Proskauer

A Green Light for Screen Scraping? Proceed With Caution…

August 24, 2017

Court Issues Injunction Barring Blocking of Scraping and Holds CFAA Likely Doesn't Apply

Websites make information available to clients, users, customers and subscribers. Data aggregators, investors, competitors and others are always thinking of new and productive ways to use that data – typically, uses other than those for which the data is being made available. "Screen scraping" is one of the main technical tools used to harvest data from websites for such uses, and the federal Computer Fraud and Abuse Act (the "CFAA"), 18 U.S.C. §1030, has been one of the main legal tools used by website owners to challenge those scraping activities.

While the law relating to screen scraping is unclear, a recent landmark decision from the Northern District of California, *hiQ Labs, Inc. v. LinkedIn, Corp.*, 2017 WL 3473663 (N.D. Cal. Aug. 14, 2017), appears to limit the applicability of the CFAA as a tool against scraping. Indeed, in granting injunctive relief against LinkedIn's blocking of hiQ's scraping activities, the *hiQ* court noted that, by invoking the CFAA, "[c]ompanies could prevent competitors or consumer groups from visiting their websites to learn about their products or analyze pricing." While the *hiQ* decision suggests that, at least in some circumstances, scraping of publicly available websites does not give rise to a cause of action under the CFAA, scrapers beware – the road may still have some rough patches ahead.

The *hiQ* Opinion

The *hiQ* case involves LinkedIn's challenge to hiQ's scraping of LinkedIn public profile data. Upon receipt of a cease and desist letter from LinkedIn alleging, among other things, hiQ's civil liability under the CFAA, hiQ sought a preliminary injunction barring LinkedIn from blocking hiQ's access to LinkedIn public profiles. Significantly, LinkedIn sent the cease and desist letter to hiQ after years of tolerating hiQ's access and use of its data; in fact, hiQ's business model of employee data analysis, is wholly dependent on crunching LinkedIn data that users have elected to publish publicly.

The key question concerning the applicability of the CFAA in this case was whether, by continuing to access public LinkedIn profiles after LinkedIn explicitly revoked permission to do so, hiQ has "accessed a computer without authorization" within the meaning of the CFAA.

The court issued a preliminary injunction, finding that the balance of equities favored *hiQ*, and distinguished the Ninth Circuit precedent in *Facebook, Inc. v. Power Ventures, Inc.*, 844 F.3d 1058 (9th Cir. 2016), that held that a commercial entity that accesses a website after permission has been explicitly revoked can, under certain circumstances, be civilly liable under the CFAA. The *hiQ* court found none of the data in the prior cases interpreting the CFAA relied on by LinkedIn (including *U.S. v. Nosal (Nosal II)*, 844 F.3d 1024 (9th Cir. 2016)[1], which involved an employer's computer network) was publicly available data but, rather, a portion of a website (or employer database) protected by a user password. Limiting the reach of the Ninth Circuit's prior holdings, the court expressed "serious doubt" as to whether LinkedIn's revocation of permission to access the public portions of its site renders hiQ's access "without authorization" within the meaning of the CFAA. In the court's view, the CFAA was intended instead to deal with "hacking" or "trespass" onto private, often password-protected mainframe computers, and the judge was "reluctant" to expand its scope absent convincing authority. According to the court, the "broad interpretation" of the CFAA advocated by LinkedIn, if adopted, "could profoundly impact open access to the Internet."

LinkedIn also argued unsuccessfully that hiQ, as a LinkedIn member, is bound by its user agreement and its prohibitions on scraping activities. The court rather superficially noted that LinkedIn had terminated hiQ's user status and failed to demonstrate that hiQ's aggregation of data from LinkedIn's public profiles is dependent on its status as a LinkedIn user. Thus, the allegation of breach of contract was left largely unaddressed.[2]

Is this the end of the CFAA as a tool against scraping (in the Northern District of California)?

Before viewing this as a green light for scraping, readers should note the following about the opinion:

Because of the importance of this issue, and some of the narrow distinctions from precedential case law made by the court in reaching its conclusions, this decision seems ripe for appeal.

- The decision is limited to only the question of whether injunctive relief was appropriate under the specific facts of the case. It is not a clear holding that the CFAA does not apply to scraping.
- The decision focuses on sites which make data "publicly available." In this case, the data was viewable by anyone, without the need for a password. As the court summed up its holding: "Where a website or computer owner has imposed a password authentication system to regulate access, it makes sense to apply a plain meaning reading of 'access' 'without authorization' such that 'a defendant can run afoul of the CFAA when he or she has no permission to access a computer or when such permission has been revoked explicitly.' But…in the context of a publicly viewable web page open to all on the Internet, the 'plainness' of the meaning of 'access' 'without authorization' is less obvious. Context matters."
- The decision does not address the breach of contract claims based on website terms of use, as well as claims based in copyright, trespass or other causes of action.
- It does not address situations where the "robots.txt" file is ignored, or where a scraper is acting in a misleading way or otherwise concealing its identity. It should be recalled, however, that the Ninth Circuit in *Power Ventures* stated, in dicta, that simply bypassing an IP address block, without more, would not constitute unauthorized use.
- It does not address factors such as the interference with a website's sale or other authorized distribution of the data being scraped.

As it now stands, the *hiQ* opinion's holding and dicta offers a positive trend for those interested in scraping, yet leaves unanswered questions that present risk. Nonetheless, website owners seeking to block scrapers must evaluate this decision to understand what contractual and technical measures, if any, a site might undertake to thwart unwanted scraping of public-facing web content.

[1] The *Power Ventures* and *Nosal* decisions are discussed at length in a prior post on our New Media and Technology Law blog.

[2] The court also considered hiQ's argument that LinkedIn was unfairly leveraging its power in the professional networking market for an anticompetitive purpose, warranting injunctive relief. It found that hiQ made a plausible inference that LinkedIn terminated hiQ's access to its public member data in large part because it wanted exclusive control over that data for its own business purposes and to eliminate a competitor in the data analytics field. Still, the court made clear that LinkedIn may ultimately demonstrate it was not motivated by anticompetitive purposes, rather a desire to preserve user privacy preferences and its users' trust, or was lawfully protecting its own long-standing data analytics services.

Related Professionals

- **Jeffrey D. Neuburger**