Washington International Law Journal

Volume 30 | Number 1

12-28-2020

Two Decades of Laws and Practice Around Screen Scraping in the Common Law World and Its Open Banking Watershed Moment

Han-Wei Liu
Monash University

TWO DECADES OF LAWS AND PRACTICE AROUND SCREEN SCRAPING IN THE COMMON LAW WORLD AND ITS OPEN BANKING WATERSHED MOMENT

Han-Wei Liu[†]

Abstract: Screen scraping—a technique using an agent to collect, parse, and organize data from the web in an automated manner—has found countless applications over the past two decades. It is now employed everywhere, from targeted advertising, price aggregation, budgeting apps, website preservation, academic research, and journalism, to name a few. However, this tool has raised enormous controversy in the age of big data. This article takes a comparative law approach to explore two sets of analytical issues in three common law jurisdictions, the United States, the United Kingdom, and Australia. As the first step, this article maps out the trajectory of relevant laws and jurisprudence around screen scraping legality in three common law jurisdictions—the United States, the United Kingdom, and Australia. Specifically, the article focuses on five selected issue areas within those jurisdictions—"digital trespass" statutes, tort, intellectual property rights, contract, and data protection. Our findings reveal some level of divergence in the way each country addresses the legality of screen scraping. Despite such divergence, one may see a sea change amid the trend of data-sharing under the banner of "Open Banking" in coming years. This article argues that to the extent that these data sharing initiatives enable information flow between entities, it could reduce the demand for screen scraping generally, thereby bringing some level of convergence. Yet, this convergence is qualified by the institutional design of data sharing schemes—whether or not it explicitly addresses screen scraping (as in Australia and the United Kingdom) and whether there is a top-down, government-mandated data-sharing regime (as in the United States).

Cite as: Han-Wei Liu, *Two Decades of Laws and Practice Around Screen Scraping in the Common Law World and its Open Banking Watershed Moment*, 30 WASH. INT'L L.J. 28 (2020).

INTRODUCTION

Text and data mining are, broadly speaking, overarching terms covering a range of techniques to extract useful information and explore patterns in data that might not be identified otherwise.[1] One popular technique is "screen scraping"—also known as "web scraping," "data scraping," "web data extraction," or "web data mining"—which refers to constructing "an agent to download, parse, and organize data from the web in an automated manner."[2] Put differently, screen scraping uses a software agent to mimic browsing interactions between web servers and people.[3]

[†] Dr. Han-Wei Liu, Lecturer (Assistant Professor), Monash University, Australia. The author is grateful for Tiana Moutafis and Lily Raynes for excellent research assistance.

[1] *See generally* RONEN FELDMAN & JAMES SANGER, THE TEXT MINING HANDBOOK: ADVANCED APPROACHES IN ANALYZING UNSTRUCTURED DATA (2006).

[2] The terms "web scraping" and "web crawling" are sometimes used interchangeably. Some data scientists remark that although the difference is vague, the term "crawler" means that a "program's ability to navigate web pages on its own, perhaps even without a well-defined end goal or purpose, endlessly exploring what a site or the web has to offer." SEPPE VANDEN BROUCKE & BART BAESENS, PRACTICAL WEB SCRAPING FOR DATA SCIENCE: BEST PRACTICES AND EXAMPLES WITH PYTHON 3, 155 (2018).

[3] Daniel Glez-Peña et al., *Web Scraping Technologies in an API World*, BRIEFING IN BIOINFORMATICS 788, 789 (2014) (describing web scraping as "[s]tep by step, the robot accesses as many

The practice is nothing new.[4] Screen scraping technology was used in the "account aggregation" services that emerged in the United States in the late 1990s.[5] These services enable clients to view account information from various institutions in one place.[6] They may either collate financial data (e.g., from deposit accounts, credit accounts, and managed funds accounts)[7] or non-financial data (e.g., from email accounts and frequent flyer accounts).[8] This business model has since diffused throughout Europe and the Asia-Pacific.[9] As early as 2000, for instance, Australia had seven firms providing data aggregation services—among them financial institutions, a stockbroker, and an app development company.[10] This era also marked the emergence of search engines such as Google, which use scraping bots that pull small amounts of data (i.e., the search terms entered) to link a user to relevant webpages.[11]

Screen scraping has since been applied in different contexts. It is now used for targeted advertising,[12] price aggregation,[13] budgeting apps,[14] website preservation,[15] academic research,[16] journalism,[17] and more.[18] Analytic start-ups draw insights for industries by scraping public data,[19]

Web sites as needed, parses their contents to find and extract data of interest and structures those contents as desired").

[4] Jeffrey Kenneth Hirschey, *Symbiotic Relationships: Pragmatic Acceptance of Data Scraping*, 29 BERKELEY TECH. L.J. 897, 903 (2014).

[5] Australian Securities & Investment Commission (ASIC), CONSULTATION PAPER 20: ACCOUNT AGGREGATION IN THE FINANCIAL SERVICES SECTOR 1 (2001) [hereinafter ASIC CONSULTATION PAPER 20].

[6] *Id.* at 9.

[7] *Id.* at 19.

[8] *Id.* at 22.

[9] *See, e.g.*, Hiroshi Fujii et. al., *E-Aggregation: The Present and Future of Online Financial Services in Asia-Pacific* (MIT Sloan School of Management, Working Paper CISL#2002-06), http://web.mit.edu/smadnick/www/wp/2002-06.pdf.

[10] ASIC CONSULTATION PAPER 20, *supra* note 5, at 16–17.

[11] Hirschey, *supra* note 4, at 898. "Bots" refer to an automated program designed to carry out a specific task or simulate a human activity. Paris Martineau, *What is a Bot?*, WIRED, (Nov. 16, 2018) https://www.wired.com/story/the-know-it-alls-what-is-a-bot/.

[12] Myra F. Din, *Breaching and Entering: When Data Scraping Should Be a Federal Computer Hacking Crime*, 81 BROOK. L. REV. 405, 408 (2015).

[13] *Id.* at 408.

[14] *Id.*; Tess Macapinlac, *The Legality of Web Scraping: A Proposal*, 71 FED. COMM. L.J. 399, 402 (2019).

[15] Andrew Sellars, *Twenty Years of Web Scraping and the Computer Fraud and Abuse Act*, 24 B.U. J. SCI. & TECH. 372, 374 (2018).

[16] Macapinlac, *supra* note 14, at 402.

[17] Jacquellena Carrero, *Access Granted: A First Amendment Theory of Reform of the CFAA Access Provision*, 120 COLUM. L. REV. 131, 146 (2020).

[18] *See, e.g.*, Sellars, *supra* note 15, at 374.

[19] Ioannis Drivas, *Liability for Data Scraping Prohibitions under the Refusal to Deal Doctrine: An Incremental Step toward More Robust Sherman Act Enforcement*, 86 UNIV. CHI. L. REV. 1901, 1903–04 (2019).

while Fintech[20] firms purchase data made available by aggregators to develop new products and services.[21] There is, therefore, an increasingly complex symbiotic relationship between scrapers and data hosts, with many scraping services benefitting both parties.[22] Given their everyday use for a wide range of commercial and non-commercial purposes, scraping bots are estimated to account for nearly a quarter of all internet traffic.[23]

However, screen scraping can be controversial. It can be detrimental to the data host and consumer.[24] Scraping is parasitic when it undercuts a website's revenue by republishing data without requiring users to view supporting advertisements.[25] It can facilitate copyright infringement at scale[26] or even impact the data host's services by overloading servers.[27] Screen scraping can also raise privacy concerns for consumers if it collects identifiable information or enables new forms of surveillance.[28] In the banking context—where login credentials may be shared to allow the scraping of account data—there are additional concerns relating to cybersecurity, data breach, and liability allocation for unauthorized transactions.[29] These problematic applications of screen scraping have led to litigation against scrapers—most notably in the United States.

This article's aim is two-fold. First, it seeks to maps the trajectory of relevant laws and jurisprudence around the legality of screen scraping in three common law jurisdictions and contrasts how one may challenge it differently. Second, it assesses the extent to which a new development—Open Banking—may affect screen scraping's legal landscape

More specifically, regarding the trajectory of relevant laws, Section II of this article focuses on five selected issue areas—digital trespass (or hacking) statutes, tort, intellectual property rights (IPR), contractual rights, and privacy/data protection in the United States, United Kingdom, and

---

[20] Fintech is a contraction of "financial technology" that refers to technology-enabled financial solutions. *See* Brian Hurh et. al., *Consumer Financial Data Aggregation and the Potential for Regulatory Intervention*, 71 CONSUMER FIN. L.Q. REP. 20, 21 (2017).

[21] *See* Douglas W. Arner et al., *The Evolution of Fintech: A New Post-Crisis Paradigm?*, 47 GEO. J. INT'L L. 1271, 1271 (2016).

[22] Hirschey, *supra* note 4, at 897–98.

[23] Macapinlac, *supra* note 14, at 402; Drivas, *supra* note 19, at 1903–04.

[24] Nabeel Hasan Saeed, *Good or Evil? What Web Scraping Bots Mean for Your Site?*, IMPERVA, (Apr. 18, 2016) https://www.imperva.com/blog/web-scraping-bots/ ("Database scraping can be used to steal intellectual property, price lists, customer lists, insurance pricing and other datasets that would require an effort prohibitively tedious for humans, but perfectly within the range of what bots routinely do."); *see also* Hirschey, *supra* note 4, at 898–99.

[25] *Id.*

[26] Sellars, *supra* note 15, at 374–75.

[27] Hirschey, *supra* note 4, at 898–99.

[28] *Id.*; Sellars, *supra* note 15, at 374–75.

[29] ASIC Consultation Paper 20, *supra* note 5, at 46; BASEL COMMITTEE ON BANKING SUPERVISION, REPORT ON OPEN BANKING AND APPLICATION PROGRAMMING INTERFACES 12, 14 (Nov. 2019).

Australia.[30] This article argues that the use of tort law, in the form of a "trespass to chattels" claim, is more likely to succeed in the United States than in the United Kingdom or Australia. The Computer Fraud and Abuse Act is also a handy tool for litigating against scrapers in the United States, despite the vague and evolving concept of "authorized" access. In the other two jurisdictions, similar legislation is either absent (in Australia) or has not been applied for this purpose (in the United Kingdom). By contrast, intellectual property infringement claims are more likely to succeed in the United Kingdom given the existence of a "database right," which does not exist in the other two states. There is room for claims based on contractual rights (as derived from a website's Terms of Use) in all three common law jurisdictions. However, in Australia, such claims may be the "first line of defense" against screen scraping given the absence of a hacking statute or database right.[31] Finally, scraping personally identifiable information (PII) may breach privacy/data protection in Australia and the United Kingdom, but no comprehensive data privacy legislation exists in the United States at the federal level yet.

The article continues in Section III, arguing that despite the jurisdictional divergence in how this technology is treated, Open Banking initiatives' rise in recent years could moderate concerns and bring a certain level of convergence. This article's analysis shows that to the extent Open Banking mandates or facilitates data sharing, it could reduce the need for screen scraping. This is especially so in the European context—and even more so if the United Kingdom's Smart Data initiative expands these data-sharing principles beyond the financial sector. Conversely, the financial data-sharing environment is less clear in the United States, which lags in building up Open Banking. Australia lies in the middle of these two extremes: it has a comprehensive Consumer Data Rights (CDR) regime that can theoretically reduce screen scraping needs.[32] But, given the fact that it imposes no ban on screen scraping (unlike the European Union or the United Kingdom), it has a loophole for data miners to work around the new regime and continue scraping data.

---

[30] Two caveats are in order. First, while screen scraping also raises antitrust or competition law concerns, these issues are not the focal point of this paper because they are complicated enough to be addressed in different scholarship. Second, for the purpose of this article, we use the terms "privacy" and "data protection" interchangeably, while acknowledging that they may be understood differently across contexts. For instance, in the context of the United States, the term privacy is read by the courts broadly enough to cover not only data protection but a wide range of rights, such as the right to be free from unreasonable search and seizure by governments and right to make private decisions like abortion or contraception. Hence, some suggest that it is more precise to use the term "data protection." COLIN J. BENNETT, REGULATING PRIVACY: DATA PROTECTION AND PUBLIC POLICY IN EUROPE AND THE UNITED STATES 12–14 (1st ed. 1992).

[31] Lesley Sutton et. al., *Screen Scraping: Legal or Not?*, GILBERT & TOBIN (Apr. 14, 2020), https://www.gtlaw.com.au/insights/screen-scraping-legal-or-not.

[32] The Treasury Laws Amendment (Consumer Data Right) Act 2019 (Cth) (Austl.).

I. CONTROVERSY AROUND SCREEN SCRAPING: A DIVERGENT COMMON LAW WORLD

Legal claims to prevent screen scraping vary between the United States, United Kingdom, and Australia. The United States has developed more jurisprudence in the area. However, there is still a relatively small amount of legal scholarship addressing screen scraping—especially from the comparative law perspective. Notable aspects of each jurisdiction's legal claims are discussed below.

A. The United States' Approach Towards Screen Scraping

While screen scraping is not explicitly addressed in the United States' legislation, it has been heatedly debated in considerable case law. The following discussion illustrates this development, dividing four major claims into sub-sections: (i) contravention of the Computer Fraud and Abuse Act (CFAA), (ii) trespass to chattels, (iii) compilation copyright infringement, and (iv) breach of contract.[33]

1. Contravention of the Computer Fraud and Abuse Act. — In 1986, the United States enacted the CFAA, which amended various parts of 18 U.S.C. §1030 ("Fraud and Related Activity in Connection with Computers").[34] This cybersecurity statute was later expanded to allow for civil liability,[35] creating an avenue for relief to harmed individuals seeking compensatory damages, or injunctive relief and other equitable remedies.[36] The CFAA's centerpiece is its prohibition on hacking, which occurs when a person "intentionally accesses a computer without authorization or exceeds authorized access and thereby obtains…information from any

[33] Arguably, there are other causes of actions that are less common, like trademark infringement, unfair competition, misappropriation, intentional interference with contractual relationship, and trade secret-related claims. *See, e.g.*, Hirschey, *supra* note 4, at 903; Vlad Krotov & Leiser Silva, *Legality and Ethics of Web Scraping* 3 (2018); Amber Zamora, *Making Room for Big Data: Web Scraping and an Affirmative Right to Access Publicly Available Information Online*, 12 J. BUS. ENTREPRENEURSHIP & L. 203, 205 (2019).

[34] Tess Macapinlac, *The Legality of Web Scraping: A Proposal*, 71 FED. COMM. L.J. 399, 403 (2019); Samantha Jensen, *Abusing the Computer Fraud and Abuse Act: Why Broad Interpretations of the CFAA Fail*, 36 HAMLINE L. REV. 81, 89 (2014).

[35] Computer Abuse Amendments Act of 1994, Pub. L. No. 103-322, 108 Stat. 2097; Myra F. Din, *Breaching and Entering: When Data Scraping Should Be a Federal Computer Hacking Crime*, 81 BROOK. L. REV. 405, 416 (2015); Kathleen C. Riley, *Data Scraping as a Cause of Action: Limiting Use of the CFAA and Trespass in Online Copying Cases*, 29 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 245, 266–67 (2019); Jensen, *supra* note 34, at 85; Andrew Hernacki, *A Vague Law in A Smartphone World: Limiting the Scope of Unauthorized Access Under the Computer Fraud and Abuse Act*, 61 AM. U. L. REV. 1543, 1550 (2012).

[36] 18 U.S.C. § 1030(g) (1984). Parties may obtain relief under these provisions if they demonstrate that they suffered a loss during a one-year period aggregating to at least $5,000 in value. *Id.* While there are other grounds for a civil action, they do not seem to arise in scraping cases. *See* Andrew Sellars, *Twenty Years of Web Scraping and the Computer Fraud and Abuse Act*, 24 B.U. J. SCI. & TECH. L. 372, 376 (2018).

protected computer." [37] Yet, the term "authorization" is left undefined, and the United States' courts have varying interpretations of it, as detailed in the sections that follow. In this regard, the best way to understand the trajectory of screen scraping jurisprudence under the CFAA is through the analytical framework offered by Andrew Sellars.[38]

(a) Evolving Judicial Interpretation of "Authorization." — The CFAA was introduced before web scrapers or the internet ever existed. Therefore, the act does not explicitly refer to screen scraping. [39] Nonetheless, the CFAA has been invoked in litigation against scraping since the early 2000s.[40] Courts' approaches can roughly be divided into four phases.[41] After broad application in Phase I, which involved a decade of litigation, the CFAA's interpretation shifted to a narrower reading in the late 2000s—Phase II.[42] Phase III began in the mid-2010s, wherein CFAA's reading expanded again.[43] Recent decisions—Phase IV—have narrowed the CFAA more, making it harder to stop scrapers from accessing public websites.[44]

In Phase I—roughly from the turn of the millennium to 2009—courts adopted an expansive view of the CFAA,[45] under which a website only had to point to a mechanism that indicated the scraper's access was "unauthorized," whether contractual, technical or otherwise.[46] Any signal of a website's disapproval could have provided sufficient notice to scrapers that subsequent access would be "unauthorized" and breach the Act.[47] These signals included breaching a term of service,[48] accessing a public website after express warnings to stay away,[49] and even the complaint

---

[37] 18 U.S.C. § 1030(a)(2) (1984); The term "protected computer" broadly encompasses any computer that affects interstate or foreign commerce or communication. 18 USC § 1030(e)(2)(B); H. MARSHALL JARRETT ET AL., PROSECUTING COMPUTER CRIMES 4 (2nd ed. 2015).

[38] Andrew Sellers is Director of the Technology Clinic at the Boston University School of Law. Sellars, *supra* note 15, at 377.

[39] Macapinlac, *supra* note 14, at 404, 412, 422.

[40] *See, e.g.*, eBay, Inc. v. Bidder's Edge, Inc., 100 F. Supp. 2d 1058, 1069 (N.D. Cal. 2000); Sellars, *supra* note 15, at 388.

[41] Sellars, *supra* note 15, at 379–81.

[42] *Id.*

[43] *Id.*

[44] *Id.* at 380–81.

[45] Phase I refers to period from the *eBay v. Bidder's Edge* decision in 2000 to the *LVRC Holdings LLC v. Brekka* decision in 2009. *eBay, Inc.*, 100 F. Supp. at 1070; LVRC Holdings LLC v. Brekka, 581 F.3d 1127 (9th Cir. 2009).

[46] Sellars, *supra* note 15, at 379; "Technical" barriers include click-through agreements, IP address blockers and robot exclusion protocols. Drivas, *supra* note 19, at 1904–05. Contrast this with "non-technical" measures such as website "terms and conditions" and cease-and-desist letters. *Id.*

[47] Sellars, *supra* note 15, at 394.

[48] Southwest Airlines Co. v. Farechase, Inc., 318 F. Supp. 2d 435, 439 (N.D. Tex. 2004) (terms of use banned "any deep-link, page-scrape, robot, spider or other automatic device, program, algorithm or methodology which does the same things.").

[49] *Id.* at 439–40.

filing itself.[50] In two cases, the court suggested that a well-deployed terms of use notice on the website could adequately signal the extent of users' authorized access on the site.[51] Overall, there was no limitation on what could inform authorization, with Sellars identifying the decade as "a very uncertain time for web scrapers."[52]

As Sellars observed, Phase II started in 2009, when the United States' courts began adopting a narrower view of the CFAA.[53] Courts rejected claims against scrapers where a website merely placed restrictions on their website's data *usage*, rather than limitations on site *access*, and interpreted the scope of a scraper's authorization by referencing code-based controls (rather than those set by contract or principles of duty).[54] With this understanding, courts denied the use of a website's terms of use to support a CFAA claim, as such contractual terms usually only imparted "use restrictions"—that is, limiting what can be done with the information after one arrives rather than "access restrictions."[55] In *Cvent v. Eventbrite*, although the website's terms of use prohibited competitors from accessing information, it had not taken any meaningful steps to block its competitors from doing so.[56] Web scraping of the publicly available site was thus not a CFAA breach, as anyone, including competitors, could search and access the plaintiff's information at will.[57] This higher threshold of liability was reinforced by a 2010 decision in *Facebook v. Power Ventures*. There, the court clarified that terms of use and cease-and-desist notices were insufficient by themselves to show liability on the scraper's part.[58] The

---

[50] Sellars, *supra* note 15, at 394 n.163 (quoting Register.com v. Verio, 126 F. Supp. 2d 238, 249 (S.D.N.Y. 2000)) ("[I]t is clear since at least the date this lawsuit was filed that Register.com does not consent to Verio's use of a search robot[.]").

[51] *Id.* at 395 (citing *Zefer*, 318 F.3d at 63; *Healthcare Advocates*, 497 F. Supp. 2d at 649).

[52] *Id.* at 393, 395.

[53] *Id.* at 380.

[54] *Id.* at 379, 396. Under the code-based interpretation of "authorization," user authorization is based on the operation of the computer system. Access would be unauthorized, and thus unlawful, if the user purposefully circumvents code-based protections (i.e., computer passwords) to gain access to or use the device in a way that would otherwise not be accessible. This approach can be dated back to the earlier CFAA cases such as *United States v. Morris*. *United States v. Morris*, 928 F.2d 504 (2d Cir. 1991). On this issue, see, e.g., Katherine Mesenbring Field, *Agency, Code, or Contract: Determining Employees' Authorization under the Computer Fraud and Abuse Act*, 107 Mich. L. Rev. 819, 825 (2009) and Orin S. Kerr, *Cybercrime's Scope: Interpreting "Access" and "Authorization" in Computer Misuse Statutes*, 78 N.Y.U. L. REV. 1596, 1642 (2003) [hereinafter Kerr, *Cybercrime's Scope*].

[55] Sellars, *supra* note 15, at 379, 398.

[56] Cvent, Inc. v. Eventbrite, Inc., 739 F. Supp. 2d 927, 932–33 (E.D. Va. 2010) ("Cvent's website, including its CNS database, is therefore not protected in any meaningful fashion by its Terms of Use or otherwise").

[57] *Id.*

[58] Facebook, Inc. v. Power Ventures, Inc., No. C 08-05780 JW, 2010 WL 3291750, at *11 (N.D. Cal. July 20, 2010). The Court discussed how imposing criminal liability on the basis of TOU or a cease-and-desist letter would grant the data host the ability to define the scope of federal criminality, which it found "constitutionally untenable." *Id.* Users cannot have adequate notice of what actions will or will not expose them to criminal liability given that a website administrator "can unilaterally change the rules

court explained that using a website's terms of use to determine authorization would "create a constitutionally untenable situation in which criminal penalties could be meted out on the basis of violating vague or ambiguous terms of use."[59] Instead, the primary issue was whether the scraper evaded technical or code-based barriers in accessing the information.[60] In short, while scrapers still faced potential liability in Phase II (2009–13), they could successfully defend a lawsuit by arguing that the plaintiff's mechanism was merely a "use restriction" or that the authorization mechanism should have been more code-based to have legal effect.[61]

The advent of Phase III in 2013 reversed this narrowing trend—the courts broadened the CFAA's interpretation with the revocation theory.[62] Beginning with *Craigslist v. 3Taps*, Craigslist filed a lawsuit against a company that scraped, aggregated, and republished its advertisements.[63] While 3Taps argued that everyone was authorized to access Craigslist, a public website, the court nevertheless found that Craigslist had revoked 3Tap's default authorization by sending multiple cease-and-desist letters and blocking its IP addresses.[64] The court found these measures to be effective notices of revocation and thus subsequent scraping was a violation of the CFAA.[65] Under the revocation theory, a website could establish liability if it demonstrated that it "revoked" access to the scraper at some point and that the scraper knew they had notice of the revoked access but continued to access the site.[66] This theory broadened what could constitute "without authorization" under the CFAA. Now, any action by a

---

at any time and are under no obligation to make the terms of use specific or understandable to the general public." *Id.* The court contrasted this to a scraper evading technical or code-based barriers: such access crosses a clear demarcation that has been erected by the website administrator "to restrict the user's privileges within the system," and a person applying the technical skill necessary to overcome such a barrier "will almost always understand that any access gained through such action is unauthorized." *Id.* Thus, accessing a computer network or website in a manner that overcomes technical or code-based barriers is "without permission" or "without authorization," and may subject a user to liability under the Act. *Id.*

[59] *Id.*

[60] *Id.* This case was decided based on California Penal Code § 502, which is California's analogue to the CFAA. *Id.* at 7. It was held that the fact that Power Venture's scraping activities breached Facebook's TOU did *not* mean that they had contravened the statute. *Id.* at 12. Yet, to the extent that Facebook could prove that Power Ventures circumvented technical barriers, it could be held liable for violating the statute. *Id.* Two years later, Facebook was granted summary judgment after it showed that Power Ventures did indeed circumvent technical barriers by deliberately evading IP address blocks. *Facebook, Inc. v. Power Ventures, Inc.*, 844 F. Supp. 2d 1025, 1038 (N.D. Cal. 2012).

[61] Sellars, *supra* note 15, at 401.

[62] *Id.* at 380.

[63] Craigslist, Inc. v. 3Taps, Inc., 942 F. Supp. 2d 962, 966 (N.D. Cal. 2013) ("Craigslist alleges that 3Taps copies (or "scrapes") all content posted to Craigslist in real time, directly from the Craigslist website.").

[64] *Id.* at 969–70.

[65] Sellars, *supra* note 15, at 410.

[66] *Id.* at 380.

website that signaled revocation of a user's access, whether done in an access-based or code-based manner, could be used as evidence of unlawful conduct in violation of the CFAA.[67] For example, courts have found CFAA violations based on a direct demand to stop accessing the website,[68] the website imposing an IP address block,[69] or even the contents of a website's terms of use.[70] They held that these actions served as a valid notice of access revocation.[71] But the United States' courts' refocus on revocation seems to sideline technical control issues (e.g., IP-address blocks), an issue highlighted in *Power Ventures II*.[72] This revocation-based theory, "[re]opened the door to a wide array of authorization mechanisms that previously had been narrowed away."[73]

In 2017—arguably Phase IV's beginning, but it is too early to definitively say—the courts began rejecting this broader revocation-based reading.[74] The first case to do so was *hiQ Labs v. LinkedIn*,[75] where hiQ Labs scraped data from public LinkedIn profiles to offer business analytics.[76] LinkedIn sent a cease-and-desist letter to the defendant and imposed an IP block on it.[77] Despite the facts bearing a striking resemblance to *3Taps*, the court found the defendant's scraping was not "access without authorization" in violation of the CFAA.[78] In reaching its conclusion, the court distinguished *Power Ventures II*, reasoning that the data there was not "public" because login credentials were required to access Facebook's content.[79] The court also seemed to signal disagreement

---

[67] *Id.* at 402.

[68] CouponCabin LLC v. Savings.com, Inc., No. 2:14-CV-39-TLS, 2016, 2016 WL 3181826, at *4 (N.D. Ind. June 8, 2016); Sellars, *supra* note 15, at 405.

[69] CouponCabin LLC v. Savings.com, Inc., No. 2:14-CV-39-TLS, 2017 WL 83337, at *3 (N.D. Ind. Jan. 10, 2017); Sellars, *supra* note 15, at 405 (as to one party who did not receive a direct communication, "[r]evocation of website access would have been sufficient to give the Defendants constructive notice that they were without authorisation to act as they allegedly did").

[70] *See* QVC, Inc. v. Resultly, LLC, 99 F. Supp. 3d 525 (E.D. Pa. 2015) ("[J]ust as a cease-and-desist letter would put a publisher on notice that its actions were prohibited, VigLink's Terms of Service . . . put Resultly on notice that QVC prohibited web-crawling"); DHI Group, Inc. v. Kent, No. 16-cv-1670, 2017 WL 4837730 (S.D. Tex. Oct. 26, 2017); Sellars, *supra* note 15, at 405.

[71] *See* Sellars, *supra* note 15, at 405 ("With the focus placed on "revocation," questions about the legal impacts of technical controls like user accounts or IP and MAC address filtering all fell away in favor of an analysis which asked whether the website owner used a technical control to signal a revocation of access . . . .").

[72] Facebook, Inc. v. Power Ventures, Inc. (*Power Ventures II*), 844 F.3d 1058, 1068 (9th Cir. 2016); Sellars, *supra* note 15, at 406.

[73] Sellars, *supra* note 15, at 404.

[74] *Id.* at 381.

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

[76] *Id.* at 1104.

[77] *Id.*

[78] Sellars, *supra* note 15, at 408; *hiQ Labs, Inc.* 273 F. Supp. 3d at 1114–18.

[79] *hiQ Labs, Inc.*, 273 F. Supp. 3d at 1109 ("[n]one of the data in *Facebook* or *Nosal II* was *public* data") (emphasis in original).

with the result in *3Taps* after looking at the CFAA's legislative history.[80] To support its reasoning, the court referenced Professor Orin Kerr's seminal work, *Norms of Computer Trespass*, which draws an analogy to trespass law to read into the scope of "authorization" and sees the website as "inherently open."[81] Lastly, the court also considered public policy, noting that assigning CFAA liability when someone accesses a website in breach of a written instruction would allow website owners to block users for improper purposes such as anti-competition.[82]

A year after *LinkedIn*, a group of scholars and journalists who used scraping in their research filed suit in *Sandvig v. Sessions*.[83]That case expressly took the narrow view of the CFAA seen in *3Taps*, finding that only code-based controls—rather than use restrictions—should be the basis of CFAA liability because the public should have a general right to access publicly-facing websites.[84] In doing so, the court recognized that scraping "is merely a technological advance that makes information collection easier; it is not meaningfully different from using a tape recorder instead of taking written notes." [85] These cases either do not address the "revocation" line of cases or attempt to fit their analysis into them, but in a manner that would seemingly make it "far more difficult to stop a scraper from accessing a website available to the general public, even if told to stop by the website in question."[86]

This narrow view would "perfectly align the CFAA with the technical realities of web scraping," which should not be thought of as inherently more invasive or dangerous than a person or web browser.[87] However, there are vital issues that have not been addressed, such as

---

[80] *hiQ Labs, Inc.*, 273 F. Supp. 3d at 1109 ("The CFAA must be interpreted in its historical context, mindful of Congress' purpose. The CFAA was not intended to police traffic to publicly available websites on the Internet . . . .").

[81] *Id.* at 1111–13. For a detailed analysis, see Orin Kerr, *Norms of Computer Trespass* 116 COLUM. L. REV. 1143, 1153, 1163 (2016) ("[A] person who connects a web server to the Internet agrees to let everyone access the computer much like one who sells his wares at a public fair agrees to let everyone see what is for sale").

[82] *hiQ Labs Inc.*, 273 F. Supp. 3d at 1119. On public interests, hiQ argued that "private party should not have the unilateral authority to restrict other private parties from accessing information that is otherwise available freely to all" and this can raise "serious constitutional questions" for it would allow private parties to decide "who gets to participate in the marketplace of ideas located in the 'modern public square' of the Internet." *Id.* LinkedIn rejected this view, contending that screen scraping can raise privacy concern and more crucially, "if its users knew that their data was freely available to unrestricted collection and analysis by third parties for any purposes, they would be far less likely to make such information available online." *Id.*

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

[84] *Id.* at 13. The court also referred to Professor Kerr's work, noting that "code-based restrictions, which 'carve[ ] out a virtual private space within the website or service that requires proper authentication to gain access,' remove those protected portions of a site from the public forum." *Id.* For a detailed analysis of Orin Kerr's work in this regard, see Kerr, *supra* note 81.

[85] *Sandvig v. Sessions*, 315 F. Supp. 3d at 1 (D.D.C. 2018).

[86] Sellars, *supra* note 15 at 381.

[87] *Id.* at 412–13, 415.

resolving the tension between the "revocation" cases (in Phase III) and the more recent cases (arguably Phase IV) that find a general right to access.[88] Thus, the United States' courts will need to develop a more coherent approach going forward.[89]

(b) Reflections. — The CFAA has long been criticized for its "disproportionate punishments," "vague definitions," and "overbroad" terms.[90] While the Act was initially aimed at criminal hackers, [91] its failure to define several key terms like "access" and "authorization" have permitted its application in situations where no hacking actually occurred.[92]

Critiques of the CFAA can be boiled down to two competing narratives.[93] First, anyone can access any website. Second, website owners could place limitations, caveats, or barriers to access. The hard dilemma facing policymakers is when and where to draw the line to keep the cyberspace open without overly undercutting the scope of "authorization" under the CFAA. There are proposals to address these issues.[94] The first option is to read the CFAA narrowly and within its unique context—that of computer networks and the internet—with contextual meanings, therefore, given to critical terms like "exceeds authorized access."[95] Legal practitioner Kathleen Riley has argued that courts should create a "judicial presumption of authorization" in CFAA cases involving public websites or valid login information.[96] Only by way of showing that a user "hacked" the website or otherwise had no permission to use the login credentials could this presumption be overcome.[97] Under this narrow reading, circumventing measures like IP address blocks would *not* be considered a CFAA violation. Instead, the scope of the CFAA would be properly limited to hacking.[98] This would square with the statute's original purpose.[99]

---

[88] *Id.* at 413.

[89] *Id.*

[90] Hernacki, *supra* note 35, at 1554; Macapinlac, *supra* note 14, at 404, 412; Riley, *supra* note 35, at 271, 299; Jensen, *supra* note 34, at 84; Jonathan Keim, *Updating the Computer Fraud and Abuse Act*, FEDERALIST SOCIETY (Oct. 2015), https://fedsoc.org/commentary/publications/updating-the-computer-fraud-and-abuse-act-1 (last visited July 9, 2020); Carrero, *supra* note 17, at 134.

[91] Riley, *supra* note 35, at 267, 272.

[92] Riley, *supra* note 35, at 271.

[93] Kerr, *supra* note 81, at 1161.

[94] *See, e.g.*, Riley, *supra* note 35, at 290.

[95] *Id.* at 245, 291. *See also* Jensen, *supra* note 34, at 81; Hernacki, *supra* note 35, at 1543; Carrero, *supra* note 17, at 170.

[96] Riley, *supra* note 35, at 245, 294–95.

[97] *Id.*

[98] *See, e.g.*, Riley, *supra* note 35, at 294–95 (arguing that "the presumption of authorization to access a public website can only be overcome by a showing that a user did not have permission…or that user 'hacked' the website.").

[99] *Id.* at 296; Hernacki, *supra* note 35, at 1574.

A more straightforward option is amending the CFAA to clarify the ambiguous terms and add exceptions around data scraping.[100] While courts have so far been reluctant to limit the CFAA's scope, adding more specific definitions of terms like "authorization" and "access" would assist them in doing so.[101] An explicit carve-out rule for scraping public information would "limit the pool of defendants to true bad actors and allow the activities of data aggregators to continue," "prevent large companies from using the courts for anti-competitive purposes," and "honor the traditions of openness upon which the Internet was built."[102] Unfortunately, efforts to amend the statute have thus far failed. In 2013, a bill known as "Aaron's Law" was introduced to amend the CFAA to add that a violation of terms of service could not be prosecuted under the Act and penalties would be made more suited to the crime.[103] Despite praise from organizations like the Electronic Frontier Foundation, the bill never passed.[104] Many scholars are calling for the CFAA's modernization to "reflect the significant technological changes that have occurred since 1986."[105] Instead, web scraping would be more appropriately examined by doctrines that police the *use* of information, as in the case of copyright law.[106] Notably, the United States Supreme Court may weigh in on this perennial issue in *LinkedIn*, should it grant review. While this case is still pending, the recent Open Banking movement that facilitates data sharing may play a role in managing these ramifications too.

2. Trespass to Chattels. — Trespass to chattels offers website operators an alternative avenue for relief in the United States.[107] Generally, this state law claim may be committed by "intentionally . . . using or intermeddling with a chattel in possession of another" when "the chattel is

---

[100] Riley, *supra* note 35, at 291; Hernacki, *supra* note 35, at 1581; Zamora, *supra* note 33, at 224–26.

[101] Riley, *supra* note 35, at 300–01.

[102] Zamora, *supra* note 33, at 224–25; Carrero, *supra* note 17, at 135 (noting that currently, the CFAA "crudely lumps together different forms of scraping that have different motivations and implications for social values.").

[103] Aaron's Law Act of 2013, H.R. 2454, 113th Cong. (2013).

[104] Carrero, *supra* note 17, at 166; Macapinlac, *supra* note 14, at 405. Aaron's Law was introduced by members of Congress in response to the prosecution of Aaron Swartz, who scraped the contents of the JSTOR academic article database for a research project and was charged with eleven violations of the CFAA as a result. Kieren McCarthy, *'Aaron's Law's Back on the Table to Bring Sanity to U.S. Hacking Laws*, REGISTER (Apr. 23, 2015), https://www.theregister.com/2015/04/23/congress_reintroduces_aarons_law/. In 2015, it was reintroduced in both the Senate and House, though these efforts were not fruitful either. Aaron's Law Act of 2015, S. 1030, 114th Cong. (2015); Aaron's Law Act of 2015, H.R. 1918, 114th Cong. (2015). For a recount, see, e.g., Kieren McCarthy, *'Aaron's Law's Back on the Table to Bring Sanity to U.S. Hacking Laws*, REGISTER (Apr. 23, 2015), https://www.theregister.com/2015/04/23/congress_reintroduces_aarons_law/; s*ee also* Indictment, United States v. Swartz, No. 1:11-cr-10260 (D. Mass. July 14, 2011).

[105] Macapinlac, *supra* note 14, at 422; Riley, *supra* note 35, at 300–01.

[106] Sellars, *supra* note 15, at 388; Riley, *supra* note 35, at 305.

[107] Zamora, *supra* note 33, at 220.

impaired as to its condition, quality, or value, or… the possessor is deprived of the use of the chattel for a substantial time."[108]

It was first applied to the digital context in the 1990s[109] and subsequently extended to a scraping case in *eBay v. Bidder's Edge*.[110] This case laid out the legal standard for trespass to chattels claims in the web scraping context: the plaintiff must demonstrate (1) that the defendant intentionally and without authorization interfered with the plaintiff's possessory interest in the computer system, and (2) that the defendant's unauthorized use damaged the plaintiff.[111] In its application to scraping cases, the tort seems to undergo at least two developmental stages. While courts were initially willing to apply the doctrine even where there was no physical damage to the digital property,[112] physical damage has recently become a requirement for trespass to chattels claims.[113]

The first stage began with *Bidder's Edge*, where eBay successfully sued Bidder's Edge to stop it from scraping its website.[114] eBay sued under a trespass to chattels claim, arguing that Bidder's Edge intermeddled with eBay's servers without authorization, resulting in them "free-riding" on the time, effort, and money that eBay had invested to create its system.[115] While the increased traffic on eBay's server caused by Bidder's Edge scraping alone was insignificant (i.e., comprising less than 2% of the total capacity),[116] the court found potential future harm in the possibility that other data aggregators would scrape the website and collectively burden servers.[117] Similarly, the court in *Southwest Airlines v. Farechase* found scraping flight information from the airline's website constituted a trespass to chattels.[118] While Southwest could not prove it had endured physical harm or deprivation, the scraper's use was unauthorized and deceived Southwest customers who mistakenly believed they had contracted with

---

[108] Riley, *supra* note 35, at 265 (citing RESTATEMENT (SECOND) OF TORTS §§ 217(b), 218(b)–(c) (AM. LAW INST. 1965); Register.com, Inc. v. Verio, Inc., 356 F.3d 393, 438 n.58 (2d Cir. 2004)).

[109] Din, *supra* note 12, at 432 (noting that in the 1990s, trespass to chattels was applied to cases involving devices that overused phone and email networks, diminishing their functionality).

[110] eBay, Inc. v. Bidder's Edge, Inc., 100 F. Supp. 2d 1058, 1065, 1069–72 (N.D. Cal. 2000); Riley, *supra* note 35, at 265 (arguing that trespass to chattels is "commonly argued in data scraping cases, under the theory that a defendant's scraping interfered with a plaintiff's use of its website and servers by consuming intangible resources such as network and server capacity. These harms are often acknowledged to be minimal.").

[111] Zamora, *supra* note 33, at 220.

[112] *See, e.g.*, *Thrifty-Tel, Inc v. Bezenek*, 54 Cal. Rptr. 2d, 473 (holding that "the electronic signals generated by the Bezenek boys' activities were sufficiently tangible to support a trespass cause of action.").

[113] Din, *supra* note 12, at 433.

[114] *Id.* at 435.

[115] eBay, Inc. v. Bidder's Edge, Inc., 100 F. Supp. 2d 1058, 1065, 1070-71 (N.D. Cal. 2000).

[116] *Id.* at 1064.

[117] *Id.* at 1071-72; Hirschey, *supra* note 4, at 919.

[118] Southwest Airlines Co. v. Farechase, Inc., 318 F. Supp. 2d 435, 442 (N.D. Tex. 2004).

Southwest.[119] As a result, the court found Farechase's scraping activity wrongfully interfered with Southwest's use and possession of its website.[120] In both cases, the websites subject to scraping were publicly accessible, but the scraping still constituted a trespass as it was sufficiently outside the scope of the sites' permitted uses.[121]

In 2003, courts began requiring that scrapers physically interfere with the use or operation of a computer before assigning liability.[122] One court reasoned that gathering data from a public website, without more, is insufficient to fulfill a trespass action's harm requirement.[123] That the same year, the Supreme Court of California employed similar reasoning in *Intel v Hamidi*.[124] Although *Hamidi* was not a scraper case,[125] the court held that to invoke a trespass to chattel claim successfully, the plaintiff would need to prove that "a legally protected interest was damaged,"[126] making a minor interference with server usage was insufficient to make out actionable harm.[127] As these cases indicate, web scraping has become less actionable under a trespass to chattels theory.[128]

Nonetheless, trespass to chattel is still a viable legal option when a scraper causes actual harm.[129] Courts have found actual harm to include "overburdened networks, lost space, threats to business reputation and goodwill with customers, threats of similar future conduct, intermeddling with servers without authorization, wrongful interference with use or possession, and free-riding on data hosts' investments."[130] One commenter suggested that while the *Hamidi* and *Ticketmaster* decisions were less flexible in their harm determinations, [131] they "involved different

---

[119] *Id.* at 422; Din, *supra* note 12, at 436.

[120] *Southwest Airlines Co.*, 318 F. Supp. 2d at 442; Din, *supra* note 12, at 436.

[121] *eBay*, 100 F. Supp. 2d at 1070; *Southwest Airlines Co.*, 318 F. Supp. 2d at 442.

[122] Din, *supra* note 12, at 436.

[123] Ticketmaster Corp. v. Tickets.Com, Inc., No. 2:99-cv-07654-HLH-VBK, 2003 WL 21406289, at *3 (C.D. Cal. Mar. 7, 2003). The court rejected the argument that "mere use of a spider to enter a publicly available web site to gather information, without more, is sufficient to fulfill the harm requirement." *Id.* Tickets.Com employed a web crawler to extract factual information (event, date, time, ticket pricing, URL) from the public webpages of Ticketmaster, and then organized the information in their own format to display it on its own page. *Id.* at 2.

[124] *See generally* Intel Corp. v. Hamidi, 71 P.3d 296 (Cal. 2003) (holding that a former Intel employee sending disruptive emails to current employees did not circumvent any techincal security measures or physically damage the computer systems).

[125] *Id.*

[126] *Intel*, 71 P.3d at 300.

[127] "However, the court left open the possibility that a greater interference, perhaps crashing a website's server, may still be an actionable harm under trespass to chattel." Hirschey, *supra* note 4, at 915.

[128] Zamora, *supra* note 33, at 223.

[129] *Register.com*, 356 F. 3d at 444 (affirming that trespass to chattel is still successful against harmful scrapers).

[130] Din, *supra* note 12, at 438.

[131] Intel Corp. v. Hamidi, 71 P.3d 296 (Cal. 2003); *Ticketmaster*, 2003 WL 21406289, at *3. In *Ticketmaster*, the scraper only compiled public pricing data (without a significant load on the server)

considerations than most harmful scraping cases" and do not significantly impact the potential trespass to chattels claim, generally.[132] Consequently, trespass to chattels continues to provide an alternative to situations where CFAA liability cannot be established, especially in cases involving less pervasive scraping that still damages the data host.

There are criticisms against applying trespass to chattels to screen scraping. For instance, Professor Riley argues that claims under this theory involve "fundamental misunderstandings of the subject matter of online property rights."[133] Put differently, the tort's application in this manner implies the existence of real property in cyberspace: a "deeply flawed" analogy considering that computers and servers are chattels, rather than real property. For this reason and others, some commentators like Riley have begun considering copyright as a more promising tool for addressing web scraping cases.[134]

3. Compilation Copyright Infringement. — Data hosts could bring copyright claims against scrapers if the scraped content involved meets the copyright infringement claim's requirements. Web scraping by its nature involves copying, which is a component of copyright infringement.[135] Moreover, copyright claims are appealing due to their availability of substantial damages and a period of copyright protection.[136] For example, in *Craigslist v. 3Taps,* the court held that Craigslist successfully acquired an exclusive license to the copyright in users' advertisements for a short period.[137]

Yet, it can be challenging, as a matter of practice, to establish a copyright claim under the United States law because a database will only benefit from copyright protection if it is "sufficiently creative."[138] A

---

and provided a hyperlink that transferred users directly to the Ticketmaster website for purchase. *See* Din, *supra* note 12, at 438.

[132] Din, *supra* note 12, at 438.

[133] Riley, *supra* note 35, at 286.

[134] *Id.* at 305.

[135] Zamora, *supra* note 33, at 215.

[136] Hirschey, *supra* note 4, at 910–11.

[137] Craigslist Inc. v. 3Taps Inc., 942 F. Supp. 2d 962, 974–76 (N.D. Cal. 2013). This finding has been criticized by commentators as anti-competitive. Eric Goldman, for instance, remarked that: "It's a terrible and anti-competitive practice for a classified advertising website to claim exclusive copyright interests in its advertisers' ad copy. Read literally, advertisers violate Craigslist's copyright interests by displaying their ad copy at any other online publication. Want to simultaneously post a photo of an item for sale on eBay and Craigslist? Craigslist's position is that you would infringe its copyright by doing so." Eric Goldman, *Craigslist's Anti-Consumer Lawsuit Threatens to Break Internet Law*, FORBES (May 23, 2013, 11:50 AM), https://www.forbes.com/sites/ericgoldman/2013/05/23/craigslists-anti-consumer-lawsuit-threatens-to-break-internet-law/?sh=283cba573e39.

[138] Hirschey, *supra* note 4, at 906–07; Feist Publications, Inc. v. Rural Telephone Service Co., Inc., 499 U.S. 340, 344–45 (1991) (quoting Harper & Row, Publishers, Inc. v. Nation Enterprises, 471 U.S. 539, 556 (1985) ("The most fundamental axiom of copyright law is that '[n]o author may copyright his ideas or the facts he narrates'"). *See also* U.S. COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 101, § 308.2 (3rd ed. 2017) ("If the Office determines that a work possesses sufficient creativity, it will register the claim and issue a certificate of registration.").

plaintiff must also be able to assert ownership and negate a fair use defense, for which courts weigh several factors: (1) the purpose and character of the use, (2) the nature of the protected work, (3) the amount of the work used, and (4) the market value of the use.[139] Many commercial data scrapers have successfully asserted a fair use defense in response to copyright claims.[140] Unlike the European Union or United Kingdom, there is no direct legal protection for databases.[141] Overall, copyright claims for data scraping have not had particular success in the United States.[142]

4. Contract and Data Privacy Claims. — Data hosts also argue that scraping is a breach of contract when it has been explicitly prohibited in the websites' terms of use.[143] Facebook,[144] LinkedIn,[145] eBay,[146] Twitter,[147] Craigslist,[148] TripAdvisor,[149] and IMDB[150] have all prohibited scraping in their terms of use.[151] For a breach of contract argument to succeed, the website user must enter an explicit agreement with the data host to comply with these policies.[152] This was the case in *Register.com v.*

---

[139] Zamora, *supra* note 33, at 215–16.

[140] *Id.*, at 216; *see, e.g.*, *Ticketmaster*, 2003 WL 21406289, at *5 (holding that the scraping of the plaintiff's ticket purchasing platform to acquire event information was protected from a copyright claim by the fair use defense. Even though the use was for a commercial purpose and only slightly transformative, only the plaintiff's aggregated non-copyrightable information was put on display, and the defendant's final product did not damage the market value of the plaintiff's product).

[141] Hirschey, *supra* note 4, at 906–07.

[142] Riley, *supra* note 35, at 264, 276; Zamora, *supra* note 33, at 220.

[143] Krotov & Silva, *supra* note 33, at 3.

[144] *Terms of Service*, FACEBOOK (Oct. 22, 2020), https://www.facebook.com/terms.php ("You may not access or collect data from our Products using automated means (without our prior permission) or attempt to access data you do not have permission to access.").

[145] *User Agreement*, LINKEDIN (Aug. 11, 2020), https://www.linkedin.com/legal/user-agreement ("You agree that you will not . . . develop, support or use software, devices, scripts, robots or any other means or processes (including crawlers, browser plugins and add-ons or any other technology) to scrape the Services or otherwise copy profiles and other data from the Services . . . .").

[146] *User Agreement*, EBAY (June 18, 2020) https://www.ebay.com.au/help/policies/member-behaviour-policies/user-agreement?id=4259 ("You agree that you will not use any robot, spider, scraper or other automated means to access the eBay services for any purpose without our express written permission.").

[147] *Terms of Service*, TWITTER (June 18, 2020) https://twitter.com/en/tos ("[S]craping the Services without the prior consent of Twitter is expressly prohibited . . . .").

[148] *Terms of Use*, CRAIGSLIST (Dec. 29, 2017) https://www.craigslist.org/about/terms.of.use/en ("You agree not to copy/collect CL content via robots, spiders, scripts, scrapers, crawlers, or any automated or manual equivalent (e.g., by hand).").

[149] *TripAdvisor Terms, Conditions and Notices*, TRIPADVISOR (Feb. 15, 2018), https://tripadvisor.mediaroom.com/us-terms-of-use ("[Y]ou agree not to . . . access, monitor . . . [or] copy . . . any Content of the Services . . . using any robot, spider, scraper or other automated means or any manual process for any purpose . . . without our express written permission . . . .").

[150] *Conditions of Use*, IMDB (Dec. 3, 2020) https://www.imdb.com/conditions ("You may not use data mining, robots, screen scraping, or similar data gathering and extraction tools on this site, except with our express written consent as noted below.").

[151] Riley, *supra* note 35, at 257–78.

[152] For example, by clicking a checkbox. Krotov & Silva, *supra* note 33, at 3.

Verio, where notice of Register's terms of use bound Verio to the agreement, resulting in Register's successful contract breach claim.[153]

Unlike the United Kingdom, the United States does not have comprehensive data privacy legislation at the federal level.[154] While some privacy-related rights are mandated in state statutes, most do not broadly regulate the collection and use of personal data.[155] However, privacy-related rights are changing significantly in some states: notably, the recent California Consumer Privacy Act. It has implemented restrictions that require companies collecting PII to disclose the "categories of personal information to be collected and the purposes for which the categories of personal information shall be used."[156] Therefore, scraping may be implicated by this new statute, which came into effect in January 2020.[157] These changes may further expose data scraping to additional legal actions in the United States' courts.

B. United Kingdom Approach

Across the Atlantic, laws surrounding web scraping are relatively unclear and untested.[158] Nonetheless, while scraping is not addressed explicitly in most legislation, website owners have attempted to shoe-horn established causes of action into this new area.[159] Depending on the particular circumstances, web scraping could infringe IPRs,[160] breach a contract,[161] violate the Computer Misuse Act,[162] or contravene data protection legislation.[163] In contrast to the United States, the trespass to chattels claim has never been applied to electronic interferences and has

---

[153] *Register.com*, 356 F.3d at 403 ("[The defendant] was offered access to information subject to terms of which [it was] well aware. [Its] choice was either to accept the offer of contract, taking the information subject to the terms of the offer, or, if the terms were not acceptable, to decline to take the benefits."); Zamora, *supra* note 33, at 224.

[154] *See, e.g.*, Paul M. Schwartz, *Preemption and Privacy*, 118 YALE L.J. 902, 930 (2009) (discussing in detail the different paths taken by the European Union and United States in terms of data protection and supporting the argument that "[a]n omnibus federal privacy law would be a dubious proposition . . . .").

[155] Alex Sisto & Herbert Swaniker, *Scraping the Barrel? Legal Issues Arising From Data Scraping*, CLIFFORD CHANCE (Nov. 30, 2018), https://talkingtech.cliffordchance.com/en/data-cyber/data/scraping-the-barrel-.html.

[156] CAL. CONSUMER PRIVACY ACT OF 2018 CIV. CODE §§ 1798.100 & 1798.110 (West 2020).

[157] *Id.*

[158] Steven James, *Screen Scraping and Web Harvesting: The Legal Issues*, E-COMMERCE L. & POL'Y 13, 13 (2011).

[159] *Id.*

[160] Arezou Rezai, *Web Crawling and Screen Scraping—The Legal Position*, PARIS SMITH (Feb. 6, 2017), https://parissmith.co.uk/blog/web-crawling-screen-scraping-legal-position/ (last visited July 9, 2020).

[161] Janet Nikova, *To Scrape or Not to Scrape?*, ROCKET LAWYERS (Oct. 3, 2019) https://www.rocketlawyer.com/gb/en/blog/to-scrape-or-not-to-scrape/.

[162] James, *supra* note 158, at 14.

[163] *Id.* at 13.

scarcely been used in any other context.[164] Furthermore, the inherent differences between the United States' and the United Kingdom's law of trespass may prevent such a development in the future. Trespass to chattels is only actionable in the United States where there has been some "damage," while trespass to chattels in the United Kingdom may be actionable *per se*—that is, even in the absence of damage.[165] Without this damage requirement to limit the action's scope, "cyber-trespass" in the United Kingdom would be applicable to too many digital situations to be practical or useful.[166] As a result, data hosts can rely on other claims like IPR infringement, contract claims, and data privacy actions.[167]

1. IPRs: Copyright and Database Right Infringement. — The most relevant IPRs in this context are copyright and database rights. Scraping may amount to copyright infringement if (1) significant portions of text are scraped, and (2) the text is from a creative source.[168] For example, in *Public Relations Consultants Association v. NLA*[169] the United Kingdom Supreme Court found that scraping headlines from a news website and subsequently hyperlinking them to the original articles amounted to copyright infringement.[170] The Court reasoned that news headlines did require a certain degree of creative input that make them susceptible to copyright suites.[171]

Given these possible hurdles associated with copyright infringement claims, database rights are more likely to protect against data scraping in practice.[172] This *sui generis* right arises from the European Database Directive and has been enacted in the United Kingdom through the Copyright and Rights in Databases Regulations 1997 (CRDR).[173] No "creative" aspect is needed, as database rights automatically subsists if there has been "substantial investment in obtaining, verifying, or presenting the contents" of the database—that is, in searching for material to include in the database, checking such material, and keeping it updated

---

[164] *Id.* at 15; Darren Read, *Should the English Legal System Adopt the US Law of Cyber-trespass?* 8 *SCRIPTED* 46 (2011); Mary W. S. Wong, *Cyber-trespass and 'Unauthorized Access' as Legal Mechanisms of Access Control: Lessons from the US Experience*, 15 INT'L J. L. & INFO. TECH. 90, 91, 94–95 (2007).

[165] Wong, *supra* note 164, at 94 (However, the application of trespass to chattels within the context of web scraping is not settled in the United Kingdom due to the scarcity of cases.).

[166] *Id.*

[167] Nikova, *supra* note 161; James, *supra* note 158, at 13.

[168] Rezai, *supra* note 160.

[169] Public Relations Consultants Association Ltd. v. The Newspaper Licensing Agency Ltd., [2013] 18 UKSC (appeal taken from [2011] EWCA Civ. 890).

[170] Public Relations Consultants Association Ltd. v. The Newspaper Licensing Agency Ltd., [2013]. 18 UKSC 305, 314–16.

[171] Rezai, *supra* note 160.

[172] *Id.*

[173] Copyright and Rights in Databases Regulations 1997, SI 1997/3032, art. 13, ¶ 1 (UK).

over time.[174] Databases can include a collection of profiles on a social website, an individual's blog, and any other collection with systematically-arranged items that are individually accessible.[175] The right is infringed when a person extracts or re-utilizes "all or a substantial part" of a database's contents.[176] Database rights, if they exists in a database, will therefore preclude many forms of unauthorized data scraping.[177] While a fair-dealing provision exists to negate liability, it only applies if the scraper is a lawful database user, who provides attribution or extracts data for a research-related non-commercial purpose.[178] Database rights are thus one of the most common claims brought by data hosts against scrapers.[179] Nevertheless, the Court of Justice of the European Union (CJEU)'s decisions in *British Horseracing Board v. William Hill*[180] and *Football Dataco v. Yahoo!*[181] set a "very high threshold" for the "substantial investment" requirement.[182] Thus, even the CRDR is not a sure protection for data hosts.

2. Contractual Restrictions in the Website's Terms of Use. — Alternatively, scraping could be prohibited with contractual restrictions.[183] If the user agrees to a website's terms of use that include an express limitation on data scraping, but the user then scrapes information, the website owner may be able to make a claim against the user for breach of

---

[174] Maarten Truyens & Patrick Van Eecke, *Legal Aspects of Text Mining*, 30 COMPUT. L. & SEC. REV. 153, 160 (2014). The database right subsists for 15 years from when the making of the database was completed. Copyright and Rights in Databases Regulations 1997, SI 1997/3032, art. 17, ¶ 1.

[175] "Database" is defined as "a collection of independent works, data or other materials which—(a) are arranged in a systematic or methodical way, and (b) are individually accessible by electronic or other means." Copyright and Rights in Databases Regulations 1997, SI 1997/3032, art. 6, ¶ 1.

[176] Copyright and Rights in Databases Regulations 1997, SI 1997/3032, art. 16; Note that 're-utilized' is understood as making the contents of the database available to the public by any means. Andrés Guadamuz & Diane Cabell, *Data Mining in UK Higher Education Institutions: Law and Policy*, 4 QUEEN MARY INTELL. PROP. REV. 3, 11 (2014). Also, Maarten and Van Eeke argued that "Evaluating whether a part is indeed substantial can be performed quantitatively (in relation to the total size of the database) and/or qualitatively (i.e., by measuring the scale of the human, technical or financial investment). Hence, even when only a small part of the entire database is extracted, this may represent a qualitatively substantial part e for example when the affected part constitutes the core part of the database or the part containing the most useful information." *See* Truyens & Van Eeke, *supra* note 174, at 161.

[177] James, *supra* note 158, at 13; Guadamuz & Cabell, *supra* note 172, at 12.

[178] Copyright and Rights in Databases Regulations 1997, SI 1997/3032, art. 20; Truyens & Van Eeke, *supra* note 174, at 161.

[179] James, *supra* note 158, at 13.

[180] C-203/02, British Horseracing Board Ltd. v. William Hill Organization Ltd. 2004 E.C.R. I-10415 (finding that to extend protection to a database, parties must show there has been "qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents . . . .").

[181] C-604/10, Football Dataco Ltd. v. Yahoo! UK Ltd., 2012 E.C.R. 0000 (finding it is irrelevant to consider the intellectual effort and skill that went into creating the original data; the key tenant for protection is whether there is originality expressed in selecting or arranging the data).

[182] Guadamuz & Cabell, *supra* note 172, 12–14; *see* James, *supra* note 158, at 14 (arguing that after the British Horseracing Board decision, website owners will have to demonstrate a substantial investment in presenting and displaying the date, which is a high threshold).

[183] Rezai, *supra* note 160.

contract.[184] In establishing this, the website owner must show that its terms of use are enforceable and have been breached.[185] However, they may struggle to do so because: (1) terms and conditions are often just optional links rather than terms expressly agreed to, and (2) automated scraping bots can simply bypass the terms and conditions, rather than "reading" and understanding them as a human would.[186] Where terms and conditions do not have to be accepted by scrapers, or else are not sufficiently brought to one's attention, it is difficult to establish that a contract has been formed.[187] One possible fix might be requiring each user to agree expressly with the terms and conditions before using the site. Yet, this solution could be commercially impractical and damage the user experience.[188] While there is no clear precedent on whether website terms form binding contracts in the United Kingdom, a 2015 case decided in the CJEU held that screen scraping could be effectively prohibited in a website's terms and conditions.[189] Overall, contract claims could prohibit scraping where a website owner could not otherwise rely on IPRs to protect their data.[190]

3. Contravention of the Computer Misuse Act. — The Computer Misuse Act of 1990 is analogous to the United States' CFAA in targeting hacking. It provides that a person is guilty of a criminal offense if they knowingly cause "a computer to perform any function with intent to secure access to any program or data held in any computer or to enable any such access to be secured" when such access is unauthorized.[191] But as of this writing, the courts have not determined whether data scraping constitutes a breach of this Act.[192]

4. Protection of Personal Data Under the GDPR/United Kingdom Data Protection Act of 2018. — If the information being collected includes

---

[184] *Id.*; Nikova, *supra* note 161.

[185] Nikova, *supra* note 161.

[186] James, *supra* note 158, at 14.

[187] *Id.*

[188] *Id.* at 15.

[189] *See* Case C-30/14, Ryanair Ltd. v. PR Aviation BV, 2015 E.C.R. 0000 (holding that where a database is not subject to Directive 96/9's application, the terms and conditions may function as a contractual limitation on screen scraping).

[190] Nikova, *supra* note 161.

[191] Computer Misuse Act 1990, c. 18, § 1 (UK), https://www.legislation.gov.uk/ukpga/1990/18/section/1; James, *supra* note 158, at 14.

[192] James, *supra* note 158, at 14; Sisto & Swaniker, *supra* note 155. According to practitioner Clare Francis, however, "the business should consider whether their 'screen scraping' of others' content breaches the [Computer Misuse] Act," given the relevant court cases in other parts of European Union. *Unauthorised 'Screen Scraping' May Breach Computer Misuse Act, Says Experts in Wake of Italian Court Ruling*, OUT-LAW NEWS (June 7, 2013), https://www.pinsentmasons.com/out-law/news/unauthorised-screen-scraping-may-breach-computer-misuse-act-says-expert-in-wake-of-italian-court-ruling (last visited Dec. 16, 2020).

"personal data,"[193] then the collector must comply with the European Union's General Data Protection Regulation (GDPR)—as implemented via the United Kingdom's Data Protection Act 2018 (DPA 2018).[194] Under this legislation, screen scraping of personal data is only lawful when done under one of six legal bases.[195] One of these basis is the consent of the data subject, which must be freely given, related to a specific purpose, informed, and unambiguous.[196] Scrapers will experience difficultly demonstrating that they have obtained an individual's consent. Arguably, scrapers may instead attempt to rely on the "legitimate interests" basis for processing, which is assessed concerning purpose, necessity, and balance between interests.[197] Alex Sisto and Herbert Swaniker remarked that this basis "is not a panacea"—rather, it entails considering the interests of the business against those of the individual, taking into account the reasonable expectations of the latter.[198] The purpose of scraping is, therefore, important in determining its legality under the GDPR/DPA 2018.[199] For instance, if a business scrapes data to compile a marketing list that is sold to third parties, it is unlikely that the individuals on that marketing list reasonably expected their personal data to be used in such a way.

Furthermore, even if a valid basis for processing is found, subsequent processing must be limited to that which is fair, proportionate, and necessary.[200] This principle needs careful consideration by businesses intending to scrape websites, because their software usually gathers data in bulk.[201] Therefore, in the context of personal data, it may be "very hard to prove that invisible scraping is fair and transparent," as the GDPR/DPA 2018 is seeking to protect people from "invisible processing."[202]

C. The Australian Approach

Australia currently has no laws that expressly prohibit or address screen scraping.[203] Neither has the practice been addressed to any significant degree by the courts, leading one commentator to remark that

---

[193] European Parliament and Council Regulation 2016/679, art. 4, 2016 O.J. (L 119) 1 (defining "personal data").

[194] *Id.* art. 1, 2; Data Protection Act 2018, c. 12, §§ 1, 2 (UK), https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted.

[195] European Parliament and Council Regulation 2016/679, art. 6, 2016 O.J. (L 119).

[196] *Id.* art. 4(11) (defining "consent"); Rezai, *supra* note 160.

[197] *What is the 'Legitimate Interests' Basis?*, ICO, https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/legitimate-interests/what-is-the-legitimate-interests-basis/ (last visited July 9, 2020).

[198] Sisto & Swaniker, *supra* note 155.

[199] *Id.*

[200] European Parliament and Council Regulation 2016/679, art. 5, 2016 O.J. (L 119).

[201] Sisto & Swaniker, *supra* note 155.

[202] Sisto & Swaniker, *supra* note 155.

[203] Sutton et al., *supra* note 31. Yet, there is one small exception: The Spam Act 2003 prohibits the harvesting/scraping of email addresses from websites (but not any other type of scraping). Spam Act 2003 (Cth) pt 3 (Austl.).

Australia "has not specifically considered web scraping in either a judicial or legislative context."[204] Despite this, several existing frameworks could be utilized to bring a claim against screen scraping.[205] In the absence of database rights—as in the United Kingdom—and any common law precedent on the "hacking" statutory provisions—as in the United States—contract claims' have been called the "first line of defense" against screen scraping in Australia.[206] In contrast to the United States, it is unclear whether trespass to chattels will be expanded to the digital domain by Australian courts.[207] Like the United Kingdom, it does not seem necessary to show damage for this tort to succeed.[208] Showing damage would lead to an overly-broad law if extended to cyberspace.[209] Mary Wong,[210] notes that while trespass to chattels has "experienced something of a renaissance in the US," there has been little judicial activities on this front in other common law countries.[211] We now consider the potential claims in turn.

1. Contractual Restrictions in the Website's Terms of Use. — Where a website's terms of use specifically prohibit screen scraping, the website owner could bring a claim against scrapers for breach of contract.[212] However, this presents the same difficulties discussed above in the United Kingdom context: in addition to expressly prohibiting screen scraping, the terms of use must be considered an enforceable agreement between the website owner and the user.[213] Knowledge and acceptance of the contract terms are a clear pre-condition to the use of the website.[214] By contrast, "browse-wrap" agreements, where the terms of use are available for viewing somewhere on the site, and no active acceptance is required, are much less clear-cut in terms of enforceability.[215] It might be that only click-wrap will allow a claim against a screen scraper for breach of contract.[216] In short, Australian courts have yet to consider the divergence of application in detail.[217]

---

[204] Adrian Agius, *Legal Perspectives on Scraping Data From the Modern Web*, 91 COMPUTERS & L. 9, 11 (2017).

[205] Sutton et al., *supra* note 31.

[206] *Id.*

[207] Trevor Jeffords, *What Is "Screen Scraping" and Is It Lawful in Australia?*, 44 COMPUTERS & L. 24, 24 (2001); Sutton et al., *supra* note 31.

[208] Jeffords, *supra* note 207, at 25.

[209] *Id.*

[210] Mary Wong is the Senior Policy Director at ICANN (The Internet Corporation for Assigned Names and Numbers) and a specialist in copyright, Internet, and international intellectual property law. *Biography, Mary Wong*, ICANN, https://www.icann.org/profiles/366. Prior to joining ICANN, Mary Wong was a tenured law professor at a top-tier intellectual property law school in the United States. *Id.*

[211] Wong, *supra* note 164, at 91.

[212] Sutton et al., *supra* note 31.

[213] *Id.*; Nikova, *supra* note 161.

[214] Sutton et al., *supra* note 31.

[215] *Id.*

[216] *Id.*; Sutton et al., *supra* note 31.

[217] Sutton et al., *supra* note 31.

2. Copyright Infringement. — Alternatively, copyright can exist in website content where there is an "original literary work" that satisfies the requirements of section 32(1) of the *Copyright Act 1968* (Cth).[218] If this is the case, substantially copying data from the website without the authority of its owner may infringe section 3 of the Act. In *Nominet UK v Diverse Internet Pty Ltd*, the respondents used data mining techniques to extract and collate the details of registrants listed on the applicant's databases.[219] The Federal Court found that copyright existed in the databases and that this had been infringed by data mining, with the applicant therefore entitled to declaratory and injunctive relief.[220]

Whether computer-generated work can be protected by copyright law has been controversial in Australian jurisprudence. The *Copyright Act* lacks specific provisions that address the use of digital technologies.[221] This was directly at issue in *Telstra Corp Ltd v Phone Directories Co Pty Ltd*,[222] which dealt with the subsistence of copyright in the White and Yellow Pages telephone directories published by Telstra's subsidiary. At first instance and on appeal to the Full Court of the Federal Court, it was held that the directories were not protected by copyright because they were computer-generated works lacking the requisite human authorship.[223] After the input of data, it was the Genesis Computer System (GCS) that checked the information for accuracy and applied the rules relating to fonts, color schemes, spacing of words and entries, etc. to generate the form of the directories.[224] The court thus concluded that any protectable expression originated from GCS program rather than from any human authors.[225] The court was not persuaded that human supervision of the computer system—in terms of selecting, customizing, and maintaining the program—was as authorial.[226] One judge considered whether human supervision of the automated system could be protected, while another analogized the

---

[218] *Nominet UK v Diverse Internet Pty Ltd* (2004) 63 IPR 543, 570 (Austl.). Such literary works can include compilations. Copyright Act 1968 (Cth) pt 2 s 10(1). A simple compilation of data (*e.g.*, a list of prices) would not generally pass the originality threshold—rather, there must be some reduction of the database to a material form, and some intellectual effort in the creation of that material. *See IceTV Pty Ltd v Nine Network Australia Pty Ltd* (2009) 239 CLR 458 (Austl.).

[219] *Nominet UK v Diverse Internet Pty Ltd* (2004) 63 IPR at 545 (Nominet UK provided a registry database for UK-based Internet domain names, as well as a searchable database derived from this register (the WHOIS Database)).

[220] *Id.* at 576.

[221] Anne Fitzgerald & Tim Seidenspinner, *Copyright and Consumer-Generated Materials – Is It Time to Reboot the Discussion About Authorship?*, 3 VICT. U. L. & JUST. J. 47, 53 (2013).

[222] *Telstra Corporation Ltd v Phone Directories Company Pty Ltd* (2010) 194 FCR 142 (Austl.).

[223] *Id.* at 146. This was an application of the "strong" version of the reasoning in *IceTV*, particularly that of Justice Gummow's judgment. David Lindsay, *Protection of Compilations and Databases after* IceTV*: Authorship, Originality and the Transformation of Australian Copyright Law*, 38 MONASH U. L. REV. 17, 39 (2012).

[224] Fitzgerald & Seidenspinner, *supra* note 221, at 59–60.

[225] Lindsay, *supra* note 223, at 39.

[226] *Telstra Corp Ltd v Phone Directories Co* (2010) 194 FCR at 178–79 (Austl.).

program to "plane with its autopilot engaged… flying itself."[227] The court then held that controlling the software does not necessarily equate to controlling the form, and operating the software is insufficient absent some independent intellectual effort directed to the shape of that material form.[228] Authorship will be "denuded" even if the computer-generated content would have, but for the computer generation, received copyright protection.[229] This "strict and probably undesirable divide" between human-authored and computer-generated works has received considerable criticism,[230] with academics claiming that it is "at odds with long-established precedent,"[231] "runs counter to the principle of technology,"[232] and "imposes an unnecessary technological restriction on the copyright system."[233] Jani McCutcheon,[234] for instance, argues that there is no convincing reason for denying copyright protection to material based solely on its computer generation, given that such material has the same potential to confer the social benefits rewarded by copyright as any other material, and that many computer-generated works are simply too complex for human creation.[235] She also notes many common law countries have introduced provisions protecting a computer-generated work, including the United Kingdom and New Zealand.[236] Authorship may be afforded to the humans responsible for selecting the particular software over other alternatives, which gives effect to a particular desired form and is similar to the skill and effort of selecting extracts for a compilation.[237] This recognizes the practical reality that complex productions do not just create or arrange their material form,[238] is consistent with the copyright policy,[239]

---

[227] *Telstra Corporation Ltd v Phone Directories Company Pty Ltd* (2010) 194 FCR 142 (Austl.). The Full Court, consisting of Justices Keane, Perram, and Yates, rejected the appeal in three independent yet concurring judgments. *Id.* Justice Keane focused on the lower court's finding that the Directories were compiled by the automated processes rather than individuals. *Id.* at 153–54. Similarly, Justice Perram found that the Directories were not original works and analogized the Genesis Computer System (GCS) to an autopilot plane. *Id.* at 176–79. While Justice Yates considered the role of human intervention in the computer system, he also determined that the key activities of selecting, ordering, and arranging the listings were performed by the GCS program rather than human authors. *Id.* at 190–91.

[228] Jani McCutcheon, *The Vanishing Author in Computer-Generated Works: A Critical Analysis of Recent Australian Case Law*, 36 MELB. U. L. REV. 915, 927, 929–33 (2013).

[229] *Id.* at 966–67.

[230] *Id.*

[231] Lindsay, *supra* note 223, at 19.

[232] Fitzgerald & Seidenspinner, *supra* note 221, at 64.

[233] *Id.* at 62.

[234] Jani McCutcheon is an Associate Professor at the University of Western Australia. *Profile: Associate Professor Jani McCutcheon*, U. W. AUSTL., https://www.uwa.edu.au/profile/jani-mccutcheon.

[235] McCutcheon, *supra* note 228, at 955, 957.

[236] *Id.* at 956 n.184 (others include Ireland, Hong Kong, India, and South Africa).

[237] *Id.* at 943 (noting how this is already recognized as conferring originality on compilations).

[238] *Id.* at 942–44.

[239] *Id.* at 953.

and aligns with the realities of how materials are now created in the digital environment.[240]

An issue with relying upon Australian copyright law is that any breach will thus turn on the type of information being scraped, rather than the actual scraping itself.[241] Further, in most cases the legal action is only taken once the scraped data resurfaces, rendering the mechanism reactionary in nature.[242] Conversely, relying on contractual provisions in the terms of use may offer "a more proactive means" of dealing with screen scraping.[243]

3. Contravention of Cybercrime Act and Related State Legislation. — Some types of data are protected under the federal *Cybercrime Act 2001* and similarly worded legislation that sets forth several computer-related offences.[244] One such offense is intentionally "caus[ing] any unauthorised access to… restricted data" with the knowledge that such access is unauthorized.[245] "Restricted data" means data protected by an access control system, like a password.[246] As such, the majority of screen scraping will not breach these provisions, as the data being scraped is not "restricted data"–it is data on publicly available websites not protected by security or access control systems.[247] Unlike the CFAA in the United States, only criminal liability exists under these cybersecurity statutes. One question remaining is whether measures like IP-blocking—illustrated in the United States case *eBay v Bidder's Edge*—constitute an "access control system," so website content would be considered "restricted data" in the Australian context. It remains to be seen how this argument would play out in Australian courts.

4. Protection of "Sensitive Data" Under the Privacy Act. — If web scraping includes the collection of "sensitive information" about an individual such as biometric information, political opinions, or religious beliefs, then Australian Privacy Principle 3 under the *Privacy Act 1988 (Cth)* applies.[248] This provision requires the organization to gather such information that obtains the relevant individual's consent.[249] It would be

---

[240] Fitzgerald & Seidenspinner, *supra* note 221, at 63.

[241] Agius, *supra* note 204.

[242] *Id.*

[243] *Id.*

[244] TONY KRONE, AUSTL. INST. OF CRIMINOLOGY, HACKING OFFENCES, HIGH TECH CRIME BRIEF NO. 5 at 1 (Jan. 1, 2005), https://www.aic.gov.au/sites/default/files/2020-05/htcb005.pdf.

[245] *Cybercrime Act 2001* (Cth) s 478.1 (Austl.); *Crimes Act 1958* (Vict) s 247G (Austl.); *Crimes Act 1900* (NSW) s 308H (Austl.).

[246] *Cybercrime Act 2001* (Cth) s 478.1(3) (Austl.); *Crimes Act 1958* (Vict) s 247G(3) (Austl.); *Crimes Act 1900* (NSW) s 308H(3) (Austl.).

[247] Sutton et al., *supra* note 31; Stephen von Muenster, *Risky Business – Is Screen Scraping Legal?*, ADVERTISING COUNCIL AUSTRALIA (March 18, 2017), https://advertisingcouncil.org.au/news/risky-business-is-screen-scraping-legal/.

[248] *Privacy Act 1988* (Cth) s 6 (Austl.) (defining "sensitive information").

[249] *Id.* at sch 1, pt 2, para 3.3.

infeasible for a scraper to obtain explicit consent from each relevant individual where they are collecting information in bulk.[250] However, they could potentially contend that due to the gathered information's public nature, consent was implicit.[251] It remains to be seen if this would be upheld in court, given the robust protections afforded under the *Privacy Act*.[252]

Some practitioners like Lesley Sutton[253] have illustrated this with the example of Clearview AI, a surveillance start-up scraping three billion images of individuals from third party websites, including Facebook, Google, and LinkedIn, and then using these images to train its surveillance tool.[254] The biometric data from the images fall within the ambit of "sensitive information" under the *Privacy Act 1988*, meaning that the organization shall obtain either explicit or inferred consent from each photographed individual.[255] While explicit consent was not given, Clearview AI could argue that there was inferred consent, given the public nature of the information obtained.[256]

II. SCRAPING NO MORE? ASSESSING THE ROLE OF OPEN BANKING INITIATIVES

In each of the jurisdictions examined, there are, at least in theory, multiple avenues for data hosts to seek relief against scrapers. Some legal options are more difficult to pursue than others, depending on the country. Other remedies have received inconsistent interpretations between the common law jurisdictions, like the term "authorization."[257] Such a divergence can be moderated with the emergence of the "Open Banking" movement.

Starting from the Directive 2015/2366 on Payment Services in the Internal Market—known as PSD II in the European Union—countries across the world have or are contemplating a new framework to govern data sharing among different players in the financial market. These Open Banking schemes require or encourage—depending upon different models taken by each jurisdiction—banks to share consumer-permissioned banking data with third parties securely, in a form, typically through a standardized method of communication that enables data flow between

---

[250] Sutton et al., *supra* note 31.

[251] *See, e.g.*, *id.* (information placed on social media sites).

[252] *Id.*

[253] Lesley Sutton is currently a partner at Gilbert & Tobin's Technology & Digitial group in Sydney, Australia, where she specializes in cybersecurity and data privacy law. *Our People*, GILBERT & TOBIN, https://www.gtlaw.com.au/people/lesley-sutton.

[254] *Id.*

[255] *Id.*

[256] *Id.*

[257] In the United States, the term "authorization," as used in the CFAA, has been interpreted in different ways by different courts.

systems and facilitates its use called application programming interface (API).[258] Given that Open Banking can facilitate data sharing, countries could arguably reduce the demand for screen scraping. This will turn on the degree of "openness" of the Open Banking regime adopted in each jurisdiction.

Each of the three jurisdictions examined herein sits at different points of the spectrum. Although the United States has a relatively large body of disputes on screen scraping in general, it lags behind its common law counterparts and "has a lot of catching up to do" in the context of data-sharing in the financial market, with no legislation and little guidance on the matter.[259] On the other end of the spectrum lies the United Kingdom model—which is based on the European Union PSDII and mandates the use of APIs for the largest banks and placed restrictions on screen scraping by third parties with few exceptions.[260] Somewhere in the middle is Australia's "Consumer Data Right" (CDR), under which banks will be required to provide access to customer data via an API—though unlike the European/United Kingdom model, there is currently no restriction on screen scraping.[261] To have a deeper understanding of Open Banking initiatives' implications for screen scraping, we now examine these regimes in detail.

A. The United States Approach

The United States is not alone in its "market-driven" approach to an Open Banking regime. Singapore, Hong Kong, and New Zealand also leave the adoption of APIs up to banks themselves. However, the United States authorities have so far minimally involved themselves in initiatives to support API development or facilitate a move away from screen scraping

---

[258] Directive (EU) 2015/2366 of the European Parliament and of the Council of 25 November 2015 on Payment Services in the Internal Market, Amending Directives 2002/65/EC, 2009/110/EC and 2013/36/EU and Regulation (EU) No 1093/2010, and Repealing Directive 2007/64/EC, 2015 O.J. (L 337) 35 [hereinafter PSD II]. Open Banking involves third parties (*e.g.*, Fintech) accessing a customer's account using customers' personal security credentials, and in some cases, initiating transactions on their behalf. DEPT. OF THE TREASURY (CTH) (AUSTL.), REVIEW INTO OPEN BANKING: GIVING CUSTOMERS CHOICE, CONVENIENCE AND CONFIDENCE 51 (Dec. 2017) https://treasury.gov.au/sites/default/files/2019-03/Review-into-Open-Banking-_For-web-1.pdf [hereinafter REVIEW INTO OPEN BANKING]. For an overview of different regulatory models, see generally BASEL COMMITTEE ON BANKING SUPERVISION, REPORT ON OPEN BANKING AND APPLICATION PROGRAMMING INTERFACE (Nov. 2019), https://www.bis.org/bcbs/publ/d486.pdf (last visited Dec. 15, 2020).

[259] Steve Boms, *US Way Behind the Curve on Open Banking*, AMERICAN BANKER (Sept. 21, 2018) https://www.americanbanker.com/opinion/us-way-behind-the-curve-on-open-banking.

[260] Commission Delegated Regulation (EU) 2018/389 of 27 November 2017 Supplementing Directive (EU) 2015/2366 of the European Parliament and of the Council with Regard to Regulatory Technical Standards for Strong Customer Authentication and Common and Secure Open Standards of Communication, art. 33, 2018 O.J. (L 69) 23, 39 [hereinafter RTS].

[261] *Competition and Consumer Act 2010* (Cth) s 56BD (Austl.); Joseph Brookes, *Fintechs Get "Screen Scraping" Green Light From Australian Regulators*, WHICH-50 (Mar. 3, 2020), https://which-50.com/fintechs-get-screen-scraping-green-light-from-australian-regulators/.

in the financial market.[262] Commentators have noted that the United States is among the "least likely" of global governments to enact Open Banking regulation, particularly given its more complex regulatory system.[263] The United States has at least eight financial services regulatory bodies, some of which would have to weigh in on such an initiative, compared to just two in the United Kingdom and one in Australia.[264] Others remarked that due to a more competitive retail banking market, the United States lacks one of the significant incentives for Open Banking seen in Europe and Australia.[265] A report released by the United States Treasury in 2018 seems to confirm the view that no federal policy will recommend that APIs be promoted and screen scraping discouraged in "a solution developed by the private sector."[266]

The United States' authorities recognize the need to move away from screen-scraping to more secure access methods. The Treasury's report stated that screen scraping "increases cybersecurity and fraud risks," provides a Fintech application with "significantly more data than needed," and leads to liability issues whereby a bank may be liable for a loss even where screen scraping was used without the bank's knowledge.[267] It also notes that "a significant amount of data is still obtained through screen-scraping."[268] While prescriptive requirements have not been issued, several regulatory bodies and industry initiatives have sought to provide guidance and standardization for API adoption.[269] Among them, the Consumer Financial Protection Bureau (CFPB) has issued guidelines for the access and use of consumer data, consisting of more high-level principles, such as means by which consent should be obtained, rather than the exhaustive detail of the PSD II.[270] Industry groups have created frameworks to create

---

[262] *Open Banking Around the World: Towards Cross-Industry Data Sharing Ecosystem*, DELOITTE, https://www2.deloitte.com/global/en/pages/financial-services/articles/open-banking-around-the-world.html (last visited Aug. 19, 2020).

[263] Susan Pandy, *Developments in Open Banking and APIs: Where Does the US Stand?*, Fed. Reserve Bank of Boston 5 (Brief, March 17, 2020), https://www.bostonfed.org/publications/payment-strategies/developments-in-open-banking-and-apis-where-does-the-us-stand.aspx.

[264] Pandy, *supra* note 263, at 4–5; Boms, *supra* note 259.

[265] Bob Hedges, *Consumer Data in an API-Enabled World*, THE CLEARING HOUSE, https://www.theclearinghouse.org/banking-perspectives/2017/2017-q4-banking-perspectives/articles/open-banking (last visited July 9, 2020).

[266] U.S. DEP'T. OF THE TREASURY, A FINANCIAL SYSTEM THAT CREATES ECONOMIC OPPORTUNITIES NONBANK FINANCIALS, FINTECH, AND INNOVATION 34-35 (2018).

[267] *Id.*

[268] *Id.* at 28.

[269] *Open Banking: US is Next*, PWC FINANCIAL CRIMES UNIT (Apr. 2018), https://www.pwc.com/il/he/bankim/assets/2018/Open%20banking-US%20is%20next.pdf (last visited July 9, 2020) [hereinafter PwC Open Banking].

[270] *Consumer Protection Principles: Consumer-Authorized Financial Data Sharing and Aggregation*, CONSUMER FINANCIAL PROTECTION BUREAU (Oct. 19, 2017), https://iapp.org/media/pdf/resource_center/cfpb_consumer-protection-principles_data-aggregation.pdf; Chris Wood, *How Does Open Banking Apply to US Banks?*, NORDIC APIS (Apr. 2, 2019), https://nordicapis.com/how-does-open-banking-apply-to-us-banks/.

common standards for the use of APIs in the sector.[271] Instead of regulatory pressure, it seems that widespread API use (and the correlative decrease in screen scraping) will be subject to market players voluntarily adopting such standards.[272]

Numerous United States' banks are moving away from screen scraping by developing API-based offerings, in contractual partnerships with third parties, to enhance digital services and gain competitive advantage.[273] Large financial institutions like JP Morgan Chase, Wells Fargo, and Bank of America actively promote API based offerings.[274] However, absent regulatory-driven API standards, screen scraping remains prevalent as a way to offer Fintech services without having to enter contractual agreements.[275] For banks, the significant capital investment required to create and maintain an API may mean that their use is confined to larger banks.[276] This may be partially alleviated by the increasing penetration of technology firms that provide an API on behalf of a bank, thereby facilitating Open Banking and decreasing the use of screen scraping in the United States' "hands-off" approach.[277]

B. The United Kingdom Approach

Upon implementing PSD II in 2018, the United Kingdom became the first nation to offer a government-led Open Banking program.[278] Generally, the PSD II requires that banks provide access to a customer's data for authorized third parties, provided that the customer's explicit consent is obtained.[279] The third-party may then use this access simply to compile account information, or else may initiate a payment from the customer's account.[280] This is intended to benefit consumers by increasing competition, improving security, and facilitating Fintech development.[281]

---

[271] Pandy, *supra* note 263, at 5; Wood, *supra* note 270.

[272] Wood, *supra* note 270.

[273] *Open Banking Around the World*, *supra* note 262; Pandy, *supra* note 263, at 5.

[274] REVIEW INTO OPEN BANKING, *supra* note 258, at 127–28.

[275] *Open Banking Around the World*, *supra* note 262; Hedges, *supra* note 265.

[276] Boms, *supra* note 259.

[277] Wood, *supra* note 270; Pandy, *supra* note 263, at 5; PwC Open Banking, *supra* note 269.

[278] *See* Fernando Zunzunegui, *Digitalisation of Payment Services* 15 (Working Paper Series 5/2018, Ibero-American Institute for Law and Finance, 2018). Implemented in part though the *Retail Banking Market Investigation Order 2017*, issued by Competition and Markets Authority (CMA) (U.K.), and the *Payment Services Regulations 2017* (U.K.). COMPETITION AND MARKETS AUTHORITY, THE RETAIL BANKING MARKET INVESTIGATION ORDER 2017, (UK) [hereinafter UK CMA Order].; The Payment Services Regulations 2017, SI 2017/752 (UK).

[279] PSD II, *supra* note 258, arts. 2, 64, 66–67.

[280] Respectively, known as "Account Information Service Providers" ('AISPs) and "Payment Initiation Service Providers" (PISPs). PSD II, *supra* note 258, arts. 66–67.

[281] *See* Zunzunegui, *supra* note 278, at 27; European Commission Press Release IP/15/5792, European Parliament Adopts European Commission Proposal to Create Safer and More Innovative European Payments (Oct. 8, 2015), https://ec.europa.eu/commission/presscorner/detail/ro/IP_15_5792 (last visited July 9, 2020); Hedges, *supra* note 265.

The PSD II also requires banks and third-party providers (TPPs) to implement various data security controls.[282]

While the PSD II seeks to make screen scraping redundant as more firms begin to use open APIs for data-sharing, the European Union's Directive itself does not expressly prohibit it.[283] Instead, methods of access are regulated in the associated Regulatory Technical Standards (RTS).[284] Banks are required under the RTS to ensure access and prepare an interface for third party providers—either by creating a dedicated API or by modifying their existing interface to enable TPPs to identify themselves (as required under the PSD II).[285] Modifying an existing interface can be seen as screen scraping in a "new, modified form" and has sometimes been referred to as "screen scraping plus."[286] Since September 2019, when the RTS went into effect, TPPs' access to accounts must take an authorized form.[287]

If a bank creates an API for data access, screen scraping by TPPs is usually permitted only where this dedicated interface is unavailable or else not performing to the required standard.[288] This "fallback provision" resulted from the controversy about the role of screen scraping. The Euro Banking Association (EBA)'s original draft prohibited the practice, reasoning that TPPs would violate the obligation to identify themselves and would gain access to information over what was necessary to provide their service.[289] After stakeholders lobbied against this total ban, the European Commission introduced the fallback provision into a later draft

---

[282] Such third parties are either "Payment Initiation Service Providers" (PISPs) or "Account Information Service Providers" (AISPs), collectively known as "third-party providers" (TPPs). For example, parties shall have adequate internal control mechanisms, risk management procedures, and incident response measures. *See* PSD II, *supra* note 258, arts. 5, 95 and 96; The Payment Services Regulations 2009, SI2009/209 (UK), arts. 5, 98, 99, sched 2. "Strong customer authentication" must also be used for data access by third parties. PSD II, art. 97; The Payment Services Regulations 2017, SI 2017/752, art. 100 (UK) [hereinafter "PSR"].

[283] REVIEW INTO OPEN BANKING, *supra* note 258, at 125–26.

[284] *See* RTS, supra note 260, arts. 30–36.

[285] In the UK, the nine largest banks (CMA9) are obliged to prepare an open API rather than merely rely on "screen scraping plus." These banks are listed in the CMA Order, including Barclays, HSBC, Lloyds, Nationwide Building Society, the Royal Bank of Scotland Group (including NatWest and Ulster Bank), Bank of Ireland (UK), AIB Group UK, Santander and Danske. *See* UK CMA Order, *supra* note 278, art. 3.1.1; RTS, *supra* note 260, art. 31; *see also* PSD II, *supra* note 258, arts. 66, 67 (a third-party provider is obliged to identify itself).

[286] Adam Polanowski & Przemyslaw Gruchala, *Can a User's Account be Accessed Through Screen Scraping*? NEWTECH LAW (Mar. 15, 2019), https://newtech.law/en/can-a-users-account-be-accessed-through-screen-scraping/ (last visited July 9, 2020).

[287] RTS, *supra* note 260, art. 38; Zunzunegui, *supra* note 278, at 29.

[288] RTS, *supra* note 260, arts. 32–33 (dedicated interfaces must "offer at all times the same availability and performance, including support, as the interfaces made available to the payment service user.").

[289] P.T.J. Wolters & B.P.F. Jacobs, *The Security of Access to Accounts Under the PSD2*, 35 COMPUT. L. & SEC. REV. 29, 36 (2019).

of the RTS.[290] The EBA objected to its inclusion, arguing that banks would be forced to maintain both an API and an interface allowing "screen scraping plus," increasing costs for new providers.[291] These objections were tacitly acknowledged by the inclusion of an exemption clause.[292] Where a bank has implemented a compliant, stress-tested, and widely-used API, it may be exempted by authorities such that it is not required to allow screen scraping as a fallback option.[293] When accessing the data held by these banks, TPPs are therefore not permitted to use screen scraping under any circumstances.[294]

More recently, the United Kingdom launched the "Smart Data" initiative, which seeks to "give consumers in regulated markets the ability to safely, securely and instantly transfer their data" to third parties to facilitate cross-sector innovations.[295] Although it remains to be seen whether the Smart Data initiative will introduce similar bans on screen scraping,[296] the fact that it enables free data flows across entities would certainly help reduce the need of such techniques, thus reducing controversy.

In short, TPPs may legitimately employ "screen scraping plus," which identifies the TPP and therefore complies with PSD II requirements where a bank modifies their existing interface for this purpose rather than creating an API.[297] Where the bank instead creates an API for data access, screen scraping can only be conducted in narrow circumstances, specifically, where the API is not performing to the required standard.[298] The legality of screen scraping is even further restricted where a bank has implemented a compliant, stress-tested, and widely-used API. In such cases, an exemption to the fallback provision can be provided by the Financial Conduct Authority (FCA), ensuring that accessing bank-held data via screen scraping will always be prohibited.[299]

---

[290] *Id.*

[291] *Id.*

[292] *Id.*

[293] RTS, *supra* note 260, art. 33; Euro Banking Authority, Guidelines on the Conditions to Benefit from an Exemption from the Contingency Mechanism under Article 33(6) of Regulation (EU) 2018/389 (RTS on SCA & CSC), at 3, EBA/GL/2018/07 (Dec. 4, 2018).

[294] Wolters & Jacobs, *supra* note 289.

[295] DEP'T. FOR BUS., ENERGY & INDUS. STRATEGY, NEXT STEPS FOR SMART DATA: PUTTING CONSUMERS AND SMES IN CONTROL OF THEIR DATA AND ENABLING INNOVATION, 2020, at 11 (UK). More recently, the European Union proposed the "Data Governance Act," which features similar concepts, setting out a framework for data sharing service providers. *See* Data Governance Act, European Commission (Nov. 25, 2020), https://ec.europa.eu/digital-single-market/en/news/data-governance-act.

[296] *Id.* at 11 (noting however that screen scraping is a "riskier" method of data sharing).

[297] RTS, *supra* note 260, arts. 30, 31; *see also* PSD II, *supra* note 258, arts. 66, 67.

[298] RTS, *supra* note 260, art. 33.

[299] *Id.*; Wolters & Jacobs, *supra* note 289.

C. Australian Approach

While the European Union and United Kingdom pioneered the prescriptive approach, Australia has since adopted its own comprehensive Open Banking regime as part of the broader Consumer Data Right (CDR).[300] It requires that the largest banks provide "accredited recipients" with access to customer data (upon that customer's request) by July 2020.[301] This requirement will eventually extend to all banks and eventually apply to other sectors of the economy.[302] The CDR implements similar security measures to those in the European Union/United Kingdom framework and creates its privacy protection mechanism.[303]

Yet, despite the CDR's similarity with its European Union/United Kingdom counterpart, screen scraping's legality is less evident in Australia, where legislation remains silent on the issue. Rather than prohibiting or endorsing the practice, a government review recommended that Open Banking should aim to make it redundant by facilitating more efficient data transfer mechanisms.[304] More recently, the Australian Securities and Investments Commission (ASIC) expressed that it has no intention to prevent screen scraping.[305]

While screen scraping financial accounts appears unrestricted, there is some uncertainty about liability. In contrast to the European Union/United Kingdom's liability framework for PSD II, which shifts the burden to service providers and requires that consumers receive a refund except in limited circumstances,[306] Australia does not yet have a specific regime in place to allocate liability in the Open Banking ecosystem.[307] The more general *ePayments Code* may find it challenging to accommodate screen scraping practices. In providing their login details to a TPP for

---

[300] *Treasury Laws Amendment (Consumer Data Right) Act 2019* (Cth) (Austl.). This Act amended the *Competition and Consumer Act 2010* (Cth) (Austl.), *Australian Information Commissioner Act 2010* (Cth) (Austl.), and the *Privacy Act 1988* (Cth) (Austl.) to create the Consumer Data Right. *Id.* at sch 1.

[301] *Competition and Consumer Act 2010* (Cth) s 56BD (Austl.); Press Release, Australian Competition and Consumer Commission, Consumer Data Right Timeline Update (Dec. 20, 2019), https://www.accc.gov.au/media-release/consumer-data-right-timeline-update; Press Release, Australian Competition and Consumer Commission, Consumer Data Right Goes Live for Data Sharing (July 1, 2020), https://www.accc.gov.au/media-release/consumer-data-right-goes-live-for-data-sharing.

[302] Explanatory Memorandum, Treasury Laws Amendment (Consumer Data Right) Bill 2019 (Cth) 7–8 (Austl.) [hereinafter CDR Bill Explanatory Memorandum].

[303] *Compare* PSD II *supra* note 258, arts. 95–97 *and* RTS, *supra* note 260, art. 35, *with Competition and Consumer (Consumer Data Right) Rules 2020* (Cth.) sch 1 pt 1.7, sch 2 pt 2.2 (Austl.) (requiring "strong customer authorization," secure network and systems with encryption techniques, and implement internal control mechanisms and incident management).

[304] REVIEW INTO OPEN BANKING, *supra* note 258, at 133–35.

[305] Brookes, *supra* note 261.

[306] PSD II, *supra* note 258, art. 73; PSR, *supra* note 282, art. 7, ¶ 76; *Trust in Open Banking: Negotiating Data Liability Between Banks and TPPs*, FINEXTRA (Nov. 22, 2019), https://www.finextra.com/newsarticle/34820/trust-in-open-banking-negotiating-data-liability-between-banks-and-tpps.

[307] REVIEW INTO OPEN BANKING, *supra* note 258, at 65.

screen scraping, a consumer could be in breach of security requirements and potentially lose their protection under the Code, thus becoming liable for any losses that occur.[308] This was identified as an issue in the government's Review Into Open Banking, which observed that it is "debatable whether all customers are aware of precisely what they've done in providing their login details in this way," with the style of some requests ensuring that "customers might not even be aware they have given their login details to someone other than their bank."[309]

Given such issues and the mandate on banks to provide data access, consumer groups have made arguments in favor of banning screen scraping. They argue that such practices run counter to every other piece of government security advice,[310] undermine the consumer data right's goals,[311] and could result in consumer liability for loss.[312] Further, it would provide little incentive for some Fintech players to seek accreditation if they could instead rely on screen scraping with financially vulnerable people, thus continuing to engage with non-CDR accredited entities bound by lower privacy protections.[313] This could potentially create a "two-tiered" Fintech system and undermine the CDR regime's success in ensuring consumer protection and confidence.[314] On the other hand, Fintech groups have pointed to the possible anti-competitive effects associated with a screen scraping ban, with such a ban only seeming feasible when the CDR regime has matured.[315] This is especially so considering that the Australian economy heavily relies on screen scraping as a cost-effective tool.[316] The government review has also noted its role as an "important market-based check" on the design of the CDR framework.[317]

Overall, while it seems like screen scraping is legal as a technique running parallel to the CDR scheme, it is controversial and could be subject

---

[308] *Id.* at 51; *ePayments Code* 2016 (Cth) ch C cl 11.2, 12 (Austl.).

[309] REVIEW INTO OPEN BANKING, *supra* note 258, at 52.

[310] Karen Cox & Gerard Brody, Financial Rights Legal Centre and the Consumer Action Law Centre, Submission No 36 to Senate Select Committee on Financial Technology and Regulatory Technology, Parliament of Australia, *Inquiry into the FinTech and RegTech Sectors* (Dec. 2019) 15, https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Financial_Technology_and_Regulatory_Technology/FinancialRegulatoryTech/Submissions.

[311] *Id.* at 16.

[312] *Id.* at 14.

[313] *Id.* at 16–17.

[314] *Id.*

[315] FinTech Australia, Submission No 19 to Senate Select Committee on Financial Technology and Regulatory Technology, Parliament of Australia, *Inquiry into the FinTech and RegTech Sectors* (Dec. 2019), https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Financial_Technology_and_Regulatory_Technology/FinancialRegulatoryTech/Submissions [hereinafter FinTech Australia, Submission No 19].

[316] FinTech Australia, Submission to Treasury of Australia, *Review into Open Banking in Australia* (Sep. 2017) 6, https://treasury.gov.au/sites/default/files/2019-03/c2017-t224510_FinTech_2.pdf.

[317] REVIEW INTO OPEN BANKING, *supra* note 258, at 83–84.

to change, with various stakeholders arguing for or against a ban.[318] There is also uncertainty as to liability associated with the practice.[319]

CONCLUSION

Screen scraping has emerged as a major legal battlefield in the age of big data. Built upon the comparative approach, this article maps out the trajectory of relevant laws and jurisprudence in the United States, United Kingdom, and Australia and critically examines and demonstrates the nuanced divergence in how screen scraping could be treated in five issue areas—"digital trespass" statutes, tort law, intellectual property, contractual rights, and privacy and data protection—differently in these three common law jurisdictions. Despite the divergences, the recent Open Banking movement could help moderate the concerns about screen scraping, thereby bringing some level of convergence. To the extent that Open Banking mandates or facilitates data sharing, it can reduce the need for screen scraping. This is especially so in the United Kingdom/European Union context—and even more so if the United Kingdom's Smart Data initiative expands these data-sharing principles beyond the financial sector.

By contrast, it is much less clear in the United States, for it lags in terms of Open Banking. Australia sits somewhere in the middle: it has a comprehensive CDR regime that could theoretically reduce the need for screen scraping but given the fact that it imposes no ban on the practice, it has a loophole for data miners to work around the new regime and continue scraping data.

In short, with the emerging trend of data sharing, one could witness a sea change in the screen scraping legal landscape. Insofar as data sharing schemes enable information flow between entities, one would expect some level of convergence. Such a convergence, however, is qualified by the institutional design of data sharing schemes—whether or not it explicitly

---

[318] FinTech Australia, Submission No 19, supra note 316, at 35 ("[T]he government should acknowledge that screen scraping provides a secure, economical, accessible and accepted system by which fintechs can and do seek information."); Raiz Invest, Submission No 29 to Senate Select Committee on Financial Technology and Regulatory Technology, Parliament of Australia, *Inquiry into the FinTech and RegTech Sectors* (Dec. 24, 2019) 5, https://www.aph.gov.au/DocumentStore.ashx?id=2bc13fcf-d9d4-4258-922c-563068d8092b&subId=676435 ("[S]creen scraping needs to remain a valid data collection method for the foreseeable future for many reasons from reducing fraud in the BECS system, to facilitating start-ups delivering products to the Australian market, and providing an alternative to Open Banking to level the competitive issues created by the CDR."); Steve Brown, Submission No 13 to Senate Select Committee on Financial Technology and Regulatory Technology, Parliament of Australia, *Inquiry into the FinTech and RegTech Sectors* (Dec. 19, 2019) 2, https://www.aph.gov.au/DocumentStore.ashx?id=ddbd95ba-88a0-4693-885e-935b27991048&subId=675308 ("DDC is a critical mechanism to empower consumers and facilitate competition, valued by consumers, is secure and cost-effective, and is making a significant contribution to the competitive dynamics in the current market.").

[319] *See* Brookes, *supra* note 261 (The Australian Securities and Investments Commission's acting executive director of financial services recently told a Senate Committee that there is "no evidence of which we're aware of any consumer loss from screen scraping . . . .").

addresses screen scraping (as in the case of Australia and the United Kingdom) and whether there is a top-down, government-mandated data sharing regime (as in the case of the United States).