# THE DEMISE OF THE CFAA IN DATA SCRAPING CASES

JENNIE E. CHRISTENSEN*

## INTRODUCTION TO DATA SCRAPING

Data scraping (or "scraping") is a relatively new technological process with widespread legal implications.[1] In e-commerce, retailers use scraping to monitor competitor pricing; lawyers use scraping to find relevant prior judgements; and recruiters use scraping to collect lists of top profiles. Despite this range of applications, however, data scraping exists in a realm of legal uncertainty.

The purpose of this Note is three-fold. First, this Note aims to provide a basic overview of what scraping is. There is a dearth of basic understanding amongst lawyers and law students alike regarding data scraping fundamentals which is problematic because many attorneys will represent clients in coming years who are either engaging in scraping themselves or who are the targets of scrapers. As such, it will become important for attorneys to understand what scraping is and how to address related issues of cybersecurity as part of their corporate legal guidance. Second, this Note reviews the ways in which recent courts have addressed the issue of scraping. There is currently no statute designed to specifically address scraping and, as such, plaintiffs have looked to a wide range of claims and statutes to challenge the behavior including unjust enrichment, copyright infringement, trespass to chattels, and the Computer Fraud and Abuse Act (CFAA). This portion of the analysis will focus on the CFAA, the primary statute applied against scrapers, and organize the courts' approach in applying this law into four rough time periods. Finally, this Note hypothesizes the future trajectory of data scraping law, one which may include the demise of the CFAA. Through an analysis of the history of CFAA applications, this Note argues that the CFAA is ill-suited for data scraping cases and that future courts may be reluctant to apply it, meaning that, unless a new statute is found to adequately challenge data scraping, scrapers may soon have free reign to scrape publicly accessible data.

[*] J.D. Candidate, University of Notre Dame Law School, 2020. I would like to extend a special thanks to Professor Patricia Bellia for sparking my interest in cyberlaw and for supporting my research interests in cyberlaw-related topics.

[1] *See generally* Andrew Sellars, *Twenty Years of Web Scraping and the Computer Fraud and Abuse Act*, 24 B.U. J. SCI. & TECH. L. 372, 373 (2018); *see also Website Monitoring*, ENVTL. DATA & GOVERNANCE INITIATIVE, https://envirodatagov.org/website-monitoring/ (last visited Nov. 24, 2019).

### *A. What is Scraping?*

Data scraping has only recently gained serious legal attention, though its origins date back to the creation of the World Wide Web (i.e., the internet).[2] The earliest version of the internet was difficult to search; it was just a collection of File Transfer Protocols (FTPs).[3] To navigate the FTPs, users needed to search for specific, shared files.[4] As such, searching the FTPs was both difficult and time consuming.[5] As the internet continued to grow, users began to seek an efficient way to find and copy specific information. To make the file search process easier, automated programs, now known as "web crawlers" or "bots," were first created in 1994 to find and organize this distributed data.[6] Since then, bots have become a common way to help users find and copy data in an automated fashion.[7]

Data scraping as we know it today involves two steps: First, bots go to an internet site. Next, the bots copy specific data out of the page and put it into a separate spreadsheet or database.[8] In most instances, this second step entails a computer script sending "tailored queries" to a website in an attempt to "retrieve specific pieces of content."[9] These requests are typically automated, meaning data can be gathered quickly, repeatedly, and simultaneously by various bots, resulting in mass data aggregation.

### *B. The Purpose of Scraping*

Scraping has countless applications and is used in just about every major industry. One of the most familiar applications to the average internet user is internet browsing.[10] Google and other search engines rely on crawlers to systematically scan and analyze web pages "to index those sites for searching."[11] What this means is that anytime an individual conducts a Google search and peruses the Google results for a match, they are relying on the work product of bots.[12] As Andrew Sellars, Lecturer at Boston University (BU)

[2] *See* Frank Jennings & John Yates, *Scrapping over Data: Are the Data Scrapers' Days Numbered?*, 4 J. INTELL. PROP. L. & PRAC. 120, 120 (2009).

[3] *See* BARRY M. LEINER ET AL., INTERNET SOC'Y, BRIEF HISTORY OF THE INTERNET: 1997 (2017), https://www.internetsociety.org/wp-content/uploads/2017/09/ISOC-History-of-the-Internet_1997.pdf.

[4] *See id.*

[5] *See id.*

[6] Tobias Knecht, *A Brief History of Bots and How They've Shaped the Internet Today*, ABUSIX (Sept. 21, 2016), https://www.abusix.com/blog/a-brief-history-of-bots-and-how-theyve-shaped-the-internet-today.

[7] *See id.*

[8] Sellars, *supra* note 1, at 373–74.

[9] *Id.* at 373.

[10] *See* Zachary Gold & Mark Latonero, *Robots Welcome? Ethical and Legal Considerations for Web Crawling and Scraping*, 13 WASH. J.L. TECH. & ARTS 275, 280 (2018).

[11] *Id.*

[12] *See id.*

School of Law and Director of the BU/MIT Technology and Cyber Law Clinic, describes:

> [Data scraping] can be used to preserve websites, help identify and extract data for analysis, [and] aggregate information from disparate sources . . . . Its use can help competition by lowering startup information barriers, enable consumers to find deals and discounts in online services, identify and correct issues of algorithmic bias, and introduce new forms of humor and playfulness.[13]

Yet, scraping can also be used for other questionable and less appealing purposes. For instance, scraping is often used to gain a competitive advantage. After scraping data from someone else's website, companies can aggregate and then repackage that same data to generate their own products and services. The data would be presented as an entirely new product or service even though it was initially generated by (and taken from) another party. Scraping can also be used to:

> [F]acilitate an invasion of one's sense of privacy, expose content that a website host wished instead to remain hidden, facilitate copyright infringement at scale, enable new forms of surveillance, or help people cheat in online trivia games.[] Given its utility, the technique [of data scraping] has been adopted widely. One company estimates that about a quarter of all current web traffic comes from web scrapers.[14]

## I. THE LEGALITY OF SCRAPING

There are numerous ways a company might attempt to prohibit the scraping of its data such as via captchas, a computer program intended to distinguish human from machine input; website-based structural challenges (e.g., dynamic websites); blocking the automated signal; or, if these technical measures prove ineffective, via a cease and desist letter.[15] Yet, there are many ways for scrapers to circumnavigate these measures[16] and most scraping cases arise after the scraper has repeatedly attempted to circumvent the website's technical protections.

If the scraping persists, plaintiffs may decide to file suit.[17] In doing so, they will need to decide what specific rights and laws have been violated. There

[13] Sellars, *supra* note 1, at 374 (footnotes omitted); *see, e.g.*, Klint Finley, *Twitter 'Joke Bots' Shame Human Sense of Humor*, WIRED (Aug. 22, 2013, 9:30 AM), https://www.wired.com/2013/08/humor-bots/ ("One of the funniest people on Twitter isn't a person at all. It's a bot called @Horse_ebooks.").

[14] Sellars, *supra* note 1, at 374–75 (footnotes omitted).

[15] *See* Shaumik Daityari, *Protect Your Site Against Web Scraping*, JSCRAMBLER (Mar. 21, 2017), https://blog.jscrambler.com/protect-your-site-against-web-scraping/.

[16] *See How to Prevent Getting Blocked While Scraping*, SCRAPEHERO, https://www.scrapehero.com/how-to-prevent-getting-blacklisted-while-scraping/ (last visited Oct. 10, 2019).

[17] *See* Daityari, *supra* note 15.

is currently no data scraping specific statute or doctrine[18] and, as such, courts have had to "cobble together a body of case law to fill the gaps."[19] The plaintiff will thus have many possible claims to choose from: trespass to chattels; unjust enrichment; breach of contract; copyright infringement; violation of the Digital Millennium Copyright Act; or a violation of the CFAA.[20] It is most common for website owners to challenge scraping based on violations of the CFAA or "analogous state claims."[21] This section begins with a brief overview of non-CFAA claims and concludes with an in-depth review of the CFAA's applicability in data scraping cases, including a review of two decades of relevant case law and a prognostication of the statute's suitability in future cases.

### *A. Trespass to Chattels*

"'A trespass to a chattel may be committed by intentionally . . . using or intermeddling with a chattel in the possession of another[,]' when 'the chattel is impaired as to its condition, quality, or value, or . . . the possessor is deprived of the use of the chattel for a substantial time.'"[22] In data scraping cases, trespass to chattels is brought under the theory that the scraping process interfered with the plaintiff's network and server capacity.[23] However, trespass to chattels is not a particularly popular claim in data scraping claims because the harms of "intermeddling" are "often acknowledged to be minimal."[24] Consequently, courts are often reluctant to grant preliminary injunctions based on ongoing trespasses to chattels.

For example, "[i]n one of the earliest cases challenging unwanted data scraping, *eBay, Inc. v. Bidder's Edge, Inc*., eBay successfully used a trespass to chattels theory to obtain a preliminary injunction against an auction aggregator" that had been using bots to "compil[e] a database of eBay's auction listings."[25] In invoking a trespass to chattels claim, eBay argued that the defendant's access used "valuable bandwidth" and "should be thought of as equivalent to sending in an army of 100,000 robots a day to check the prices in

[18] *See* Paven Malhotra et al., *What Courts Have Said About the Legality of Data Scraping*, LEGALTECH NEWS (Jul. 20, 2017), https://www.keker.com/Templates/media/images/010071701%20Keker.pdf.

[19] *Id.*

[20] *See id.*

[21] Margaret A. Esquenet et al., *The Computer Fraud and Abuse Act and Third-Party Web Scrapers*, FINNEGAN (Oct. 29, 2018), https://www.finnegan.com/en/insights/the-computer-fraud-and-abuse-act-and-third-party-web-scrapers.html.

[22] Kathleen C. Riley, Note, *Data Scraping as a Cause of Action: Limiting Use of the CFAA and Trespass in Online Copying Cases*, 29 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 245, 265 (2018) (alteration in original) (quoting RESTATEMENT (SECOND) OF TORTS §§ 217(b), 218(b)–(c) (AM. LAW. INST. 1965)).

[23] *See id.*; *see, e.g.*, Craigslist Inc. v. 3Taps Inc., 942 F. Supp. 2d 962, 980 (N.D. Cal. 2013).

[24] Riley, *supra* note 22, at 265.

[25] Linda Henry, *Data Scraping, Bots and First Amendment Rights*, JDSUPRA (Oct. 11, 2017), https://www.jdsupra.com/legalnews/data-scraping-bots-and-first-amendment-80879/.

a competitor's store."[26] The *eBay* court granted the preliminary injunction but only after noting that it "disagreed with the metaphor"[27] and after adding that "courts rarely grant preliminary injunctions based on ongoing trespasses to chattels."[28]

### *B. Unjust Enrichment*

Unjust enrichment occurs when one person is enriched at the expense of another in circumstances that the law perceives to be unjust.[29] When this occurs, the law imposes an obligation upon the recipient to make restitution.[30] In scraping cases, an unjust enrichment claim is founded on the notion that the scraper is copying information that it did not create or work to develop.[31] In *ShopLocal LLC v. Cairo, Inc*.,[32] for example, ShopLocal sued Cairo, a competitor, when Cairo accessed, then republished, advertisements created by ShopLocal. ShopLocal asserted a claim of unjust enrichment, while Cairo moved to dismiss the unjust enrichment claim on the basis of failure to state a claim, contending that ShopLocal should not be permitted to recover for both breach of contract and unjust enrichment in the same action.[33] Ultimately, the court denied Cairo's motion.

### *C. Breach of Contract*

Companies often attempt to limit scraping of their data through their website's terms and conditions. When scrapers disobey these terms, companies might bring a breach of contract claim in conjunction with CFAA claims. Zillow's terms of use, for example, "prohibit automated queries, specifically 'screen and database scraping, spiders, robots, [and] crawlers,'" but make an exception for search engines where the scraping is fair use.[34] Recent courts have found that a scraper assents to the website's terms of use anytime it accesses a website, regardless of whether it clicks a button specifically agreeing to the website's terms.[35]

[26] eBay, Inc. v. Bidder's Edge, Inc., 100 F. Supp. 2d 1058, 1065, 1071 (N.D. Cal. 2000).

[27] Riley, *supra* note 22, at 266.

[28] *Id.*

[29] *See generally* CHARLES MITCHELL ET AL., GOFF & JONES: THE LAW OF UNJUST ENRICHMENT (8th ed. 2011).

[30] *See id.*

[31] *See generally* QVC, Inc. v. Resultly, LLC, 159 F. Supp. 3d 576 (E.D. Pa. 2016).

[32] ShopLocal LLC v. Cairo, Inc., No. 05 C 6662, 2006 U.S. Dist. LEXIS 7768 (N.D. Ill. Feb. 27, 2006).

[33] *See id.*

[34] Riley, *supra* note 22, at 257 (alteration in original) (quoting *Zillow Terms of Use*, ZILLOW, https://www.zillow.com/corp/Terms.htm (last updated Sept. 10, 2019)).

[35] *See, e.g.,* Register.com, Inc. v. Verio, Inc., 356 F.3d 393, 403–04 (2d Cir. 2004).

### *D. Copyright Infringement*

In the United States, intellectual property protects original works of fixed authorship but not ideas, concepts, discoveries, or "pure information or facts."[36] To this end, the Supreme Court noted in *Feist Publications, Inc. v. Rural Telephone Service Co*. that "[b]ecause copyright is about expression and originality, . . . 'sweat of the brow' [alone] does not entitle a work to copyright protection."[37] Further, while "specific selection and arrangement of . . . facts may be copyrightable," the "raw facts and data themselves are not."[38]

Given the thin protections extended to data, copyright infringement is not a particularly popular claim in data scraping cases. When the claim is brought, it is often alleged and dismissed.[39] For example, in *Ticketmaster Corp. v. Tickets.com, Inc.*, a case involving the scraping of ticket information, the court denied the plaintiff's injunction on its copyright claim based on the fact that the plaintiff was "attempting to find a way to protect its expensively developed basic information from what it considers a competitor," an improper use of copyright law.[40] In a later decision on the same matter, the court accepted that Ticketmaster's website was copyrightable but determined that Ticket.com's spidering activity was fair use.[41]

As exhibited in *Ticketmaster*, in response to a claim of copyright infringement, the scraping defendant might argue that its scraping falls under fair use. Fair use, found in § 107 of the Copyright Act,[42] is a doctrine which holds that use of a copyrighted work may be transformative (and thereby protected) if it "adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message."[43] "Because copying is so fundamental to the functioning of the Internet . . . courts have sometimes found that online services that involve extensive copying—such as search engines—are fair use."[44]

[36] Riley, *supra* note 22, at 261; *see also* 17 U.S.C. § 102 (2018).

[37] Riley, *supra* note 22, at 261 (quoting Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 352–54 (1991)).

[38] *Id.*; *see also Feist Publ'ns, Inc.*, 499 U.S. at 347 ("[F]acts do not owe their origin to an act of authorship. The distinction is one between creation and discovery: The first person to find and report a particular fact has not created the fact; he or she has merely discovered its existence.").

[39] *See, e.g.*, eBay, Inc. v. Bidder's Edge, Inc., 100 F. Supp. 2d 1058, 1072 (N.D. Cal. 2000) ("BE argues that the trespass claim . . . 'is similar to eBay's originally filed but now dismissed copyright infringement claim . . . .'").

[40] Ticketmaster Corp. v. Tickets.com, Inc., No. CV99-7654-HLH (BQRx), 2000 U.S. Dist. LEXIS 12987, at *10–11 (C.D. Cal. Aug. 10, 2000), *aff'd*, 2 F. App'x 741 (9th Cir. 2001).

[41] *See id.*

[42] *See* Riley, *supra* note 22, at 262; 17 U.S.C. § 107 (2018).

[43] Riley, *supra* note 22, at 262 (quoting Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 579 (1994)).

[44] *Id.*; *see, e.g.*, Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146, 1166 (9th Cir. 2007).

### *E. Violation of the Digital Millennium Copyright Act*

Companies create technical measures to protect scraping of their website by measuring traffic for unusual behavior and scrapers. By virtue of circumventing an IP block or traffic monitoring software, scrapers can be liable under the Digital Millennium Copyright Act (DMCA), a 1998 United States copyright law which states that "[n]o person shall circumvent a technological measure that effectively controls access to a [copyrighted] work."[45] The DMCA is "a much broader statute than the CFAA, but is specifically tailored to protect its underlying property right, that of copyright."[46] In essence, the statute criminalizes the act of circumventing an access control, regardless of whether there is actual copyright infringement.

### *F. Violation of the Computer Fraud and Abuse Act*

The CFAA[47] has become one of the primary legal tools used by website owners to challenge illegal scraping activities. In 1984, five years before the invention of the World Wide Web, Congress passed the Counterfeit Access Device and Computer Fraud and Abuse Act (CADCFAA),[48] which was soon amended to create the CFAA. "[T]he CFAA was originally envisioned as an anti-hacking or computer trespass statute"[49] but has since been applied to much broader internet offenses. This section provides a brief overview of the CFAA, follows with an overview of the past two decades of the statute's application in data scraping cases, and concludes with the hypothesis that the CFAA will inevitably be phased out of data scraping cases.

#### i. Introduction to Key Statutory Provisions

The CFAA creates two forms of improper access: "(1) obtaining access without authorization; and (2) obtaining access with authorization but then using that access improperly."[50] The statute's broadest and most controversial provision, § 1030(a)(2)(C) (also referred to as "the Access Provision"), creates liability for anyone who "intentionally accesses a computer without authorization or exceeds authorized access, and thereby obtains information . . . from any protected computer."[51] These short, seemingly simple phrases have been at the center of debate in dozens of data scraping cases.

[45] 17 U.S.C. § 1201(a)(1)(A) (2018).

[46] Riley, *supra* note 22, at 299.

[47] 18 U.S.C. § 1030 (2018).

[48] Counterfeit Access Device and Computer Fraud and Abuse Act of 1984, Pub. L. No. 98-473, § 2102(a), 98 Stat. 2190 (1984) (codified as amended at 18 U.S.C. § 1030 (2018)); *see also* Riley, *supra* note 22, at 266.

[49] Riley, *supra* note 22, at 267.

[50] Musacchio v. United States, 136 S. Ct. 709, 713 (2016).

[51] 18 U.S.C. § 1030(a)(2)(C) (2018).

To begin, it will be helpful to understand how the CFAA defines "computer." The CFAA covers "protected computers," and in practice, "protected computer" has come to refer to virtually all internet-connected devices, including cellphones, due to the interstate nature of most internet communication.[52] This broad interpretation of "protected computer" and virtually every other provision of the CFAA has partially fueled the statute's widespread use in data scraping cases.

Next, to "'exceed[] authorized access,' one must have permission to access the computer at issue . . . ."[53] That is, the "exceed[ing] authorized access" provision refers to "an inside job."[54] Most data scraping cases do not involve insider jobs, and as such, it is the other part of § 1030(a)(2)(C), "without authorization," that has received the most focus. This part of the definition requires that the information at issue be information "'that the accesser is not entitled so to obtain or alter' . . . . '[a]nd, in context, the most "sensible reading of 'entitled' is as a synonym for 'authorized.'"'"[55] Thus, most CFAA claims in data scraping cases are brought under the contention that the scraping activity violated § 1030(a)(2)(C) of the CFAA because the scraper "intentionally accesse[d] a computer without authorization . . . and thereby obtain[ed] . . . information from [a] protected computer."[56] One of the primary difficulties in interpreting this phrase, however, has been the fact that "the language at issue was written in 1986—before the World Wide Web or websites existed."[57] As will be explored hereafter, interpretation of this phrase has differed drastically over the past two decades, largely in line with the expansion of the web.

### ii. History of CFAA Applicability

Over the past two decades, courts have varied drastically in their interpretations of the CFAA's key provisions. In data scraping cases, the debate has focused specifically on the Access Provision and how narrowly or broadly it should be interpreted. Some plaintiffs have likened "without access" to a physical world approach in furtherance of their claim. For example, in *hiQ Labs, Inc. v. LinkedIn Corp.*,[58] during oral arguments before the United States District Court for the Northern District of California, hiQ argued that social

[52] STEPHEN D. GANTZ, *IT Audit Drivers*, *in* THE BASICS OF IT AUDIT 129 (2014).

[53] Sandvig v. Sessions, 315 F. Supp. 3d 1, 23 (D.D.C. 2018) (quoting 18 U.S.C. § 1030(e)(6)).

[54] *Id.*; *see, e.g.*, United States v. Nosal (*Nosal I*), 676 F.3d 854, 858 (9th Cir. 2012).

[55] *Sandvig*, 315 F. Supp. 3d at 23 (quoting *Nosal I*, 676 F.3d at 858; Hedgeye Risk Mgmt., LLC v. Heldman, 271 F. Supp. 3d 181, 194 (D.D.C. 2017)).

[56] Ramirez v. SupportBuddy Inc., No. 17 CV 5781 (VB), 2018 U.S. Dist. LEXIS 76257, at *8 (S.D.N.Y. May 4, 2018) (quoting Jet One Grp., Inc. v. Halycon Jet Holdings, Inc., No. 08-CV-3980 (JS) (ETB), 2009 U.S. Dist. LEXIS 72579, at *5 (E.D.N.Y. Aug. 14, 2009)).

[57] *Sandvig*, 315 F. Supp. 3d at 24 (citing Orin S. Kerr, *Norms of Computer Trespass*, 116 COLUM. L. REV. 1143, 1161 (2016)).

[58] hiQ Labs, Inc. v. LinkedIn Corp., 273 F. Supp. 3d 1099 (N.D. Cal. 2017).

media sites such as LinkedIn are the modern equivalent of the town square and that allowing LinkedIn to choose who can access the site constituted a violation of the First Amendment.[59] LinkedIn drew a different analogy of access and described the information on their site like books in a public library.[60] Under this lens, the public library allows users to borrow books but only in compliance with specific library policies (i.e., LinkedIn's privacy policies and terms of service).[61]

Beyond this physical world approach, there are countless other ways to frame access.[62] For example, Professor Orin Kerr contends that only access-circumventing code-based restrictions should be a valid basis for CFAA claims.[63] Kerr reasons that limiting the scope of the CFAA in this way would allow for a balancing of internet freedom and data privacy.[64] Yet, this approach risks strongly favoring scrapers because scrapers rarely face more code-based restrictions than a regular human user would.[65]

In addition to the characterizations set forth by Kerr and other scholars, the history of the CFAA applicability in data scraping cases can be broken into four rough phases; these phases largely represent differences in the court's interpretation of these provisions. Rather than the "single, incoherent mess [that] some scholars have suggested," these phases merely represent ebbs and flows in how the court has weighted the rights of scrapers versus those of the scraped.[66]

The following discussion of these ebbs and flows relies heavily on research conducted by Professor Andrew Sellars, working in conjunction with researchers at BU and MIT.[67] Professor Sellars and his research team strove to compile the totality of CFAA-scraping cases. In total, the BU-MIT team found approximately sixty-one cases, most of which are based on commercial disputes, thirty-nine of which go beyond procedural questions and directly analyze the substantive claims, and only twelve of which are appellate cases.[68]

[59] *See id.* at 1115–16.

[60] *See* Alison Frankel, HiQ v. LinkedIn*: Does First Amendment Limit Application of Computer Fraud Law?*, REUTERS (Aug. 1, 2017, 5:17 PM), https://www.reuters.com/article/us-otc-linkedin/hiq-v-linkedin-does-first-amendment-limit-application-of-computer-fraud-law-idUSKBN1AH59X; Nicholas Iovino, *LinkedIn Lawsuit Against Data-Scraper Has Wide Implications*, COURTHOUSE NEWS SERV. (July 28, 2017), https://www.courthousenews.com/linkedin-lawsuit-data-scraper-wide-implications.

[61] *See User Agreement*, LINKEDIN, https://www.linkedin.com/legal/user-agreement (last visited Nov. 24, 2019).

[62] *See, e.g.*, Orin S. Kerr, *Cybercrime's Scope: Interpreting "Access" and "Authorization" in Computer Misuse Statutes*, 78 N.Y.U. L. REV. 1596 (2003).

[63] *See id.* at 1643.

[64] *See id.* at 1649.

[65] *See* Sellars, *supra* note 1, at 379–80.

[66] *Id.* at 412.

[67] *See id.* at 372.

[68] *See id.* at 388–90. It should be noted that, of the sixty-one opinions identified, about a third stem from just four underlying disputes: Facebook, Inc. v. Power Ventures, Inc., 252 F. Supp.

Their work allows us, here, to review four periods of case law and, in turn, to anticipate the future direction of the CFAA's applicability in data scraping cases. What follows is a review of these four rough phases, with special emphasis on recent court holdings.

### *1. First Phase: Broad Application of the CFAA*

The early 2000s presented a highly uncertain time for data scrapers as "courts embraced virtually all theories of authorization . . . . *any* mechanism could be used to determine that the scraper's access was unauthorized and therefore in violation of the statute."[69] In fact, during this time, a website restriction on the use of information could even "retroactively make the scraper's access unauthorized."[70] Scrapers could be found liable so long as the website could point to any mechanism which signaled that the access had been unauthorized, "be that contractual, technical, or otherwise."[71] For example, in one instance, the court found "that the filing of the complaint in the case itself served to signal that subsequent access was unauthorized, thus giving grounds for a preliminary injunction in the very same case."[72] Given such holdings, the first nine years of data scraping were "a very uncertain time for web scrapers."[73]

### *2. Second Phase: The Implementation of Technological Controls*

Beginning in the late 2000s, about a decade after the first scraping activities, an influential wave of cases began to adopt a narrow view of the CFAA.[74] During this second phase of time, courts began to deny claims against scrapers where "websites merely placed restrictions on the *use* of the data hosted

3d 765 (N.D. Cal. 2017) (a decade-long dispute between Facebook and social network aggregator Power.com); CollegeSource, Inc. v. AcademyOne, Inc., 597 F. App'x 116 (3d Cir. 2015) (where CollegeSource sued its competitor, AcademyOne, for a number of claims, including violating the CFAA, after AcademyOne bots copied course information from CollegeSource's site); Tamburo v. Dworkin, 974 F. Supp. 2d 1199, 1206 (N.D. Ill. 2013) (a series of claims brought by a scraper of dog pedigree databases against hosts who described him as a "thief"); EarthCam, Inc. v. Oxblue Corp., No. 1:11-cv-02278-WSD, 2012 U.S. Dist. LEXIS 191822 (N.D. Ga. Mar. 26, 2012) (litigation involving trade secrets between rival security camera companies).

[69] Sellars, *supra* note 1, at 393; *see, e.g.*, Ticketmaster L.L.C. v. RMG Techs., Inc., 507 F. Supp. 2d 1096, 1113 (C.D. Cal. 2007); EF Cultural Travel BV v. Zefer Corp., 318 F.3d 58, 63 (1st Cir. 2003); *and* Healthcare Advocates, Inc. v. Harding, Earley, Follmer & Frailey, 497 F. Supp. 2d 627, 649 (E.D. Pa. 2007).

[70] Sellars, *supra* note 1, at 394.

[71] Sellars, *supra* note 1, at 379.

[72] *Id.* at 394 (citing Register.com, Inc. v. Verio, Inc, 126 F. Supp. 2d 238, 249 (S.D.N.Y. 2000)); *see also Register.com*, 126 F.Supp. 2d at 249 ("It is clear since at least the date this lawsuit was filed that Register.com does not consent to Verio's use of a search robot . . . .").

[73] Sellars, *supra* note 1, at 395.

[74] *See* Jonathan Mayer, *The "Narrow" Interpretation of the Computer Fraud and Abuse Act: A User Guide for Applying* United States v. Nosal, 84 GEO. WASH. L. REV. 1644, 1645 (2016); *see, e.g.*, LVRC Holdings LLC v. Brekka, 581 F.3d 1127 (9th Cir. 2009); *see also* United States v. Nosal (*Nosal I*), 676 F.3d 854, 863 (9th Cir. 2012).

on their site as opposed to restrictions on *access* to a website, and looked more towards code-based controls to interpret the scope of a scraper's authorization."[75] Of course, "[t]his approach tended to benefit web scrapers."[76]

At the same time, courts also began to focus more carefully on the presence of technical controls (rather than contract-based controls) along with whether imposed restrictions were "access restrictions" or more closely resembled "mere 'use restrictions.'"[77] Use-based restrictions, which restrict types of use,[78] were viewed as overly-expansive and created the possibility for a slippery slope problem, wherein any use outside of intended, outlined, or required use would represent a CFAA violation.[79]

Two cases particularly highlight the courts' thinking at this time: *LVRC Holdings LLC v. Brekka* in 2009 and *United States v. Nosal (Nosal I)* in 2012. In *Brekka*, an employer brought suit against its employee who, in the course of his employment, sent emails from his work computer to his personal computer.[80] Here, the court rejected the plaintiff-employer's argument that the employee's behavior constituted of access without authorization based on duty-based obligations and found, instead, that Brekka had not violated the CFAA because he had permission to both use the computer and to access the documents.[81]

In a somewhat similar set of circumstances, in *Nosal I*, an employer argued that an employee's access to its customer database was unauthorized when the access exceeded the purpose for which the employee was hired.[82] Here again, the court found that "such 'use restrictions' were improper grounds for liability under the CFAA" and suggested a concern in dicta "that [the plaintiff's] reasoning would also bar claims based on violations of restrictions memorialized in websites' terms of use."[83] The narrowed approach employed by the *Nosal I* court and other courts during this time period ultimately created a favorable landscape for scraping. In fact, five cases against scrapers during

[75] Sellars, *supra* note 1, at 379.

[76] *Id.* at 379–80.

[77] *Id.* at 396 (quoting *Nosal I*, 676 F.3d at 863–64); *see also* Wentworth-Douglass Hosp. v. Young & Novis Prof. Ass'n., No. 10-cv-120-SM, 2012 U.S. Dist. LEXIS 90446, at *13 (D.N.H. June 29, 2012) (questioning whether a purported "access restriction" is just a "use restriction" in disguise: "[D]enominating limitations as 'access restrictions' does not convert what is otherwise a use policy into an access restriction").

[78] *See, e.g.*, *Nosal I*, 676 F.3d at 866–67 (Silverman, J., dissenting) (where the court found that such "use" restrictions were improper grounds for liability under the CFAA).

[79] *See id.*

[80] *See* LVRC Holdings LLC v. Brekka, 581 F.3d 1127, 1129−30 (9th Cir. 2009).

[81] *See id.* at 1135.

[82] *See Nosal I*, 676 F.3d at 856.

[83] Sellars, *supra* note 1, at 397.

this phase "were either adjudicated in the scraper's favor or dismissed,"[84] with some courts noting an inability to find CFAA liability.[85]

### *3. Third Phase: Revocation Theory*

*Brekka* inadvertently opened the door for a new wave of lawsuits.[86] Although the *Brekka* court ultimately ruled in favor of the scraping defendant, the court opined whether related circumstances would have constituted a CFAA violation[87]— dicta which would ultimately impact the next five years of data scraping case law. For example, the court noted that defendant Brekka would have violated the CFAA if he had either continued use after the employer had rescinded permission[88] or if he had accessed plaintiff-company's information after he left the company.[89] Courts quickly "seized" on this "extraneous language,"[90] resulting in a sharp turn from the trend of narrowing.

*Craigslist Inc. v. 3Taps Inc.*[91] was the first case to exemplify this "revocation" period.[92] Here, Craigslist, an online classified ads website, brought suit against a series of services that facilitated its listing processes. The plaintiff argued that these service providers had engaged in unauthorized use because they violated Craigslist's terms of use, failed to abide by Craigslist's cease and desist letters, and circumnavigated IP blocks.[93] Citing *Brekka*, the *3Taps* court observed that "computer owners have the power to revoke the authorizations they grant,"[94] and because Craigslist had explicitly revoked access, any access by 3Taps thereafter was unauthorized.[95] Of course, this narrowed perspective of the CFAA opened the door for a series of new cases in which courts found CFAA violations where the scraper continued access despite clear and unequivocal articulations that access had been revoked.[96] For

[84] *Id.* at 398–99; *see, e.g.*, EarthCam, Inc. v. OxBlue Corp., 49 F. Supp. 3d 1210, 1245 (N.D. Ga. 2014); Bilotta v. Citizens Info. Assocs., LLC, No. 8:13-cv-2811-T-30TGW, 2013 U.S. Dist. LEXIS 197367, at *11 (M.D. Fla. Dec. 20, 2013); CollegeSource, Inc. v. AcademyOne, Inc., 597 F. App'x 116, 131 (3d Cir. 2015); Snapt Inc. v. Ellipse Commc'ns Inc., 430 F. App'x 346, 353 (5th Cir. 2011); Cvent, Inc. v. Eventbrite, Inc., 739 F. Supp. 2d 927, 940 (E.D. Va. 2010).

[85] *See* Sellars, *supra* note 1, at 399.

[86] *See Brekka*, 581 F.3d 1127.

[87] *See id.* at 1135.

[88] *See id.* at 1136.

[89] *See id.*

[90] Sellars, *supra* note 1, at 401.

[91] Craigslist Inc. v. 3Taps Inc., 942 F. Supp. 2d 962, 969–70 (N.D. Cal. 2013).

[92] *See* Sellars, *supra* note 1, at 402.

[93] *See Craigslist*, 942 F. Supp. 2d at 969–70.

[94] Sellars, *supra* note 1, at 404 (quoting Craigslist Inc. v. 3Taps Inc., 964 F. Supp. 2d 1178, 1183 (N.D. Cal. 2013)).

[95] *See id.*

[96] *See, e.g.*, CouponCabin LLC v. Savings.com, Inc., No. 2:14-CV-39-TLS, 2017 WL 83337, at *3 (N.D. Ind. Jan. 10, 2017).

example, in one case, *CouponCabin LLC*,[97] the court went so far as to find that the act of blocking an IP-address alone "served as effective notice of revocation of access,"[98] even without other warnings or indications of revocation. There, the *CouponCabin* court noted that "[r]evocation of website access would have been sufficient to give the Defendants constructive notice that they were without authorization to act as they allegedly did."[99] Yet, such a restrictive approach was problematic in the sense that it potentially puts any user who received a "website is down" notification at risk of violating the CFAA.

In *United States v. Nosal* (*Nosal II*), the Ninth Circuit attempted to clarify the meaning of "without authorization" within the context of the CFAA.[100] The defendant in this case, David Nosal, had resigned from his position at an executive search and recruiting company, Korn/Ferry International.[101] In the process of departing from the company, Nosal had agreed not to compete with Korn/Ferry for one year.[102] Yet, a few months after leaving, Nosal recruited three of the company's current employees to help him start a competing business.[103] The employees, using their company usernames and passwords, downloaded a significant volume of highly confidential and proprietary data from company computers, including source lists, names, and contact information for executives.[104] Nosal then used this information to start his own contracting business. When Korn/Ferry received a tip advising that Nosal was conducting his own business in violation of his non-compete agreement, the company filed suit. The *Nosal II* court ultimately found that Nosal had violated the CFAA when he used login credentials to gain access to his former employer's computer systems after his credentials had been "affirmatively revoked."[105] The court reasoned that because Nosal's authorization had been revoked when he left the company, his actions therefore constituted accessing a protected computer without authorization.

Finally, one of the most noteworthy cases to uphold revocation theory was *Facebook, Inc. v. Power Ventures, Inc.*. In *Power Ventures*, the court held 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."[106] In this case, the defendant operated a site that extracted and aggregated users' social networking information from Facebook. The defendant had obtained this

[97] *See id.*

[98] Sellars, *supra* note 1, at 405.

[99] *CouponCabin*, 2017 WL 83337, at *3; *cf.* hiQ Labs, Inc. v. LinkedIn Corp., 273 F. Supp. 3d 1099, 1113 (N.D. Cal. 2017) ("[T]he Court has serious doubt whether LinkedIn's revocation of permission to access the public portions of its site render[ed] hiQ's access 'without authorization' within the meaning of the CFAA.").

[100] United States v. Nosal (*Nosal II*), 828 F.3d 865, 868 (9th Cir. 2016).

[101] *See id.* at 870.

[102] *See id.*

[103] *See id.*

[104] *See id.* at 870–71.

[105] *Id.* at 868.

[106] Facebook, Inc. v. Power Ventures, Inc., 828 F.3d 1068,1077 (9th Cir. 2016).

data by accessing password-protected Facebook member profiles.[107] Upon realizing the presence of the defendant's bots, Facebook sent a cease and desist letter demanding that Power Ventures stop accessing information on users' pages.[108] However, Power Ventures refused to cease its scraping operations, so Facebook filed suit in the Northern District of California.[109] When looking at the issue of CFAA applicability, the court found that Power Ventures had, in fact, violated the CFAA by continuing to access Facebook's servers without authorization after receiving written notice from Facebook demanding to cease access.[110] Together, these cases stood for the proposition that, once access had been clearly revoked, any continued access made by the scraper could be deemed a violation of the CFAA.

### *4. Fourth Phase: The Demise of the CFAA*

Recent decisions have once again narrowed the CFAA on the basis of the public interest, First Amendment claims, recognition of the technical similarities between web scraping and web browsing, and arguments that the CFAA was never intended to apply to scraping. This fourth phase has been exemplified by two recent cases: *hiQ Labs, Inc. v. LinkedIn Corp.*, a case brought when LinkedIn blocked hiQ's access to data from its site, and *Sandvig v. Sessions*, where a claim was brought by a group of scholars who used scraping in their research and sought a First Amendment protection to scrape data, contending that it implicated their rights "to record or preserve information."[111] What follows is an exploration of the arguments that scrapers have recently succeeded on, with a focus on these two particular cases.

#### a. The Public Forum and the First Amendment

In recent years, courts have focused carefully on whether the scraped data was public or private. Successful CFAA claims against data scraping have framed the scraping of password-protected (or private) data as "access without authorization," and thereby actionable under the CFAA.[112] Likewise, CFAA claims have been unsuccessful when the scraper was successful in showing that the data scraped was publicly available.[113] Yet, whether data is deemed "public" is a question deeply intertwined with First Amendment concerns, as the principle that the "'government has no power to restrict expression because of its message, its ideas, its subject matter, or its content' applies with full force

[107] *See id.* at 1072–73.

[108] *See id.* at 1073.

[109] *See id.*

[110] *See id.* at 1077.

[111] Sandvig v. Sessions, 315 F. Supp. 3d 1, 15 (D.D.C. 2018); hiQ Labs, Inc. v. LinkedIn Corp., 938 F.3d 985, 989 (9th Cir. 2019).

[112] *See, e.g.,* Facebook, Inc. v. Power Ventures, Inc., 844 F.3d 1058, 1067 (9th Cir. 2016).

[113] *See, e.g., hiQ Labs, Inc.*, 938 F.3d at 1003–04.

in a traditional public forum."[114] Of course, while the First Amendment provides broad protections, a hacker cannot legally break into a Gmail account and copy the account-holder's emails, nor can one "legally reach into someone else's mailbox and open her mail."[115] Further, "[t]he First Amendment does not give someone the right to breach a paywall on a news website any more than it gives someone the right to steal a newspaper."[116] For data to truly be deemed "private" or protected from scraping, effort must be taken to remove the data from the public forum.

Recent cases, including *Sandvig v. Sessions*[117] and *hiQ Labs, Inc. v. LinkedIn Corp.*,[118] have focused on whether the data was public or private. In *Sandvig v. Sessions*, a claim was brought by a group of scholars and journalists who used scraping in their research and were aware that such techniques were blocked by the terms of service of many websites they scraped from.[119] The scholars sought First Amendment protection to scrape, contending that it "implicated their rights to record or preserve information, and to publish the information that they found."[120] In reviewing the constitutional claim at hand, the *Sandvig* court focused heavily on the public nature of the scraped data, noting that "heightened First Amendment scrutiny [was] appropriate"[121] because the plaintiff's "Access Provision both limit[ed] access to and burden[ed] speech in the public forum that is the public Internet."[122] The court added that, even though the defendant had not based its primary claim on speech, a dispute is "'subject to First Amendment scrutiny' if 'it restricts access to traditional public fora.'"[123]

In attempting to distinguish the difference between public versus private data, the *Sandvig* court rejected a physical world approach, noting that the public internet is more of a "metaphysical"[124] space than a geographic one and, further, is "too heavily suffused with First Amendment activity . . . to sustain a direct parallel to the physical world."[125] The court further added that scraping activity "plausibly falls within the ambit of the First Amendment,"[126] writing:

> [T]he information plaintiffs seek is located in a public forum. Hence, plaintiffs' attempts to record the contents of public websites for

[114] *Sandvig*, 315 F. Supp. 3d at 29 (quoting McCullen v. Coakley, 573 U.S. 464, 477 (2014)).

[115] *Id.* at 13.

[116] *Id.*

[117] *See id.* at 1.

[118] hiQ Labs, Inc. v. LinkedIn Corp., 273 F. Supp. 3d 1099 (N.D. Cal. 2017).

[119] *See Sandvig*, 315 F. Supp. 3d at 8.

[120] Sellars, *supra* note 1, at 411.

[121] *Sandvig*, 315 F. Supp. 3d at 29.

[122] *Id.*

[123] *Id*. (quoting McCullen v. Coakley, 573 U.S. 464, 476 (2014)).

[124] *Id.* at 12 (quoting Rosenberger v. Rector of Univ. of Va., 515 U.S. 819, 830 (1995)).

[125] *Id.* at 13.

[126] *Id.* at 15.

> research purposes are arguably affected with a First Amendment interest.
>
> . . . .
>
> . . . "[A]bsent any evidence that the speech [would be] used to gain a material advantage," plaintiffs' false speech on public websites retains First Amendment protection, and rendering it criminal does not appear to advance the government's proffered interests. Hence, plaintiffs have plausibly alleged an as-applied First Amendment claim, and the motion to dismiss that claim will be denied.[127]

Ultimately, the *Sandvig* court adopted a particularly narrow view of the CFAA, holding that scraping falls outside of the CFAA's reach when it is in the public forum and thus protected by the First Amendment.[128]

In a separate case, *hiQ Labs, Inc. v. LinkedIn Corp.*,[129] hiQ challenged a cease and desist letter it received after scraping LinkedIn's data. HiQ labs filed a lawsuit in the Northern District of California, requesting a declaration that its scraping of LinkedIn's data was lawful[130] and that LinkedIn's blocking of hiQ's scrapers "constitute[d] a violation of free speech under the California Constitution."[131] Here, the court considered whether, "by continuing to access public LinkedIn profiles after LinkedIn ha[d] explicitly revoked permission to do so, hiQ ha[d] 'accesse[d] a computer without authorization' within the meaning of the CFAA."[132] In arguing that hiQ had, in fact, violated the CFAA, LinkedIn relied primarily on *Power Ventures*[133] and *Nosal II*.[134] Although both cases upheld CFAA liability, the fact that the data in those cases involved private data, rather than public data, did not go unnoticed. As the *hiQ* court wrote, both *Power Ventures* and *Nosal II* involved "unauthorized intruders reach[ing] into what would fairly be characterized as the private interior of a computer system not visible to the public."[135] The *hiQ* court distinguished *Power Ventures* and *Nosal II* from the case at hand by explaining that LinkedIn was attempting to apply the CFAA to the defendant's aggregation and downloading of *public* data, not private data. Given the public nature of LinkedIn's data, the court ruled in hiQ's favor, noting that it had "serious doubt whether LinkedIn's revocation of permission to access the public portions of its site render[ed] hiQ's access 'without authorization' within the meaning of the CFAA."[136]

[127] *Id.* at 16, 30 (third alteration in original) (citations omitted) (quoting United States v. Alvarez, 567 U.S. 709, 723 (2012)).

[128] *See id.* at 34.

[129] hiQ Labs, Inc. v. LinkedIn Corp., 273 F. Supp. 3d 1099 (N.D. Cal. 2017).

[130] *See id.* at 1104.

[131] *Id.* at 1103.

[132] *Id.* at 1108 (third alteration in original).

[133] Facebook, Inc. v. Power Ventures, Inc., 844 F.3d 1058, 1067 (9th Cir. 2016).

[134] United States v. Nosal (*Nosal II*), 828 F.3d 865 (9th Cir. 2016).

[135] hiQ Labs, Inc. v. LinkedIn Corp., 273 F. Supp. 3d 1099, 1109 (N.D. Cal. 2017).

[136] *Id.* at 1113.

Following the ruling, many practitioners were hesitant to interpret the Northern District of California's holding as a green light for data scraping. That was, of course, until the Ninth Circuit affirmed the district court's ruling.[137] On September 9, 2019, the Ninth Circuit found that "the legislative history . . . of the CFAA . . . support[ed] the district court's distinction between 'private' computer networks and websites, protected by a password authentication system and 'not visible to the public,' and websites that are accessible to the general public."[138]

### b. The Technical Similarities Between Web Scraping and Web Browsing

Scrapers have also successfully argued for a narrowed view of the CFAA based on the technical similarities between web scraping and web browsing. The foundation of this argument has been the fact that "[c]opying is essential to the functioning of the Internet."[139] When a user watches a video online, for instance, a copy of that movie is "temporarily stored on the computer's random access memory (RAM), or cached on or near the central processing unit (CPU)."[140] Web browsers work in a similar way: they cache data from websites and then store short-term copies of the website content.[141] Somewhat confusingly, courts have held work in RAM to "constitute the copying required for copyright infringement," whereas caching, in contrast, has been deemed "transformative use and thus not infringement."[142] Copying is also a critical component of searching in the sense that anytime a user searches something on Google, Google's web crawler fetches web pages and caches the sites, a process involving "copying [the pages] in their entirety, and keeping a backup of the website's content in case it becomes unavailable."[143]

The *Sandvig* court firmly articulated the fact that scraping is not unlike many other internet-based, or even manual, copying activities. Judge Bates of the United States District Court for the District of Columbia opined:

> That plaintiffs wish to scrape data from websites rather than manually record information does not change the analysis. Scraping

[137] *See* hiQ Labs, Inc. v. LinkedIn Corp., 938 F.3d 985, 1005 (9th Cir. 2019).

[138] *Id.* at 1003.

[139] Riley, *supra* note 22, at 251.

[140] *Id.*

[141] *See id.* at 252.

[142] *Id. Compare* Quantum Sys. Integrators, Inc. v. Sprint Nextel Corp., 338 F. App'x 329, 337 (4th Cir. 2009) (holding that temporary storage in RAM was copying sufficient for copyright infringement), *and* DocMagic, Inc. v. Ellie Mae, Inc., 745 F. Supp. 2d 1119, 1148 (N.D. Cal. 2010) ("The Ninth Circuit has held that loading the program into the computer's RAM constitutes an act of 'copying' for the purposes of copyright law."), *with* Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146, 1162 (9th Cir. 2007) (holding that caching was not copying required to allege copyright infringement), *and* Field v. Google Inc., 412 F. Supp. 2d 1106, 1118 (D. Nev. 2006) (holding that Google search's caching of copyrighted works was fair use).

[143] Riley, *supra* note 22, at 252; *see, e.g.*, *Googlebot*, GOOGLE, https://support.google.com/webmasters/answer/182072?hl=en (last visited Nov. 24, 2019).

> is merely a technological advance that makes information collection easier; it is not meaningfully different from using a tape recorder instead of taking written notes, or using the panorama function on a smartphone instead of taking a series of photos from different positions.[144]

Similarly, both the *hiQ* district court and Ninth Circuit compared data scraping to far less controversial forms of copying. As hiQ pointed out, scraping is "a common method of gathering information, used by search engines, academic researchers, and many others."[145] According to hiQ, "letting established entities . . . decide who can scrape that data from otherwise public websites gives those entities outsized control over how such data may be put to use."[146] In an internet world where copying is the norm, hiQ contended that it seemed odd to allow one form of bots while banning another, noting that LinkedIn did allow its site to be accessible via public search engines such as Google; sites which function using bots.[147] After weighing both of these public interest arguments, the Ninth Circuit ultimately found in favor of hiQ.

### c. The Purpose of CFAA

A recent focus on the CFAA's original intent and whether the statute is suitable for application in any kind of scraping case has also led to outcomes favorable to scrapers. In *Sandvig*, the court noted that the legislative history of the CFAA indicates that Congress was interested in passing the Access Provision to "prevent the digital equivalent of theft," not web scraping.[148] The court added that, based on the legislative history, Congress "viewed exceeding authorized access as the digital equivalent of being allowed into a house but entering a room within it that the owner has declared to be off-limits."[149] In addition to the fact that the language at issue was written before the World Wide Web, the *Sandvig* court ruled that the amendment history of the statute did not suggest that Congress intended for this original conception to change.[150] Likewise, in *hiQ Labs*, the district court pointed to the fact that the CFAA was never intended to be applied to public internet data, writing, "[t]he CFAA was not intended to police traffic to publicly available websites on the Internet—the Internet did not exist in 1984."[151] Together, these courts, and other recent rulings, have favored scraping activities based on the fact that the CFAA was never intended to apply to scraping.

[144] Sandvig v. Sessions, 315 F. Supp. 3d 1, 16 (D.D.C. 2018).

[145] hiQ Labs, Inc. v. LinkedIn Corp., 938 F.3d 985, 1004 (9th Cir. 2019).

[146] *Id.* at 1005.

[147] *See id.* at 990.

[148] *Sandvig*, 315 F. Supp. 3d at 29.

[149] *Id.* at 24.

[150] *See id.*

[151] hiQ Labs, Inc. v. LinkedIn Corp., 273 F. Supp. 3d 1099, 1109 (N.D. Cal. 2017).

## CONCLUSION

For the past twenty years, data scraping has existed in a realm of legal uncertainty. Amidst the absence of a statute specifically designed to address scraping, complainants seeking to challenge scraping activities have turned to the CFAA as their primary avenue for doing so. However, recent decisions have narrowed the applicability of the CFAA in data scraping cases on the basis of the public interest and First Amendment claims; a recognition of the technical similarities between web scraping and web browsing; and based on the fact that the CFAA is ill-suited for cases involving data scraping. For these reasons, the Second, Fourth, and Ninth Circuit courts of appeals, along with numerous district courts,[152] have all preferred to interpret the CFAA narrowly. This trend will likely continue.

Moving forward, courts will be increasingly likely to consider data public and thus protected by the First Amendment. Even where data is password-protected, the CFAA may eventually be found to be deemed inapplicable based on courts' focus on the CFAA's original purpose. Given this fact, the CFAA may be phased out in all cases involving data scraping, even where the data is password-protected. If this becomes the case, challenging scraping may require bringing other types of claims against scrapers such as unjust enrichment or claims under the DMCA—claims which, historically, have not been particularly successful. Perhaps sooner rather than later, the legal landscape will bestow scrapers with a green light to copy website data without fear of CFAA liability.

[152] *See Sandvig*, 315 F. Supp. 3d at 22.