Notre Dame Law Review Reflection

Volume 96 | Issue 4 Article 3

5-19-2021

Biopolitical Opportunities: Between Datafication and Governance

Orly Lobel
Warren Distinguished Professor of Law, University of San Diego

BIOPOLITICAL OPPORTUNITIES: BETWEEN DATAFICATION AND GOVERNANCE

*Orly Lobel**

A popular Government, without popular information, or the means of acquiring it, is but a Prologue to a Farce or a Tragedy; or, perhaps both.

—James Madison[1]

Julie Cohen's dazzling tour de force *Between Truth and Power* asks us to consider the new ways powerful actors extract valuable resources for gain and dominance.[2] Cohen in particular warns that "the universe of personal data as a commons [is] ripe for exploitation."[3] Cohen writes that "if protections against discrimination, fraud, manipulation, and election interference are to be preserved in the era of infoglut, regulators will need to engage more directly with practices of data-driven, algorithmic intermediation and their uses and abuses."[4] I read *Between Truth and Power* as not only a compelling account of the contemporary transformations of law and technology but also a call to action. This Essay takes up Cohen's challenge by considering ways in which governments can engage in new forms of governance to leverage the very same biopolitical data extracted by private actors for profit purposes in service of public goals of fairness, equality, and distributive justice. In particular, the Essay describes several current contexts that demonstrate how datafication can, and indeed should, be employed to aid regulatory research, enforcement, and accountability. The three examples I focus on are: first, current developments in labor market information flows that are attempting to address salary inequities, labor market concentration, and bias; second, the

* Warren Distinguished Professor of Law, University of San Diego.

1 Letter from James Madison to W.T. Barry (Aug. 4, 1822), *in* 9 THE WRITINGS OF JAMES MADISON 103, 103 (Gaillard Hunt ed., 1910).

2 *See generally* JULIE E. COHEN, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF THE INFORMATIONAL CAPITALISM (2019).

3 *Id.* at 51.

4 *Id.* at 200.

technique of scraping data off platforms in service of regulatory compliance; third, the issue of monitoring and tracking viral spread during a global pandemic. I argue that if we are to take Cohen's framework seriously, then policymakers have no choice but to identify opportunities within disruptive technological changes and to mirror, rather than attempt to block, these innovations.

In their book *She Said*, Pulitzer Prize-winning journalists Jodi Kantor and Megan Twohey describe a meeting they had with Chai Feldblum, then Equal Employment Opportunity Commission (EEOC) Commissioner.[5] Feldblum told the reporters "'[w]e know internally who the companies are that have the most charges' . . . . [b]ut the agency was prohibited from making that information public. Before taking a job, a woman could not check with the EEOC to see what kind of record the prospective employer had on harassment."[6] The journalists concluded that the United States had a "system for muting sexual harassment claims, which often enabled the harassers instead of stopping them."[7] But could we leverage today's technologies, online connectivity, and data against such systems of silence? Could we imagine the endless opportunities, beyond the imminent risks, that datafication offers to longstanding democratic goals and challenges?

Between Truth and Power rejects the conventional story that law lags behind markets, always struggling to keep up in the face of disruption and innovation. Rather, Cohen shows us that law is already being mobilized, helping powerful actors engender the changes we see now.[8] This time around, law is being reconfigured, substantively and performatively, de facto and de jure, to help reshape, and profit from, what Cohen calls the biopolitical public domain.[9] I agree with this description. Precisely for the reason that law is in the mix at the very beginning of any paradigmatic shift, lawmakers and scholars must be creative and forward thinking in finding windows and shaping the law in ways that promote public ends.

I. WAGE DISCRIMINATION AND INFORMATION FLOW REFORMS

In a new article, *Knowledge Pays: Reversing Information Flows and the Future of Pay Equity*, I argued that recent law reforms in the field of wage discrimination are focused on correcting a longstanding information imbalance in the labor market: employers demand secrecy from their employees about compensation and rarely reveal the pay scale of the company to prospective and current employees; at the same time, however, employers extract information from employees about their salary histories as well as

5 JODI KANTOR & MEGAN TWOHEY, SHE SAID: BREAKING THE SEXUAL HARASSMENT STORY THAT HELPED IGNITE A MOVEMENT 80 (2019).

6 *Id.* at 80–81.

7 *Id.* at 81.

8 *See* COHEN, *supra* note 2, at 8, 139–40.

9 *See id.* at 48–49.

share wage information with other businesses.[10] This information imbalance has long prevented employees from detecting pay discrimination and knowing whether their talents are adequately valued.[11]

New reforms in state law and federal caselaw are targeting these information asymmetries by banning employers from asking job candidates about their salary histories, relying on salary histories, or prohibiting employees from discussing their salaries with other workers or third parties.[12] Some law reforms go a step further by requiring employers to provide a pay scale for a position[13] and, even more impactfully, requiring employers to annually report data about employee race, gender, and ethnicity to the EEOC.[14] Self-reported numbers are easily manipulated, and therefore we might imagine a role for digital platforms in creating more systematic transparency, which in turn can mobilize workers and aid regulators with enforcement. Platforms such as LinkedIn, Glassdoor, Salary.com, and SalaryExpert crowdsource salary information, helping employees in their job searches and negotiations.[15] Glassdoor, for example, has a pay data tool called *Know Your Worth* that dynamically analyzes trends to provide an increasingly accurate estimate of a position and an employee's market value.[16]

---

10 *See* 120 COLUM. L. REV. 547, 549, 589–90 (2020).

11 *See id.* at 549, 558.

12 *See id.* at 550, 567–87.

13 *See* CAL. LAB. CODE § 432.3(c).

14 *See* Amy Conway, Stephanie Scheck & Carroll Wright, *End to EEO-1 Component 2 Pay Data Reporting for Now . . .*, JD SUPRA (March 2, 2020), https://www.jdsupra.com/legalnews/end-to-eeo-1-component-2-pay-data-20353/; *Notice of Proposed Changes to the EEO-1 Report to Collect Pay Data from Certain Employers*, U.S. EEOC, https://www.eeoc.gov/employers/notice-proposed-changes-eeo-1-report-collect-pay-data-certain-employers (last visited Apr. 17, 2021). A "new California law somewhat addresses pay transparency by extending—from two years to three—an employer's obligation to maintain records of wages and pay rates, job classifications, and other terms of employment, though the records are kept confidential unless they are ordered in discovery." Lobel, *supra* note 10, at 591 n.287; *see* CAL. LAB. CODE § 1197(e).

15 Benjamin Arendt, *Glassdoor? Google? LinkedIn? Any Which Way, the Future of Recruiting Is Transparency*, TALENT DAILY (June 6, 2018, 4:16 PM), https://web.archive.org/web/20190403021951/https://www.cebglobal.com/talentdaily/glassdoor-google-linkedin-any-which-way-the-future-of-recruiting-is-transparency/.

16 *Know Your Worth,* GLASSDOOR, https://www.glassdoor.com/Salaries/know-your-worth.htm (last visited Apr. 17, 2021); *see* Susan Adams, *How Companies Are Coping with the Rise of Employee-Review Site Glassdoor,* FORBES (Feb. 24, 2016, 3:49 PM), https://www.forbes.com/sites/susanadams/2016/02/24/how-companies-are-coping-with-the-rise-of-employee-review-site-glassdoor/; Jillian Kramer, *Are You Worth More This Year than You Were in 2018?*, GLASSDOOR (Jan. 11, 2019), https://www.glassdoor.com/blog/are-you-worth-more-this-year; Queenie Wong, *Are You Getting Paid Enough? LinkedIn Launches Salary Comparison Tool*, MERCURY NEWS (Nov. 3, 2016, 6:27 AM), https://www.mercurynews.com/2016/11/02/are-you-getting-paid-enough-linkedin-launches-salary-tool/.

Nobel laureate Gary Becker described in his research on labor market inequities the way in which secrecy was the fuel for continued inequality.[17] Secrecy prevents employees from knowing whether they earn less than their peers, as well as whether they are alone in wage theft due to misclassification, unpaid overtime, and other noncompliance. For years, economists have estimated billions in lost wages due to imperfect information.[18] With the use of crowdsourced platforms, secrecy norms are changing, and for millennials, "[p]ay confidentiality has been eroding for years."[19] This is in large part due to access to online data and social networks.

Beyond detection of salary inequities, data mining can help uncover, and tame, persisting biases and narratives that contribute to inequality in the workplace. An example of such efforts is a service named Textio, which mines through job searches and labor market advertisements and discovers how certain phrases used in ads of job openings, such as military analogies like "mission critical," can result in fewer women applicants.[20] Many employers want to increase the diversity of their workforce, and this kind of analysis can provide them with important tools. Textio has identified more than twenty-five thousand phrases that indicate gender bias.[21] Words like "top-tier," "aggressive," "coding ninja," "fast-paced work environment," and sports terms, like the military jargon, decrease women applicants, while words like "partnerships" and "passion for learning" attract more women.[22] This kind of data analytics service is an example of potentially creating more inclusive job listings using new technological capabilities.

In my article *The Law of the Platform*, I argued that the normative challenges facing policy for a digital platform era are not new:

---

17 *See* GARY S. BECKER, THE ECONOMICS OF DISCRIMINATION 9–18 (2d. ed. 1971); *see also* PAY SECRECY AND WAGE DISCRIMINATION, INST. FOR WOMEN'S POL'Y RSCH. (2014), https://iwpr.org/wp-content/uploads/2020/09/Q016.pdf ("While there may be no direct link between pay secrecy and pay inequality, pay secrecy appears to contribute to the gender gap in earnings.").

18 *See* Richard A. Hofler & Kevin J. Murphy, *Underpaid and Overworked: Measuring the Effect of Imperfect Information on Wages*, 30 ECON. INQUIRY 511, 512, 525 (1992); Yannis M. Ioannides & Linda Datcher Loury, *Job Information Networks, Neighborhood Effects, and Inequality*, 42 J. ECON. LITERATURE 1056, 1056 (2004); Alexandre Mas, *Does Transparency Lead to Pay Compression?* 5 (Nat'l Bureau of Econ. Rsch., Working Paper No. 20558, 2014), http://www.nber.org/papers/w20558.

19 Howard Risher, *Pay Transparency is Coming*, 46 COMP. & BENEFITS REV. 3, 3 (2014).

20 *See* Claire Cain Miller, *Can an Algorithm Hire Better Than a Human?*, N.Y. TIMES (June 25, 2015), https://www.nytimes.com/2015/06/26/upshot/can-an-algorithm-hire-better-than-a-human.html.

21 *See* Emily Peck, *Here Are the Words that May Keep Women from Applying for Jobs*, HUFFPOST (June 2, 2015, 2:24 PM), https://www.huffpost.com/entry/textio-unitive-bias-software_n_7493624.

22 *See* T.L. Andrews, *Just a Few Words Can Increase Female and Minority Job Applicants by More than 20%*, QUARTZ (July 11, 2017), https://qz.com/1023518/just-a-few-words-can-increase-female-and-minority-job-applicants-by-over-20/; Miller, *supra* note 20.

> These emerging challenges of equity and identity on the platform reveal the inevitable points of tension that policymakers have always faced. Balancing equality and anonymity, inclusion and credibility, and safety and privacy is not a new legal challenge. At the same time, the platform presents new opportunities for monitoring and compliance in order to reach a desirable delicate balance. Technology-based monitoring can detect misbehavior in more accurate and fine-tuned ways than broad-brush rules that risk stifling experimentation and growth.[23]

Technology has the advantage of reducing the viability of claims that such tradeoffs between our normative commitments to efficiency and equality, privacy and accountability, and so forth are too insurmountable, incommensurable, unknowable, or too costly to achieve. It helps uncover which companies have traditionally hidden behind formal job descriptions and divisions between positions. Software devoted to data analytics for inclusion and core values is a frontier that scholars should celebrate and study.

II. SCRAPING FOR THE COMMON GOOD

In *Between Truth and Power*, Cohen explains how the law enables both data harvesting and data enclosure—a performative process—via platforms extracting massive amounts of information from users, turning the data into profitable resources, and, in turn, claiming ownership over this extracted information.[24] But what if governments, or private actors in service of public ends, could similarly collect and make use of data? Legal entrepreneurship, as Cohen refers to it,[25] is not an exclusive activity of the for-profit sector, and even within the for-profit sector, startups can profit from targeting the regulatory arbitrage that other startups have been profiting from, in aid of regulatory compliance.

Examples of such entrepreneurial spirit are new data companies that offer their services to cities to monitor short term rentals.[26] These companies have been scraping listings off platforms like Airbnb and VRBO and comparing these listings with formally licensed short-term rentals, in compliance with local laws.[27] They are finding massive amounts of noncompliance.[28] These startups, for example Host Compliance and STR Helper, offer cities a way to recover unpaid taxes by using similar data analytics that other companies use to serve other ends.[29] According to the CEO of

23 101 MINN. L. REV. 87, 165 (2016).

24 COHEN, *supra* note 2, at 44–45.

25 *Id.* at 9, 25.

26 Tom Banse, *Pacific Northwest Cities Hire Outside Vendors to Police Airbnb-Type Rentals*, NW NEWS NETWORK (Aug. 22, 2018), https://www.nwnewsnetwork.org/post/pacific-northwest-cities-hire-outside-vendors-police-airbnb-type-rentals.

27 *See id.*

28 *See id.*

29 *See id.*

Host Compliance, Airbnb is engaged in "'a city-by-city, block-by-block guerrilla war' against local governments."[30] Using startups that focus their energies on data analytics in service of the law is one way to offset the power of such guerrilla tech wars.

These sorts of initiatives by competing market actors collect data by "scraping"—gathering data from other platforms through automation. Scraping is a technique increasingly used not only for profit but for public ends, including accountability and compliance.[31] Journalists are increasingly scraping data for investigative exposés.[32] In 2016, ProPublica published a story about Amazon's pricing algorithm that used simulations of purchases after scraping Amazon product listings.[33] *The Atlanta Journal-Constitution* developed scrapers for a 2017 Pulitzer Prize-nominated national investigation called "Doctors & Sex Abuse."[34] Scraping was essential for the investigation since the reporters' requests for public records to medical boards and regulatory agencies were mostly not fulfilled.[35] Instead, the reporters scraped board orders from public websites finding more than one hundred thousand disciplinary documents.[36]

Researchers are also increasingly turning to scraping to conduct studies in many different contexts. One prominent study that used scraping was conducted by the Harvard Business School on racial discrimination on Airbnb.[37] The researcher found that Airbnb hosts were sixteen percent less

30 Paris Martineau, *Inside Airbnb's 'Guerrilla War' Against Local Governments*, WIRED (Mar. 20, 2019, 7:00 AM), https://www.wired.com/story/inside-airbnbs-guerrilla-war-against-local-governments/.

31 *See* D. Victoria Baranetsky, *Data Journalism and the Law*, COLUM. JOURNALISM REV. (Sept. 19, 2018), https://www.cjr.org/tow_center_reports/data-journalism-and-the-law.php.

32 *See* Julia Angwin & Surya Mattu, *Amazon Says It Puts Customers First. But Its Pricing Algorithm Doesn't*, PROPUBLICA (Sept. 20, 2016, 8:00 AM), https://www.propublica.org/article/amazon-says-it-puts-customers-first-but-its-pricing-algorithm-doesnt; *see also* Sophie Chou, *To Scrape or Not to Scrape: Technical and Ethical Challenges of Collecting Data Off the Web*, STORYBENCH (Apr. 4, 2016), http://www.storybench.org/to-scrape-or-not-to-scrape-the-technical-and-ethical-challenges-of-collecting-data-off-the-web/; Shelly Tan, *Five Data Scraping Tools for Would-Be Data Journalists*, KNIGHT LAB (Mar. 20, 2014), https://knightlab.northwestern.edu/2014/03/20/five-data-scraping-tools-for-would-be-data-journalists/.

33 *See* Angwin & Mattu, *supra* note 32.

34 *See How the Doctors & Sex Abuse Project Came About*, ATLANTA J.-CONST., http://doctors.ajc.com/about_this_investigation/?ecmp=doctorssexabuse_microsite_stories (last visited Apr. 17, 2021); *Finalist: The Atlanta Journal-Constitution Staff*, PULITZER PRIZES, https://www.pulitzer.org/finalists/staff-187 (last visited Apr. 17, 2021).

35 *See How the Doctors & Sex Abuse Project Came About*, *supra* note 34.

36 *See id.*; *see also* Carrie Teegardin & Danny Robbins, *Still Forgiven*, ATLANTA J.-CONST. (2018), http://doctors.ajc.com/still_forgiven/?ecmp=doctorssexabuse_microsite_nav.

37 *See* Benjamin Edelman, Michael Luca & Dan Svirsky, *Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment*, AM. ECON. J.: APPLIED ECON., Apr. 2017, at 1.

likely to accept requests from guests with African American names compared to identical requests from guests with distinctly white names.[38]

The caselaw concerning the legality of web scraping has changed since the early 2000s when the first web scraping cases were litigated.[39] Scraping poses possible liability under the Computer Fraud and Abuse Act ("CFAA").[40] The CFAA has a broad anti-hacking provision against whoever "intentionally accesses a computer without authorization or exceeds authorized access, and thereby obtains . . . information from any protected computer."[41] Many courts use the CFAA as a gauge to determine whether particular web scraping should be allowed. Specifically, courts consider whether a website has granted actual "access," within the Act's meaning, and the data's purpose. The question of whether the practice is lawful is, unsurprisingly, evolving, and unsettled. In *Sandvig v. Sessions*, the ACLU represented scholars and media organizations that were scraping platform information to identify algorithmic bias.[42] The ACLU argued that criminalizing a violation of a term of service under 1032(a)(2)(c) of the CFAA would chill research and reporting: "Refraining from conducting their research, testing, or investigations constitutes self-censorship and a loss of Plaintiffs' First Amendment rights."[43] In 2018, United States District Judge John D. Bates wrote, "scraping plausibly falls within the ambit of the First Amendment."[44] As the law continues to shape in relation to the new wave of datafication, supporting research, journalism, and efforts to detect and enforce policy become a top priority. In an even more recent case, the Ninth Circuit Court of Appeals ruled that a startup, hiQ Labs, can legally scrape publicly available data from LinkedIn despite LinkedIn's argument that the scraping violates user privacy.[45] LinkedIn claimed that the scraping was effectively hacking.[46] HiQ Labs monitors workforce and labor market trends through user profile data, predicting, among other things, when employees are likely to leave their jobs

38 *Id.* at 2.

39 *See, e.g.*, EF Cultural Travel BV v. Explorica, Inc., 274 F.3d 577, 579 (1st Cir. 2001); eBay, Inc. v. Bidder's Edge, Inc., 100 F. Supp. 2d 1058, 1066 (N.D. Cal. 2000); Tess Macapinlac, *The Legality of Web Scraping: A Proposal*, 71 FED. COMM. L.J. 399, 402, 407–08 (2019); Edward Roberts, *Is Web Scraping Illegal? Depends on What the Meaning of the Word Is*, IMPERVA (Sept. 17, 2018), https://www.imperva.com/blog/is-web-scraping-illegal/. For an example of an international web scraping case, see generally Case C-30/14, Ryanair Ltd. v. PR Aviation BV, ECLI:EU:C:2015:10 (Jan. 15, 2015).

40 *See* 18 U.S.C. § 1030(a)(4).

41 18 U.S.C. § 1030(a)(2).

42 *See* 315 F. Supp. 3d 1, 8–9 (D.D.C. 2018).

43 Complaint at 2, 32, Sandvig v. Sessions, 315 F. Supp. 3d 1 (D.D.C. 2018) (No. 1:16-cv-01368).

44 *Sandvig*, 315 F. Supp. 3d at 15.

45 HiQ Labs, Inc. v. LinkedIn Corp., 938 F.3d 985, 994, 1005 (9th Cir. 2019).

46 *See id.* at 992 (LinkedIn warned "if hiQ accessed LinkedIn's data in the future, it would be violating state and federal law.").

and where skills shortages are likely to occur.[47] Judge Marsha Berzon rejected the privacy claims of LinkedIn, writing:

> [T]here is little evidence that LinkedIn users who choose to make their profiles public actually maintain an expectation of privacy with respect to the information that they post publicly, and it is doubtful that they do. . . . And as to the publicly available profiles, the users quite evidently intend them to be accessed by others.[48]

Moreover, the court held that the data was not the property of LinkedIn, but of the users themselves.[49] It also noted that blocking hiQ would force the business to close.[50] At the same time, the court wrote, "LinkedIn could satisfy its 'free rider' concern by eliminating the public access option, albeit at a cost to the preferences of many users and, possibly, to its own bottom line."[51] The court said that in the current design context of the platform, the data was not private, defining private information as "information delineated as private through use of a permission requirement of some sort."[52] These factors led the court to conclude "when a computer network generally permits public access to its data, a user's accessing that publicly available data will not constitute access without authorization under the CFAA."[53] But when a party circumvents safeguards, like a username and password, to gain access to a computer it would likely constitute "without authorization."[54] Though a limited victory, there is opportunity to proactively evolve our privacy and computer regulations to allow access to user generated data.

Similarly, in *Cvent, Inc. v. Eventbrite, Inc.*, the data Eventbrite allegedly took from Cvent's website was "publicly available on the Internet, without requiring any login, password, or other individualized grant of access."[55] While Cvent's Terms of Use prohibited web scraping through denial of unauthorized third-party access, users were not required to manifest assent to the Terms of Use, for example by clicking "I agree," before accessing the database, and anyone, including competitors, could access Cvent's data.[56] In *Fidlar Technologies v. LPS Real Estate Data Solutions, Inc.*, Fidlar developed "software for county offices to manage public land records."[57] LPS, a real

47 *See id.* at 991.

48 *Id.* at 994–95.

49 *See id.* at 995.

50 *See id.* at 996–97.

51 *Id.* at 995.

52 *Id.* at 1001.

53 *Id.* at 1003.

54 *Id.*; *see also* DHI Grp., Inc. v. Kent, No. CV H-16-1670, 2017 WL 4837730, at *5 (S.D. Tex. Oct. 26, 2017) (Company employed anti web-scraping measures like blocking IP addresses and using a firewall.).

55 739 F. Supp. 2d 927, 932 (E.D. Va. 2010).

56 *Id.*; *see also* Craigslist Inc. v. 3Taps Inc., 942 F. Supp. 2d 962, 974 (N.D. Cal. 2013) ("'[A]ll rights' language" in the terms of service "relates specifically to enforcement rights—not rights to the content of the posts.").

57 Fidlar Techs. v. LPS Real Est. Data Sols., Inc., 810 F.3d 1075, 1077 (7th Cir. 2016).

estate data analytics company, used one of Fidlar's products to gather real property data and "designed a 'web-harvester' . . . to download county records en masse" from the eighty-two county databases it subscribed to (and paid fees to).[58] LPS never had a direct contract with Fidlar.[59] Characteristics of Fidlar's services suggested that downloading records through another program, like a web-harvester, was permissible.[60] Fidlar did not implement any encryption, and the data was accessible through other third-party applications.[61] Thus, access through the front-end may have been limited, but "was completely open on the back-end."[62]

Importantly, courts may view companies who attempt to ban web scraping as doing so for anticompetitive goals. LinkedIn's interest of stopping hiQ's web scraping was not enough "to outweigh hiQ's interest in continuing its business, which depends on accessing, analyzing, and communicating information derived from public LinkedIn profiles."[63] The court in *hiQ* noted that if companies with "vast amounts of public data, are permitted selectively to ban only potential competitors from accessing and using that otherwise public data, the result . . . may well be considered unfair competition."[64] But some courts seem to infer wrongdoing when companies are direct competitors.[65]

When the use of web-scraping is for academic purposes, *Sandvig v. Sessions* suggests that web scrapers can assert the First Amendment as a defense.[66] In *Sandvig*, researchers were planning to use web scraping as part of a research paper.[67] The court held "[t]he First Amendment does not give someone the right to breach a paywall on a news website any more than it gives someone the right to steal a newspaper."[68] But, "[h]ere, plaintiffs . . . seek only to prevent the government from prosecuting them for obtaining or using information that the general public can access."[69] Further, the court noted that if a human who creates a bot can read and interact with a website, then the bot should also be allowed access:

58 *Id.* at 1078.

59 *See id.* ("Fidlar was not a party to any of the contracts between LPS and the individual counties.").

60 *Id.* at 1082.

61 *Id.* at 1082–83.

62 *Id.* at 1083.

63 HiQ Labs, Inc. v. LinkedIn Corp., 938 F.3d 985, 995 (9th Cir. 2019).

64 *Id.* at 998.

65 *See, e.g.*, CouponCabin LLC v. Savings.com, Inc., No. 2:14-CV-39, 2017 WL 83337, at *5 (N.D. Ind. Jan. 10, 2017) (holding it reasonable to infer "[defendant] had constructive notice as to the website's Terms and Conditions given that it is a business entity in direct competition with the Plaintiff").

66 *Sandvig v. Sessions*, 315 F. Supp. 3d 1, 15–16 (D.D.C. 2018).

67 *See id.* at 8–9.

68 *Id.* at 13.

69 *Id.* at 17.

> The website might purport to be limiting the identities of those entitled to enter the site, so that humans but not robots can get in. *See Star Wars: Episode IV—A New Hope* (Lucasfilm 1977) ("We don't serve their kind here! . . . Your droids. They'll have to wait outside."). But bots are simply technological tools for humans to more efficiently collect and process information that they could otherwise access manually. *Cf. Star Wars: Episode II—Attack of the Clones* (Lucasfilm 2002) ("[I]f droids could *think*, there'd be none of us here, would there?").[70]

Ultimately, in an age where data extraction is the key to tech's future, web scraping law should be designed to enable more transparency, research, accountability, and competition.

III. DATA BETWEEN HEALTH AND PRIVACY: A GLOBAL PANDEMIC REQUIRES A DIFFERENT BALANCE

Can governments use data reporting technology to further goals, like achieving public health during the COVID-19 pandemic by combatting the virus's widespread transmission? Or does monitoring and tracking citizens too greatly infringe on personal privacy? These questions have become anything but theoretical during the current global pandemic. And the variances in the response to these questions by countries around the world are a reminder that any such line that we strike between competing public goals is a normative and political decision that evolves throughout time and space. Countries like Taiwan, Israel, Singapore, and South Korea did far more to contact trace and minimize the spread of infections than did the United States government, resulting in far fewer needless deaths.[71]

Traditional contact tracing is done manually through in-person interviews at a medical facility and often involves asking patients where they have been, or with whom they have come in contact. For a fast-moving virus such as COVID-19, this method is ineffective. Technology has an advantage over manual reporting and can be used to save countless lives through an implementation process to help track citizens' contact and the spread of the virus. Smartphones and apps are tracking our movement and sharing the most personal details of who we are, who we talk to, and where we have been. This information can be funneled into a database and distributed to the public in order to inform them of potential contact with an infected individual. Of course, personal privacy must be protected. As recently stated by the ACLU, "[w]e need a sober consideration of the risks and tradeoffs of

70 *Id.* at 27.

71 *See, e.g.*, Justin Fendos, *PART I: COVID-19 Contact Tracing: Why South Korea's Success Is Hard to Replicate*, GEO. J. INT'L AFFS. (Oct. 12, 2020), https://gjia.georgetown.edu/2020/10/12/parti-covid-19-contact-tracing-why-south-koreas-success-is-hard-to-replicate/.

such a system so that it protects not only the fundamental right to health, but also our rights of privacy and free association."[72]

There have been various response strategies around the globe as each country has its own battle between the health and wellness of its citizens versus individual privacy concerns. In China, for example, the largest tech companies, Alibaba Group, Tencent Holdings, and Baidu Inc., developed algorithms to create ratings on their respective apps.[73] These apps were required by the government throughout checkpoints in cities in order to restrict travel based on the ratings: green—unrestricted travel; yellow—seven-day quarantine; and red—fourteen-day quarantine.[74] The apps require basic information to sign up such as name, phone number, home address, and national identity card number.[75] Once registered, the apps ask about health status, travel history, and contacts diagnosed with the virus.[76]

In Israel, "the Health Ministry launched a voluntary app" in March "called Hamagen—The Shield in Hebrew," which was downloaded by 1.5 million people out of a population of nine million.[77] Hamagen, a GPS location-based app, focused on voluntary information provided by coronavirus patients, but was deemed inaccurate by the ministry's chief information officer.[78] Meanwhile, Israel has also used Shin Bet security, Israel's internal security service, to track infected citizens using mobile data, despite any privacy concerns.[79] Israel's Supreme Court ruled "a suitable alternative, compatible with the principles of privacy, must be found."[80] In Australia, a poll showed that over half of Australians supported the CovidSafe app, but only 16 percent had actually downloaded it.[81] Even Australia's top

72 DANIEL KAHN GILMOR, PRINCIPLES FOR TECHNOLOGY-ASSISTED CONTACT-TRACING 11 (2020), https://www.aclu.org/report/aclu-white-paper-principles-technology-assisted-contact-tracing.

73 Naomi Xu Elegant & Clay Chandler, *When Red is Unlucky: What We Can Learn from China's Color-Coded Apps for Tracking the Coronavirus Outbreak*, FORTUNE (Apr. 20, 2020, 6:30 AM), https://fortune.com/2020/04/20/china-coronavirus-tracking-apps-color-codes-covid-19-alibaba-tencent-baidu/.

74 *Id.*

75 *Id.*

76 *Id.*

77 Steven Scheer & Tova Cohen, *Israel Extends Coronavirus Cell Phone Surveillance by Three Weeks*, REUTERS (May 5, 2020, 5:43 AM), https://www.reuters.com/article/us-health-coronavirus-israel-surveillanc/israel-extends-coronavirus-cell-phone-surveillance-by-three-weeks-idUSKBN22H11I.

78 *Id.*

79 *Id.*

80 *Id.*

81 Max Koslowski, *Half of Us Say We Support the COVIDSafe App, But Only 16 Percent Have Downloaded It*, SYDNEY MORNING HERALD (May 3, 2020, 12:00 AM), https://amp.smh.com.au/politics/federal/half-of-us-say-we-support-the-covidsafe-app-but-only-16-per-cent-have-downloaded-it-20200501-p54p53.html.

coronavirus advisor with the World Health Organization said she would not download the app due to privacy concerns.[82]

In the United States, similar contact tracing app technology has been created but with limited receptivity. For example, only 1.5 percent of the total population of North and South Dakota had actually downloaded the apps released in their states.[83] In Utah, less than one percent of the population downloaded an available tracing app.[84] Google and Apple teamed up quickly to launch software updates to introduce an exposure notification through their application programming interface.[85] This update would support COVID-19 contact tracing apps from public health authorities.[86] This software update's goal was to "reduce the spread of the virus, with user privacy and security central to the design."[87] Rather than automatically tracking users' private information, users still need to download a contact tracing app in order for an exposure notification to come through the phone. Similar to many other updates, certain features can always be turned off by going into the settings.

CONCLUSION

From sharing labor market information to data scraping to digital contact tracing, both governments and private actors can employ legal entrepreneurship to promote public ends, including health and safety, equality, and fraud detection. As in the past, such ends must constantly be balanced against other public values, including privacy. But identifying these biopolitical opportunities is key to expanding the capabilities of extracting massive amounts of information from digital life beyond the scope of private profit-driven sectors' current use and toward shared public goals. The tradeoffs, benefits, and costs of datafication not only evolve over time, but also may be strikingly different from context to context. During a health crisis, technology can be used to save countless lives through an implementation process to help track the spread of a deadly virus. The balance between preserving privacy and health, freedom of speech, equality, and many other democratic values will continue to be at the heart of any new technological capability. Cohen's landmark book on the policy of data warns against abuses of power and imbalances in our normative commitments. But it is also a call

82 *Id.*

83 Elliot Setzer, *Contact-Tracing Apps in the United* States, LAWFARE (May 6, 2020, 4:08 PM), https://www.lawfareblog.com/contact-tracing-apps-united-states.

84 *Id.*

85 *See Apple, Google Collaborate on COVID-19 Contact-Tracing Apps*, WASH. INTERNET DAILY (Apr. 13, 2020), https://washingtoninternetdaily.com/article/view?BC=bc_606f9c963657e&search_id=447000&id=143193&p=1.

86 *Id.*

87 *Id.*

for a more positive vision, one which can only be constructed through engagement with new capabilities and recognition of the opportunities that data can offer.