Publicly Available Information and its Protection on the Internet in Light of *hiQ Labs v. LinkedIn Corporation*

ABE Junko

Introduction

On 29th July 2020, Japan's Personal Information Protection Commission (PPC)[1] published a statement about its administrative action against two businesses which were alleged to be illegally processing personal

1) The Personal Information Protection Commission (PPC) is an independent organization with a council system comprised of one chairperson and eight members, with the role of ensuring proper processing of personal information taking into account its utility, founded based on the Act of Protection of Personal Information (APPI) (Act No. 57 of 2003). The PPC functions pursuant to APPI and the Act on the Use of Numbers to Identify a Specific Individual in Administrative Procedures (Act No. 27 of 2013). In Japan, statutes of personal information protection are divided between private and public sectors, and the PPC has oversight over the former, the latter, the Act on the Protection of Personal Information Held by Administrative Organs (Act No. 58 of 2003) and the Act on the Protection of Personal Information Retained by Independent Administrative Institutions (Act No. 59 of 2003) are governed by the Minister of the Ministry of Internal Affairs and Communications. In August 2020, a

information on their websites [2]. On these websites, personal information relating to bankruptcy is disclosed without consent from the individuals named, and the purpose of this use of the data is not notified to those individuals or published publicly. This exposure of personal data on the internet violates the Act on the Protection of Personal Information (APPI), which requires business operators to give notice to relevant individuals at the time of data acquisition and to obtain their consent when they provide the personal information to third parties; the PPC issued an order to the businesses to suspend immediately the websites based on Article 42 (2) of APPI [3].

The text of the Act on the Protection of Personal Information explicitly sets out the importance of the utility of personal information as well as individuals' rights and interests [4]. It is generally agreed that the aim of this

---

taskforce to review the Japanese personal information protection system established within the Cabinet Secretariat published its interim statement that the current divided oversight areas shall be integrated into one statute all under the oversight areas of the PPC, which will be more reinforced as an integrated authority for the protection of personal information. *See* the taskforce's statement on the website of the Cabinet Office (posted in August 2020):
https://www.cas.go.jp/jp/seisaku/kojinjyoho_hogo/ (Last visited on 11th September 2020).

2) Personal Information Protection Commission (news press posted on 29th July 2020): https://www.ppc.go.jp/news/press/2020/200729kouhou/ (Last visited on 11th September 2020).

3) The orders were made by "public services" since the places of both businesses are unknown. As for the public services, *see* article 98 of Civil Code (Act No. 89 of 1896) (Minpō), and Code of Civil Procedure (Act No. 109 of 1996) at Article 110 *et seq.*

4) *See* art 1 of the Act on the Protection of Personal Information.

act is primarily to protect an individual's rights and interests rather than to promote the utility of personal information including "the creation of new industries and the realization of a vibrant economic society and an enriched quality of life for the people of Japan" [5]. It is true that the free flow of personal data is in the public interest, but this needs to be balanced against individuals' interests and rights.

It should be noted that, according to the PPC's statement, the information itself was already published in an official gazette (paper and online alike), and that the business operators collected and used the information based on the gazette.

A key question is therefore whether such "publicly available" information should be protected against later publication. Should there be no rational privacy expectation in relation to such a publicly available information due to its first publication? The real world in today's digital era can make this question more complex. Indeed, it seems that publicly viewable information on the internet should still be protected from later insensitive exposure, but how, why and to what types of information should this protection apply?

Following the case of *hiQ Labs., v., LinkedIn Corporation*, this paper examines problems relating to and presented by the internet, especially those relevant to the Computer Fraud Abuse Act (CFAA) and the First Amendment of the Constitution of the United States of America, and offers consideration of the protection of the publicly available information on the

5) Katsuya Uga, Legal System of Personal Information Protection 83－84 (Yuhikaku, 2019).

internet in Japan[6].

1 The Internet as a "Modern Public Square"

1.1 The Internet-Access to Free Speech

The internet is a useful tool to access and exchange information. In the United States Supreme Court, the cyberspace was referred to as the "vast democratic forums of the Internet" in general[7], and social media in particular. In *Packingham v. North Carolina*[8], for example, Justice Kennedy, who delivered the opinion of the Court, characterized the internet as the "modern public square".

In this case, the North Carolina statute making it a felony for a registered sex offender "to access a commercial social networking website where the sex offender knows that the site permits minor children to become members to create or maintain personal web pages"[9] was in issue. Considering the constitutionality of the statute in light of the Free Speech Clause of the First Amendment[10], Justice Kennedy concluded it as invalid.

He mentioned this case as one of the first to deal with the relationship between the First Amendment and the modern internet, and went on to pay attention to the extreme care of the access to social media when

---

6) Various problems are concerned in this point, and this paper does not argue about the Fourth Amendment problems, libel or slander, or copyrights law like "Digital Millennium Copyright Act".

7) Reno v. American Civil Liberties Union, 521 U.S. 844, 868 (1997).

8) Packingham v. North Carolina, 582 U.S. __, 137 S.Ct. 1730 (2017).

9) N. C. Gen. Stat. Ann. §§14－202.5 (a), (e) (2015). The definition of the "commercial social networking web site" therein is provided in §14−202.5 (b).

10) U.S. Const. amend. I.

analyzing the scope of the protection by the First Amendment. While it may be obvious that the purpose of the legislature is to protect minors and (potential) sexual victims from abuse, according to Justice Kennedy, it is still necessary to take the rights of the First Amendment into account carefully in the social media. Even convicted criminals might have legitimate interests in access to "the world of ideas", he said [11]. After all, Justice Kennedy pointed out that fully excluding access to the social media is inconsistent with the users' legitimate exercise of their First Amendment rights. The State's interest to keep sex offenders away from potential vulnerable victims was not considered to be sufficient to discharge its burden to show the necessity or legitimacy of the legislation.

The United States Supreme Court recognized the necessity to take special caution when it comes to the scope of the First Amendment in *Packingham*. This case reflects the fact that open access to the internet in this modern era is of significance in the context of the First Amendment [12], and the internet can be "the world that all may enter without privilege or prejudice accorded by race, economic power, military force, or station of birth" and a "place" where anyone may express his/her own beliefs without fear of being forced into silence or conformity [13].

---

11) However, Justice Kennedy added that this opinion should not be interpreted in a way to ban a State from making "more specific" laws than the one at issue in this case.

12) Jamie L. Williams, *Automation is not "Hacking": Why Courts Must Reject Attempts to Use the CFAA as an Anti-Competitive Sword*, 24 B.U. J. Sci. & Tech. L. 416, 416–417 (2018).

13) John Perry Barlow, A Declaration of the Independence of Cyberspace,

On the other hand, it is still uncertain how much free speech on the internet should be protected by the First Amendment in the courts. A few years after from *Packingham*, the Pennsylvania Supreme Court stated in *Carr v. Commonwealth Department of Transportation* [14] that the interests of the employee of the state department did not outweigh those of the department to terminate its employment contract based on the "rant" she posted on a closed Facebook group [15].

Justice Mundy, stating the opinion of the Court, considered the question "whether the speech of the employee, Rachel Carr, could reasonably be said to adversely affect the state Department's interest as an employer". In light of the precedents of *Sacks* [16] and *Pickering* [17], Justice Mundy decided that Carr's speech about the safety of a particular bus driver does not outweigh the Department's interest in the safety of the public. While *Pickering* and

---

Electronic Frontier Found (posted on 8th February 1996): https://www.eff.org/cyberspace-independence (Last visited on 11th of September 2020).

14) Carrv.Pa. Dep't of Transp., 3 MAP 2019; 200 A.3d 435 (Pa. 2019).

15) Rachel Carr was hired by the appellant, the Department of Transportation (hereinafter "Department") as a seasonal/non-permanent employee and started to work in March 2016. When she was off duty at home, she posted a rant about school bus drivers to say they broke traffic laws on her Facebook account, which was available only for the closed Facebook group ("Creeps and Peeps"). On Carr's Facebook profile, her employee status was identified as a Roadway Programs Technician employed by the Department. Three members of the private group forwarded the Carr's comment of complaints to the Department's Facebook site, and it was then sent to human resources office of the Department. Subsequently, she was fired due to her inappropriate behavior on Facebook.

16) Sacks v. Dep't of Pub. Welfare, 465 A.2d 981 (Pa. 1983).

17) Pickering v. Bd. of Educ. of Twp. High Sch. Dist., 391 U.S. 563 (1968).

Sacks both concerned employees who had specialized knowledge about matters of public concern from their positions and experiences, Justice Mundy said that Carr's comment was not related to a matter of public concern. Instead, Carr's complaint was based on her personal observation of a particular bus driver, indifferent from an explanation of safety concern as a Department employee.

Since the Carr's post was of limited public interest and was harmful to the Department, Justice Mundy reversed the Commonwealth Court order supporting Carr's argument [18].

The question raised in *Bartnicki v. Vopper* [19] was whether the First Amendment protects the disclosure of the contents of illegally intercepted communications. The title Ⅲ of the Omnibus Crime Control and Safe Streets Acts of 1968 [20], as amended [21], bans the interception of wire, electronic, and

---

18) If such a dismissal by the public authority can be legitimate based on a private post on social media, it could be also justifiably argued that such authorities have legitimate interests to monitor employees' private life on the internet. It may be that the possible risk resulting from this monitoring should have been considered in the balancing test between the Department and the employee in this case.

19) Bartnicki v. Vopper, 532 U.S. 514 (2001).

20) 18 U.S.C. §§2510-22.

21) As amended by the Electronic Communications Privacy Act ("ECPA") (Pub. L. 99-508), the Communications Assistance to Law Enforcement Act ("CALEA") (Pub. L. 103-414), Antiterrorism and Effective Death Penalty Act of 1996 ("Antiterrorism Act") (Pub. L. 104-132), USA PATRIOT Act (Pub. L. 107-56), USA PATRIOT Additional Reauthorization Amendments Act of 2006) (Pub. L. 109-178), FISA (Foreign Intelligence Surveillance Act) Amendments Act of 2008 (Pub. L.110-261), FISA Sunsets Extension Act (Pub.

oral communications in general.

The provision applied to any person who willfully intercepts such communications [22] and to any person who, knowing or having reason to know that the information was obtained through illegal interception, intentionally discloses its contents [23].

The Court accepted that the respondents had reasons to know that the interception was illegal, and that the disclosure itself broke the statute. However, it was found by the Court that the respondents did not contribute to the illegal interception, that they obtained the contents information lawfully, and that the communications were related to the matter of public concern [24]. The precedents left open the question of whether the government should punish the disclosure if those who disclose information acquire it lawfully from a source who obtains it unlawfully.

L. 112-3) PATRIOT Sunsets Extension Act of 2011 (Pub. L. 112-14). These statutes are codified, *inter alia*, at 18 U.S.C. §2510, *et seq.*

22) 18 U.S.C. §2511 (1)(a).

23) 18 U.S.C. §2511 (1)(c). In Bartnicki, the Pennsylvania State Education Association, a union representing the high school teachers, had collective-bargaining negotiations with the school board, and the petitioner Batrtnicki as the union's chief negotiator talked with the president of the local union on the cellular phone about the status of the negotiation including the timing of a proposed strike. Then, this call was intercepted unlawfully by an unidentified person, and respondent Vopper, who was a radio commentator, played a tape of the intercepted conversations on his radio program.

24) The United States Supreme Court recognized the press' right to publish information which dealt with great public concern but obtained through stolen documents by a third party. New York Times Co. v. United States, 403 U.S. 713 (1971). The Court did not focus on the fact that the documents were stolen, but the nature of the documents of public concern.

In *Bartnicki*, Justice Stevens, delivering the opinion of the Court, examined whether the disclosure of the information originally obtained illegally should be restricted based on the statutes at issue. Justice Stevens states this question as "both novel and narrow", and that this was the first time the Supreme Court of the United States had had to deal with it. As for the question of whether the application of the statutes to the circumstances in this case violated the First Amendment, while considering the fact that the respondents did not contribute to the unlawful interception, that they obtained the tapes lawfully, and that the conversations' subject matter was a matter of public concern, and also the basic purposes of the statutes as protecting privacy of wire, electronic, and oral communications, Justice Stevens held that in balancing act in this case, privacy matters gave way to the interests of disclosing matters of public importance. Even though the conversations were intercepted unlawfully, the fact that those who discloses the information obtained it legally should be important, because Justice Stevens held that disclosing and publishing information were a kind of "speech" which the First Amendment protects, and that the privacy interests of the statutes in question cannot justify their restrictions on free speech in this case.

If the internet is thought of as a modern public square, one may argue for public employees to express their personal opinions on matters of public concern on social media such as Facebook or Twitter without fear of spreading of them, especially based on *Packingham* [25]. For an integrated understanding of these cases, it may be necessary to consider whether

25) *See* Electronic Privacy Information Center (posted on 21st May 2020): https://epic.org/2020/05/pa-supreme-court-says-state-ca.html (Last visited on 11th of September 2020).

accessing information has a different meaning from expressing some contents on the internet. What kind of legal protection, then, should be granted for the access to the internet?

1.2 Open Access on the Internet v. CFAA?

One of the main reasons that *Packingham* regarded the internet as "modern public square" was the open access to information it affords [26]. Open access has become more important as the internet has been an ever-growing data source which is said to be the largest on the planet, and the resources are critical for various areas including journalists, academics, businesses and people's daily lives. On the other hand, it is also pointed out that open access is endangered by recent companies' efforts to prevent their competitors from using automated scripts to access "publicly available" information on their websites by using the Computer Fraud Abuse Act ("CFAA") [27]. One feature of CFAA is that it makes it a crime to access another person's computer "without authorization" and the intention of Congress to enact the statute was to criminalize malicious break-ins of "private" computer systems [28]. If a company tries to apply CFAA to the public available information on the open web in order to block its competitors, then it can impair benefits from open access on the internet, and this can harm the open web [29]. But how can you apply CFAA, which was to protect "private" computer system, to "publicly available" information on the open web? What values "to be protected" does such information have on the internet as "modern public square"? For this question, the

---

26) Williams, *supra* note 12, at 416-417.

27) 18 U.S.C. § 1030 (2012); *see also id.*, at 417.

28) Williams, *supra* note 12, at 417.

29) *Id.*, at 417-418.

interpretation of terms "without authorization" under CFAA is key.

Among many data protection laws in the United States, CFAA is one of the most notable and controversial [30]. CFAA was established in 1986, and while its original aim was to punish people who gain access to federal computer without authorization, now it is viewed as a computer trespass law [31]. One of the most problematic points is that Congress has not defined the term "without authorization", and it has been quite difficult for the courts to formulate a uniform understanding of what types of access are to be interpreted as, "without authorization", under CFAA [32]. As §1030 (a)(2)(c), which is regarded as the most controversial part of the act, provides that anyone who intentionally accesses a computer without authorization, or exceeds authorized access to obtain information from any protected computer, is liable, the scope of CFAA can be, in fact, based on the interpretation of "authorization" or "authorized access" by the Court [33]. Except for some easy cases such as hacking, which is clearly considered unauthorized access, in deciding whether it access is with "authorization", it is necessary to consider who are accessing the data, and whose data has been accessed [34].

---

30) Brent W. McDonough, *Data Scraping and the Computer Fraud and Abuse Act: An Analysis of When Unwanted Digital Access Should Implicate an "Antihacking" Statute*, 2019 Rutgers U.L. Rev. Comments 62, 62–63 (2019).

31) *Id.*, at 63. CFAA has been amended as efforts to catch up with the development of the internet.

32) *Id.*

33) *Id.*, at 63–64.

34) *Id.*, at 64.

It is also important to consider who can allow or withdraw access to data on a website [35]. Given that almost all internet services are constructed based on someone else's computer system, accessing the website always involves accessing a computer owned by someone else [36]. If owners of computers can use CFAA to monitor those who may access data on websites and how those users access such data, this can be risky for open access to publicly available data on the web [37]. When arguing about the scope of CFAA, it is important to remember to consider what the internet society can bring now and in the future [38], the maintenance of open access should be valuable in this sense [39].

In *LVRC Holdings LLC v. Brekka* [40], where the Ninth Circuit reviewed a claim by an employer saying that a former employee accessed a work computer "without authorization", the court described the meaning of "authorization" under 18 U.S.C. § 1030 (a)(2) as "permission or power granted by an authority" [41]. In this case, the former employee was authorized to use the computer while he was employed at LVRC, and then accessed the work computer to send the documents to his personal email in order to use them in a future competing business.

The court went on to state that the plain language of "authorization"

---

35) Williams, *supra* note 12, at 420.

36) *Id.*

37) *Id.*, at 420-421.

38) Michael J. Madison, *Authority and Authors and Codes*, 84 GEO. WASH. L. REV. 1616, 1620 (2016).

39) Williams, *supra* note at 421.

40) LVRC Holdings LLC v. Brekka, 581 F.3d 1127 (9th Cir. 2009).

41) Brekka, 581 F.3d at 1133.

under the statute may depend on actions taken by the employer, and that the employee did not access a computer "without authorization" because he had the authorization to use the computer by the employer [42].

In *United States v. Nosal (Nosal II)* [43], the court developed the meaning of "accessing a protected computer without permission" to find that "once authorization to access a computer has been affirmatively revoked, the user cannot sidestep the statute by going through the back door and accessing the computer through a third party".

The court held that CFAA was violated under the "without authorization" orovision of 18 U.S.C. § 1030 (a) (4) when a former employee used a current employee's credentials to access the company's internal database.

In this case, David Nosal, a former employee of the global executive search firm, Korn/Ferry International, as a high-level regional director, had a plan to leave the company after he was passed over for a promotion and to launch a competitor with associates. When he was still an employee at

---

42) Brekka, 581 F.3d at 1134–1135.

43) United States v. Nosal, 844 F.3d 1024 (9th Cir. 2016). This was the second time for the Ninth Circuit to examine the scope of CFAA (18 U.S.C. § 1030) for David Nosal. In the second case (Nosal II), the court considered the meaning of "knowingly and with intent to defraud *accessing a computer without authorization*" under the first prong of § 1030 (a) (4). Finding the distinction between "access restrictions" and "use restrictions", Nosal II distinguished this case from the former case, United States v. Nosal (Nosal I) (676 F.3d 854 (9th Cir. 2012)), which considered the use restrictions under § 1030 (a) (4) of CFAA. It is, however, pointed out that these two terms (access restrictions and use restrictions) can be used interchangeably, and there can be no difference between them in practice. *See* Williams, *supra* note at 426–427.

Korn/Ferry, he started to download confidential information from "Searcher", the internal database of data of over one million executives, to use it for his startup competitor, by using his own username and password, but this violated Korn/Ferry's confidentiality and computer use policy. Also, Nosal and his compatriots continued to access the firm's internal database using current employee's login credentials.

The Ninth Circuit confirme that a former employee's credentials to access computer were revoked and that Nosal (and his associates) acted "without authorization" under CFAA when they accessed computer data with a current employee's credentials. Since the authorization can be granted only by the computer owner (i.e. Korn/Ferry International), not by a current employee for the purpose of CFAA, the court held that Nosal's access had no possible source of authorization because his credentials were revoked when he left the company [44].

In *Facebook Inc., v. Power Ventures* [45], the Ninth Circuit held that Power Ventures (hereinafter "Power") violated CFAA after it received a letter from Facebook because it then accessed Facebook computers without permission [46].

---

44) Nosal was also found to be convicted because he accessed the trade secret and stole it intentionally under Economic Espionage Act of 18 U.S.C. § 1832 (a) in this case.

45) Facebook Inc., v. Power Ventures, 844 F.3d 1058 (2016).

46) Facebook also sued against Power Ventures based on the Controlling the Assault of Non-Solicited Pornography and Marketing Act of 2003 ("CAN-SPAM"), but this allegation was denied.

The defendant, Power, was operating a social networking website (Power.com) which accessed Facebook users' data to use it for its promotion campaign. The Power system allowed individuals who already have any accounts on social networking websites to login to Power.com and create a new account. The users of Power.com could then use contacts from various social networking websites on a single page because Power.com aggregated the users' information on different websites.

Facebook required its users to give consents to its terms and use policy and requested individuals to register at Facebook before using its services. The users could connect with other Facebook users as "friends" on the Facebook website.

Even after Facebook sent a cease and desist letter and blocked one of Power's IP addresses, Power continued to access Facebook computers. Facebook cited its terms and use policy to object the Power's access. Noticeably, Facebook had never revoked its users who spontaneously provided their passwords with Power, though it tried to limit its users access.

The Ninth Circuit found that the access of Power to Facebook's computers was without valid authorization under CFAA after it received the cease and desist letter from Facebook, though it also admitted that Facebook users could give permission to third parties, including Power, to gain access to their own accounts. Even though the users continued to provide such permissions to Power, the access to Facebook computers by Power was regarded as illegal under CFAA after it received the letter from Facebook, the court said.

1.3 The District Court and Ninth Circuit's Decisions of *hiQ Labs v. LinkedIn Corporation* and Data Scraping

LinkedIn, founded in 2002 and acquired by Microsoft in 2016, is a largest worldwide network for professionals with more than 540 million users in the world. Compared to LinkedIn, hiQ Labs is much smaller, and is a startup founded in 2012: a data science company applied to human capital through information from public data sources [47].

The company, hiQ, uses two software tools, "Keeper" and "Skill Mapper". These solutions can help HR and business leaders to keep their key ability about the employees, and to acquire talents and manage teams with difficulty to build workforces, based on publicly available data. These tools are reinforced by data "scraped" from LinkedIn's user information.

When using automated scripts to access publicly available information, the process of automatically loading and reading vast amounts of data on the web, to be analyzed later, is often recognized as "data scraping" [48].

Data scraping involves taking content from a website, called a "data host", through a piece of software, extracting vast amounts of data by using automated programs, often called "bots", for the user's specific purposes [49]. The data host is not necessarily aware of data scraping, and even in

---

47) hiQ Labs, Inc.,: https://www.hiqlabs.com/ (Last visited on 11th September 2020).

48) Williams, *supra* note 12, at 418. It explains that many people still use the term "scraping", rather than something more technically descriptive such as "screen reading" or "web reading" since the choice of terminology "scraping" has discussed only among engineers, never been widely debated.

49) Marissa Boulanger, *Scraping the Bottom of the Barrel: Why It is No Surprise*

cases where the host is aware, the activity is not necessarily approved. Data scraping can be divided into two categories in general, "harmful or beneficial", and harmful data scrapers can cause problems including spamming email accounts or network crashes of the website [50]. Also, the information which users intend to keep private can be stolen by harmful data scrapers [51]. The difference between harmful and beneficial data scrapers is important because lawsuits may occur when it comes to harmful data scraping.

Although LinkedIn asked hiQ to stop data scraping and tried to prevent hiQ from accessing the data with various blocking techniques, hiQ continued its data scraping activities. After LinkedIn issued a cease and desist letter to terminate hiQ's automated data collection, hiQ initiated the action in the Northern District of California, claiming that LinkedIn's action threatened its business and that its scraping was lawful under CFAA [52]. The court held that hiQ was entitled to a preliminary injunction, finding that hiQ would likely go out of business if LinkedIn prevented hiQ from scraping. As for the LinkedIn users' privacy interests, the court found it unclear whether "LinkedIn members who decide to set their profiles to be publicly viewable expect much privacy at all in the profiles they post" and held that "those expectations are uncertain at best, and in any case, LinkedIn's own actions

---

That Data Scrapers Can Have Access to Public Profiles on LinkedIn, 21 SMU Sci. & Tech. L. Rev. 77, 77–78 (2018). The number of the use of online bots has been increasing.

50) *Id.*, at 78. Beneficial data scrapers can lead users to tools such as search engines and price aggregators.

51) *Id.*

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

do not appear to have zealously safeguarded those privacy interests" [53].

One of the key problems in this case was whether hiQ's access to the LinkedIn's user publicly available data violated the CFAA. The underlying reasoning of the court was that, due to the vagueness of the CFAA text, there should not be a finding of violations of the CFAA when the access is to data publicly available, even though the data host explicitly states it does not want to be accessed in certain way or by certain entity or individuals [54].

On the other hand, the hiQ's First Amendment claim that LinkedIn is a public forum failed. The grounds on which hiQ was entitled to its preliminary injunction was not the First Amendment, but the challenging of the applicability of the CFAA and state competition law.

In September 2019, the Ninth Circuit affirmed the District Court's preliminary injunction for hiQ [55]. Regarding the meaning of the CFAA's "without violation" under 18 U.S.C. §1030 (a)(2), the court referred to *Nosal Ⅱ*, stating that "without authorization is a non-technical term that, given its plain and ordinary meaning, means accessing a protected computer without permission" [56], and citing the provision's legislative history to make clear that the prohibition by this term is "properly understood to apply only to private information – information delineated as private through use of a permission requirement of some sort" [57]. Citing the article by Orin S. Kerr,

---

53) hiQ Labs, v. LinkedIn, 273 F. Supp. at 1119.

54) McDonough, *supra* note 30, at 67.

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

56) United States v. Nosal, 844 F.3d 1024, 1028 (9th Cir. 2016).

57) hiQ Labs, Inc., 938 F.3d at 1001.

the court pointed out that some authentication requirement like a password gate is necessary to make barriers to closed spaces on the web. Kerr argued that the nature of the web space is "inherently open" and the authorization line under CFAA should be recognized only "when access is gained by bypassing an authentication requirement" like password gate.

In addition, the court weighed the public interest which hiQ claimed as maximizing the free flow of data on the internet against LinkedIn [58].

The two Ninth Circuit decisions in *Facebook* and *Nosal II* held access to computers to be in violation of CFAA. But these decisions were criticized on the basis that the Ninth Circuit had relied on "poorly" reasoned password sharing decisions, and that this approach made "blurred lines" about liability under CFAA [59].

In *Nosal II* and *Facebook*, the access in question was made via a "password gate". There was no "break-in" because legitimate and valid login credentials were used to access the data, which were stored behind a code-based authentication barrier [60]. The panel in the *hiQ* case found that the LinkedIn users accepted their personal information were publicly viewable and access to such publicly available information was not held to be "without authorization" under CFAA.

The Ninth Circuit's decisions in *Facebook* and *Nosal II* doubtlessly

58) *See* Orin S. Kerr, Norms of Computer Trespass, 116 Colum. L. Rev. 1143, 1161 (2016).

59) *See* Williams, *supra* note 12, at 419.

60) *See Id.*, at 427.

presupposed that access "without authorization" under the statute of CFAA was obvious, that organizations such as LinkedIn have authority to rescind access using data scraping to their websites, and that any violation of the rescission may cause the CFAA. Indeed, the Ninth Circuit in *hiQ* had to distinguish the case from these precedents and it did with using a password gate theory by citing the Kerr's argument [61]. It should be noted that the panels in *Nosal Ⅱ* and *Facebook* found no evidence that defendants circumvented the technological access barriers in practice.

CFAA cases including alleged illegality of web scraping by competitors are not new. It can, in fact, be noted that the Ninth Circuit Court in *hiQ* case showed "renewed" effort to find CFAA liability for automated access by declaring that the purpose of CFAA was punishing hacking in *hiQ* case [62]. However, it can be still argued whether authentication requirements such as password gates can be an effective benchmark to decide whether or not the access in issue should be regarded as "without authorization", given the Ninth Circuit decisions found no evidence of circumvention of technological access obstacles, and it is not certain that automated data scraping scripts can always discern the will of website owners.

Based on the Ninth Circuit decisions in *Facebook* and *Nosal Ⅱ*, it seems that computer owners possibly have exclusive power to permit or revoke access. If they provide authentication requirements such as password gates, access "with authorization" needs to be pursuant to requirements under CFAA, even if users permit third parties to share their own data (like in

61) McDonough, *supra* note 30, at 68.

62) Williams, *supra* note 12, at 419.

Facebook case in 2016). If the purpose of CFAA was punishing hacking in *hiQ*, then, if someone uses a user's password with his/her permission, would the access give rise to liability (and would it be criminal) under CFAA? These decisions of the Ninth Circuit can be questioned about this problem. If open access to the internet should be protected by the free speech of the First Amendment, how is the CFAA's protection for computer owners compatible with the interests of internet users under the First Amendment?

2 No Harm for the Publicly Available Data on the Internet?

2.1 Questions around "Who" and "Whose" Data

The legality of web scraping is still an open question [63]. This is the question of whether third parties are allowed to use automated scripts to access to, and collect mass data exposed to social media for targeting without the service platforms' or the users' consent [64]. Indeed, the Ninth Circuit decisions have not properly solved who has the ownership of personal data publicly available on social media platforms [65].

A novel legal question raised by *hiQ Labs, Inc. v. LinkedIn Corp.*, was that of who has control over the publicly viewable data placed on a social

63) Caitlin E. Jokubaitis, *There and Back: Vindicating the Listener's Interests in Targeted Advertising in the Internet Information Economy*, 42 Colum. J.L. & Arts 85, 111 (2018).

64) *Id.*, at111. Cambridge Analytica used data scraping to access and collect user data of Facebook without users' express consent in spring 2018, and it has been pointed out that these mass data were used to manipulate the some elections including the "Brexit" referendum in 2016 in the United Kingdom. Pre-existing anti-fraud, privacy and copyright laws tend to be used to challenge the use of data scraping, but the ownership of the data disclosed on social media platform is still uncertain.

65) *See Id.*

media platform [66].

Data scraping is not new technology, but the legal situation about it is not clear. Civil claims [67] have been filed against harmful data scraping founded on trespass to chattels, breach of content, copyright violations and CFAA violations, though the U.S. courts have failed to show a uniform approach [68].

Trespass to chattels is a tort doctrine, but was hardly used until the late 1990s due to its vagueness, as opposed to the well-known trespass to land [69]. Courts have started to use the trespass to chattels doctrine for cases related to computers since the 1990s, and this change is often connected with two significant developments. One is reduction of the harm requirement, and the other involves the expansion of the definition of the physical contact to include transmission of electronic signals [70]. In sum, courts have forbidden the intrusion of an individual into another computer through electronic signals under a new version of the doctrine created by the courts: "trespass to computers" [71].

---

66) Marissa Boulanger, *Scraping the Bottom of the Barrel: Why It is No Surprise That Data Scrapers Can Have Access to Public Profiles on LinkedIn*, 21 SMU SCI. & TECH. L. REV. 77, 77 (2018).

67) This does not mean that CFAA is strictly limited to civil cases, and certain actions can be dealt with as criminal cases under CFAA. *See* Orin S. Kerr, Vagueness Challenges to the Computer Fraud and Abuse Act, 94 MINN. L. REV. 1561, 1587 (2010).

68) Boulanger, *supra* note 66, at 78-79.

69) Omri Rachum-Twaig & Ohad Somech, *Breaking into an Empty House: A Theory of Remedies for CFAA Unauthorized Access to Non-Proprietary Information*, 82 ALB. L. REV. 555, 560 (2019).

70) *Id.*, at 561.

71) *Id.*

In particular, the mitigation of the harm requirement has been key for the change in the approach of the courts[72]. The new trespass to computers doctrine has been interpreted expansively by many courts now in applying it to computer-based situations[73]. Courts basically still recognize the right of website owners to selectively restrict who can access their sites and to allow them to complaint on the ground of potential or indirect harm to their businesses[74]. More recently, courts tend to include CFAA discussions into the courts arguments, but at the same time, it is also pointed out that another narrative has been developed[75].

In *NetApp Inc., v. Nimble Storage Inc.*[76], the Northern District of California said, "in order for the taking of information to constitute wrongdoing, the information must be "property" as defined by some source of positive law", and the court mentioned that if the information were "property" based on a positive law (either Uniform Trade Secrets Act or the Copyrights Act), then the claim would be preempted. This narrative suggests that the application of the trespass to computer doctrine should be limited when it comes to scraping activity by re-introducing the harm requirement and

---

72) *See e.g.*, Thrifty-Tel, Inc v. Bezenek, 54 Cal. Rptr. 2d 468 (Cal. Ct. App. 1996); CompuServe Inc. v. Cyber Promotions, Inc., 962 F. Supp. 1015 (S.D. Ohio 1997); eBay, Inc. v. Bidder's Edge, Inc., 100 F. Supp. 2d 1058 (N.D. Cal. 2000); Register.com, Inc. v. Verio, Inc., 126 F. Supp. 2d 238 (S.D.N.Y. 2000); Intel Corp. v. Hamdi, 114 Cal. Rptr. 2d 244 (Cal. Ct. App. 2001).

73) Rachum-Twaig & Somech, *supra* note 69, at 562.

74) *Id.*

75) *Id.* (The case of the Central District of California is explained here; to state that copyright law preempted, the claim of the aggregation of information publicly viewable on a website was refused).

76) NetApp Inc, v. Nimble Storage, Inc, 41 F. Supp. 3d 816 (N.D. Cal. 2014).

focusing on the nature of the information being accessed as non-property [77]. In respect of the application of the trespass to chattels doctrine to data scraping, courts suggest two different perspectives. The first intimates that "the owner" of a server is allowed to restrict access to its site by providing notice, and the other severely restricts the application of the doctrine due to the nature of the publicly available information as non-property.

Although it involved the context of CFAA, Northern District Court of California stated in *hiQ Labs, Inc. v. LinkedIn Corp.*, that personal property with resultant injury is necessary, to apply the trespass to chattels doctrine to a case. According to Richard Epstein, the harm requirement is presented only in the trespass to chattels doctrine, compared to the land context, and this is because only the trespass to land theory deals with the underlying question of ownership, and so, the chattels doctrine does not involve boundary disputes [78]. Regarding the recent trend in the use of the trespass to chattels doctrine by the courts, it is noted that the creation of legal borders has been denied when there exists no legally protected interests on the website at issue [79].

2.2 Publicly Available Data and Rational Expectation for Privacy

It seems that waivers and consent at least lessen the possibilities of violations of privacy rights [80]. But can this be a valid assumption in

---

77) Rachum-Twaig & Somech, *supra* note 69, at 563.

78) *See* Richard Epstein, Centennial Tribute Essay: Cybertrespass, 70 U. CHI. L. REV. 73, 78 (2003).

79) Rachum-Twaig & Somech, *supra* note 69, at 564-565.

80) Ron Brown, Robots, *New Technology, and Industry 4.0 in Changing Workplaces. Impacts on Labor and Employment Laws*, 7 AM. U. BUS. L. REV. 349, 377 (2018).

considering the development of technology?

The internet is characterized as an "illusionary" space due to its anonymity and privacy [81]. This perspective warns that the social media presents a risk partly because users do not pay enough attention to their privacy, and are supposed to be fragile to protect their own privacy [82]. It is not clear how federal laws can be effective in protecting people from such risks. There are, in fact, federal and state statutes covering intentional interceptions of electronic communications and hacking, and unauthorized access to electronic information. However, these laws cannot be effective to prevent the harms if people who try to intercept such transmissions have no idea about them. Additionally, arguing about the illegality of the electrical interception of private communications may not be an effective solution once it has been made public on the internet [83].

Privacy interests are sometimes used to impose a burden on the exercise of the First Amendment, but the Supreme Court [84] has suggested that privacy is not readily a viable ground for restrictions on the application of the First Amendment [85]. Regarding the constitutional importance of the dissemination of truthful information, the Court showed concern over

---

81) Cheryl B. Preston, *Lawyers' Abuse of Technology*, 103 Cornell L. Rev. 879, 893-894 (2018).

82) Indeed, except for computer experts, for example, many users may not know the systems and have less knowledge about who can track their IP address or identify the MAC number of their computer, and so on. *See* Preston, *id.*, at 894.

83) *See id.*, at 894-895.

84) Central Hudson Gas & Elec. v. Public Svc. Comm'n, 447 U.S. 557 (1980).

85) Jokubaitis, *supra* note 63, at 114; *see also* U.S. W., Inc. v. FCC, 182 F 3d. 1224, 1234-35 (10th Cir. 1999).

restricting the public dissemination of commercial information [86]. In this case, Justice Blackmun, writing for the majority opinion, articulated about the meaning of the right to advertise as a "reciprocal right to receive the advertising". He showed the significance of the interests in the free flow of commercial information for a particular consumer and society alike, noting "(the particular consumer's) interest may be as keen, if not keener by far, than his interest in the day's most urgent political debate". For society, Justice Blackmun said that even entirely commercial advertisement by individuals may have general public interest. Thus, to the question of "whether a State may completely suppress the dissemination of clearly truthful information about entirely lawful activity, fearful of that information's effect upon its disseminators and its recipients", Justice Blackmun answered in the negative.

However, it should not be overlooked that protection of privacy for individuals can be on a different basis than that of speech on matters of public concern [87].

If the Court applies CFAA to this case because it intends to prevent unauthorized data scraping, this can mean that it may reinforce the idea that such a certain method of data collection and transmission is not covered by the scope of protectable speech under the First Amendment [88].

---

86) Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, 425 U.S. 748 (1976).

87) Jokubaitis, *supra* note 63, at 114.

88) *Id.*, at 112.

2.3 Cases in Japan

The PPC issued an order to suspend immediately the operations of two businesses which illegally disclosed personal information such as names and addresses in bankruptcy cases on their websites in July 2020, and this is the first time for the PPC to order the suspension. According to the paper published by the PPC, these businesses collected such personal information from the Official Gazette (called "Kanpō") [89] and disclosed it online without publishing publicly nor notifying that purpose to the people concerned. Additionally, these disclosures were not based on consent of them. These disseminations of personal information violated the Act on the Protection of Personal Information (APPI) [90]. The PPC also gave notice that it would take these businesses to court, relying on APPI, if they did not comply with this order by 27th August 2020 [91].

In relation to this case, the Japan Federation of Bar Association (JFBA) already published an opinion paper and submitted it to the Prime Minister, Minister of Finance, and the Commissioner of the PPC [92]. The paper can be summarized by two points. The first asserts that the information about

89) The public notice about bankruptcy information published on Kanpō can be meaningful in that a service may be substituted by a public notice, and notification of judicial decision can be deemed to be done with a public notice. *See* article 10 of Bankruptcy Law (Act No. 75 of 2004); *see also* article 10 of Civil Rehabilitation Act (Act No. 225 of 1999). A public notice published on Kanpō is thought of as useful and easy way to notify necessary information to interested people including creditors in bankruptcy cases.

90) Articles 18 and 23 (1) of the Act on the Protection of Personal Information.

91) Article 84 of the Act on the Protection of Personal Information.

92) Japan Federation of Bar Association (Posted on 16 July 2020): https://www.nichibenren.or.jp/document/opinion/year/2020/200716_4.html (Last visited on 11th September 2020).

bankruptcy should be categorized as "Special care-required personal Information" by the Cabinet order of APPI [93]. The other requires the government to take technical measures to block automated programs accessing and collecting such information from Kanpō online, which is free and publicly available.

Even before this case, concerns for the disclosure of bankruptcy information had been raised in society. One of the most noticeable cases, which the JFBA's opinion paper also referred to, occurred in March 2019. The website at issue accessed and collected comprehensively information about bankrupts from Kanpō, and compiled mapping databases to be published online.

This website was criticized by people including those whose information was exposed, because they said that such disclosure of personal information violated the personal interests such as privacy. To solve this problem, the PPC issued administrative advice to the website, stating that personal information should not be sent to third parties without consent from principals of the information. This advice does not legally bind the website, but the website was "spontaneously" closed [94]. Nevertheless, there are still similar websites on the internet, and information about bankruptcy has been disseminated in a way inconsistent with APPI.

---

93) Cabinet Order to Enforce the Act on the Protection of Personal Information (Cabinet Order No. 507 of 2003).

94) As for the fact about this case, *see* the Opinion Paper of the JFBA: https://www.nichibenren.or.jp/library/pdf/document/opinion/2020/opinion_200716_4.pdf.

The JFBA pointed out that the problems have become more serious in this today's digital era where information can easily be digitalized and exposed to more people. Thus, it is necessary to stop the illegal proliferation of digitalized personal information on the internet, and to update the law itself to deal with these problems[95].

The JFBA argued that it was necessary to regulate processing of personal information by business operators at the time of its acquisition, and stated that the law should ban such acquisitions without consent from the principals of the data, and that the data of bankruptcy should be regarded as "Special care-required personal Information" under APPI[96]. Given that disclosure of the bankruptcy information can invade the privacy of the principals and bring about stigma against them, the JFBA noted that it can "to cause unfair discrimination, prejudice or other disadvantages to the principal", which the JFBA argued should be subject to rules appropriate for this special type of personal information. The JFBA focused strongly on the importance and necessity of the legal restrictions on the business operators processing the information relating to bankruptcy.

However, it should not be overlooked that the system of a public notice is a significant tool to notify necessary data about bankruptcy to creditors.

---

95) *See id.*

96) Article 2(3) of the Act on the Protection of Personal Information. It provides that "'Special care-required personal information' in this Act means personal information comprising a principal's race, creed, social status, medical history, criminal record, fact of having suffered damage by a crime, or other descriptions etc. prescribed by cabinet order as those of which the handling requires special care so as not to cause unfair discrimination, prejudice or other disadvantages to the principal".

It is true that the protection of privacy or general personal interests for the relevant individuals should be considered carefully, so as to take the right balance between the interests of creditors and bankrupt individuals. This public notice should not hinder bankrupts from rehabilitating from economic difficulty. The JFBA stated that the restriction on the publication of bankruptcy information on Kanpō should be limited to minimum necessary for that purpose. As the JFBA presumed that business operators publishing bankruptcy data use "automated programs" to access, collect and compile the information publicly available on Kanpō, it said that taking technical safeguards against such automated programs to access and collect it can be justified as minimum requirements under the proper balancing test, stating that the purpose of such operators in publishing the bankruptcy data cannot be a justification.

When the Supreme Court of Japan argued about the constitutionality of the "Basic Resident Register Network", it admitted that citizens' liberty in private life should be protected against the exercise of public authority under Article 13 of the Constitution, and that any individual has the liberty which protects his/her own personal information from being disclosed or published to any third parties without good reason [97].

The Basic Resident Register Network (called "Juki-Net") was introduced by the revision of the Basic Resident Register Act in 1999 [98], to construct

---

97) Saikō Saibansho [Sup. Ct.] Mar. 6, 2008, Hei 19 (o) no. 403, 62 Saikō Saibansho Minji Hanreishū [Minshū] 665 (Japan) (citing the case of Saikō Saibansho [Sup. Ct.] Dec. 24, 1969, Sho 40 (a) no. 1187, 23 Saikō Saibansho Keiji Hanreishū [Keishū] 1625 (Japan)).

98) Act No. 133 of 1999.

a network of basic resident registers to share the matters of basic resident registers among the State and across various municipalities. The appellants alleged that this Network system was unconstitutional because it violated the appellees' privacy right and other rights of person under Article 13 of the Constitution in that such administrative organs collect, manage or use their personal information. By "Juki-Net", names, addresses, genders, birthdates, resident certificate codes, and some personal identifiable information are to be shared among the State and municipal administrative organs.

Although the feature of Juki-Net in question should be its collecting, managing or using of the personal information by the public authority, the Court described the scope of the protection under Article 13 as individual's liberty protecting his/her personal data not to be disclosed or published, and thus it concluded that public authorities should not disclose or publish personal data of residents without good reason, based on its proper management of the personal information of residents which was "already collected", but that it may not even matter whether the principals do not consent to such disclosure or publication.

The Supreme Court recognized the legal interests of privacy as protecting personal matters from being disclosed without good reason [99]. In this case, the criminal records of the appellant [100] were provided after one

99) Saikō Saibansho [Sup. Ct.] Jan. 31, 2017, Hei 28 (kyo) no. 45, 71 Saikō Saibansho Minji Hanreishū [Minshū] 63 (Japan).

100) The appellant was arrested in 2011 because of paying for child prostitution and was punished by a fine. The fact of his arrest was broadcasted in the media on the same date of his arrest, and all or some of the coverage were posted on the websites many times.

puts his name and the prefecture name where he lives on a search engine box on the internet (Google). The appellant sued against the "search service provider", who performed searches on the website for users' requests and produced the search results or URLs which are codes to identify particular websites to them for deleting the search results relating to his criminal records based on his personal right or interests.

The Court admitted legal privacy founded on precedents, but also mentioned the value of the expression act of the search service provider, considering that the search results were provided pursuant to the provider's policy relating to the production of the results, even though these processes of broadly collecting, compiling and providing publicly available information on the internet websites are executed automatically.

Moreover, the Court also recognized the role of such a search service provider as a significant foundation for free flow of data on the internet in the modern era, helping the public by providing the information on the internet and acquiring necessary information from vast amounts of data therein.

Considering the value and roles of search service providers, according to the Court, whether the provider's act to provide information involving the appellant's privacy is illegal or not should depend on the balancing of legal interests in not disclosing criminal records and those in providing the information including the URLs as search results. In relation to this balancing test, the Court stated that the appellant can justify his allegation against the search service provider only when it is proved to be clear that the former interest obviously outweighs the latter.

Conclusion

This paper considers application of CFAA in the United States in comparison with recent problems relating to illegal disclosure of personal information of bankruptcy in Japan.

Access to the internet can be addressed by the First Amendment of the Constitution of the United States, but the reasoning of *Carr* case in the Supreme Court of Pennsylvania raises questions about the meaning of "matter of public concern". The question of whether "purely" personal opinions about some matter of public concern should be regarded as "private" should not so easily be decided.

Truly, the utility of information including personal data on the internet space is of significance, and the consent from people concerned plays an important role in ensuring a proper balance is struck and that there is protection of individuals' rights and interests. However, it has been argued the consent can be effective in rapid developments of technology in modern period. It should be clear whether automated processing of case amount of personal information publicly available like data scraping on the internet should be covered by Free Speech under Article 21 of the Constitution of Japan.

It seems that the Supreme Court of Japan focuses on the publication of information without good reason as the important point to think about individuals privacy protections, but it should be unclear that such an idea based on divided boundaries between public and private spheres is valid in internet society. Whether publication (publicly available) should be

key to examine the protection of information and/or data subjects' rights and interest should be considered because later publication should not be justified just because it is already published before. What rights and interests should be regarded as legally protected for subjects needs to be clear. This is a problem of the interpretation of rights and interests under Article 1 of APPI, which sets out the purpose of this Act. In the case I consider in this paper, I feel the traditional idea of "consent" of the people concerned is not enough or essential for the its meaning of the substantive rights and interests, and that the conception of the right to control over their information needs to be reconsidered.

(Former Assoc. Prof. Daito Bunka Univ.)