ROPES & GRAY

ALERT

Hedge Funds

September 7, 2017

hiQ Labs, Inc. v. LinkedIn Corp.: A Federal Court Weighs in on Web Scraping, Free Speech Rights, and the Computer Fraud and Abuse Act

Attorneys
Edward G. Black
Patrick J. Reinikainen

In recent years, a number of firms in a variety of industries have utilized automated research methods, including web scraping tools and certain forms of artificial intelligence such as bots, to gather information from a variety of sources. For example, asset managers use these tools for research purposes, and third-party vendors use them to obtain information, which they then re-package and sell to investors. Although this practice is increasingly common, questions remain regarding potential exposure to liability for both users of these tools and website hosts who implement technological barriers to prevent their use and protect information.

With that said, a U.S. District Judge in the Northern District of California has issued a novel opinion in *hiQ Labs, Inc. v. LinkedIn Corp.*, with potentially far-reaching impact on the use of automated research tools. In granting a preliminary injunction guaranteeing a company the right to scrape data, the court found that the public nature of the information sought potentially vitiates the application of the federal Computer Fraud and Abuse Act's ("CFAA") civil and criminal provisions and other legal restrictions on scraping and similar forms of data harvesting. In reaching its decision, the court even suggested, albeit without specifically holding, that serious questions exist as to whether there is a free speech right under the California State Constitution to access and obtain information that has already been made publicly available on the internet.

Background

Plaintiff, hiQ Labs, Inc. ("hiQ"), brought a federal action against Defendant, LinkedIn Corp. ("LinkedIn"), the popular business and professional social network, asserting claims under California common law, California's Unfair Competition Law, and the California State Constitution. *hiQ Labs, Inc. v. LinkedIn Corp.*, No. 17-CV-03301-EMC, 2017 WL 3473663, at *1 (N.D. Cal. Aug. 14, 2017). At the heart of the dispute is hiQ's business model, which consists of selling public information to businesses about their employees. *Id.* at *1. In practice, hiQ collects that information exclusively from employees' public LinkedIn profiles by using web scraping or harvesting tools, a method of utilizing computer software to automatically extract data from websites. *Id.* at *1-2. The company then conducts an analysis of the information for businesses to provide key employment-related data, including, for example, whether employees are at a high risk of leaving a company or being recruited by another business, as well as summaries of skills possessed by certain employees. *Id.* at *1. LinkedIn eventually sent hiQ a cease and desist letter, alleging that hiQ's unauthorized data scraping violated LinkedIn's user agreement, which prohibited certain methods of data collection from the site, including scraping. *Id.* at *2.

Following failed efforts of resolving the dispute out of court, LinkedIn implemented technical barriers to prevent hiQ from accessing the site, which included systems intended to identify, monitor, and block data scraping activity. *Id.* hiQ responded by bringing an action in U.S. District Court for the Northern District of California, initially seeking a temporary restraining order against LinkedIn and moving for an order to show cause why a preliminary injunction should not be issued. *Id.* at *2. After the court held a hearing on the TRO, hiQ's preliminary motion was eventually converted to a motion for a preliminary injunction by way of a standstill agreement. *Id.* hiQ

also sought a declaration from the court that its actions would not violate, among other laws, the Digital Millennium Copyright Act (DMCA), the Computer Fraud and Abuse Act (CFAA), the California Penal Code, and the common law of trespass to chattels. *Id.* In response, LinkedIn argued that hiQ's web scraping efforts, in addition to violating LinkedIn's terms and conditions of use, threatened the privacy of LinkedIn users. *Id.* at *3-4. LinkedIn also contended that hiQ's conduct violated the CFAA and, consequently, that hiQ's state law claims were preempted by federal law. *Id.* at *4. Rejecting LinkedIn's arguments, the court ultimately granted hiQ's motion for a preliminary injunction, barring LinkedIn from continuing to prevent hiQ's access to the public user profiles. *Id.* at *13-14. In doing so, the court provided some of the most extensive analysis to date regarding both access to public information on the internet and the legal frameworks governing the practice of scraping web data.

The Court's Key Findings

In conducting its preliminary injunction analysis, the court first concluded that the likelihood of irreparable harm from allowing LinkedIn to prevent hiQ from continuing to access the information weighed heavily in favor of hiQ, as the company's business model relied entirely on accessing that information. *Id.* at *2-4. The court thus found irreparable harm in the fact that the company would likely be forced to shut down if it were prevented from obtaining the information. *Id.* at *4. In the second step of the court's analysis, it also held that hiQ had raised "serious questions" going to the merits as to one of its claims. *Id.* at *12. In reaching its holding, however, the court was first required to address a threshold question of whether hiQ's claims were preempted by federal law under the CFAA. *Id.* at *4-9

The CFAA imposes civil and criminal liability where an individual "intentionally accesses a computer without authorization or exceeds authorized access, and thereby obtains ... information from any protected computer." 18 U.S.C. § 1030(a)(2)(C). As the court noted, the statute "provides two ways of committing the crime of improperly accessing a protected computer: (1) obtaining access without authorization; and (2) obtaining access with authorization but then using that access improperly." *Id.* at *4 (internal citation omitted). Thus, in addressing hiQ's conduct, the court analyzed whether "hiQ . . . 'accesse[d] a computer without authorization' within the meaning of the CFAA." *Id.* The court first rejected the proposition that "LinkedIn's revocation of permission to access the public portions of its site renders hiQ's access 'without authorization'" under the CFAA. *Id.* at *8. The court then went on to also reject LinkedIn's argument that hiQ's use of web scraping methods constituted "access" obtained "without authorization," stating that there were "'serious questions' as to the applicability" of the Act to hiQ's conduct given that the information at issue was publicly available. *Id.* at *9. In particular, the court made the key observation that "[a] user does not 'access' a computer 'without authorization' by using bots, even in the face of technical countermeasures, when the data it accesses is otherwise open to the public." *Id.* at *8. The court spurned a broad reading of the CFAA, stating that the "application of the CFAA to the accessing of websites open to the public would have sweeping consequences well beyond anything Congress could have contemplated; it would 'expand its scope well beyond computer hacking.'" *Id.* at *6. (citation omitted). Relying on an analogy to support its ruling, the court also invoked the law of trespass, finding that LinkedIn sought "to prohibit hiQ from viewing a sign publicly visible to all," suggesting that hiQ's conduct also did not amount to a trespass sufficient to give rise to liability under the CFAA. *Id.* at *7. In summary, while the court concluded that LinkedIn could use anti-bot measures to prevent intrusions in the future, it expressed doubt, at least at the preliminary injunction stage, that the CFAA applied to hiQ's scraping methods where the information obtained was publicly available. *Id.* at *8-9.

Having found that hiQ's claims were not preempted under the CFAA, the court then went on to address the potential merits of hiQ's claims under state law. The court first addressed hiQ's free speech claim under the California State Constitution. *Id.* at *10-11. hiQ argued that LinkedIn is a public forum, and therefore hiQ had a free speech right "to access that marketplace on equal terms with all other people and that LinkedIn's private property rights in controlling access to its computers cannot take precedence." *Id.* at *10. Although recognizing that the California Constitution affords broader free speech protection than the First Amendment to the U.S. Constitution, the court found that hiQ had failed to raise "serious questions" regarding whether LinkedIn's conduct violated hiQ's free speech rights under state law. *Id.* at *11. Specifically, the court declined to hold that publicly accessible websites could constitute "public

fora" for purposes of free speech protection given "the potentially sweeping implications . . . and the lack of any more direct authority." *Id.* at * 11. However, in its analysis, the court discussed at length the California Supreme Court's decision in *Robins v. Pruneyard Shopping Ctr.*, 23 Cal. 3d 899, 905, 153 Cal. Rptr. 854, 592 P.2d 341 (1979) *aff'd*, 447 U.S. 74, 100 S. Ct. 2035, 64 L. Ed. 2d 741 (1980), in which the court held that California's "guarantee of free expression may take precedence over the rights of private property owners to exclude people from their property." *Id.*at *10. In *Pruneyard*, the California Supreme Court issued a decision, later affirmed by the U.S. Supreme Court, concluding that the State Constitution protected political speech at a private shopping mall, specifically noting "the importance of the shopping mall as a public forum and center of community life." *Id.* at *10 (citing *Pruneyard* 23 Cal. 3d at 910). After questioning the salience of *Pruneyard*, the court ultimately found the comparison between a mall and the internet to be problematic with respect to free speech rights. *Id.* at *10. The court, however, also specifically limited its ruling on hiQ's public forum-based free speech claim to the preliminary injunction stage by stating, "***at this juncture***, the Court has doubts about whether *Pruneyard*may be extended wholesale into the digital realm of the Internet," thus leaving open the possibility that hiQ could succeed on the merits of its claim at a later time. *Id.* at * 10 (emphasis added).

The court then addressed hiQ's unfair competition claim under California Unfair Competition Law, Cal. Bus. & Prof. Code §§ 17200 *et seq.*, finding that – unlike its other claims – hiQ had raised "serious enough questions" regarding the potentially anti-competitive nature of LinkedIn's conduct to support the issuance of a preliminary injunction. *Id.* *11-12. In so finding, the court pointed to evidence that LinkedIn may have revoked hiQ's access to the site in an effort to eliminate hiQ as a competitor in the data analytics field. *Id.* at *11-12. The court then found that hiQ's remaining promissory estoppel claim failed to raise such serious questions as to the merits to support the issuance of a preliminary injunction. *Id.* at *12. In the final part of its analysis, the court concluded that the public interest warranted the issuance of a preliminary injunction, noting that providing private parties "the blanket authority to block viewers from accessing information publicly available on its website for any reason, backed by sanctions of the CFAA, could pose an ominous threat to public discourse and the free flow of information promised by the Internet." *Id.* at *13. The court thus granted hiQ's motion for a preliminary injunction, barring LinkedIn from continuing to prevent hiQ's access to the public LinkedIn user profiles. *Id.* at *13. The court's Order, among other things, also required LinkedIn to affirmatively remove any technical barriers that were preventing hiQ's access within 24 hours. *Id.*

Best Practice Takeaways

The *hiQ* court's decision leaves open the question of whether hiQ will ultimately succeed on the merits of its claims.[1] Moreover, other courts may view much of the court's analysis as dicta and therefore decline to follow its approach moving forward. With that said, the opinion provides a basis for at least some best practice takeaways relating to legal restrictions on scraping and other data harvesting methods moving forward. First, websites employing terms and conditions of use aimed at preventing unauthorized access to information, whether by means of scraping or circumventing security measures, should consider the efficacy of those terms. In particular, the *hiQ* decision strongly suggests that allegations of "unauthorized access" under the CFAA would have little weight, whether asserted affirmatively or as a federal preemption defense, if the information obtained is otherwise publicly available. In addition, web hosts should consider whether barring users from accessing information on a website based on violations of terms and conditions of use could potentially give rise to civil liability under state or federal law, where information sought is publicly available. By the same token, companies using web scraping methods should be aware that if the information accessed or obtained is not public on its face, a court could impose liability under state or federal law, including the CFAA. Moreover, independent of private litigation, the *hiQ* court's decision could leave the door open for possible civil or criminal enforcement proceedings by the SEC or DOJ based on the CFAA, the federal Wire Fraud Statute (18 U.S.C. § 1343), and other federal securities laws, rules and

[1] It also appears that an appeal was taken by LinkedIn to the Ninth Circuit Court of Appeals on September 5, 2017, which could result in an affirmance or reversal of the District Court's decision. *hiQ Labs, Inc. v. LinkedIn Corporation* (Ninth Circuit U.S. Court of Appeals 0:17-cv-16783). The matter appears to be fully briefed with oral arguments scheduled for March 15th.

regulations.[2] Given the court's findings in *hiQ*, such enforcement proceedings could likely turn on the particular conduct involved or method employed to obtain the data, the terms and conditions of use aimed at preventing misuse of information, and, perhaps most importantly, whether the information is non-public. Finally, the *hiQ* decision leaves unanswered the question of whether the internet is a public forum equivalent to the private mall at issue in the *Pruneyard* case, the answer to which could have an enormous impact on free speech jurisprudence and internet regulation.

[2] Note that, while potential liability in the context of civil litigation was addressed by the *hiQ* court, other sources of liability, including SEC Rule 10b-5 (17 C.F.R. 240.10b-5) and the Wire Fraud Statute, have arisen in context of government enforcement actions, specifically in instances where individuals have used hacking methods to obtain material nonpublic information. *See, e.g.*, *S.E.C. v. Dubovoy, et al.*, Civil Action No. 2:15-cv-06076-MCA-MAH (D.N.J. filed Aug. 10, 2015) (amended Aug. 23, 2015); *see also S.E.C. v. Dorozhko*, 574 F.3d 42 (2d Cir. 2009). Query whether future cases may involve similar factual scenarios in which outsiders either use web scraping tools, other forms of artificial intelligence, or weaknesses in a website security measure to obtain material nonpublic information or proprietary data without authorization of a website host.