uncertainty around whether notice must be given in writing (such as through a cease-and-desist letter), or whether notice can be achieved through other means, such as technical barriers that block the scraper's access.

[10] 844 F.3d at 1038-39.

[11] hiQ, 938 F.3d at 989-92.

[12] Id. LinkedIn users can adjust privacy settings to determine whether their profiles are publicly accessible.

[13] Id. at 992.

[14] Id.

[15] Id.

[16] Id. at 1000-01.

[17] Id. at 1001.

[18] Id. at 1000, 1002.

[19] Id. at 1003.

[20] Id. at 989.

[21] Id. at 995-99.

[22] Id. at 1004.

[23] Id. at 995.

[24] Id.