

Clyde Cabico

Professor DRT

October 14, 2024

Case Analysis 3

Jason Tan exposes the ethical concerns of data scraping¹ in the confines of LinkedIn. For those unfamiliar, the author clearly describes data scraping as analogous to automating the process of gathering book prices individually, using technology to extract targeted information from LinkedIn's website for collection and use. A clear distinction is made regarding its legality and debatable ethical side that may lead to malicious activities. Using innovative software programs and technology has been shown to improve work output and ensure quality. However, in the article, Tan explains, "Data scraping, in its essence, is not illegal. However, LinkedIn's position is that unauthorized scraping violates its Terms of Service and is thus not allowed on its platform," showcasing that although data scraping is not in violation of any laws, private entities can further implement restrictions on its services that one must abide by or face penalties. One of the examples involved hiQ, which courts ruled in favor of, seeing no lawful violation of their data scraping practices on LinkedIn. Ultimately, hiQ fell victim to the violation set forth by LinkedIn's Terms of Use and User Agreement governing data scraping.

¹ https://en.wikipedia.org/wiki/Data_scraping - Using program software to extract desired information for collection

In this case analysis, I will argue that using scraped data from LinkedIn for developing training materials, according to virtue ethics, is not morally permissible. While it may serve practical purposes, it fails to align with the virtues of justice, judgment, and integrity, as it disrespects users' expectations of privacy. Virtue ethics is more than a set of individual virtues; it is about developing character by consistently acting in ways that align with moral integrity. This framework promotes doing the right things for the right reasons, which requires cultivating virtues such as justice, integrity, and respect through habitual practice. Ethical decision-making should not just satisfy practical objectives, but should contribute to a trustworthy character. Meaning that, organizations handling data must strive to consistently consider ethical implications to cultivate a reputation for respecting user privacy.

Michael Zimmer's case study "But the data is already public," shows how the complexities of privacy in a seemingly public environment create ethical challenges. Zimmer's analysis focuses on Facebook, where researchers failed to consider user' privacy expectations despite data being publicly accessible. Tan's literature, however, looks at LinkedIn, where data scraping is legal but restricted by LinkedIn's terms. In both cases, the disregard for user privacy illustrates a failure to uphold virtues central to virtue ethics. Also, both Facebook and LinkedIn have publicly available information on display and Zimmer introduces the key concept that, although the data is publicly accessible, it does not mean it is ethically acceptable to use it without permission. A critical flaw was overlooked in the research, "researchers concede that one RA might have different access to a student's profile than a different RA." (Zimmer) Researchers overlooked how privacy settings and

user expectations regarding data visibility vary depending on the context. Researchers failed to account for nuanced privacy controls, highlighting a broader ethical issue—public access does not equate to universal permission. The theory of virtue ethics in Zimmer and Tan's pieces of literature circulates on the concept of justice, which requires fairness and respect for others. In both cases, an organization that values integrity would assess user expectations of privacy, acknowledging that a public accessibility does not imply permission. Both social media platforms are publicly accessible, though people cannot ignore or disregard the user's expectations of the privacy of their data. "we have not accessed any information not otherwise available on Facebook," (Zimmer) Displaces the moral theory of virtue ethics, the researchers fail to acknowledge their failure in overlooking to protect users' identities. The LinkedIn article, like Zimmer's case study, features how data scraping, though, "isn't a violation of the CFAA as defined by U.S. law," (Tan) crosses ethical boundaries by exploiting information without respect for terms of service or the users' expectations.

Organizations undermine virtue ethics in terms of integrity and adherence to ethical standards when they take advantage of data-scraping alert systems by setting delays between actions and limiting scraping activity, as mentioned in the text. (Tan) A virtuous organization would act with integrity by adhering to ethical standards and following rules or provisions. For organizations, virtue ethics would entail upholding integrity and justice by obtaining user consent, even if data scraping is legally permissible. This ethical stance respects the dignity and privacy of individuals, fostering a virtuous culture within corporate settings. Exceeding baseline standards, organizations should reach for the maxima of

obtaining permission or consent from a user with a clear purpose for its use. Returning to the prompt, the HR department's decision to scrape data without consent violates the virtues of justice, integrity, and adherence to ethical standards. The implied consent of LinkedIn users is blatantly ignored. In virtue ethics, actions should be driven by the intention to do good, respecting others' privacy as an end in itself rather than a legal formality. By scraping LinkedIn data without consent, HR departments disregard this ethical duty, prioritizing efficiency over respect for user trust. Yet, although public, LinkedIn users may not expect that their publicly shared information may be used for other means than what it was originally intended for.

O'Neil's literature, "Weapons of Mass Destruction," revolves around the concept that predictive algorithms in software and technology in law enforcement are not only biased but also disproportionately affect certain communities. O'Neil states, "nuisance crimes are endemic to many impoverished neighborhoods," clearly explaining that the use of these crime-predicting tools exacerbates the divide between the wealthy and the poor. Although helpful to hiring managers, bias is created like LinkedIn data scraping. In both instances, the benefits undermine the ethical consequences, as the use without considering broader context leads to unfair outcomes. Oneil explains that "it would take remarkable restraint not to let loads of nuisance data flow into their predictive models," that the tools that aid law enforcement precincts that are undermanned are controlled through bias and reinforce existing inequalities, targeting already disadvantaged communities. In the same context, LinkedIn data scraping, while beneficial for creating efficient training materials, risks embedding biases in hiring practices by pulling data that

reflects only certain segments of the population or professional backgrounds.

Organizations operating with a virtuous character would weigh such ethical consequences carefully, emphasizing fairness and respect for individual rights over efficiency gains. More importantly, the violation of privacy and trust is a key concern, as individuals did not consent to having their information used in this way. In both cases, the tools may serve a practical purpose, but they do so at the cost of compromising moral principles: justice and integrity. Also, both cases fail to account for the full ethical implications, such as the reinforcement of systematic biases, and the negation of virtuous ethical decisions that should guide decision-making in both law enforcement and corporate environments. O'Neil also highlights that through the previously mentioned, data misuse can occur. O'Neil supports this possibility with the example of the unconstitutional searches and seizures that occurred in New York under Bloomberg's administration. The fact that the tools that can help direct law enforcement to patrol probable areas of crime are at law enforcement's full control of modification, can ultimately lead to biased policing and the over-surveillance of minority communities. With LinkedIn data scraping, the HR department's use of scraped data for hiring or training could lead to biased hiring decisions and disproportionately represent certain demographics or professional backgrounds. This parallels the harm done by predictive policing, where the benefits are outweighed by the ethical consequences of reinforcing discrimination and violating trust and integrity.

The stance that an HR department of a mid-sized private company wants to use data scraped from LinkedIn to develop training materials for new Hiring Managers brings forth ethical challenges. These challenges are explained thoroughly in the three articles,

Zimmer, especially. The researchers sought to produce new knowledge for the community of sociology. Yet, in doing so, they overlooked the ethical implications of using publicly available data without proper consent. Congruent, the HR department's reliance on scraped LinkedIn data disregards the context in which users share their information, prioritizing convenience over ethical considerations. The purpose is for a good cause: training and efficiency, but justice and integrity are the expenses. Likewise, in the military, character-building is central to developing trust and respect. Only through daily repetition can such character be built. Similarly, organizations must consistently practice virtuous ethics consistently to develop ethical habits that respect privacy to build trust with users. Core virtues are violated and a more ethical approach should be sought after. Some may argue that, because the data is public and the actions lawful, using it for training purposes is justified. Not only that, but the purpose serves to benefit training and knowledge, can be argued. Such rebuttal overlooks the broader ethical responsibility companies have toward respecting user privacy and trust. Ultimately, virtue ethics calls for actions that respect the inherent dignity of individuals. While public data may be legally accessible, a commitment to virtue requires organizations to consider the ethical implications and potential harm caused by their actions. Again, although an HR department of a mid-sized private company wants to use data scraped from LinkedIn to develop training materials for new Hiring Managers is lawful, trust between users and platforms is tarnished.

Works Cited

- O'Neil, Cathy. "Civilian Casualties: Justice in the Age of Big Data." *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown, 2016, pp.
- Tan, Jason. "Data Scraping on LinkedIn: Legality, Consequences & Best Practices." *Engage AI*, <https://engage-ai.co/linkedin-data-scraping-legality-consequences-best-practices/>. Accessed 14 Oct. 2024.
- Zimmer, Michael. "But the Data is Already Public: On the Ethics of Research in Facebook." *Ethics and Information Technology*, vol. 12, no. 4, 2010, pp. 313-325.
DOI:10.1007/s10676-010-9227-5.