Northwestern Journal of Technology and Intellectual Property

Volume 13 | Issue 3 Article 2

2015

Using the Computer Fraud and Abuse Act to Secure Public Data Exclusivity

Nicholas A. Wolfe

N O R T H W E S T E R N
JOURNAL OF TECHNOLOGY
AND
INTELLECTUAL PROPERTY

Hacking the Anti-Hacking Statute: Using the Computer Fraud and Abuse Act to Secure Public Data Exclusivity

Nicholas A. Wolfe

2015 VOL. 13, NO. 3

Northwestern Journal of Technology and Intellectual Property
Volume 13, Number 3 (2015)

Hacking the Anti-Hacking Statute: Using the Computer Fraud and Abuse Act to Secure Public Data Exclusivity

By Nicholas A. Wolfe*

ABSTRACT

In August, 2015, hackers exposed approximately 33 million user records associated with the extra-marital affair website Ashley Madison. The hackers made this data available to the public through torrents and other file sharing protocols. This data became instantly irresistible to the media and suspicious spouses everywhere. However, is accessing the user records illegal under the Computer Fraud and Abuse Act? While many legal scholars agree that accessing or publishing this data is not likely a violation of the Computer Fraud and Abuse Act, the United States Attorney's office does not necessarily see it that way.

"Once you download or distribute hacked information without specific permission or a fair use license, you've exposed yourself to potential criminal liability under the Computer Fraud and Abuse Act," says a representative of the Chicago U.S. Attorney's office. "An individual who retweets or forwards a link to a website containing hacked information could potentially be viewed as an accessory to the hack after the fact."

A "hack after the fact" not only leads to criminal penalties but a civil cause of action under the Act, which is quickly becoming a leading statute in U.S. cybersecurity law.

This Article describes problems inherent in the Act when compared with modern web-based applications and how savvy civil litigators are "hacking" the Computer Fraud and Abuse Act for their own purposes, namely as a para-copyright tool. This "hack" is accomplished by exposing two vulnerabilities: (1) the literal application of the term "access controls" encompassing token controls; and (2) the mere facial review of loss declarations. For example, by taking advantage of these two vulnerabilities, attorneys for Craigslist were able to secure exclusivity to the publicly-available advertisements on its website.

This Article's solution to the vulnerabilities is to build in reference to data security standards and define the type of data protectable under the Act, specifically private and confidential data.

* General Counsel at BitTitan, Inc.

TABLE OF CONTENTS

INTRODUCTION

¶1 Work smarter, not harder. Perhaps no other saying better captures the era of hyper-productivity and automation in which we live. Titles such as 'Top Ten Hacks to Avoid Paywalls,' 'Five Ways You're Wasting Your Time,' and 'One Weird Trick' fly across our computer screens on a commoditized basis.[1] Tips and tricks that automate our lives and help us get more done, faster – better living through automation. However, as these shortcut solutions (colloquially referred to as "hacks") get better and automation advances, the question arises: When does working smarter cross the line into cheating? Or put differently, when do "hacks" rise to the level of computer hacking?

¶2 The Computer Fraud and Abuse Act (CFAA) was drafted to draw the line between "hacks" and hacking. Drafted in 1986 and amended with a frequency similar to iOS updates, there are nine ways to violate the CFAA.[2] This article covers just one. Subsection 1030(a)(2)(C) provides that anyone who (1) intentionally accesses without or in excess of authorization (2) information (3) that causes a plaintiff at least $5,000 loss in a 1-year period is engaged in hacking under the CFAA.[3] Access without or in excess of authorization is generally interpreted with reference to circumvention of some access control. Concerning the $5,000 loss requirement, the plaintiff is generally required to submit a supporting declaration.[4]

¶3 At its core, the CFAA is intended to deter the exploitation of computer system vulnerabilities that cause damage to the computer system. Ironically, in so doing, the CFAA has exposed two of its own vulnerabilities in the face of a dynamic technological landscape:

(1) Literal application of "access control" to encompass any access control, including mere token controls;[5]
(2) Cursory review of loss declarations to include any first party expense, whether incurred reasonably or unreasonably.[6]

¶4 The consequences of these vulnerabilities are widespread. For example, if you are reading this article in a Chrome browser and were to open a new tab and navigate to the *Seattle Times* website, read your maximum article limit, and then press Ctrl + U to view the source code and read one additional article in HTML form, the *3Taps* court would likely interpret your actions as hacking under the CFAA.[7]

¶5 These two vulnerabilities enable unchecked application of a powerful criminal statute as "as a tactical tool to gain business or litigation advantage," particularly as a para-copyright tool to secure exclusivity to otherwise publicly accessible data.[8]

---

[1] *See, e.g.*, Thorin Klosowski, *Bypass Paywalls and Other Blocks with a Few Google Proxy Servers*, LIFEHACKER (Jul. 16, 2013, 6:30 AM), http://lifehacker.com/use-google-as-a-proxy-server-to-bypass-paywalls-and-oth-799030304.

[2] 18 U.S.C. § 1030 (2013).

[3] 18 U.S.C. § 1030(a)(2)(C).

[4] 18 U.S.C. § 1030(c)(4)(A)(i).

[5] *See generally* Int'l Airport Ctrs., L.L.C. v. Citrin, 440 F.3d 418 (7th Cir. 2006).

[6] *See generally*, Craigslist, Inc. v. 3Taps Inc., 942 F. Supp. 2d 962 (N.D. Cal. 2013).

[7] *Id.*

[8] *Id.* at 969, n.8.

¶6 This article proposes a cure for these two vulnerabilities by first restricting application of Subsection 1030(a)(c)(2) to "protected data" only. The First, Fourth, and Ninth circuits follow this approach already.[9] The Seventh, Fifth, and Eleventh Circuits disagree.[10] Second, to add clarity to the meaning of protected data, this article aims to distinguish between protected and unprotected data with reference to access controls and to exclude mere token access controls, achieved by defining in the context of good cybersecurity practices. Notably, an access control that by default admits all traffic should be scrutinized more thoroughly than an access control that by default denies all traffic because of the potential for selective and tactical wielding of the CFAA as a sword against competitors. Third, this article proposes the elimination of cursory review of loss declarations by applying a three-factor test to better limit plaintiff abuse. Currently, a plaintiff need only hire an expensive forensic IT consultant to meet its requirement under the statute, whether any actual harm in fact occurred or was likely to occur.

I. THE *3TAPS* DECISION

¶7 The *3Taps* decision most clearly captures the two CFAA vulnerabilities discussed in this article. Reading the decision, one may be reminded of that proverbial frog in a pot of water, slowly reaching the boiling point. The judicial analysis seems logical degree by degree, but the ultimate decision is hard to swallow. 3Taps, Inc. operates an apartment-listing website that displays available listings geographically.[11] One of 3Taps' largest sources of data for apartment listings was Craigslist.[12] Craigslist is an online classified advertisement web site on which users post a variety of classified advertisements, including apartment listings.[13] 3Taps developed a web scraping software to pull listings from Craigslist in real time.[14] Upon realizing how 3Taps was accessing its website, Craigslist sent a cease and desist notice and implemented an IP block.[15] 3Taps used a proxy server to circumvent the IP block and continued to gather apartment listing data directly from Craigslist.[16] Craigslist commenced a lawsuit against 3Taps in which the complaint was eventually amended to include a CFAA claim and its California state law counterpart.[17] In denying 3Taps' motion to dismiss the CFAA claim, the court extended the CFAA to cover public data when such data is behind an IP block and combined with a cease and desist notice.[18] On the surface, this decision seems well-reasoned – Craigslist asked 3Taps to stop using Craigslist data and 3Taps not only refused but used computer technology to

[9] *See* WEC Carolina Energy Solutions L.L.C. v. Miller, 687 F.3d 199, 206-07 (4th Cir. 2012); United States v. Nosal, 676 F.3d 854, 859 (9th Cir. 2012); EF Cultural Travel BV v. Explorica, Inc., 274 F.3d 577, 579 (1st Cir. 2001).

[10] *See* United States v. John, 597 F.3d 263, 271-73 (5th Cir. 2010); United States v. Rodriguez, 628 F.3d 1258 (11th Cir. 2010); Int'l Airport Ctrs., L.L.C. v. Citrin, 440 F.3d 418 (7th Cir. 2006).

[11] *See 3Taps*, 942 F. Supp. 2d at 966.

[12] *Id.*

[13] *Id.*

[14] *Id.*

[15] *Id.*

[16] *Id.* at 967.

[17] *Id.*

[18] *Id.* at 969, n.8.

circumvent the IP block.[19] However, when applied to modern web browsing contexts, the decision loses its logical footing.

II. WEB SCRAPING

¶8 To understand modern web browsing, one must first understand how web scraping fits in. Before the concept of web scraping, there was web crawling. Web crawling enables the copying and transposing of massive data sets into machine-readable and analyzable formats.[20] The process of web crawling shares many similarities with web scraping. For instance, like web crawling, web scraping involves programmatic browsing of web pages and targeted data collection.[21] Moreover, both methods collect data through end-user accessible network ports.[22] Unlike web crawling, however, web scraping does not obtain permission from each website. Instead, web scraping employs additional technologies to mimic human browsing and delve deeper into each website.[23]

¶9 Both technologies are methods to grab large sums of data, fast and without human interaction. Indeed, certain web scraping technologies like the browser add-on that enables users to find free Pacer[24] resources represent a bright spot in web scraping and its potential to advance the progress of science and the useful arts.[25] Conversely, exploitative forms such as device or user impersonation designed to obtain private (often confidential) data are less than inspiring, but nevertheless are generally limited in application to the gathering of private or confidential data.[26]

¶10 Modern web browsing involves both web scraping and web crawling. Accordingly, in a cybersecurity or trade secret protection context, it is critical to define the scope of accessible data with reference to both. That is, it would be ineffective data protection to restrict crawling but not scraping. For instance, the USPTO website restricts crawling of patents and trademarks by including a disallow instruction to crawlers in its public directory in a file labeled "robots.txt" and restricts scraping by using a form that requires human interaction before accessing data, also known as a CAPTHCA form.[27] To

---

[19] *Id.* at 967.

[20] *See* Sys. and Method for Providing Dynamic User Info. in an Interactive Display, U.S. Patent No. 7,343,567 (filed Apr. 25, 2003) (issued Mar. 11, 2008); Andrew Hogue & David Karger, *Thresher: Automating the Unwrapping of Semantic Content from the World Wide Web*, PROCEEDINGS OF THE 14TH INTERNATIONAL CONFERENCE ON WORLD WIDE WEB. ACM (2005), *available at* http://dl.acm.org/citation.cfm?id=1060762.

[21] Seyed M. Mirtaheri et al., *A Brief History of Web Crawlers*, PROCEEDINGS OF THE 2013 CONFERENCE OF THE CENTER FOR ADVANCED STUDIES ON COLLABORATIVE RESEARCH (2013), *available at* http://www.site.uottawa.ca/~bochmann/Curriculum/Pub/2013%20-%20A%20brief%20history%20of%20Web%20crawlers.pdf.

[22] *Id.*

[23] Bjorn Hartmann, Scott Doorley & Scott R. Klemmer, *Hacking, Mashing, Gluing: Understanding Opportunistic Design*, 7.3 Pervasive Computing IEEE, 46, 46-54 (2008), *available at* http://bjoern.org/papers/hartmann-pervasive2008.pdf; (notably, web scraping employs methodologies such as executing Javascript, submitting forms, impersonating a browser, and ignoring robots exclusionary headers).

[24] RECAP the Law, About, https://www.recapthelaw.org/about/ (last visited Mar. 17, 2014).

[25] Making the writing of law review articles, for example, faster and less expensive.

[26] *See* Bjorn Hartmann et al., *Hacking, Mashing, Gluing: Understanding Opportunistic Design*, 7.3 Pervasive Computing IEEE, 46, 47-48 (2008), *available at* http://bjoern.org/papers/hartmann-pervasive2008.pdf.

[27] *See* USPTO.gov website, Public Application Information Retrieval (PAIR) gateway at

summarize, effectively restricting access to data in the modern web browsing context necessarily involves attention to web scraping.

III. UNAUTHORIZED ACCESS AND MERE TOKEN ACCESS CONTROLS

¶11 The first vulnerability of the CFAA as highlighted in the *3Taps* decision concerns access controls. This Section explores the notable absence of a definition for the term access controls and how that absence is handled among the circuits, which is often by analogy to other legal doctrines such as the law of trespass. This section then explores how courts ignore web technology when applying the CFAA. Instead, many courts opt to evaluate the mental state of the defendant or consider the issue of access control in a vacuum and without regard to industry standards, leading to an ever-expanding scope of hacking that includes innocuous and useful web technologies.[28] More specifically, this approach expands the definition of access controls to encompass mere token controls, which in the context of an anti-hacking statute operate instead as use controls. By encompassing use controls in its interpretation of the CFAA, the *3Taps* court inadvertently increased potential for many of us to be categorized as hackers.

A. Policy-based Access Controls

¶12 The first element of subsection 1030(a)(2)(C), unauthorized access, is not clearly defined in the statute.[29] Instead, the CFAA focuses on authorization in terms of permission and the scope thereof. Not surprisingly, some courts tend to interpret this element in terms of trespass doctrines.[30] Accordingly, courts often apply policy-based safeguards, such as terms of use, computer use restrictions, and even an employee's general duty of loyalty, to trigger liability under the CFAA.[31] Notably, the Seventh, Fifth, and Eleventh Circuits still follow this broad interpretation.[32] The policy-based safeguard approach completely ignores the technology underlying access controls and instead shifts the focus onto the mental states of the plaintiff and defendant. In *Brekka*, the Ninth Circuit held that an employee's violation of her duty of loyalty is not an appropriate proxy for triggering liability under the CFAA.[33] The court reasoned that this interpretation inappropriately ignored the nature of access.[34] Because the CFAA is all about computer hacking, the nature of the access, particularly in the context of modern technology, should play an important role. How better to define the scope of authorization than to explore the context in which the data was protected? Otherwise, the court fails to completely measure the scope of authorization, and risks unjustifiably expanding the scope of the CFAA.

---

http://portal.uspto.gov/pair/PublicPair.

[28] *See* United States v. Nosal, 676 F.3d 854, 856 (9th Cir. 2012).

[29] Subsection 1030(e)(6) defines what it means to "exceed access," but nowhere in the statute is the term "access" defined.

[30] eBay, Inc. v. Bidder's Edge, Inc., 100 F. Supp. 2d 1058, 1062 (N.D. Cal. 2000).

[31] Int'l Airport Ctrs., LLC v. Citrin, 440 F.3d 418, 420 (7th Cir. 2006);

[32] *See* United States v. John, 597 F.3d 263, 271-73 (5th Cir. 2010); United States v. Rodriguez, 628 F.3d 1258 (11th Cir. 2010); *Citrin*, 440 F.3d 418 (7th Cir. 2006).

[33] LVRC HOLDINGS LCC v. Brekka, 581 F.3d 1127, 1135 (9th Cir. 2009).

[34] A violator's state of mind is already well accounted for in the CFAA.

B. Narrowing the Scope under Nosal

¶13 In *Nosal*, the Ninth Circuit further narrowed the scope of unauthorized access to better incorporate evaluating how the data was protected by holding that unauthorized access (a) applies only to the circumvention of technological safeguards and (b) does not apply in the context of use controls.[35] Here, "use controls" means the exercise of control over use of the data after such data has already been accessed and collected, which is chiefly handled by trade secret and misappropriation law doctrines.[36] The Court reasoned that narrowing application of the CFAA to circumvention of technological barriers better captures the Statute's anti-hacking intent and cited its decision in *Brekka* for support.[37] The Fourth Circuit adopted a similar stance in *WEC Carolina Energy Solutions LLC v. Miller*, which mirrors the court's reasoning in *Nosal*.[38] Moreover, the court used the online dating scene to illustrate the untenable impact of importing use controls for CFAA liability: many people would be in violation of the CFAA with outdated online profiles that do not adhere to the most recent terms of use policies on the sites.[39] Nevertheless, and perhaps more impactful than the relief provided to online singles everywhere, the court held that the CFAA was not intended to apply as a misappropriation tool.[40] Notably, the court failed to provide any test or guidelines for when a technological safeguard or access control would meet the standard under *Nosal*.

¶14 Taking advantage of what the court in *Nosal* left unclear, the defendant in *3Taps* used a proxy server in circumvention of an internet protocol address block (IP block) initiated on otherwise publicly-accessible data.[41] The *3Taps* court construed the *Nosal* decision to apply to *any* technological safeguard or access controls, regardless of whether such were mere tokens of protection.[42] Grouping mere token safeguards with functional safeguards exposes a "hack" in the framework of the statute and the holding in *Nosal*. *Nosal* stands for the proposition that use controls are beyond the scope of the CFAA.[43] Further, in order to be defined as an access control it follows that the control would need to provide some access-restricting function.[44] Otherwise, a danger arises that a plaintiff need only initiate patchwork, token access controls to disguise a use control on public data.[45]

¶15 In support of this potential danger, relevant authority reveals that the lower the sophistication of the safeguard involved, the higher the tendency there is for using the CFAA as a misappropriation tool.[46] For example, in holding that the defendant's violations

---

[35] United States v. Nosal, 676 F.3d 854, 863 (9th Cir. 2012).

[36] *See Id.* at 863; Restatement (Third) of Unfair Competition § 40, cmt. a (2009).

[37] *Id.* at 865.

[38] Andrew F. Popper, *More than the Sum of All Parts: Taking on IP and IT Theft Through a Global Partnership*, 12 Nw. J. Tech. & Intell. Prop. 253 (2014), *available at* http://scholarlycommons.law.northwestern.edu/njtip/vol12/iss4/1 (citing *WEC Carolina Energy Solutions LLC v. Miller*, 687 F.3d 199, 206-07 (4th Cir. 2012).

[39] *Nosal*, 676 F.3d at 865.

[40] *Id.* at 863.

[41] *See* Craigslist, Inc. v. 3Taps Inc., 942 F. Supp. 2d 962 (N.D. Cal. 2013).

[42] *Id.*

[43] *Nosal*, 676 F.3d at 863-64.

[44] Cvent, Inc. v. Eventbrite, Inc., 739 F. Supp. 2d 927, 933 (E.D. Va. 2010).

[45] *See Nosal*, 676 F.3d at 857.

[46] *Id.* at 862 ("We remain unpersuaded by the decisions of our sister circuits that interpret the CFAA broadly to cover violations of corporate computer use restrictions or violations of a duty of loyalty . . . . These courts looked only at the culpable behavior of the defendants before them, and failed to consider the

of terms of use cannot be defined as hacking under the CFAA, the court in *Cvent* reasoned that the plaintiff's data was "not protected in any meaningful fashion by its Terms of Use *or otherwise*."[47] In other words, and similar to the rationale behind the doctrine of trade secret law, a plaintiff cannot cry foul over the use of data that it fails to effectively protect.[48]

¶16 Nevertheless, the *3Taps* court took a page from the Seventh Circuit's reasoning in *Citrin* and found solace in the fact that the defendant had clear notice that its access was unwanted (a policy-based safeguard approach).[49] Similar to the homeowner who uses an ADT sticker to safeguard herself from home invasion, the court shifted its focus from functional web security to notice; a shift that does not square with the Statute's access-centered, anti-hacking intent.[50] Notably, had the defendant decided to read Craigslist advertisements in Spanish, it would have been engaged in hacking under this logic, because translation websites often use a proxy server.[51]

IV. PUBLIC V. PRIVATE DATA UNDER THE CFAA

¶17 Tied closely to the definition of access controls is the nature of the data behind them. For instance, should data behind a mere token access control such as entering in your age or state of residency really be considered private or protected? Section IV analyzes the precedential deviation made the *3Taps* Court by applying the CFAA to publicly-accessible data, and argues that the emphasis on meaningful and technological safeguards set out in *Nosal* and *Cvent* imply that only private and protected data was intended to be covered by the CFAA. This section also examines the case law distinguishing private and public data in the context of the CFAA, and suggest a third, intermediary category of "protected data."

A. An Unjustified Leap

¶18 Perhaps the most critical step in the *3Taps* court's reasoning was "[a]ssuming that the CFAA encompasses information generally available to the public."[52] Taken by itself, this is a logical step necessary to the conclusion drawn by the court. By accepting that an IP block was a covered access control under the CFAA, the court had little option but to open the CFAA to public data.[53] To be sure, while the access control and public data questions were bifurcated in *3Taps*, these issues remain inextricably linked, particularly in the context of the Internet. Data on the internet behind token access controls might as well be publicly-available – just ask Sony Pictures.[54] The bifurcation of these questions and the assumption that the CFAA applies to public data has widespread import. For example, the

---

effect on millions of ordinary citizens caused by the statute's unitary definition of 'exceeds authorized access.'").

[47] *Cvent*, 739 F. Supp. 2d at 933 (emphasis added).

[48] Restatement (Third) of Unfair Competition § 39, cmt. a (2009).

[49] Craigslist, Inc. v. 3Taps Inc., 942 F. Supp. 2d 962 (N.D. Cal. 2013); *see also* Int'l Airport Ctrs., LLC v. Citrin, 440 F.3d 418 (7th Cir. 2006).

[50] *See Nosal*, 676 F.3d at 857.

[51] The use of Google Translate software masks the originating IP address similar to using a proxy server.

[52] *3Taps*, 942 F. Supp. 2d at 969.

[53] *Id.* at 983, n.8.

[54] Mike Masnick, *Sony Goes One Ridiculous Step Further: Threatens To Sue Twitter Over Leaked Email Screenshots* (Dec. 23, 2014), https://www.techdirt.com/articles/20141222/16125129508/sony-goes-one-ridiculous-step-further-threatens-to-sue-twitter-over-leaked-email-screenshots.shtml.

CFAA does not define the meaning of "information."[55] Congressional history indicates that the CFAA was designed to protect the privacy of data.[56] In amending the CFAA to include subsection 1030(a)(2)(C), drafters stated that "[t]he proposed subsection 1030(a)(2)(C) is intended to protect against the interstate or foreign theft of information by computer."[57] Moreover drafters stated the entire "premise of this subsection is privacy protection" and that unauthorized access under this subsection "includes mere observation of the data."[58] These three elements - theft, privacy, and mere observation – cut against the assumption that subsection 1030(a)(2)(C) applies to public data. To be sure, how can one steal public information?

B. An Emphasis on Private Data

¶19 The dichotomy between the Act's treatment of public and private data becomes even more pronounced in judicial application. In Facebook, Inc. v. Power Ventures, Inc the defendant Power Ventures developed scraping software whereby users input their Facebook login credentials to "invite their friends" to join the Power Ventures' service offering.[59] The software accessed the user's Facebook account and email contact list, thereafter sending invite emails to the private contact list without recipient permission.[60] In siding with Facebook's position, the court pointed to the password-protected information, individual to each user, and the technique of user impersonation as a disingenuous exploitation of network trust.[61]

¶20 Importantly, scraping of private data can generally be distinguished from scraping of public data in that the former involves circumvention of individualized login or some other type of obfuscation applied against all users.[62] For instance, in *EF Cultural Travel v. Explorica, Inc.*, the First Circuit held that the use of a web scraper in combination with information not readily accessible to users amounts to unauthorized access under the CFAA.[63] Here, the defendant used the services of a current employee to translate tour codes for use in the web scraping software.[64] The software would then access the website and submit tour codes and harvest related pricing data.[65] In isolation, the tour codes were meaningless. However, when combined with the knowledge of how EF Cultural Travel's website used the tour codes in connection with querying a backend database for current pricing, use of the tour codes became more exploitative in nature.[66]

---

[55] 18 U.S.C. § 1030(a)(2)(C).

[56] S. Rep. No. 104-347, at 3 (1996) ("[This] Leahy-Kyl-Grassley amendment to the National Information Infrastructure (NII) Protection Act, S. 982, would strengthen the Computer Fraud and Abuse Act, 18 U.S.C. 1030 by closing gaps in the law to protect better the confidentiality, integrity, and security of computer data and networks.").

[57] *Id.* at 6.

[58] *Id.*

[59] Facebook, Inc. v. Power Ventures, Inc. (*Facebook II*), 844 F. Supp. 2d 1025, 1028 (N.D. Cal. 2012).

[60] *Id.*

[61] *See id.* at 1028.

[62] *See* EF Cultural Travel BV v. Explorica, Inc., 274 F.3d 577, 579 (1st Cir. 2001).

[63] *Id.*

[64] *Id.*

[65] *Id.*

[66] *Id.* at 582-83

¶21 Notably, the *3Taps* court did not cite any precedent that applies the CFAA to public data.[67] Indeed, there are very few judicial applications to public data in the context of the CFAA. In *Register.com v. Verio*, the defendant developed a scraping software to harvest publicly-accessible WHOIS domain registrant email addresses.[68] Verio emailed registrants for the purposes of offering a competitive website development service.[69] Register was an ICANN registrar and operated the public database pursuant to an agreement with ICANN and its own terms of use.[70] The court focused on Verio's violation of Register's terms of use in the context of contract law, and declined to weigh in on the applicability of CFAA.[71] However, the dissenting opinion makes clear that establishing damages in this context is improbable.[72] Moreover, terms of use that purport to restrict the use of public data, but which do not require a private, individualized step to limit access are unenforceable in many jurisdictions and raise concerns equivalent to those encountered in *Feist* regarding protection of factual information.[73] For instance, in *Ticketmaster Corp. v. Tickets.com, Inc.*, a California federal district court examined the potential misappropriation claims of public data in the context of preemption under the Copyright Act.[74] Although not applying the CFAA, the court emphasized the danger for protecting factual, public data under para-copyright claims in circumvention of the Copyright Act and the longstanding principles in *Feist*.[75]

C. Protected Data

¶22 If anything, the dearth of CFAA application to protect publicly-accessible data indicates its inappropriateness under an anti-hacking statute. Notably, Senator Leahy recently proposed an amendment to limit the information protected under the CFAA to seven categories of protected information, such as passwords or personally-identifiable information.[76] By limiting applicability to protected data, Senator Leahy's bill shifts the focus back onto the access controls used to safeguard the data. More importantly, this third category of "protected data" not only synthesizes the current dichotomy between public and private data, but ensures CFAA coverage of public data which is destroyed or defaced through intrusion to a backend, protected system.

---

[67] *3Taps*, 942 F. Supp. 2d at 970 (citing Weingand v. Harland Financial Solutions, No. C-11-3109, 2012 WL 2327660 (N.D. Cal. June 19, 2012) and *Facebook II*).

[68] Register.com, Inc. v. Verio, Inc., 356 F.3d 393, 395-96 (2d Cir. 2004).

[69] *Id.* at 395.

[70] *Id.*

[71] *Id.* at 401-02.

[72] *Id.* at 440 (Parker, J., dissenting) ("To maintain a cause of action under the CFAA against Verio, Register.com must demonstrate the Verio violated the CFAA in a manner that has caused Register.com damages or losses of at least $5,000. There is nothing in the record to suggest that this has occurred.").

[73] *See* Ticketmaster Corp. v. Tickets.com, Inc., 2000 U.S. Dist. LEXIS 4553 (C.D. Cal. Mar. 27, 2000), *aff'd* 248 F.3d 1173 (9th Cir. 2001); Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991).

[74] *Id.*

[75] *Feist*, 499 U.S. at 363.

[76] S. 1897, 113th Cong. §§ 104, 107 (2014) ("Personal Data Privacy and Security Act of 2014").

V. LOSS DECLARATIONS: THREE IMPORTANT FACTORS

¶23 The second vulnerability of the CFAA as highlighted in the *3Taps* decision concerns loss declarations. This section examines the evolution of the loss requirement under the CFAA and how cursory review of loss declarations has swallowed the meaning of the once important statutory requirement. Now, loss declarations serve as more of a procedural "check box" where plaintiffs need only retain a forensic IT consultant to tell them what they already know.[77]

¶24 In order to maintain a civil action against a defendant under the CFAA, a plaintiff must demonstrate a loss to one or more persons during any one-year period aggregating at least $5,000 in value.[78] A "loss" is defined as any reasonable cost to any victim, including the cost of responding to an offense, conducting a damage assessment, and restoring the data, program, system, or information to its condition prior to the offense, and any revenue lost, cost incurred, or other consequential damages incurred because of interruption of service.[79] Prior to this definition being added to the statute in the USA PATRIOT Act of 2001, courts had difficulty analyzing claims of loss as made by business entities.[80] Eventually, the Ninth Circuit solved the problem by defining "loss" in terms of system-related harm, such as service interruption or requiring users to change passwords.[81] For example, courts have held that the mere copying of proprietary data is insufficient to establish a loss under the CFAA.[82] Instead, a plaintiff must show that there was an impairment to its computer system or data as a result of the intrusion.[83] Moreover, a plaintiff must connect the loss to the actual intrusion.[84] Notably, this language tracks closely with most cyber liability insurance policy language.[85]

A. Three Categories of Losses

¶25 Congress also intended that the loss be reasonably or necessarily incurred by the plaintiff.[86] Jury instructions often mirror this requirement.[87] In the context of web scraping and before *3Taps*, there were three types of losses recognized by relevant case law.

---

[77] *See, e.g.*, Multiven, Inc. v. Cisco Sys., Inc., 725 F. Supp. 2d 887, 895 (2010) ("Costs associated with investigating intrusions into a computer network and taking subsequent remedial measures are losses within the meaning of the statute . . . ."); Kimberlite Corp. v. John Does 1–20, No. C08-2147, 2008 WL 2264485 (N.D. Cal. June 2, 2008).

[78] 18 U.S.C. § 1030(c)(4)(A)(i)(I).

[79] *Id.* at 1030(e)(11).

[80] *See* Pub. L. No. 107-56, § 814(d)(5) ("USA Patriot Act of 2001").

[81] United States v. Middleton, 231 F.3d 1207, 1213 (9th Cir. 2000).

[82] *See* Andritz, Inc. v. S. Maint. Contractor, 626 F. Supp. 2d 1264, 1266-67 (M.D. Ga. 2009).

[83] *Id.*

[84] *See* Shurgard Storage Ctrs., Inc. v. Safeguard Self Storage, Inc., 119 F. Supp. 2d 1121, 1126 (W.D. Wash. 2000).

[85] *See* Insuring Innovation. CyberFirst® Coverage for Technology Companies, The Travelers Indemnity Company, p. 10 (2012), *available at* https://www.travelers.com/business-insurance/specialized-industries/technology/docs/CyberFirst-Suite-locked.pdf.

[86] S. Rep. No. 104-357, pt. IV(1)(E).

[87] *Middleton*, 231 F.3d at 1213 ("In determining the amount of losses, you may consider what measures were reasonably necessary to restore the data, program, system, or information that you find was damaged or what measures were reasonably necessary to resecure the data, program, system, or information from further damage.").

¶26 The first is with respect to spamming and looming violations of applicable laws, regulations, or orders. For example, the court in *Facebook II* focused on the potential violations of the CAN-SPAM Act to substantiate the plaintiff's argument that it had properly established the loss requirement under the subsection 1030(e)(11).[88] At issue in *Facebook II* was the defendant's software that requested user credentials, and then impersonated a user and emailed the user's Facebook contact lists with unauthorized communications.[89] In this regard, the looming threat of a CAN-SPAM Act (or its international anti-spam counterpart) lawsuit justifies Facebook incurring expenses in connection with a forensic IT investigation. Moreover, it is likely a covered first-party loss under most incarnations of a cyber-liability policy.[90] In contrast to a loss incurred solely to comply with the statutory loss requirement, Facebook's loss declaration more squarely fits within the purpose of the CFAA as an anti-hacking statute.

¶27 The second type of loss applicable to a web-scraping context is where a password portal has been circumvented and private data leaked. In *Successfactor*, the court agreed with the plaintiff's argument that when a password protected environment is bypassed, certain costly security measures need to be undertaken to ensure network and application integrity.[91] Moreover, in *Vanderhye v. iParadigms*, the Fourth Circuit held that investigation into the possibility of a technical glitch in the system qualified as a loss under the CFAA, because such glitch impacted the way in which the plaintiff secured private data.[92] Significantly, each of these cases involved a password-protected portal.[93]

¶28 The third type of recognized loss in the context of web scraping demonstrates that losses related to web access must be reasonably incurred, but specifically in the context of trespass to chattels. In the famous case *Bidder's Edge*, the plaintiff claimed harm to its system based on the potential for unchecked aggregation of access and irreparable harm caused thereby.[94] In overruling this decision, the court in *Intel v. Hamidi* held that the potential for harm is not sufficient proof of harm.[95] The Fifth Circuit followed suit, stating that web scraping of a publicly-available website, without more, is insufficient to fulfill the harm requirement (in the context of trespass to chattels).[96]

B. A New Category

¶29 The *3Taps* Court made no effort to categorize the loss incurred by Craigslist within relevant case law.[97] Craigslist's loss declaration cites an internal investigation into the harm caused by the defendant's intrusion, well in excess of the $5,000 threshold.[98] The court's

---

[88] *See, e.g.*, *Facebook II*, 844 F. Supp. 2d at 1028; Am. Online, Inc. v. LCGM, Inc., 46 F. Supp. 2d 444 (E.D. Va. 1998).

[89] *Facebook II*, 844 F. Supp. 2d at 1028.

[90] *See* Rawlings, Philip. "Cyber Risk: Insuring the Digital Age." Queen Mary School of Law Legal Studies Research Paper 189, p. 21 (2015).

[91] SuccessFactors, Inc. v. Softscape Inc., 544 F. Supp. 2d 975 (N.D. Cal. 2008) (internal citations omitted).

[92] AV ex rel. Vanderhye v. iParadigms, LLC, 562 F.3d 630, 645 (4th Cir. 2009).

[93] *Id.* at 634; *SuccessFactors*, 544 F.Supp. 2d at 975.

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

[95] Intel Corp. v. Hamidi, 71 P.3d 296 (Cal. 2003)

[96] White Buffalo Ventures, LLC v. Univ. of Tex., 420 F.3d 366 (5th Cir. 2005).

[97] *See generally 3Taps*, 942 F. Supp. 2d at 968.

[98] First Amended Complaint at 35, Craigslist v. 3Taps INC., 942 F. Supp. 2d 962 (N.D. Cal. 2013) (No.

primary interest centered on whether defendant had access the information without authorization, without more than a cursory glance at the nature of the loss involved.[99] Notably, California Penal Code section 502 does not require a similar showing of loss.[100] Regardless, the loss requirement under the CFAA is anything but a meaningless requirement.

¶30 What is interesting about the *3Taps* decision is that the court dealt with the CFAA application to public data with nothing more than literal application of the statute; whereas, the decisions discussed previously devote considerably more attention to the nature of the data and the loss incurred in maintaining data security.[101] By applying the statute literally, courts streamline application of the CFAA, which may have benefits in circumstances where the loss is intangible or reputational in nature. To be sure, conducting a forensic investigation to confirm that no other data was leaked has value in and of itself, even if the investigation merely confirms what the plaintiff already knew. Nevertheless, it bears mentioning that outside of *3Taps*, the case history on the harm caused by a proxy server is nonexistent. A proxy server does not change the nature of access (or scope thereof), but instead changes the identity of the person accessing (similar to lying about your date of birth to circumvent a self-attested age restriction), which is a critical distinction in the context of public data.[102] In this regard, the circumvention of access controls that by default deny all traffic versus those that by default admit all traffic can be distinguished to better analyze a loss declaration. Indeed, releasing data to the public has certain inherent risks, risks that are not affected by the identity of the person accessing. That Craigslist needed to investigate the use of a proxy server seems disingenuous at best.[103] At worst, such investigations represent the perfunctory circumvention of a meaningful statutory requirement. Notably, Craigslist did not require passwords or the CAPTCHA tool used by many U.S. government websites to prevent robotic access to public data.[104]

¶31 A cursory application of the loss requirement leads to the potential for punishing innocuous behavior. For instance, had the Defendant translated a Craigslist advertisement in Spanish, it would have caused the same harm to Craigslist computers. To avoid this result, courts can add meaning to the loss requirement by: (1) analyzing the nature of the data and its value to the plaintiff; (2) distinguishing between access controls which by default deny all traffic and those that by default admit all traffic; and (3) analyzing the reasonableness of incurring expenses under the circumstances. By adding meaning to the loss requirement, courts not only avoid punishing innocuous activity, but incent proper web security.

CV 12-03816 CRB).

[99] *See generally 3Taps*, 942 F. Supp. 2d at 968 (The court states, "[t]he parties agree that the requirements of both statutes are functionally identical." This is incorrect. *See* Cal. Penal § 502, *infra* note 100).

[100] Cal. Penal Code § 502(e)(1) (Any damage or loss is recognized under the Statute).

[101] *Cf. Facebook II*, 844 F. Supp. 2d at 1039.

[102] *See* Cui, Lawrence, *Method and Apparatus for Proxy Server Cookies*, US Patent No. 6,910,180 (2000) ("Not only does the invention solve the problems of browsers that cannot handle cookies, the invention also protects the privacy of surfers by hiding their identities.").

[103] First Amended Complaint at 35, Craigslist v. 3Taps INC., 942 F. Supp. 2d 962 (N.D. Cal. 2013) (No. CV 12-03816 CRB).

[104] *See* United States Patent and Trademark Office, Public PAIR, http://portal.uspto.gov/pair/PublicPair (last visited Sept. 8, 2015).

VI. CHALLENGES AND PROPOSAL

***

¶32 When does working smarter cross the line into cheating? Answering that question largely involves a normative analysis, which can only be done in context. For example, we can all agree that using your mobile device to win at trivia night is likely cheating. On the other hand, the outcome seems different when connecting to Wi-Fi in court to win your trial.

¶33 The purpose of this article was to identify two vulnerabilities in the CFAA as it applies in the modern web-browsing context, namely as against web scraping. More broadly, however, this article isolates a problem inherent in applying an anti-hacking statute in a vacuum and in a purely literal sense. For example, access controls that poorly control access are viewed in the same light as those that do so effectively.[105]

¶34 It bears mentioning that the arguments advanced in this article are not intended to require a plaintiff under the CFAA to have *effective* safeguards or access controls. Indeed, the effectiveness of web security is in the eye of the hacker, and not all hackers are created equal. Nor does the author intend that the CFAA should not be applied to public data which are defaced, destroyed, or damaged by hacking attacks; rather, the CFAA should apply to any data behind reasonable access controls. By focusing on the reasonableness of the access controls, the CFAA is narrowed to protect certain data types, which prevents the type plaintiff abuse identified in this article.

¶35 Recent legislative proposals aim to solve this problem. Aaron's Law proposes to codify much of the *Nosal* decision and represents a step in the right direction, but also contains a flaw common in the many amendments to the CFAA.[106] Applying a hacking statute to cover specific scenarios is necessarily backward-looking in the face of dynamic technology. In other words, past trends do not indicate future results. Indeed, this flaw is evident in the *3Taps* decision.[107]

¶36 Instead, Senator Leahy's proposal to identify seven categories of protected information represents a forward-looking solution to the vulnerabilities in the CFAA.[108] By defining the CFAA in terms of the data it seeks to protect, not only does this proposal avoid the CFAA being used as a misappropriation tool, but it prevents the CFAA from being applied to punish innocuous activity, which in some cases represents advancement of the sciences and useful arts.

¶37 In the interim, this article proposes that courts first limit application of the CFAA to "protected data." "Protected data" in this context means data behind access controls that are reasonable under the circumstances and in accordance with good web security practices. By combining the data type analysis (public v. private) with the access control analysis, this proposal better synthesizes relevant case law as well as statutory intent. Bifurcation of this analysis tends to confuse the issue and opens the analysis to include mere token access controls, as demonstrated in the *3Taps* decision.[109]

---

[105] *See 3Taps*, 942 F. Supp. 2d at 983, n.6.
[106] H.R. 2454, 113th Cong. (2013) ("Aaron's Law Act of 2013").
[107] *See 3Taps*, 942 F. Supp. 2d at 983, n.6.
[108] S. 1897, 113th Cong. §§ 104, 107 (2014) ("Personal Data Privacy and Security Act of 2014").
[109] *See 3Taps*, 942 F. Supp. 2d at 983, n.8.

¶38 In addition to limiting applicability to protected data, this article proposes that loss declarations should be reviewed using the follow three-factor test: (1) nature of the data and value to plaintiff; (2) type of access control and default scope of authorization; and (3) reasonableness of losses incurred. By evaluating under these three factors, courts effectively add meaning back to an important statutory element and place themselves back in the equation to better curtail plaintiff abuse.

¶39 Common in the web software industry is the use of application programming interfaces or APIs to connect to distinct software platforms.[110] Best industry practices dictate that APIs should be developed in a dynamic and evolvable fashion to prevent simple changes from devastating the overall purpose of the software.[111] In this regard, the CFAA acts as an API to the normative definition of computer hacking, and can be redesigned to be dynamic and evolvable simply by incorporating industry context.

[110] Dig, Danny, and Ralph Johnson. "The Role of Refactorings in API Evolution." Software Maintenance, 2005. ICSM'05. Proceedings of the 21st IEEE International Conference on. IEEE, 2005.
[111] *Id.*

Northwestern Journal of Technology and Intellectual Property