COMMENTS

Van Buren v. United States: The Supreme Court's Ruling on the Fate of Web Scraping—"Access" to Discovery or Detention?

Madison Addicks*

I. INTRODUCTION

On June 3, 2021, the Supreme Court of the United States affirmed a police officer did not violate the Computer Fraud and Abuse Act (CFAA) when he accessed the department's database, which he was authorized to do, but for an improper purpose.[1] This landmark decision overruled the Eleventh Circuit's holding that the officer violated the CFAA by accessing the database for an "inappropriate reason."[2]

---

* © 2022 Madison Addicks, Managing Editor, Volume 24, *Tulane Journal of Technology and Intellectual Property*, J.D. Candidate 2022, Tulane University Law School; B.S. Marketing, B.A. Telecommunication and Film, The University of Alabama, 2018. The author wishes to thank her parents, Jeffery and Sharon Addicks, and sisters, Jessica and Allison Addicks, for their guidance and encouragement. The author would also like to thank her fellow *Tulane Journal of Technology and Intellectual Property* members for their help in preparation of this Comment for publication.

1. Van Buren v. United States, 141 S. Ct. 1648, 1662 (2021).
2. United States v. Van Buren, 940 F.3d 1192, 1208 (11th Cir. 2019).

The Supreme Court granted certiorari to answer a divided question among circuit courts: whether one who accesses a computer system and discovers information contrary to its intended use is a criminal.[3] The Supreme Court heard oral argument on November 30, 2020, providing hints and speculation as to why they took this case.[4] The Court's ruling not only defined the language of the statute, but addressed numerous policy concerns, ultimately invalidating a limitation on how individuals use their computers and phones for everyday use.[5]

To fully understand the impact of this decision, this Comment begins by outlining the background and legislative intent behind the Computer Fraud and Abuse Act in Part II. Part III addresses the circuit split among courts to interpret the "exceeds authorized access" prong of Section 1030(a)(2), contrasting the broad and narrow readings used by courts. Part IV discusses the impacts arising from the decision, specifically on the practice of web scraping. This Comment concludes with industry-specific implications from the *Van Buren* decision and the dilemma to balance First Amendment interests with privacy rights.

II. BACKGROUND

What started as the Comprehensive Crime Control Act of 1984 evolved into one of the most far-reaching crime statutes to date.[6] Congress added Section 1030 to the Act in 1986 to protect computer systems from those with unauthorized access.[7] Today, the statute is referred to as the Computer Fraud and Abuse Act.[8] The legislation serves as a barricade for computer hacking, or at least that was certainly Congress's intention.[9] Some assert that its enactment was the result of lawmakers viewing the 1983 film *War Games*, featuring the realistic depiction of a tech-whiz kid breaking into a United States defense

3. Kevin M. Cloutier & David M. Poell, *U.S. Supreme Court Case Preview—Van Buren v. United States: Does Use of a Computer for an "Improper Purpose" Violate the Computer Fraud and Abuse Act?*, 10 NAT'L L. REV. 121 (Apr. 30, 2020), http://www.natlawreview.com/article/us-supreme-court-case-preview-van-buren-v-united-states-does-use-computer-improper.

4. Tr. of Oral Argument, Van Buren v. United States, 140 S. Ct. 2667 (2020) (No. 19-783).

5. Van Buren v. United States, ELEC. PRIVACY INFO. CTR., http://epic.org/amicus/cfaa/van-buren/ (last visited Mar. 17, 2021).

6. *CFAA Background*, NAT'L ASS'N CRIM. DEF. LAWYERS (Mar. 10, 2020), http://www.nacdl.org/Content/CFAABackground.

7. S. REP. NO. 9-432, at 9 (1986).

8. *See* 18 U.S.C. § 1030.

9. S. REP. NO. 9-432, at 9 (1986).

agency's computer system.[10] Whether the inspiration actually came from Hollywood writers or not, the Computer Fraud and Abuse Act was a necessary stride in an age of increasing computerization and technological challenges.[11]

Congress continued to broaden the scope of the Act by drafting amendments.[12] Notably, the 1994 amendment redefined the scope of liability from a criminal act to a civil cause of action.[13] While the CFAA originally only penalized illegal actors in a criminal capacity, the amendment magnified the language to include multiple claims of civil wrongdoings.[14] However, the most influential amendment was the addition of Title II of the Economic Espionage Act in 1996.[15] This expanded the language of Section 1030(a)(2) from solely prohibiting unauthorized access in financial matters to unauthorized access that obtains *any* "information from any protected computer."[16] The CFAA was no longer just a governmental privacy act, but instead, a comprehensive protector from all persons engaging in illegal cyber activity. Under the Computer Fraud and Abuse Act, one who "intentionally accesses a computer without authorization or exceeds authorized access" to obtain information is committing a crime.[17]

The terms "authorization" and "authorized" are not defined in the statute, originally leading courts to supplement their interpretation with a dictionary definition.[18] Generally, "without authorization" refers to computer hackers, while "exceeds authorized access" relates to individuals who *have* access to a computer system, but who use the accessible information in a different way than intended.[19] However, Judge Kozinski in *United States v. Nosal* proposed both terms could apply to hackers: "'[w]ithout authorization' would apply to *outside* hackers (individuals who have no authorized access to the computer at all) and 'exceeds authorized access' would apply to *inside* hackers

10. Greg Pollaro, *Disloyal Computer Use and the Computer Fraud and Abuse Act: Narrowing the Scope*, 9 DUKE L. & TECH. REV. 12, 13 (2010).

11. S. REP. NO. 9-432, at 9 (1986).

12. *See* 18 U.S.C. § 1030.

13. *See* Violent Crime Control and Law Enforcement Act of 1994, Pub. L. No. 103-322, 108 Stat. 1796 (1994).

14. *See* 18 U.S.C. § 1030(g).

15. *See* Economic Espionage Act of 1996, Pub. L. No. 104-294, 110 Stat. 3488 (1996).

16. *Id.*

17. 18 U.S.C. § 1030(a)(2).

18. *See generally id.*

19. United States v. Nosal, 676 F.3d 854, 858 (9th Cir. 2012) (emphasis added).

(individuals whose initial access to a computer is authorized but who access unauthorized information or files)."[20]

The following Part focuses primarily on the latter, although the difference between the two terms is "paper thin."[21] The Computer Fraud and Abuse Act explicitly defines the term "exceeds authorized access" as accessing "a computer with authorization and [using] such access to obtain or alter information in the computer that the accessor is not entitled so to obtain or alter[.]"[22] The constitutionally vague definition led to split interpretations across all circuits.

III. CIRCUIT COURTS SPLIT ON INTERPRETING "EXCEEDS AUTHORIZED ACCESS"

Circuit courts encountered much difficulty reaching a conclusion on what exactly it meant to "exceed authorized access." In fact, prior to the Supreme Court's intervention, there was a jurisdictional split on the boundaries of authorization and how far was too far. Circuit courts ruled essentially in two different ways, utilizing a broad or narrow interpretation.

A. The First, Fifth, Seventh, and Eleventh Circuit's Broad Interpretation

The United States Court of Appeals for the First, Fifth, and Eleventh Circuits broadly interpreted "exceeds authorized access," while the Seventh Circuit maintained the same viewpoint, but utilized an agency approach in reaching that conclusion. The broad interpretation ultimately criminalized individuals with access to a company system, but who used the access for a purpose contrary to their authorization.[23] The question answered by these circuit courts was "whether 'authorized access' or 'authorization' [could] encompass limits placed on *the use* of information obtained by permitted access to a computer system and data available on that system."[24] The pro-plaintiff interpretation adopted by these circuits answered this question in the affirmative.[25]

---

20. *Id.*
21. Int'l Airport Ctrs., L.L.C. v. Citrin, 440 F.3d 418, 420 (7th Cir. 2006).
22. 18 U.S.C. § 1030(e)(6).
23. *See generally* Samuel Kane, *Available, Granted, Revoked: A New Framework for Assessing Unauthorized Access Under the Computer Fraud and Abuse Act*, 87 U. CHI. L. REV. 1437, 1447 (2020).
24. United States v. John, 597 F.3d 263, 271 (5th Cir. 2010).
25. *See* EF Cultural Travel BV v. Explorica, Inc., 274 F.3d 577 (1st Cir. 2001); *see John*, 597 F.3d at 271; *see* United States v. Rodriguez, 678 F.3d 1258, 1263 (11th Cir. 2010).

The First Circuit in *EF Cultural Travel BV v. Explorica, Inc.* presented one of the earliest interpretations of the term in 2001.[26] The court held that a tour company exceeded authorization when it gave proprietary information from its prior employer, such as source codes, to a third-party to scrape the website for company data.[27] The purpose was to obtain the competitor's price structure in order to maintain the lowest costs in the industry.[28] The First Circuit reasoned that mining its previous employer's website "reeks of use—and indeed, abuse—of proprietary information that goes beyond any authorized use of [the competitor's] website."[29]

Similarly, the Fifth Circuit in *United States v. John* held that an individual "exceed[ed] authorized access" when "the user [knew] or reasonably should [have] know[n] that he or she [was] not authorized to access a computer and information obtainable from that access [was] in furtherance of or to perpetuate a crime."[30] However, the facts of the case did not meet this burden.[31] The defendant, an account manager at Citigroup, exceeded authorized access by giving customer account information to her half-brother for purposes of incurring fraudulent charges on customer accounts.[32] The court reasoned that although the defendant had access to this customer information, it was a breach of data beyond what the employee was confined to do with the accessible information, denying the defendant's argument of "any and all" access.[33]

Further, the Eleventh Circuit, in *United States v. Rodriguez*, drew a line in the sand, noting that as long as the use of a database to obtain personal information was in furtherance of the business, an employee was within the confines of their authorization.[34] However, in that case, an employee used his access as a Social Security Administration representative to obtain sensitive personal information on seventeen people.[35] The court reasoned, and the plaintiff conceded, that his access to uncover information on various individuals was not in furtherance of

26. *See* EF *Cultural Travel*, 274 F.3d at 577.
27. *Id.* at 579.
28. *Id.*
29. *Id.* at 583.
30. *John*, 597 F.3d at 271.
31. *See id.* at 272.
32. *Id.* at 269.
33. *Id.* at 272.
34. United States v. Rodriguez, 678 F.3d 1258, 1263 (11th Cir. 2010).
35. *Id.* at 1260.

his role as a teleservice representative, ultimately violating the Computer Fraud and Abuse Act.[36]

The Seventh Circuit, in *International Airport Centers, L.L.C. v. Citrin*, shared this broad-view of liability, even though the analysis was conducted through a common law agency lens.[37] There, an employee permanently deleted files on his company computer, ridding both company data and evidence of improper conduct before quitting.[38] The court reasoned this breached "the duty of loyalty" between the employer and employee.[39] Further, "failing to disclose adverse interests void[ed] the agency relationship."[40] The Seventh Circuit ultimately did not address its interpretation of "exceeds authorized access."[41] Doing so labeled the Seventh Circuit's analysis as the agency approach, although the circuit has been grouped under this broad-interpretation umbrella due to the ultimate conclusion it reached.[42]

B. The Second, Fourth, Sixth, and Ninth Circuit's Narrow Interpretation

The narrow view was most supported by the United States Court of Appeals for the Ninth Circuit, and followed closely by the Second, Fourth, and Sixth Circuits. The increasing popularity of restricting civil and criminal liability under the CFAA spiked in the last decade.[43] A narrow interpretation of the term "exceeds authorized access" was almost religiously supported by a policy argument of avoiding criminal repercussions on those who are not criminals.[44]

The Ninth Circuit, in a decision that separated it from its sister circuits, trail blazed the idea that access to a database, no matter what the access consisted of, did not constitute an offense.[45] In *United States v. Nosal*, an ex-employee convinced his former colleagues to download

---

36. *Id.* at 1263.
37. *See* Int'l Airport Ctrs., L.L.C. v. Citrin, 440 F.3d 418, 421 (7th Cir. 2006).
38. *Id.* at 419.
39. *Id.* at 421.
40. *Id.*
41. *See generally id.* at 420-21.
42. Justin Precht, *The Computer Fraud and Abuse Act or the Modern Criminal at Work: The Dangers of Facebook from Your Cubicle*, 82 U. CIN. L. REV. 359, 362 (2014).
43. *CFAA Background*, *supra* note 6.
44. *See generally* Peter A. Crusco, *'Van Buren v. United States': 'Unauthorized Access' in the Virtual World of Expanding Federal Criminal Liability*, N.Y. L.J. (Dec. 21, 2020, 12:45 PM), http://www.law.com/newyorklawjournal/2020/12/21/van-buren-v-united-states-unauthorized-access-in-the-virtual-world-of-expanding-federal-criminal-liability/.
45. *See* United States v. Nosal, 676 F.3d 854, 864 (9th Cir. 2012).

confidential client information from the company's database and send it to him so that he could start a competing business.[46] While the court reasoned that the government may proceed with other criminal counts of indictment, this was not a breach of the Computer Fraud and Abuse Act.[47] The Ninth Circuit reasoned that the CFAA prohibits unauthorized *access*, not unauthorized *use*.[48]

In 2016, the Ninth Circuit again considered the same facts, but this time, evaluating whether the ex-employee's access to the computer, after both he and his co-conspirators were terminated and revoked of their access, were "without authorization" when they continued to use the system.[49] In *United States v. Nosal* ("Nosal II"), when the employees' roles as insiders within the company changed to outsiders, their claim changed from "exceed[ed] authorized access" to "without authorization."[50] Thus, since the employees no longer had access to the system, specifically a password system only intended for company employees, they no longer had authorization to access the information.[51]

Shortly thereafter, the Fourth Circuit held a similar view in *WEC Carolina Energy Solutions LLC v. Miller*.[52] The court held that even though an employee and his assistant may have misappropriated information when they downloaded confidential documents and customer information from their computers prior to leaving the company, they did not exceed their authorized access and violate the CFAA.[53] The court reasoned that it did not want to hold employees liable when they violated a use policy out of bad faith, as this was not the statute's intention.[54] Two years later, the Fourth Circuit in *United States v. Steele* did in fact hold a defendant liable for exceeding company access, after the defendant continued to log on to his previous employer's server to gain government contract bids for nine months after leaving the company.[55] The court differentiated the fact pattern from *WEC Carolina*, however, by clarifying that the defendant no longer worked as a

46. *Id.* at 856.
47. *Id.* at 864.
48. *Id.*; *see also* LVRC Holdings LLC v. Brekka, 581 F.3d 1127 (9th Cir. 2009) (holding the interpretation of "access" does not mean misappropriation).
49. United States v. Nosal, 844 F.3d 1024, 1028-29 (9th Cir. 2016).
50. *Id.* at 1036.
51. *Id.* at 1039.
52. *See generally* WEC Carolina Energy Sols. LLC v. Miller, 687 F.3d 199 (4th Cir. 2012).
53. *Id.* at 207.
54. *Id.*
55. United States v. Steele, 595 F.App'x. 208, 209-210 (4th Cir. 2014).

company employee, eradicating his authorization altogether.[56] Therefore, the evidence confirmed he did not "exceed authorized access" because he no longer had it in the first place.[57] The Fourth Circuit still maintained its previous ruling in *WEC Carolina* and support of the narrow view.[58]

Further, the Second Circuit in *United States v. Valle* liberally construed "exceeds authorized access" by holding that an officer did not violate the CFAA when he conducted a search on a woman to obtain her home address and date of birth on a restricted database with no law enforcement purpose.[59] The court borrowed the Ninth Circuit's reasoning in *Nosal* to avoid "unintentionally turn[ing] citizens into criminals."[60]

The Sixth Circuit recently contributed to the conversation in its decision in *Royal Truck & Trailer Sales & Service, Inc. v. Kraft*.[61] Two company employees emailed themselves copies of customer and vendor information, shortly after resigning to work for a competitor.[62] They deleted any trace of the act, rendering the data unrecoverable.[63] The court held that even though this was against company policy, this information was accessible on their company-issued computers and cell phones, authorizing them to view the information.[64] The court supported its reasoning by arguing violations to corporate policies should not result in treating employees as violent criminals.[65]

C. The Supreme Court Weighs In

The Supreme Court defined Section 1030(a)(2) in a long-awaited decision on June 3, 2021.[66] The Court granted certiorari on April 20, 2020 "to resolve the split in authority regarding the scope of liability under the CFAA's 'exceeds authorized access' clause."[67] Ultimately, the Supreme Court reversed the Eleventh Circuit's holding that a police

56. *Id.* at 211.
57. *Id.*
58. *Id.*
59. United States v. Valle, 807 F.3d 508, 523 (2d Cir. 2015).
60. *Id.* at 528; *see also* United States v. Nosal, 676 F.3d 854, 863 (9th Cir. 2012).
61. *See generally* Royal Truck & Trailer Sales & Serv., Inc. v. Kraft, 974 F.3d 756 (6th Cir. 2020).
62. *Id.* at 758.
63. *Id.*
64. *Id.*
65. *Id.* at 762.
66. Van Buren v. United States, 141 S. Ct. 1648 (2021).
67. *See id.* at 1654; *see also* United States v. Van Buren, 940 F.3d 1192 (11th Cir. 2019), *cert. granted*; Van Buren v. United States, 140 S. Ct. 2667 (2020).

officer "violated the CFAA by accessing the law enforcement database for an 'inappropriate reason.'"[68]

The facts of the case are simple. Nathan Van Buren, a police sergeant with the Cumming, Georgia Police Department, searched for a woman's license plate on the police database to confirm whether or not she was an undercover cop.[69] This act was in exchange for $6,000 from a local man named Andrew Albo, who was known for paying minors and prostitutes to spend time with him, but ultimately accusing the women of stealing money after the night was over.[70] However, unbeknownst to Van Buren, Albo recorded their conversation and turned it into the Forsyth County Sheriff's Office, gaining attention from the FBI.[71] Evidence revealed that Van Buren was trained not to use the police database for an improper purpose, defined as "any personal use."[72] Therefore, Van Buren was aware "the search breached department policy."[73] Van Buren was ultimately charged with one count of honest-services wire fraud and one count of felony computer fraud.[74] This Section solely focuses on the computer fraud charge.

In its analysis, the Eleventh Circuit followed its earlier decision in *United States v. Rodriguez*.[75] Shadowing the defendant's argument in *Rodriguez*, "Van Buren allege[d] that he [was] innocent of computer fraud because he accessed only databases that he was authorized to use, even though he did so for an inappropriate reason."[76] While the court acknowledged other opinions from its sister circuits upholding a narrow reading of the statute, and presented the coinciding policy arguments, the Eleventh Circuit promulgated that Van Buren "identified no Supreme Court or en banc decision of this Circuit that abrogates *Rodriguez*[.]"[77] The court reasoned that misusing the database to run a tag search, in exchange for money, falls strictly within a violation of the CFAA.[78]

Almost immediately, the Supreme Court expressed discernment with both parties' interpretation of "exceeds authorized access," as

---

68. *Van Buren*, 141 S. Ct. at 1653-54.
69. *Van Buren*, 940 F.3d at 1198.
70. *Id.* at 1197-98.
71. *Id.*
72. *Van Buren*, 141 S. Ct. at 1653.
73. *Id.*
74. *Van Buren*, 940 F.3d at 1198.
75. *See id.* at 1207-08; *see also* United States v. Rodriguez, 678 F.3d 1258, 1263 (11th Cir. 2010).
76. *Van Buren*, 940 F.3d at 1208.
77. *Id.*
78. *See id.*

voiced in oral argument on November 30, 2020.[79] Justice Sotomayor labeled Section 1030(a)(2) as "a very broad statute and dangerously vague," while Justice Gorsuch opined that the broad interpretation could "perhaps mak[e] a federal criminal of us all."[80] Relying on the "parade of horribles," the Court altered the focus from an employment dispute, and extended it to the person "who lies about weight on a dating website" or a law student who uses Westlaw or Lexis for personal use.[81] The Court's reaction previewed their ultimate fear in "criminalizing widespread, innocuous online-behavior," as exhibited in the opinion.[82]

Justice Alito said it plainly, "I find this a very difficult case to decide based on the briefs that we've received[,]" pitting personal privacy against "criminaliz[ing] all sorts of activity . . . ."[83] Numerous organizations and individuals filed amicus briefs to share their concerns with the Court.[84] The brief for Americans of Prosperity Foundation asserted that if the Eleventh Circuit's decision were to stand, it "could extend to violations of the fine print in website terms of service, company computer-use policies, and other breaches of contract."[85] Further, the scope of liability would turn "millions of honest, hardworking Americans into federal criminals," ultimately giving independent organizations the ability to set the law.[86] Leading computer security researchers, presented by an amicus, shared their opposition to giving the owner of data such ability to determine what is and is not beneficial to the public by labeling it as "highly risky."[87] Illuminating the scope of the decision, the brief concluded that computer systems no longer just extend to desktops at the office, but "into our homes, vehicles, and even our bodies."[88]

However, proponents of the broad interpretation argued that the system owner's right to determine the scope of each user's access would

---

79. Tr. of Oral Argument, *supra* note 4, at 19.

80. *Id.* at 48, 54.

81. *Id.* at 16.

82. Camille Fischer & Andrew Crocker, *Victory! Ruling in hiQ v. Linkedin Protects Scraping of Public Data*, ELEC. FRONTIER FOUND. (Sept. 10, 2019), http://www.eff.org/deeplinks/2019/09/victory-ruling-hiq-v-linkedin-protects-scraping-public-data.

83. Tr. of Oral Argument, *supra* note 4, at 45-46.

84. *See generally* No. 19-783, http://www.supremecourt.gov/search.aspx?filename=/docket/docketfiles/html/public/19-783.html.

85. Br. for Americans for Prosperity Foundation as Amici Curiae Supporting Pet'r, at 3, Van Buren v. United States, 140 S. Ct. 2667 (2020).

86. *Id.* at 3.

87. Br. for Computer Security Researchers et al. as Amici Curiae Supporting Petitioner, at 5, Van Buren v. United States, 140 S. Ct. 2667 (2020).

88. *Id.*

maintain the system's accuracy, security, and above all, reputation.[89] The Electronic Privacy Information Center (EPIC) argued the need for the CFAA "to be an extra check against abuse by the people entrusted to access sensitive data and systems."[90] EPIC went on to list all the modern capabilities of government computer systems, including "limit[ing] an individual's freedom to travel, . . . impact[ing] their ability to seek employment or credit, . . . restrict[ing] their access to healthcare and other essential benefits, and . . . plac[ing] them under the microscope of a law enforcement inquiry."[91] Improper access to these systems, EPIC asserted, could lead to immeasurable abuse and damage.[92]

Both sides presented compelling arguments to contribute to this landmark case. However, in a 6-3 decision, the majority ultimately sided with Van Buren's reading of the statute in avoidance of severe policy implications.[93] The Court defined "exceeds authorized access" when one "accesses a computer with authorization but then obtains information located in particular areas of the computer—such as files, folders, or databases—that are off-limits to him."[94] The Supreme Court reasoned that Van Buren had authorization to use the system to retrieve license-plate information; therefore, he did not "excee[d] authorized access," as the CFAA defines the phrase.[95] While the dissent argued that Van Buren never had a "right" to obtain *this* specific license plate information, the majority reverted to the plain language of the statute.[96] The Court relied heavily on the word "so" in its analysis of Section 1030(e)(6).[97] The statutory phrase "not entitled *so* to obtain" plainly reads as information that one "is not allowed to obtain by using a computer that he is authorized to access."[98] Justice Barrett, who delivered the opinion, went

89. Br. for The Federal Law Enforcement Officers Association as Amici Curiae Supporting Resp't, at 1-3, Van Buren v. United States, 140 S. Ct. 2667 (2020) (arguing "[p]erhaps in no single other area would the administration of justice in this country be so corrupted than if federal law enforcement computer systems were to be rendered unavailable or unreliable").

90. Br. for Electronic Privacy Information Center et al. as Amici Curiae Supporting Resp't, at 5, Van Buren v. United States, 140 S. Ct. 2667 (2020).

91. *Id.* at 4.

92. *Id.* at 4-5.

93. *Van Buren*, 141 S. Ct. at 1662.

94. *Id.*

95. *Id.*

96. *See id.* at 1663 (Thomas, J., dissenting).

97. *Id.* at 1652; *see also* 18 U.S.C. § 1030(e)(6) ("the term 'exceeds authorized access' means to access a computer with authorization and to use such access to obtain or alter information in the computer that the accessor is not entitled so to obtain or alter").

98. *Van Buren*, 141 S. Ct. at 1655.

as far to say that without the word "so," the statute *could* impose multiple restrictions on the ability to gather information on a device.[99] Therefore, regardless of Van Buren's improper motive in obtaining the information, as long as he did not breach his scope of authorized access to conduct the search, the Supreme Court held Van Buren was within the confines of protection.[100]

The Supreme Court concluded by addressing the policy arguments introduced in the amicus briefs.[101] As numerous scholars indicated leading up to the decision, the Court had the opportunity to "act as a bandage to cover the ever-growing problem of privacy . . . ."[102] However, the Court ultimately avoided such burden to "attach criminal penalties to a breathtaking amount of commonplace computer activity."[103] For example, the government's reading of the CFAA would likely criminalize any violations of a website providers' Terms of Service.[104] Justice Barrett scrutinized the government's CFAA charging policy, which plainly states prosecution "*may not* be warranted—not that it would be prohibited" if such contract or Terms of Service breach occurred.[105] The Supreme Court's reproach of the government's intent to criminalize this behavior provides significant insight into the Court's reaction to the practice of web scraping.[106] Further analysis of the Terms of Service argument is discussed in Part IV.

IV. VAN BUREN'S INFLUENCE ON WEB SCRAPING

While the Computer Fraud and Abuse Act has been at the center of employment disputes, another prominent area affected by this statute is web scraping. Web scraping involves "extracting data from a website and copying it into a structured format, allowing for data manipulation or analysis."[107] However, while this practice has the potential to harm a

99. *Id.* at 1656 (noting "[t]he modifying phrase 'so to obtain' directs the reader to consider a specific limitation on the accessor's entitlement").

100. *See generally id.* at 1657.

101. *Id.* at 1661-62.

102. Nicole Sakin & Sarah Rippy, *How the Lack of a Federal Privacy Law is Resulting in a Problematic Application of the CFAA*, IAPP (Feb. 5, 2021), http://iapp.org/news/a/how-the-lack-of-a-federal-privacy-law-is-resulting-in-a-problematic-application-of-the-cffa/.

103. *Van Buren*, 141 S. Ct. at 1661.

104. *Id.* (noting the Government's interpretation of Section 1080(a)(2) would "criminalize everything from embellishing an online-dating profile to using a pseudonym on Facebook[,]").

105. *Id.*

106. *See generally id.*

107. hiQ Labs, Inc. v. LinkedIn Corp., 938 F.3d 985, 991 n.3 (9th Cir. 2019).

website, it is mostly condemned for its violation of many websites' Terms of Service.[108]

This Part focuses on the "without authorization" prong of Section 1030(a)(2), or "accessing a protected computer without permission."[109] The non-technical term, however, implicitly overlaps with definitional undertones of "exceeds authorized access."[110] In fact, one legal scholar believes "exceeds authorized access" uses "a concept, entitlement, that is simply a synonym for authorization."[111] The Supreme Court in *Van Buren* agreed.[112] Because the clauses are consistent, liability under both "stems from a gates-up-or-down inquiry—one either can or cannot access a computer system, and one either can or cannot access certain areas within the system."[113] Therefore, its mention is largely interchangeable.

While the *Van Buren* decision will ultimately impact the practice of web scraping, it is important to consider a distinguished case that introduced this topic. *hiQ Labs, Inc. v. LinkedIn Corp.* posed the substantive question of whether "without authorization" includes web scraping data from a public website.[114] The Ninth Circuit held "it is likely that when a computer network generally permits public access to its data, a user's accessing that publicly available data will not constitute access without authorization under the CFAA."[115]

In *hiQ Labs*, a data analytics company scraped job-related information off public user profiles on LinkedIn to create "people analytics," which was later sold to business clients.[116] In its analysis of the CFAA claim, the Ninth Circuit scrutinized the language of the statute and its legislative history.[117] First, the court reasoned that the definition of "authorization" plainly means to restrict only those with access, inferring "without authorization" to include those with free access.[118] Second,

---

108. Amber Zamora, *Making Room for Big Data: Web Scraping and an Affirmative Right to Access Publicly Available Information Online*, 12 PEPP. J. BUS. ENTREPRENEURSHIP & L. 203, 204 (2019).

109. *hiQ Labs*, 938 F.3d at 999.

110. *See generally id.* at 999-1000.

111. Br. of Orin S. Kerr as Amici Curiae Supporting Pet'r, at 6, Van Buren v. United States, 140 S. Ct. 2667 (2020).

112. Van Buren v. United States, 141 S. Ct. 1648, 1658-61 (2021).

113. *Id.* at 1658.

114. *See hiQ Labs*, 938 F.3d at 1001.

115. *Id.* at 1003.

116. *Id.* at 991.

117. *See id.* at 1000.

118. *Id.*

based on the intent of Congress, and the Ninth Circuit's previous ruling in *Nosal*, the court stated that "the CFAA is best understood as an anti-intrusion statute and not as a 'misappropriation statute,'" ultimately rejecting the contract-based, or broad interpretation, of its sister circuits.[119] The Ninth Circuit did not address any First Amendment arguments in its opinion. Using policy as its sword, the court adopted the narrow interpretation in order to avoid "turn[ing] a criminal hacking statute into a 'sweeping Internet-policing mandate.'"[120] The Ninth Circuit reasoned that because hiQ Labs did not need a username or password to gain access to the users' information, this was public access that did not breach the "without authorization" criteria under the CFAA.[121]

On June 14, 2021, the Supreme Court granted a petition for a writ of certiorari, ultimately vacating the judgment and remanding the case to the Ninth Circuit "for further consideration in light of *Van Buren v. United States*."[122] As previously discussed, the Supreme Court's repugnance with the government's inclusion of potentially enforcing criminality upon those who breach a website's Terms of Service provides significant insight into how the Ninth Circuit will rule.[123] The court will likely uphold its previous decision and presuppose an individual not guilty, even if they surpass a level of permitted access, so long as it is a public website.[124] However, the repercussions of a web scraping decision are specific to themselves, and the Ninth Circuit could potentially carve out an exception to the practice. One such exception could subject those "to criminal liability as long as a company lists the infraction in its [T]erms of [S]ervice."[125] As such, the next Sections focus specifically on what this means for the e-commerce and journalism industries.

A. The Debate on Web Scraping and E-commerce

Massive brands are fed up with start-ups and industry newbies entering the marketplace and scraping years of accumulated data from their websites. Big data, or "massive quantities of information produced by and about people, things, and their interactions[,]" is coveted to both large companies with loyal customers, as well as start-ups seeking new

119. *Id.*
120. *Id.* at 1003.
121. *Id.*
122. Linkedin Corp. v. hiQ Labs, Inc., No. 19-1116, 2021 WL 2405144, at *1 (2021).
123. *See generally* Van Buren v. United States, 141 S. Ct. 1648, 1661 (2021).
124. *See generally hiQ Labs*, 938 F.3d at 1003.
125. Jacquellena Carrero, *Access Granted: A First Amendment Theory of Reform of the CFAA Access Provision*, 120 COLUM. L. REV. 131, 149 (2020).

client information quickly.[126] There is no question web scraping can produce unparalleled business results, giving companies a competitive advantage by effectively monitoring competitors, leading generations, enhancing investment opportunities, and optimizing products.[127] It is up to companies whether or not they want to take the risk.

Southwest Airlines recently pushed back after an online travel site, Kiwi.com, "engaged in the unauthorized scraping of Southwest flight and pricing data and the selling of Southwest tickets (along with allegedly charging unauthorized service fees) . . . ."[128] After sending multiple cease-and-desist letters recognizing Kiwi's practice as a violation of Southwest's Terms of Service, and only being answered by a prospective business relationship, Southwest filed suit listing numerous claims of action.[129] While it is not clear how Kiwi ultimately responded, they likely relied on *hiQ Labs*' narrow holding, which would allow scraping of public data, or here, online airfare prices.[130] The parties, however, settled the matter outside of court.[131]

Similarly, Instacart filed suit against Cornershop, a grocery delivery start-up recently acquired by Uber, after it scraped over 2,000 images and product descriptions from Instacart's website to use for its own launch.[132] This practice was against Instacart's Terms of Service.[133] Cornershop agreed to stop web scraping after a preliminary injunction was filed.[134] However, to send a message, Instacart also brought an action against Uber several months later to uncover what they knew about Cornershop's fraudulent data scraping.[135] Instacart took matters into its

126. *Id.* at 136 (citing Danah Boyd & Kate Crawford, *Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon*, 15 INFO. COMM. & SOC'Y 662, 663 (2012)).

127. Ashley, *5 Reasons Why Web Scraping May Benefit Your Business*, OCTOPARSE (Jan. 20, 2021), http://www.octoparse.com/blog/why-web-scraping-may-benefit-your-business.

128. Jeffrey Neuburger, *Southwest Airlines Sues to Stop Web Scraping of Fare Information*, JD SUPRA (Jan. 22, 2021), http://www.jdsupra.com/legalnews/southwest-airlines-sues-to-stop-web-3302549/; *see also* Southwest Airlines Co. v. Kiwi.com, Inc., No. 21-00098 (N.D. Tex. filed Jan. 14, 2021).

129. *Id.*

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

131. Southwest Airlines Co. v. Kiwi.com, Inc., No. 21-00098 (N.D. Tex. Filed Jan. 14, 2021).

132. Alison Frankel, *Instacart Goes After Uber in Data-Scraping War with Cornershop*, REUTERS (Jan. 14, 2021, 3:25 PM), http://www.reuters.com/article/legal-us-otc-instacart/instacart-goes-after-uber-in-data-scraping-war-with-cornershop-idUSKBN29J2SY.

133. *Id.*

134. *Id.*

135. *Id.*

own hands by holding these companies accountable as it awaits a decision from the Ninth Circuit.

Clearview AI was also under fire for comparable practices.[136] The start-up app Clearview AI has become one of the largest tools for law enforcement agencies to uncover criminals and victims of crimes.[137] The database holds over three billion profiles, quickly becoming an industry leader.[138] But Clearview couldn't do it alone—the data was scraped from Google, YouTube, LinkedIn, Twitter, Venmo, and Facebook.[139] These companies, however, did not stand by to support its newcomer.[140] Twitter started the trend by sending Clearview a cease-and-desist letter, calling the company out for violating their Terms of Service, which bans using its data for facial recognition purposes.[141] The next month, each of the other companies followed suit.[142] Clearview AI hid behind a First Amendment defense, claiming "that their system [was] built 'to only take publicly available information.'"[143]

Businesses using web scraping methods to gain a competitive advantage are not only hoping, but relying, on the Ninth Circuit to rule narrowly in its new decision, following in the footsteps of *Van Buren*. Yet, industry leaders with decades of information packed into the seams of its digital footprint are counting on the Ninth Circuit to diverge from *Van Buren* precedent and carve out an exception to use the CFAA as an undetectable firewall against data scraping.[144] Although unlikely, this result could potentially turn capitalistic entrepreneurs and businesses into criminals after just a few clicks.[145]

B. Web Scraping's Impact on Journalism

Reporters and news outlets have been at the forefront of the web scraping debate, expressing their trepidations of a broad reading of the

---

136. Kaixin Fran, *Clearview AI Responds to Cease-and-Desist Letters by Claiming First Amendment Right to Publicly Available Data*, JOLT DIGEST (Feb. 25, 2020), http://jolt.law.harvard.edu/digest/clearview-ai-responds-to-cease-and-desist-letters-by-claiming-first-amendment-right-to-publicly-available-data.

137. *Id.*

138. *Id.*

139. *Id.*

140. *See generally id.*

141. *Id.*

142. *Id.*

143. *Id.*

144. Andrew Sellars, *Twenty Years of Web Scraping and the Computer Fraud and Abuse Act*, 24 B.U. J. SCI. & TECH. L. 372, 412 (2018).

145. *Id.* at 412, 414-15.

Computer Fraud and Abuse Act.[146] First Amendment interests serve as the crux of the journalism industry's argument, garnering the constitutional right of "democratic self-governance, autonomy, and truth."[147] The Reporters Committee for Freedom of the Press, presented by its amicus, argued "[t]he court of appeals' interpretation of Section 1030(a)(2) [in *United States v. Van Buren*] threatens to criminalize a wide range of ordinary journalistic activity without offering any means of guarding journalists' First Amendment freedoms—raising the prospect that the 'freedom of the press could be eviscerated.'"[148]

While the *Van Buren* decision will likely ease many concerns, it has become a daily practice for reporters to use web scraping methods to uncover how "collections of data influence [people's] lives."[149] There are two reasons for this. First, web scraping helps journalists collect data at scale, giving reporters the opportunity to obtain much more information than if they collected the data manually.[150] Second, a journalist can now follow that information in real-time, whether it's tracing a poll or reviewing comment feeds.[151] By using this systematic process, some of the greatest truths within the last decade have unfolded.[152] In fact, three data journalists have won Pulitzer Prizes from their discoveries.[153]

The COVID Tracking Project, by The Atlantic, provided Americans with real-time pandemic statistics involving "testing, hospitalization, patient outcomes, [and] racial and ethnic demographic information."[154] Reveal, from The Center for Investigative Reporting, discovered that hundreds of police officers across the nation joined "closed racist, Islamophobic, misogynistic or anti-government militia groups on Facebook."[155] Reuters exposed an underground market where parents

146. Carrero, *supra* note 125, at 144-45.

147. *Id.* at 144.

148. Br. for The Reporters Committee for Freedom of the Press et al. as Amici Curiae Supporting Pet'r, at 16, Van Buren v. United States, 140 S. Ct. 2667 (2020) (citing Branzburg v. Hayes, 408 U.S. 665, 681 (1972)).

149. Carrero, *supra* note 125, at 144-45.

150. Br. for The Markup as Amici Curiae Supporting Pet'r, at 16, Van Buren v. United States, 140 S. Ct. 2667 (2020).

151. *Id.*

152. *See generally* The Markup, *Why Web Scraping is Vital to Democracy*, NEXT WEB (Dec. 28, 2020, 5:00 PM), http://thenextweb.com/syndication/2020/12/28/why-web-scraping-is-vital-to-democracy/.

153. Br. for The Markup as Amici Curiae Supporting Pet'r, *supra* note 150, at 8-9.

154. The Atlantic Monthly Group, *About the Data*, COVID TRACKING PROJECT (Mar. 7, 2021), http://covidtracking.com/about-data. (emphasis omitted).

155. Will Carless, *Hundreds of Cops Are in Extremist Facebook Groups. Why Haven't Their Departments Done Anything About It?*, REVEAL (Sept. 30, 2019), http://revealnews.org/

sold their adopted children online, later claiming that the child went missing.[156] The Trace and The Verge located individuals selling guns on an online market, whom escaped the mandatory background check and licensing requirements.[157] Most recently, one independent researcher uncovered deleted images and videos from the January 6, 2021 attack on the United States Capitol by scraping data from the social media network Parler, and sharing the behind-the-scenes content with Americans.[158]

While these stories have changed the minds of millions, not to mention saved the lives of innocent human beings, data scraping is so controversial that journalists can't help question, "[i]s this a story worth going to prison for?"[159] Journalists are not "circumventing technological barricades" or "hacking" systems to discover this knowledge.[160] Applying the Ninth Circuit's analysis in *hiQ Labs*, journalists are merely using public knowledge to expose the intricacies of human fallacy on the Internet. The Supreme Court has even labeled the Internet as a "modern public square.".[161] Prohibiting individuals and news sources from using this public forum will inhibit the speed and accuracy of journalism.[162]

Certainly, most newsgathering tactics are no longer restricted in the wake of the *Van Buren* decision.[163] However, journalists who incorporate web scraping into their work will likely not rest easy until the Ninth Circuit decision is affirmed. If the court were to carve out a narrow exception for data scraping, the decision could ultimately restrict the dissemination of information from the public's point of view—an "individual's data would receive heightened protection and individuals

---

article/hundreds-of-cops-are-in-extremist-facebook-groups-why-havent-their-departments-done-anything-about-it/ (noting the Harris County Sheriff's Office was the only department to take action against an officer's involvement).

156. Megan Twohey, *Americans Use the Internet to Abandon Children Adopted from Overseas*, REUTERS (Sept. 9, 2013), http://www.reuters.com/investigates/adoption/#article/part1 (including the discovery was made by analyzing 5,029 posts from Yahoo and Facebook groups, after scraping the data).

157. Sean Campbell & Colin Lecher, *Millions of Guns for Sale. Few Questions Asked.*, TRACE (Jan. 16, 2020), http://www.thetrace.org/2020/01/armslist-unlicensed-gun-sales-engaged-in-the-business/.

158. Grayson Clary, *Parler Wasn't Hacked, and Scraping Is Not a Crime*, LAWFARE (Feb. 1, 2021, 3:42 PM), http://www.lawfareblog.com/parler-wasnt-hacked-and-scraping-not-crime.

159. Lam Thuy Vo, *Web Scraping is a Tool, Not a Crime*, MIT TECH. REV. (Dec. 8, 2020), http://www.technologyreview.com/2020/12/08/1013440/web-scraping-van-buren-case-supreme-court-opinion/.

160. Br. for The Markup as Amici Curiae Supporting Pet'r, *supra* note 150, at 5.

161. Packingham v. North Carolina, 137 S. Ct. 1730, 1732 (2017).

162. *See generally Why Web Scraping is Vital to Democracy*, *supra* note 152.

163. *See generally* Van Buren v. United States, 141 S. Ct. 1648 (2021).

would have more privacy . . . ."[164] This is especially helpful amongst protecting one's work from commercial users, or those scraping for profit.[165] Nonetheless, it is difficult to forget about the "major potential cost to other electronic liberties."[166] Perhaps this difficulty will introduce a middle ground. Creating an exception for researchers, academics, and journalists, otherwise known as non-commercial individuals, to scrape and review data could at least take privacy interests out of the public ring and prioritize the discoverers of society.[167] Leaving commercial users susceptible to liability would inevitably invite backlash, but the means for those companies to discover data through other avenues would greatly offset the end for stand-alone discovers whom could be labeled as criminals, as consistently voiced by the Supreme Court in *Van Buren*.[168]

V. CONCLUSION

The Supreme Court's decision in *Van Buren* has provided individuals with the ability to discover information without limitations, irrelevant of motives, so long as they do not "exceed authorized access."[169] The Supreme Court heavily relied on the plain language of Section 1030(a)(2) to avoid labeling the average American as a criminal in their everyday Internet use.[170] The Ninth Circuit in *hiQ Labs* now has the opportunity to affirm its previous decision in the wake of *Van Buren*, leaving open the possibility to uncover data and truths for the benefit of society.[171] If, however, the Ninth Circuit carves out an exception for web scraping, a broad reading of the "without authorization" prong could implement a stringent learning curve on individuals and companies alike in discovering both data and information.[172] A solution that prioritizes an individual's intellectual freedom to discover public data, yet guards against corporate greed, will likely provide the greatest balance.[173] The final debate on web scraping must come to an end.

---

164. Sakin & Rippy, *supra* note 102.
165. *See generally* Carrero, *supra* note 125 at 141.
166. Sakin & Rippy, *supra* note 102.
167. *See generally* Vo, *supra* note 159.
168. *See generally Van Buren*, 141 S. Ct. at 1656-62.
169. *Id.* at 1662.
170. *See id.*
171. *See generally* Carrero, *supra* note 125, at 132-35.
172. *See generally* Sakin & Rippy, *supra* note 102.
173. *See generally* Vo, *supra* note 159.