The Ninth Circuit Rejects LinkedIn's Efforts to Block Web-Scraping of Member Public Profiles

October 21, 2019 Lee Johnston

PRACTICES Anti-SLAPP and First Amendment Rights, Technology Contracts Litigation, Media Entertainment and Sports

Social media companies ("SMC's") are constantly working to leverage data they gather from customers to develop new, innovative products and effective advertising strategies to market those products. At the same time, SMC's face threats from competitors seeking to harvest and exploit the publicly-available customer data hosted on SMC servers. On the technology side, SMC's employ increasingly sophisticated artificial intelligence (AI)-based software to prevent automated bots and web crawlers from accessing and scraping customer data from SMC websites. And, under the auspices of enforcing their own proprietary rights and their customers' privacy rights, SMC's have asserted a variety of legal claims – ranging from common law trespass and breach of contract theories to federal copyright and Computer Fraud and Abuse Action (CFAA) claims -- in an effort to shut down, or at least deter, their competitors' efforts to access and "scrape" SMC customer data.

As judges have gained a better understanding of the technology and legal issues in these cases, the viability of some of these claims has been circumscribed.[i] Nevertheless, SMC's have largely been on the offensive in this battle, primarily due to their ability to outspend their competitors, which are often start-ups lacking the resources for extended legal battles. The Ninth Circuit's September 9, 2019 decision in *hiQ Labs, Inc. v. LinkedIn Corporation*,[ii] however, suggests a more favorable future for web scraping in general, and specifically highlights the effectiveness of smaller competitors' strategy of "taking the battle" to larger SMCs rather than waiting to be sued.

HiQ Labs v. LinkedIn Corporation

In *hiQ Labs*, the Ninth Circuit affirmed the trial court's preliminary injunction barring LinkedIn from blocking or otherwise hindering hiQ's ability to "scrape" LinkedIn users' public profiles. The underlying dispute in *hiQ Labs* centered on hiQ's data analytics business model, which depends exclusively on its ability to scrape LinkedIn's users' public profile information. Using automated bots to harvest LinkedIn users' name, job title, work history and skills, hiQ applies a proprietary algorithm to this data to yield "people analytics," which it then sells to business clients to allow them to identify employees at the greatest risk of being recruited away, as well as to identify skill gaps in an employer's workforce.

LinkedIn took issue with hiQ's activities, especially because LinkedIn itself sought to develop and market its own skill-based predictive analytics product (Talent Insights) based on users' profiles. In May of 2017, therefore, LinkedIn sent hiQ a cease-and-desist letter, asserting that hiQ had violated LinkedIn's terms of use agreement, and that any future access of LinkedIn data would subject hiQ to liability under the CFAA, the Digital Millennium Copyright Act ("DMCA"), California Penal Code § 502(c) and the California common law of trespass.

Rather than taking a defensive posture, hiQ went on the offensive and filed a pre-emptive lawsuit seeking a declaration that it was legally entitled to scrape LinkedIn user profiles and that LinkedIn could not lawfully invoke the federal and state laws identified in its cease-and-desist letter. HiQ also

went a step further, and sought an injunction prohibiting LinkedIn from erecting technological barriers to hiQ's automated bots. By doing so, hiQ effectively pivoted the Court's analysis, and instead of being seen as an Internet parasite, hiQ was able to successfully argue that it was the victim of LinkedIn's heavy-handed, anti-competitive tactics.[iii] And, by posturing the case as one requiring immediate injunctive relief, hiQ highlighted its strongest argument – that LinkedIn's actions would destroy hiQ's business – and reduced its burden of proof on establishing the likelihood of success on the merits of its legal claims.[iv]

HiQ's high-risk/high return legal strategy paid off, primarily due to (1) LinkedIn's inability to argue plausibly that its users' privacy interests were harmed by hiQ's conduct and (2) the Court's concern that a finding of liability under the CFAA would expand the statute's reach beyond what Congress intended. First, as to privacy concerns, both the trial court and Ninth Circuit found it significant that LinkedIn had no proprietary interest in the factual information contained in its users' online profiles. LinkedIn users, not LinkedIn, "owned" this factual data, and voluntarily chose to make their profiles available to the public. Indeed, LinkedIn's own privacy policy stated that "any information you put on your profile and any content you post on LinkedIn may be seen by others," and warned users not to "post or add personal data to your profile that you would not want to be public."[v] Moreover, LinkedIn's professed privacy concerns were undermined by the fact that LinkedIn allowed other third-parties to access user data without its members' knowledge or consent.

The trial court and the Ninth Circuit also expressed serious concerns about LinkedIn's CFAA argument that hiQ's violation of the LinkedIn website terms of use provisions and disregard of LinkedIn's subsequent cease-and-desist letter constituted violations of the CFAA's prohibition against computer access "without authorization." As the trial court noted, LinkedIn's interpretation of the CFAA would permit a website owner to revoke the "authorization" of any person at any time, for any reason, and then pursue civil and criminal penalties against that person for merely *viewing* the website – an outcome which the trial court characterized as "effectuating the digital equivalence of Medusa."[vi] According to the trial court, allowing a private entity to effectively criminalize access to publicly viewable information, without any consideration of the website owner's reasons for denying access or an individual's possible justification for ignoring the website owner's denial of access, would be "particularly pernicious" to healthy competition and the public's right to information.[vii]

The Ninth Circuit agreed that the CFAA's prohibition against accessing a protected "without authorization" must be viewed in the context of the three types of information which exist on computers:

- Information for which access is open to the general public and permission is not required
- Information for which authorization is required and has been given; *i.e.*, username and password authentication
- Information for which authorization is required but has not been given (or, in the case of the prohibition on exceeding authorized access, has not been given for the part of the system accessed.)

According to the Ninth Circuit, the information which hiQ accessed and "scraped" fell into the first category of "computer information" for which no permission was required. As such, the court found that liability under the CFAA could not be based on LinkedIn's digital user agreement or the express revocation of hiQ's access rights contained in LinkedIn's cease-and-desist letter.[viii]

The Renewed Importance of Requiring Password Authentication of Customer/User Data for CFAA Liability

The Ninth Circuit's decision underscores the importance of user authentication systems in determining whether liability under the CFAA will be triggered. In *U.S. v. Nosal*, 844 F.3d 1024 (9th Cir. 2016) ("*Nosal II*"), the Ninth Circuit held that a former employee whose computer access rights had been terminated when he left his employer, but who had then used current employees' login credentials to access company computers and collect confidential information, had acted "without authorization" in violation of the CFAA. Nosal II, 844 F.3d at 1038. Similarly, in *Facebook v. Power Ventures, Inc*, 844 F.3d 1058 (9th Cir. 2016), the Ninth Circuit held that Power Ventures, a social networking website that aggregated social networking information from multiple platforms, had violated the CFAA by accessing Facebook users' password-protected data (e-mail/contact information) and then using that data to send mass e-mail messages as part of a promotional campaign. *Id*. at 1062-63.

Using its newly-articulated analytical framework, the Ninth Circuit in *hiQ Labs* observed that, unlike LinkedIn users' public profiles, the computer information being accessed in *Nosal II* and *Power Ventures* was "plainly" the type where authorization was generally required; *i.e.*, requiring password authentication, and that authorization had either never been given or had been revoked:

> It is likely 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. The data hiQ seeks to access is not owned by LinkedIn and has not been demarcated by LinkedIn as private using … an [username/password] authorization system.[ix]

Stay Tuned: LinkedIn's Petition for Rehearing and Rehearing En Banc

On October 11, 2019, LinkedIn filed its Petition for Rehearing and Rehearing En Banc, seeking reversal of the three-judge's panel's September 9th decision.[x]

While the outcome on LinkedIn's Petition is uncertain, one thing remains clear: the battle between SMC's and "scrapers" is far from over. Even if the panel's September 9th opinion remains intact, the Ninth Circuit made clear that SMCs and other online entities which view themselves as victims of data scrapping are not without legal recourse, noting that common law claims (e.g., trespass to chattels, unjust enrichment, conversion, breach of contract, and breach of privacy) and statutory claims (e.g., copyright infringement and misappropriation of trade secrets) may still be available.[x]

---

[i] *See, e.g., TicketMaster.com v. Tickets.com*, 2003 WL 21406289 (C.D. Cal. March 7, 2003) (dismissing copyright and trespass to chattels claims where only factual, publicly-available data was "scraped" from TicketMaster's website and re-published by Tickets.com in a different format, and Tickets.com's use of web crawler did not impact or interfere with the functionality of Ticketmaster.com's server); *Sandvig v. Sessions*, 315 F.Supp.3d 1 (D.D.C. 2018) (First Amendment interests were implicated and thus called into question the criminal prosecution of journalists under the Computer Fraud and Abuse Act for their use of automated bots to scrape data in breach of a website's terms of use agreement).

[ii] -- F.3d --, 2019 WL 4251998 (9th Cir. Sept. 9, 2019).

[iii] *Id.* at *9 (observing that LinkedIn's conduct "may well not be 'within the realm of fair competition.') (citations omitted).

[iv] *See Alliance for the Wild Rockies v. Cottrell*, 632 F.3d 1127, 1131 (9th Cir. 2011) (adopting a sliding scale approach and holding that where the party seeking an injunction establishes irreparable harm is virtually certain, it need only demonstrate that there are "serious questions

going to the merits" of its legal claims).

[v] -- F.3d --, 2019 WL 4251998 at * 5-6.

[vi] *HiQ Labs v. LinkedIn Corp*., 273 F.Supp.3d1099, 1110 (N.D. Cal. 2017).

[vii] *Id*. at 1112.

[viii] *Id*. at *12.

[ix] *-- F.3d --,* 2019 WL 4251998 at *14.

[x] *HiQ Labs v. LinkedIn Corp*., Case No. 17-16783 (9th Cir.), Dkt No. 82.

[xi] *Id*. (citing *Associated Press v. Meltwater U.S. Holdings, Inc*., 931 F.Supp.2d 537, 561 (S.D.N.Y. 2013) (holding that a software company's conduct in scraping and aggregating copyrighted news articles was not protected by fair use)).