DOXING ETHICS

Juhani Naskali, Minna Rantanen, Maria Rottenkolber, Kai K. Kimppa

University of Turku, Turku School of Economics (Finland)

juhani.naskali@utu.fi; mimaran@utu.fi; maria.m.rottenkolber@utu.fi; kai.kimppa@utu.fi

EXTENDED ABSTRACT

Doxing is a practice where a third party, i.e., one or several doxer(s), intentionally publishes personal information about another individual, the doxee or target, without consent on the Internet (Douglas, 2016; Eckert and Metzger-Riftkin, 2020). The information revealed may include victims' real names, home addresses, or telephone number, among others. Thus, doxing can be considered a user-led violation of privacy (Trottier, 2017). In public discourse, doxing frequently holds an exclusively negative connotation (Barry, 2021), and academic literature has referred to doxing as a form of "problematic speech on social media" (Fleischman and Rosenbloom, 2020). While doxing may be intuitively condemned as a form of online harassment, there may be cases in which it serves as an ethically justified means of resistance (Cheung, 2021).

Utilitarianism is an ethical theory that advocates for actions to be judged based on their consequences. It posits that the moral worth of an action lies in its ability to maximize well-being, and to promote the greatest overall happiness or utility to the greatest number of people, often referred to as the "greatest happiness principle".

This paper investigates the phenomenon of doxing to answer the question of how to ethically evaluate instances of doxing from a utilitarian perspective. Specifically, this paper aims to examine whether doxing can be categorically considered ethically problematic or whether a more nuanced understanding of doxing is needed.

A utilitarian analysis of doxing must first identify the different types of doxing actions and the relevant groups of people whom are influenced by the consequences of doxing. Then, an analysis is conducted on the utility of the consequences of the identified types of doxing actions to these groups. The analysis includes indirect consequences, such as ridicule or financial and physical harm to the target.

There are many different ways to categorize doxing (Anguita, 2021; Cheung, 2021; MacAllister, 2017; Snyder et al., 2017). For the purposes of this paper, a high-level categorization is the most fruitful. Douglas (2016) identifies three main forms of doxing: 1) deanonymizing doxing, which identifies a previously anonymous or pseudonymous person; 2) targeting doxing, which makes it easier to physically locate and contact the target, possibly increasing the risk of harassment or physical harm; 3) delegitimizing doxing, which undermines the credibility of the target.

Information that suggests that the victim has breached a social norm or committed an immoral act, for instance, can undermine the target's reputation. While targeting doxing provides the means to harass the target, delegitimizing doxing can provide a 'reason' for harassment. (Douglas, 2016.)

Special cases of doxing have been identified previously. For example, if an individual uses anonymity as a means to avoid being held accountable for wrongdoing or to mislead others, the public might have a legitimate interest in revealing this wrongdoing by uncovering the person's identity and thereby removing anonymity (Douglas, 2020). This disclosure increases the public's ability to hold the wrongdoer accountable as such doxing can help stop such wrongdoing in the future and signal to the broader community that such wrongdoing is not tolerated and has negative consequences.

The affected groups in doxing are a) the target of doxing, b) people close to the target, such as friends and family, who can be directly affected by the situation, c) the doxer, d) people who benefit from the doxing (e.g., when doxing reveals illegal misdeeds), whether legitimately or illegitimately, and e) the public at large.

In deanonymizing doxing, the target suffers the loss of protection that anonymity provides (1a). This can lead to online harassment and threats, resulting in emotional distress, fear, a sense of insecurity and loss of privacy. Additionally, their personal and professional life may be adversely affected by the reactions of their friends, families and employers, leading to reputational damage or job loss. Douglas (2020) posits that any shaming that accompanies doxing is only permissible if it is reintegrative: *"Without the possibility that those who are exposed by digital vigilantism can be reintegrated into their communities, DV risks further alienating them and reinforcing their extreme views."* From a utilitarian perspective, permanent consequences have a much stronger utility (whether negative or positive). Sometimes harassment and threats can bleed over to the close ones of the target, who can also face harassment or threats due to their association with the target (1b). The doxer presumably has a goal in mind with their doxing, which yields utility to them (1c). If there are people who directly benefit from the information released with the doxing (e.g. financial records proving people were subject to fraud would benefit the victims, as well as the investigators), they may receive positive utility from it. The public at large may benefit from doxing that combats negative behavior (1d) if the behavior stops, and/or people are less likely to repeat the behavior in the future due to it being costly.

Targeting doxing holds the same negative utility for the target of the doxing (2a), and in addition, holds the possibility of physical harm, stalking and other offline harassment, while simultaneously holding a greater chance of psychological trauma, fear and diminished quality of life. The same applies to their close ones (2b). Unless their goal is physical harm, utility to the doxer does not change (2c). In cases where the target poses a genuine threat to others (e.g. terrorism), targeting doxing can have a high positive utility to law enforcement and potential victims of the target (2d). Utility to the public is likely to suffer with the inclusion of location information, as the threat of physical harm can lead to a culture that legitimizes vigilantism, harassment and the fear they lead to. In summary, targeting doxing has a much lower net utility than deanonymizing doxing unless it leads to the prevention of great tragedies, such as in the case of preventing physical calamities such as school shootings or terrorist attacks.

Delegitimizing doxing holds a larger negative utility for the target, in comparison to deanonymizing doxing, as they further take a hit to their credibility (3a). Delegitimizing doxing is an "attempt to shame and humiliate the subject, often by portraying her as a transgressor of an established (or supposed) social norm" (Douglas, 2016), and as such, it can weigh heavily on the target, and the change of social ostracism and loss of financial opportunity is much greater. Their close ones are similarly affected negatively (3b). While the doxer might experience a stronger sense of schadenfreude compared to deanonymizing doxing, it is generally short-lived and superficial (3c). It is difficult to imagine circumstances where a group of people would be positively affected by delegitimizing doxing—surely such utility is provided by evidence or wrongdoing or the acts of the doxing target, and not from the delegitimizing (at worst, dehumanizing) the target. Even deanonymizing doxing can lead to delegitimizing the target, but delegitimizing doxing is done for the express “intention of undermining the target’s credibility, reputation and/or character” (Douglas, 2016), which brings no further utility to others (3d). Delegitimizing doxing can even harm the public in “maintaining the ‘tyranny of majority’ that concerned John Stuart Mill” (Douglas, 2016), contributing to a culture of public shaming (3e). While some members of the public can view this as a means to expose hypocrisy, such acts can have negative consequences on free expression and public discourse that ultimately leads to better understanding.

Interestingly, the group with the most variability in the utility of doxing is d) the people who possibly benefit from the doxing. If doxing prevents a school shooting, for example, the negative utility of multiple shooting victims greatly outweighs the negative utility of the doxing target’s loss of reputation and likely incarceration, were such a situation prevented. Outside of such extreme examples, doxing does not yield any positive consequences to the public, outside of deterrence to breaking social norms (general utility to e), the public), and it becomes much harder to overshadow the negative consequences to the target, especially in the case of targeted doxing.

In accordance with Douglas (2016), the ethical analysis finds that morally permissible cases of doxing need to serve the public interest without violating several side-constraints. The benefits to the public of exposing the victim’s wrongdoing need to outweigh the harms to the victim. Douglas’ original conceptual analysis was criticized by Barry (2021) to the extent that Douglas’s specific consequentialist approach remains unclear, and that it is not explicitly stated what makes a public interest “compelling”. We hope that this more detailed utilitarian analysis helps alleviate these concerns.

To conclude, balancing the potential benefits of accountability, deterrence, and public safety with the potential harms of privacy invasion, reputational damage, and emotional distress is crucial to addressing doxing in a responsible and balanced manner. For doxing to be morally permissible from a utilitarian viewpoint, it needs to bring benefit to others—either in reconciling a previous wrong or preventing future suffering.

KEYWORDS: Doxing, social media, freedom of speech, utilitarianism, ethics.

REFERENCES

Anguita, P. (2021). Freedom of expression in social networks and doxing. In *The Handbook of Communication Rights, Law, and Ethics: Seeking Universality, Equality, Freedom and Dignity*, pages 279-291. John Wiley & Sons.

Barry, P. B. (2021). Doxing racists. *The Journal of Value Inquiry*, 55(3):457- 474.

Cheung, A. (2021). Doxing and the challenge to legal regulation: When personal data become a weapon. In Bailey, J., Flynn, A., and Henry, N., editors, *The Emerald International Handbook of Technology-Facilitated Violence and Abuse*, pages 577-594. Emerald Publishing Limited.

Douglas, D. M. (2016). Doxing: a conceptual analysis. *Ethics and Information Technology*, 18(3):199-210.

Douglas, D. M. (2020). Doxing as audience vigilantism against hate speech. In Trottier, D., Gabdulhakov, R., and Huang, Q., editors, *Introducing Vigilant Audiences*, volume 259, pages 259-280. Open Book Publishers Cambridge.

Eckert, S. and Metzger-Riftkin, J. (2020). Doxxing. In Ross, K., Bachmann, I., Cardo, V., Moorti, S., and Scarcelli, M., editors, *The International Encyclopedia of Gender, Media, and Communication*, pages 1-5. Major Reference Works.

Fleischman, W. and Rosenbloom, L. (2020). Problems with problematic speech on social media. In Pelegr´ım-Borondo, J., Arias-Oliva, M., Murata, K., and Palma, A. M. L., editors, *Paradigm Shifts in ICT Ethics: Proceedings of the Ethicomp 2020*, pages 116-120.

MacAllister, J. M. (2017). The doxing dilemma: seeking a remedy for the malicious publication of personal information. *Fordham Law Review*, 85(4):2451- 2483.

Snyder, P., Doerfler, P., Kanich, C., and McCoy, D. (2017). Fifteen minutes of unwanted fame: Detecting and characterizing doxing. In *Proceedings of the 2017 Internet Measurement Conference*, IMC ’17, pages 432-444. Association for Computing Machinery. event-place: London, United Kingdom.

Trottier, D. (2017). Digital vigilantism as weaponisation of visibility. *Philosophy & Technology*, 30(1):55-72.