Doxing:
A Literature Review

Contents

Introduction..... 4

The Literature 6

Discussion 15

Conclusion 19

Bibliography 20

This literature review is the first deliverable of a project looking at pro-Russian doxing practices against pro-Ukraine foreign individuals based in Ukraine and abroad. The project, entitled “Anti-Dox: Identifying, Evaluating and Countering Disinformation in Times of War”, is supported by the European Media and Information Fund: Investigations into Disinformation Dynamics, and led by the think-and-do-thank International Centre for Counter-Terrorism (ICCT), based in the Netherlands, together with the Fundacja Reporterów (FR), a group of investigative reporters and fact-checkers based in Poland. The project aims to investigate and evaluate doxing with a view to help counter Russia’s disinformation campaign against Europe.

Introduction

The word “doxing” (sometimes “doxxing”) is made up of the words “dropping dox” whereby dox, an abbreviation of the word “document”, refers to personal information (Strandell, 2024). Doxing, or revealing personal information in the online public space with the general intent of causing harm, is increasingly being used in modern armed conflicts. For example, Ukraine's military has released private information of over 100,000 Russian soldiers, including alleged war criminals and FSB officials, in multiple doxing campaigns (Jensen and Watts, 2022). On the other hand, hackers from Russian hacker group RaHDit have published data on more than 3,000 Ukrainian Armed Forces mercenaries (*Rossa Primavera,* 27 July 2024), in addition to leaking information on 7,700 Azov soldiers (*Al Mayadeen,* 28 August 2024). Another group of Russian hackers, EvilWeb, leaked data from Ukraine’s Security Service (SBU), including IP addresses, emails, and encryption keys of SBU employees (*URA,* 29 September 2024). Finally, members of the Russian hacker project "NemeZida" revealed the identities of 800 Ukrainian Armed Forces soldiers who participated in the attack on the Kursk region, including representatives of the 82nd Airborne Assault Brigade, the 61st Mechanized Brigade of the Ukrainian Armed Forces, as well as about 200 foreign mercenaries from Israel, Armenia, Georgia, Kazakhstan, and Syria (*URA,* 29 September 2024).

In a non-conflict environment, doxing may serve the purpose of extorting, silencing, controlling, or serving the public interest (Snyder 2017, p. 438; Anderson 2021, pp. 208-9; Li 2023, p. 368). In short, the role of doxing in today’s strategies to gain or retain power over enemy actors or rival factions is prominent and more relevant than ever before. This raises questions over the nature and legitimacy of doxing, including what (and who) exactly is that doxing involves, what makes a particular case of doxing ethically acceptable, or whether the practice should be seen as a crime or as a means for anti-repression activism.

In order to shed light upon such questions, this literature review provides findings on academic discussions around doxing, from its conceptual or theoretical understanding to its real-life forms and implications. It does so by assessing a total of 17 peer-reviewed research papers published in the time span of 10 years (2014-2024). The contributions include approaches to the subject by scholars from the Social Sciences, the Data Sciences, and Public Health, located across North America, Europe, and Asia. Selecting the material involved open-source methodology (OSINT), with keywords including both scholarly and culturally sensitive vocabulary, especially in relation to state surveillance and the misuse of data sharing. For example, “dox” AND “anti-dox”, “doxing” AND “legal”, “doxing” AND “vigilantism”, or “doxing” OR “doxed” AND “security” as well as “doxing” AND “malicious” retrieved relevant sources. Due to part of the academic discussion on doxing being morality-based, namely whether it is “good” or “bad”, which is an inherently subjective assessment, research contributions were not disregarded based on their moral assessment, thereby allowing for this review to be nuanced and whole-encompassing.

As a way to complement scholarly contributions with some preliminary data on the subject of doxing, the discussion following the literature summary includes data from semi-structured interviews with individuals who have been, or are, victims of doxing. In particular, such conversations took place with combatants, humanitarian workers, and journalists active in conflict zones, specifically in Ukraine. The incorporation of real and direct testimonies to doxing allows for a more nuanced grasp of the nature and impact of the practice, and helps fill out some gaps found in the literature, namely state-sponsored and/or state-supported doxing in the context of war. Indeed, among the existing literature, the only scholars that address the subject of doxing and conflict are Jensen and Watts from Brigham Young University and the United States Military Academy, respectively. While their work illustrates the use of doxing on enemy soldiers, it does so focusing solely on Ukraine's current tactics against Russian soldiers. This analysis contributes to such work by adding evidence on pro-Russia combatants' doxing tactics against pro-Ukraine individuals in the area and abroad.

This literature review contains a summary of findings, which includes a chronological content analysis of the scholarly contributions to the subject together with data from the above-mentioned interviews. Such an analysis is followed by a brief discussion, designed to stress points of agreement and disagreement between the authors, namely around conceptual approaches to doxing, its ethical use, and its legality. It ends with a conclusion section synthesising the results of the literature review and highlighting where our project, "Anti-Dox: Identifying, Evaluating, and Countering Disinformation in Times of War", hopes to contribute to current debates on the subject. Ultimately, this analysis aims to situate the project into an evidence-based conversation in which doxing is considered a form of harmful information spread, characterised by actors employing manipulation tactics to advance political, military, or commercial goals.

The Literature

In 2014, Matthews *et al.* from the Concordia University of Edmonton, Canada, published "A study of doxxing its security implications and mitigation strategies for organizations". After defining doxing as the "overt collection, aggregation and publication of information of a targeted individual (without his/her consent) on the internet [...] with the intention of causing embarrassment, humiliation and damages'' (p.1), the article examines how doxing leads to severe threats like hacking and harassment, particularly affecting organizations' reputation and market advantage. It emphasizes the need to include doxing in risk management strategies and offers mitigation recommendations. Specifically, it proposes a model designed to mitigate doxing threats, especially in organisations at higher risk, such as small businesses preparing for acquisitions, or those with high-profile employees (p. 3). The model is structured into three layers: preventive, detection, and corrective/compensation (p. 5-8), forming together a "defence in depth" strategy (p. 3). The preventive measures emphasise the need for strong policies and procedures to safeguard sensitive information. For instance, organizations are encouraged to set up clear guidelines for managing data, conduct regular policy reviews, and enforce strict protocols on data retention (p. 5). The detection layer combines technological tools to monitor potential doxing threats. It focuses on the continuous tracking of the organization's digital footprint using OSINT tools like Maltego and FOCA (p. 6). Once doxing is detected, an IT unit should quickly assess the threat, identify compromised information, and inform leadership if needed (p. 7). Finally, containment actions are proposed to prevent further damage, such as tightening network access (p. 8).

The most prominent theoretical foundation to understand the types of doxing was published by Douglas (2016), from the Australian University of Queensland, under the title "Doxing: a conceptual analysis". The paper develops a typology of different doxing types (deanonymising, targeting, delegitimising) and a test to assess whether a doxing incident is ethically justified. *Deanonymising* refers to the release of personal information that reveals the identity of a formerly anonymous individual (pp. 203-204). *Targeting* refers to the release of information disclosing an individual's whereabouts (pp. 204-205). *Delegitimising* refers to the release of intimate personal information that undermines the victim's credibility (pp. 205-206). The information revealed in doxing incidents does not necessarily have to be new; it may already be available in scattered form across the Internet. What makes it doxing is the systematic bundling of this information and the context in which it is released (p. 205). To test whether doxing can be ethically justified, the author suggests weighing the benefit to the public against the consequences for the victim, excluding curiosity as a legitimate public interest (pp. 206-209). Based on this consideration, the author calls for even justified instances of doxing to be reduced to the minimum necessary to correct the wrongdoing (p. 208).

In 2017, Snyder *et al.* at University of Illinois in Chicago and New York University published "Fifteen minutes of unwanted fame: detecting and characterizing doxing". This article defines doxing as "an attack where a victim's private information is released publicly online" (p. 432). It quantitatively identifies doxed files on three platforms prone to doxing, namely pastebin.com, 4chan.org, and 8ch.net, by means of a computational pipeline model. Four overarching motives for doxing were extracted from the research: *competitiveness*, to demonstrate one's own doxing abilities; *revenge*, to retaliate for something the doxee has previously done to the doxer; *justice*, to avenge a wrongdoing of the doxee, although not personally against the doxer; and *political*, for the sake of a higher goal (p. 438). With regard to the measured effects for the victims, the paper finds a substantial decrease in the openness (update of privacy settings to either private or deletion of account) of doxees' social media accounts in the first two weeks following the doxing (p. 440). This indicates that victims try to minimise the potential harm following their doxing (p. 440). It is worth noting that the authors demonstrate that substantial harm from doxing does not only result from the doxing itself but also from an often-following harassment campaign (p. 442).

Chen *et al.* at The Hong Kong Polytechnic University and The University of Hong Kong published in 2018 "Doxxing victimization and emotional problems among secondary school students in Hong Kong". The paper investigates the effects that doxing victimisation causes on high school students' depression, anxiety and stress. For this purpose, the authors survey 2120 students from different high schools and socioeconomic backgrounds across Hong Kong. The research explores the correlation between the types of personal information disclosed, the identity of the perpetrators, and the platforms where doxing incidents occur, with consequent impacts on the victims' psychological health. Understanding doxing as the act of "searching for and publishing private or identifying information about a particular individual on the Internet, typically with malicious intent" (p. 1), the research finds that doxing by classmates was the most harmful and also the most common practice for this case study, along with other people known to the victims (p. 4), leading to the highest levels of anxiety and depression among victims (p. 6). This suggests that peer-driven doxing amplifies the emotional distress due to ongoing social interactions and the fear of judgment from peers (p. 5). The platform on which the doxing occurred also played a significant role in the emotional impact on victims, for example, doxing incidents on Instant Messaging apps and social networking sites were the most common, with significant correlations to increased anxiety and stress (p. 5). Based on these results, the authors highlight the importance of developing preventive educational programs to raise awareness among adolescents about online privacy and the long-term psychological risks of sharing personal information (p. 6).

In 2019, Trottier, at the Erasmus University Rotterdam in the Netherlands, released "Denunciation and doxing: towards a conceptual model of digital vigilantism". The paper develops a conceptual model of internet vigilantism based on its underlying coordination, moral,

and communicative aspects. It does not employ an empirical methodology. In this contribution, doxing is described as a form or a manifestation of "internet vigilantism", or a "set of practices to scrutinise, denounce and even leverage harm against those deemed to transgress legal and/or moral boundaries, with the intention of achieving some form of justice" (p. 197). Online vigilance can be passive, through observation, or active, through actions like liking, sharing, subscribing, or commenting (p. 204). The author argues that doxing may occur to maintain or heighten a victim's public visibility after such a victim has been identified as a perpetrator for an "offence". For example, it may happen that citizens feel that the state is not adequately addressing moral issues, and then they may take it upon themselves to scrutinise public life and react by doxing public servants (p. 205).

The book chapter "Doxxing and the challenge to legal regulation: When personal data become a weapon" by A. Cheung (2021), in J. Bailey, A. Flynn, & N. Henry (Eds.), *The Emerald international handbook of technology facilitated violence and abuse,* advocates for a more nuanced definition of doxing, particularly when it is used as a form of political pushback to expose misconduct of individuals tasked with civic duties. It focuses specifically on the rise of doxing in Hong Kong against police officers during the anti-government protests of 2019. In this work, Cheung defines doxing as "the intentional public release on the internet of personal data that can be used to identify or locate an individual without their consent" and it describes it as a practice that is gaining popularity as a method for social action (p. 578). The author argues that legal actions should target identifiable defendants to uphold natural justice, ensuring proper notice. In contrast, Hong Kong courts issued broad injunctions against vaguely defined groups involved in doxing, without specifying individuals (p. 586). So, while the courts acted to safeguard privacy, the authorities essentially ignored the rights of dissenting citizens. Therefore, there should be a reassessment of how doxing cases are handled, advocating for a more balanced approach that upholds privacy rights without infringing upon freedom of expression. Cheung concludes by suggesting that any future reforms to Hong Kong's privacy laws should consider introducing a public interest defence (p. 589), which could allow the disclosure of personal data when it is genuinely intended to hold public officials accountable for their actions. In other words, courts need to differentiate between cases where doxing serves the public interest and those where it is used to intimidate or silence individuals (p. 590- 591).

In 2022, Liu, from United States District Court for the District of Columbia, published "Doxfare as a Tool for Strategic Deterrence". Using deterrence theory, the paper deems traditional deterrence measures from the Cold War as no longer effective with regard to cyberwarfare. Instead, it argues for doxfare as an effective, legally, and ethically justified deterrence measure. Arguably, doxfare falls below an escalation threshold because it can be executed highly precisely and constitutes no direct use of force, as unlike cyber operations, it can hardly disrupt or physically impact a state's actionability (p. 75). Instead, doxfare is directed at the public opinion of those involved in hostile cyber activities in the targeted country (p. 75). Considering that

authoritarian or totalitarian regimes, which are mainly responsible for hostile cyber activities against the US, rely heavily on a favourable perception of themselves by their population, doxfare may be a particularly effective tool (p. 75). From a legal perspective, the author considers doxfare to be in accordance with US law as long as the revealed information is true (pp. 76-77). Likewise, the author considers doxfare to be in accordance with international law because it does not deprive the target state of its free will and therefore does not violate the non-intervention rule (p. 78). However, the author notes that nonetheless, doxfare could violate international law, namely the human right to privacy of the target individual(s) (p. 79). Similarly, in light of Russia's hybrid disinformation warfare, legal scholars increasingly call for replacing the "coercive" test of the non-intervention rule with a "disruptive" test, which would likely render doxfare illegal under international law (p. 79). In terms of ethical concerns, based on Douglas' (2016) doxing typology (deanonymising, targeting, delegitimising), the author considers deanonymising and delegitimising as justified because both can influence public opinion, but not targeting (p. 76).

Jensen and Watts, from Brigham Young University and the United States Military Academy, wrote on doxing in the context of the Ukraine-Russia conflict in "Ukraine Symposium - Doxing Enemy Soldiers and the Law of War" (2022). The article analyses the lawfulness of doxing enemy soldiers under international law, using Ukraine's current tactic of doxing Russian soldiers as a case study. The Third and Fourth Geneva Conventions protect prisoners of war and civilians from actions that humiliate them or violate their dignity, including their privacy. However, the doxing of active combatants seems generally lawful. A doxing operation in itself arguably falls short of the *direct violence* threshold required for an attack, meaning it does not necessitate specific precautionary measures. Moreover, the authors argue that, while broad "dead or alive solicitations" are inconsistent with international law, targeting specific combatants with the intent to kill them, even when preceded by doxing, is not inherently a violation of international law. The authors assess that doxing aligns with the principles of humanity (it does not cause unnecessary suffering) and military necessity (it serves to defeat the enemy efficiently and is not prohibited by the laws of war). Even if doxing causes mental suffering, it can be justified if it brings a concrete military advantage and does not "cruelly reveal extraordinarily sensitive personal details." The authors conclude that, with a few exceptions, doxing that is directed at active combatants is lawful under international law and can be compared to other information operations, including propaganda.

In 2023, Li and Whitworth, from the University of New South Wales and University of Sydney, published "Coordinating and doxing data: Hong Kong protesters' and government supporters' data strategies in the age of datafication", which examines the practice of doxing, or the "public exposure of private documents" with malicious intent (p. 359), by both anti- and pro-government supporters during the 2019 protests in Hong Kong. During the protests, China responded with increasingly repressive measures, including the violent suppression of protesters by the police (p.

356). In reaction, protesters began doxing police officers, and over time extended this to their families, aiming to deanonymise them and hold them accountable (p. 361). On the LIHKG social media platform as well as the anti-government Telegram channel "Dadfindboy," which was specifically set up to dox police officers, doxing was framed as a justified punishment of the police, emphasising the power disparity between the protesters and the government. This positioned doxing as a means in the struggle for a higher purpose, namely democracy (pp. 366-367). Doxing in this context also constitutes a form of datafication with significant implications: through collective doxing, police officers are positioned as a definable community and stripped of their role as representatives, becoming (physically) tangible to the protesters (pp. 367-368). Interestingly, the practice of doxing evolved over the course of the protests: as the calls for, and the revealing of, information about police officers became increasingly refined, pro-government supporters launched a similar campaign to dox protesters (p. 361, 364). On the pro-government Telegram channel "Youcangotojail," doxing was justified as a legitimate measure against protesters, who were framed as rioters, in order to maintain societal stability (p. 367). The study highlights that doxing can be employed both as a form of *sousveillance* (a way for individuals or groups to monitor the state and regain power from below) and as a form of *surveillance* (positioning oneself as an agent of the state and claiming to exercise power in its name when it is believed the government is not taking sufficient action) (p. 368).

Schoenebeck *et al.* from the University of Michigan wrote “Online Harassment: Assessing harms and remedies” in 2023, a paper examining the harm resulting from different types of online harassment, including doxing, as well as variations in its severity. By conducting surveys, the research found that doxing is rated highest in physical harm, sharing sexual photos in psychological harm, and both sharing sexual photos and receiving unsolicited photos in sexual harm (p. 6). With regard to demographic background, women and non-white individuals generally reported a higher perceived harm (p. 7). In terms of responses, survey participants reported the highest preference for user ban, followed by content removal and public listing (p. 7). Identity characteristics had a significant influence on the remedy type preference (e.g. women reported lower satisfaction with monetary compensation for doxing but higher satisfaction with content removal compared to men) (p. 7). In general, the study shows that underrepresented groups (e.g. women or people of colour) perceived the highest harm (p. 8). As such, the paper recommends including the voices of victims in remedy mechanisms, which are currently largely perpetrator-focused (p. 8). Finally, the results indicate that increasing harm correlates with increasing remedy preferences up to a certain point (p. 8). However, beyond this point, remedy preferences decrease, suggesting that some forms of online harassment (e.g. doxing) cause such severe harm that they cannot be compensated by any of the available remedies (p. 9).

In 2024, Chief Marketing Officer at Besedo Strandell published “What is doxxing and how do you prevent it?”, a short piece providing a general description of doxing and suggesting ways users can protect themselves from being doxed. In this paper, doxing is defined as a “form of

online harassment involving the publication of personal information about an individual without their consent […] such as their full name, home address, telephone number, place of work, and other sensitive information." Doxing is often underpinned by a malicious intention to threaten or intimidate someone, for instance critical journalists or political activists. The (il-)legality of doxing may often depend on the context and varies between countries: for instance, whereas doxing is not per se illegal in the US, it is strictly outlawed in Germany due to its strict privacy laws. However, enforcement is difficult as doxers often hide behind the anonymity of the internet. Consequently, the author argues that social media platforms, the places where doxing usually takes place, have a particular responsibility to protect their users. As such, the author presents three content moderation forms to counter doxing namely proactive through content moderators, automated by means of AI, and reactively by relying on users for flagging. However, in most instances, the damage is already done and actions aim more at damage limitation. Therefore, the author argues that to prevent doxing from the outset, social media platforms should give users more control over their personal information as well as engage in actively educating users of how to protect themselves against doxing.

Mukti *et al.* at the University of Muhammadiyah Sidoarjo and M. Auezov South Kazakshtan University, recently contributed to discussions on doxing with "Doxing patterns using social engineering in cyberspace" (2024). The paper aims to identify patterns behind *social engineering* that lies at the core of doxing, as well as to differentiate legal from illegal forms of doxing. For this purpose, the study employs a normative juridical approach using statutory analysis. Having defined doxing as a "crime on the internet" and a form of cyberbullying by disseminating personal data with the aim to intimidate or threaten the victim (pp. 531, 534), the author explains the mechanisms by which doxing happens. In particular, the authors focus on "social engineering" as a means to amplify the reach of the doxed information by gathering attention and inviting other users to participate in it, or as a way to lure users into voluntarily revealing personal information (p. 544). The authors argue that the purpose of doxing is decisive in assessing its legality: revealing personal information without prior consent, with the intention of harassing or extorting the victim, or if it contains sensitive content, is illegal (p. 542). In contrast, in the context of responsible news reporting by journalists, by law enforcement for legitimate investigations, or to enhance cybersecurity, it is legal (p. 542). Therefore, revealing personal information is legal if it serves the public interest and illegal if it is done to harm an individual. However, even in legitimate cases, doxing must be done proportionally, revealing only the minimum amount of information necessary to achieve the intended purpose (p. 543).

Also in 2024, Schuster *et al.* at the Technical University of Darmstadt wrote "What Makes Doxing Good or Bad? Exploring Bystanders' Appraisal and Responses to the Malicious Disclosure of Personal Information" (2024). Using a qualitative vignette study that employs two doxing scenarios (person and issue-based), the paper investigates the factors that influence the perception of whether an instance of doxing is legitimate or not. In this paper, doxing is defined

as the “disclosure of an individual’s personal information with malicious intent” (p. 116). The authors argue that, in contrast to other forms of cyber harassment, the outcome of doxing is significantly shaped by the audience’s perception of, and reaction to, the doxing incident (pp. 116-117). The perception of the legitimacy of doxing depends on the involved actors (doxer, doxee, bystander), the trajectory of the campaign, and the revealed information. In terms of actors, the doxer’s credibility in the community, as well as the perception that the doxer acted “out of heat,” contribute to a positive appraisal (p. 120). This is especially the case if, based on the severity of the doxee’s misconduct, the audience believes that the doxer even has a responsibility to share the information (p. 120). In contrast, assumed personal motives of the doxer negatively influence the appraisal (p. 120). With regard to the doxee, appraisal is mainly influenced by their importance in the community, the genuineness of the accusations, and the proportionality between the alleged misconduct and the actions taken (p. 120). Bystanders tend to approve of doxing if they are personally, emotionally involved, and disapprove of it if they have subject matter knowledge about doxing and its consequences (p. 120). Second, support over the course of a doxing campaign tends to decrease with increasingly drastic consequences for the doxee, towards outright rejection in cases of physical harassment (p. 121). Third, appraisal significantly depends on the sensitivity of the revealed information (e.g., home address vs. social media account) as well as on the perception of whether the information is revealed for the public good or out of malicious intent (p. 121). Overall, evidence suggests that doxing incidents tend to be considered rather neutrally or negatively (pp. 120-122).

Another newly incorporated contribution on doxing is that by Dannhauser from the University of California (UCLA), “Protecting the Innocent: How to Prevent the Consequences of Misidentification and Doxing by Volunteers Helping with Open Source Investigations” (2024). The paper explores crowdsourcing, which involves gathering work, information, or opinions from a large group of people who submit their data via the Internet, social media, and smartphone apps, and it studies the presumption that crowdsourcing may increase the risk of doxing. It does so based on three case studies: Europol’s *Stop Child Abuse – Trace an Object* programme, Bellingcat’s investigation of airstrikes in Afghanistan, and efforts by Trace Labs to find missing persons. It identifies best practices employed and how they can guide crowdsourcing investigations to minimise doxing risks. After sharing best practices by Europol, Bellingcat, and Trace Labs, the paper suggests that to ensure accountability during crowdsourcing, organisations can limit the public’s role in the actual investigation, as well as require users to keep track of their steps (pp. 16-17). Since the competency criterion seemingly contradicts the nature of crowdsourcing, organisations should rely on trained staff as much as possible (p. 18). Creating closed environments, as done by Bellingcat, can reduce the number of incompetent users, while growing a community like Trace Labs can progressively increase competency (pp. 18-19). Arguably, accuracy is the most difficult criterion, with peer reviews being the only valid option for organisations (p. 19). A general problem arises if the

crowdsourcing investigation is not initiated by professional organisations that cannot or do not want to employ any oversight (p. 21).

Also recent is the research paper by Naskali *et al.* at University of Turku in Finland entitled "Doxing Ethics" (2024). Here, doxing is defined as 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" (p. 101). While doxing holds a mainly negative connotation both in the public and in academia, the authors argue that there might be cases in which doxing is ethically justifiable (p. 101). According to the authors, doxing must overall yield a positive utility to be considered legitimate. To this end, the author employs Douglas's (2016) doxing typology and respectively assesses the utility for 1) the target of doxing, 2) people close to the victim, 3) the doxer, 4) people who benefit from the doxing, and 5) the public in general (p. 102). The authors generally consider deanonymising the least harmful form of doxing for the victim and people close to them (p. 102). As such, it constitutes the easiest form to justify if the public derives a positive utility from it, for instance, if the deanonymisation halts illicit and/or harmful behaviour by the doxee (p. 102). In the case of targeting doxing, the threshold for its justification is significantly higher due to the potentially much greater negative utility for the victim and people close to them (p. 102). Lastly, delegitimising doxing is generally assessed as unethical since it is not about revealing wrongdoing or preventing a tragedy but rather about harming the victim out of malice (p. 102). The negative consequences for the victim are much worse compared to deanonymising or targeting, and it is difficult to imagine circumstances in which the public would benefit from it (pp. 102-103). Even for the public, utility is negative as it may poison discourse and contribute to a "tyranny of the majority" (p. 103). All in all, the ethical justification mainly depends on weighing the negative consequences for the victim and people close to it against the benefit to the larger public (p. 103).

The latest contribution to the field is entitled "Where Are They Now?: The Costs and Benefits of Doxxing Far-Right Extremists" (2024), by Amarasingam and Galloway from Queen's University (Canada), and the Evolve Program, Organization for the Prevention of Violence, respectively. The paper explores the "immediate and long-term effects of doxxing'' based on 10 interviews with former members of the far right. The paper offers a perspective on the advantages and drawbacks of doxing, both for the targeted individuals and for society as a whole. The study highlights two definitions of doxing to show its complexity and why it's a topic of debate. On one side, doxing is often seen as something done "with malicious intent." However, Cheung (2021) suggests that it can also serve a positive purpose, like improving community safety (p. 162). Immediate consequences reported by participants were severe. Some lost their jobs, even when the information exposed wasn't entirely accurate (p. 170). Others faced threats to their physical safety, from being harassed in public to fearing for their lives (p. 171). The fallout also took a toll on family relationships, as loved ones often struggled to deal with the situation (p. 172). The long-term consequences included difficulties in finding new work and challenges in

opening up about their past to friends or new partners. When asked whether doxing was beneficial, participants had mixed views depending on their experiences (p. 176). For some, the experience triggered a change, helping them realize they were involved in something harmful and pushing them toward personal growth (p. 176, 178). One person even said, “the doxxing he experienced was necessary.” But for others, it was “the worst thing” that ever happened to them (p. 177). Based on these insights, the authors recommend that law enforcement institutions rethink using doxing as a deterrent tool (p. 179). While it might seem effective, it often leaves lasting damage on the lives of those targeted (p. 179). Instead, they suggest considering programs on, for instance, reintegration. They also urge social media platforms to stick to their terms of service, avoid sharing user data with governments, and promptly remove doxing content (p. 179).

Discussion

The act of doxing is generally understood by the assessed literature as involving the "online publication" of personal, private identifying, and sensitive information, including name, address, or whereabouts (e.g. Snyder 2017, p. 432; Jensen and Watts, 2022). In doxing, the public release of such information is done "without consent" and typically with "malicious intent", and may be used "with the intent to humiliate, harass, intimidate, punish and/or blackmail targets, or condemn certain actions and ideas" (e.g. Matthews 2014, p.1; Li, Chen 2018; Schuster, 2024, p. 116; p. 359; Amarasingam 2024; Naskali, 2024, p. 101). Overall, doxing is considered a form of "online harassment" that can cause severe harm (Schoenebeck, 2023, p. 6; Strandell, 2024).

Doxing can be conducted by state and non-state actors alike. Studies distinguish between doxing as a bottom-up tool for surveillance, which allows a state to either attain or strengthen its power through data knowledge gathering and manipulation, as well as for "sousveillance", which according to Li and Whitworth involves monitoring the state to regain agency (e.g. Li and Whitworth, 2023, p. 368). This can take place through open-source search for personal and sensitive information previously shared by a victim (Dannhauser 2024, and as we have seen in this project for the cases of Dorota Kwietniewska or Jakub Sochuiko Jarowski[1]); through hacking, or by gaining unauthorised access to online accounts, systems, and information (Matthews 2014); or through "social engineering", a process by which others are manipulated into sharing sensitive information in order use it against them later on (Mukti, 2024, pp. 536). Existing literature highlights several reasons why doxing might take place. With a few overlaps, we can identify nine main motivations: extortion, competitiveness, revenge/retribution, silencing, justice, control, reputation-building, unintentional, and public interest (e.g. Snyder 2017, p. 438; Anderson 2021, pp. 208-9; Li 2023, p. 368). The latter may be misused by undemocratic states and state-sponsored proxies or supporters to justify revealing personal information of citizens targeted as a threat to public stability, for example (Li 2023). On the other hand, it may also be used by activists against repressive measures (Strandell 2024). That is what Li calls *sousveillance*, a way for individuals or groups to monitor the state and regain power from below (Li 2023, p. 368). Naskali 2024 calls this using doxing with a "positive utility", which renders it legitimate or ethically justifiable (p. 101).

The literature examined identifies four main victims of doxing: hackers, gamers, celebrities, and combatants or soldiers in conflict zones (Snyder, 2017; Jensen and Watts, 2022). Our interview data, however, reveals that doxing in war zones is not just a weapon used against soldiers but also against journalists, humanitarian workers, and against family members of combatants and activists. Indeed, doxing appears to be a "regular occurrence" against journalists, according to a

[1] Both are foreign combatants for Ukraine who shared personal information mainly through their public Facebook pages, which was later doxed.

foreign journalist based in Kyiv who moved there to write about the conflict. This, however, failed to make a great impression as the given individual found it nothing more than a nuisance as they had already been living in a "warzone" and "my ability to be threatened is [thus] skewed." A foreign fighter recalls that doxing is "common enough, I guess". This includes the public posting of "addresses and things [about] family member[s] sometimes who live in their own country [i.e. outside of Ukraine]". Not a source of concern for this victim of doxing either, who stressed that whoever targeted them can "come and fight me." Another combatant expressed a more fatalistic approach while insisting that "all the Russians will do is kill me either way and I doubt anyone is going to try me in the U.S." A Ukrainian fighter who had spent most of his life abroad was doxed and accused of being a "Nazi" in Russian media (and pro-Russian media in Ukraine) long before 2022, when the full scale war broke out. A foreign fighter who had also been victim of doxing explained that the reasons for this type of strategy in the context of war is not necessarily to provide Russia with intelligence, but rather to "make themselves feel good" by diversifying targeting tactics against the enemy. Generally, doxing is understood by all of the interviewees as used primarily to undermine the morale and mental health of the "enemy," in this case Ukraine.

A humanitarian worker interviewed by the project team indicated that doxing may start with the simple procedure of handing over one's identity documents to a trooper patrolling the streets of a given Ukrainian town, which begs the question of how one's private information might reach pro-Russian doxxers. Doxing can also happen to family members of people involved in supporting Ukraine. The aforementioned humanitarian worker, whose spouse is Russian and published in Russian opposition outlets, experienced it firsthand when their work was "discussed" live on Russian TV. Threats and cyber bullying followed their exposure. Interestingly, the project team also learnt of at least one pro-Ukraine foreign fighter who got involved in doxing their comrades, thereby assisting pro-Russian doxing platforms, after a bad experience in Ukraine. As doxing seems to be a fairly common practice against the members of what might be called a social movement of supporters for Ukraine in general, and the foreign fighters in particular (Rekawek 2023, pp. 7-8), it is pertinent to ask why this has become the case. A foreign fighter explains that perhaps victims of doxing have not been protective of their public information, especially on social media, therefore allowing pro-Russian doxxers to access sensitive data through open-source information gathering, or OSINT. Other theories speak about the Ukrainian recruitment process of foreign individuals, which "gets hacked twice per year", or "there has long been a discussion over the possibility of a mole […] Guys that have been doxxed on telegram and twitter but they don't have any social media presence […] but it is just rumours after all."

Whereas Snyder's research (2017) argues that most victims of doxing tend to be men, as also suggested by our interviews, in fact the literature largely agrees that women are often more vulnerable to it. Schoenebeck's study, for example, demonstrates that women and other

underrepresented or minority groups, including people of colour, are particularly vulnerable to harmful doxing (2023, p. 8). In contrast, Anderson (2021) reveals that women are more frequently targeted than men (p. 214). Finally, Chen's work on the Hong Kong anti-government protests demonstrates that doxing incidents were more prevalent among female students compared to their male counterparts (2018). The consequences of doxing for this group included anxiety, depression, and physical, mental, and emotional distress not only from identity exposure, but from the fear of social judgment, for example (Chen, 2018, p. 6). The current project studies targeted profiles that do not fall under either of the three above-mentioned categories. Instead, they include soldiers, volunteer fighters, and other actors actively involved in conflict, as explained in the Introduction.

While most authors agree that doxing is generally harmful for the targeted victims, some literature points at the short and long term benefits of doxing when the context is right. Amarasingam and Galloway's research into the experiences of doxed far-right individuals (2024) illustrates that while doxing may have been utterly damaging for some individuals, including by causing them to lose their job or by having an impact upon the relationship with their family, for others doxing encouraged them to reconsider their lifestyle and ideologies, thereby starting the process of disengagement. Whereas such insights are very relevant to understand the impact of doxing, that this practice would accelerate the disengagement process away from extremism is not fully applicable to our research as the focus groups are different. Similarly, Trottier explains that doxing as a form of public vigilantism has the capacity to empower citizens with an unprecedented level of agency (2019). Where the justice system and the state fail to protect its people or treat them with fairness, citizens may use doxing to conduct public scrutiny and hold institutions accountable. Dannhauser (2024), explores "best practices" in doxing, explaining how in certain cases doxing has been used to identify harmful perpetrators or find missing people. Finally, Cheung argues that doxing may be used as a "political pushback" to punish individuals tasked with civic duties, such as law enforcement, as it occurred during the anti-government protests of 2019 in Hong Kong. Focusing on three case studies of doxing that were brought to Court, including against journalists, police officers, and students, Cheung concludes that doxing is illegitimate when used to harass or threaten the victim, but acceptable for legitimate journalistic reporting, aiming to balance freedom of press with the protection of privacy (p. 583).

Whether or not doxing is legitimate is an ongoing debate that seems to have developed after reaching the general consensus that doxing ultimately means harm. Such a discussion has evolved mainly around three elements: the perpetrator/doxer, the victim/doxee, and the context. Most of the examined literature agrees that if doxing is used as a deterrence mechanism against threats to democracy and human rights, including repression, harassment, and authoritarian regimes, meaning that it generally benefits the public good, then it is legitimate (e.g. Douglas, 2016; Liu, 2022; Schuster, 2024; Naskali, 2024). This position also includes research focused on military operations conducted by democratic states, where doxing is justified when it meets the

military necessity requirement, for example by undermining enemy morale, fighting spirit, and for operational focus, even if it causes mental suffering (Jensen and Watts, 2022). Evidently, the ethics of doxing are thus based not only on intent, but also on moral grounds. In other words, if harm is understood as being a necessary step in the process of either ensuring or bringing about good or right against wrong, then doxing is generally accepted.

Even though doxing is likely to cause long-lasting harm in most cases (Amarasingam 2024), this does not mean that the practice is illegal. In fact, the literature tends to stress that while doxing may be a form of "cyberbullying" (Mukti, 2024, pp. 531-534), it may be considered legal if the information revealed is true; if the data was obtained from open sources; if the information was revealed publicly by the victim first; or if it fulfills legitimate journalistic reporting, for instance. In reality, the (il)legality of doxing will depend on the context and country where it is taking place. Illustratively, while doxfare is considered legal in the United States unless explicit threats are involved (Liu 2022, pp. 76-77), it is outlawed in Germany due to privacy laws (Strandell, 2024). In Australia, anti-stalking laws may be applied in cases of doxing, but this remains insufficient for large-scale incidents (Anderson 2021, p. 220). Largely, European countries have stronger privacy protections under the European Convention on Human Rights, but enforcement remains inconsistent (Anderson 2021, pp. 221-222). According to Jensen and Watts (2022), under international law, doxing that is directed at active combatants is lawful and can be compared to other information operations such as propaganda. Generally, the literature is clear in pointing out that current legal frameworks fail to define doxing in a nuanced manner, or differentiate between the motives behind doxing in a systematic way. This is problematic because it prevents a consistent assessment of doxing's legality and its proportionality, in particular the way in which it might outweigh the negative consequences for the larger public benefit. Illustratively, this project has observed the misuse of personal public information for political gains during a conflict that, while respectful of international law, challenges national privacy legislation.

Conclusion

While the collection of literature analysed for the purpose of this review indicates a general agreement on the nature and impact of doxing, a few questions remain. Above all, the assessment of whether doxing is legitimate or not seems closely dependent on personal values, which means that the practice may be easily misused by repressive actors if the law does not specifically indicate what makes it a crime. For instance, how do we understand "public good" or "positive utility" and who is to define it to inform policy tackling doxing? Secondly, and similarly, the element of "malicious intent" is not always present in conceptual interpretations of doxing, which leaves room for state-actors to justify doxing against citizens under the premise of public stability, for example. Finally, while doxing involves the dissemination of harmful information for power-related advantage, existing research does not explicitly approach doxing as a radicalising strategy. It would be particularly useful if current discussions around doxing compared the challenges involved in this practice with the spread of radicalising borderline content (aka "awful but lawful") specifically. Indeed, current debates on borderline content struggle to understand when this type of content is illegal or not, and efforts to engineer a more systematic way to detect and respond to borderline information online are undergoing. Studies in doxing should then address this practice as a related problem, and try to offer more systematic remedies rather than case-by-case conclusive remarks and recommendations, including when doxing occurs by state-actors or state-supported actors in conflict settings.

This literature review hopes to contribute to a more systematic understanding of doxing for the purpose of a research project

Bibliography

Amarasingam, A., Galloway, B. "Where Are They Now?: The Costs and Benefits of Doxxing Far-Right Extremists" *Journal for Deradicalization*, no. 40 (2024).

Anderson, B., & Wood, M. "Doxxing: A Scoping Review and Typology." In *The Emerald International Handbook of Technology-Facilitated Violence and Abuse.* Emerald Publishing Limited, 2021.

Chen, Q., Chan, K. L., & Cheung, A. S. Y. "Doxxing victimization and emotional problems among secondary school students in Hong Kong." *International Journal of Environmental Research and Public Health* 15, no.12, (2018), https://doi.org/10.3390/ijerph15122665.

Cheung, A. "Doxing and the challenge to legal regulation: When personal data become a weapon". In *The Emerald international handbook of technology facilitated violence and abuse*. Emerald Publishing Limited, 2021.

Dannhauser, L. M. "Protecting the Innocent: How to Prevent the Consequences of Misidentification and Doxing by Volunteers Helping with Open Source Investigations." *Catholic University Journal of Law and Technology* 32, no. 2 (2024): 1–26.

Douglas, D. M. "Doxing: a conceptual analysis." *Ethics and Information Technology 1*8, no. 3 (2016): 199–210. https://doi.org/10.1007/s10676-016-9406-0.

Jensen, E., & Watts, S. "Ukraine Symposium - Doxing Enemy Soldiers and the Law of War." *Lieber Institute West Point,* October 31, 2022. https://lieber.westpoint.edu/doxing-enemy-soldiers-law-of-war/.

Li, Y., & Whitworth, K. "Coordinating and doxing data: Hong Kong protesters' and government supporters' data strategies in the age of datafication." *Social Movement Studies* 23, no.3 (2023): 355–372. https://doi.org/10.1080/14742837.2023.2178404.

Liu, Y. "Doxfare as a Tool for Strategic Deterrence." *Global Security and Intelligence Studies* 8, *no.*1 (2022): 69–82. https://doi:10.18278/gsis.8.1.4.

Marx, Gary T. "What's in a Name? Some Reflections on the Sociology of Anonymity." *The Information Society* 15, no. 2 (1999): 99–112. https://doi:10.1080/019722499128565.

Matthews, R. S., Aghili, S., & Lindskog, D. "A Study of Doxing, its Security Implications and Mitigation Strategies for Organizations," (2014), https://doi.org/10.7939/r3-nh05-7x95.

Mukti, A. T., Multazam, M. T., Rosnawati, E., & Kurmangaly, S. "Doxing patterns using social engineering in cyberspace." *Advances in Social Science, Education and Humanities Research,* (2024): 530-46. https://doi.org/10.2991/978-2-38476-247-7_54.

Naskali, J., Rantanen, M., Rottenkolber, M., & Kimppa, K. K. "Doxing Ethics." In *Proceedings of the ETHICOMP 2024. 21th International Conference on the Ethical and Social Impacts of ICT, 2024,* https://dialnet.unirioja.es/descarga/articulo/9326116.pdf.

Rekawek, K. A Year of Foreign Fighting for Ukraine: Catching Fish With Bare Hands?. Countering Extremism Project (2023), https://www.counterextremism.com/sites/default/files/2023-03/CEP%20Report_A%20Year%20of%20Foreign%20Fighting%20for%20Ukraine_March%202023.pdf

Schoenebeck, S., Lampe, C., & Triệu, P. "Online Harassment: Assessing harms and remedies." *Social Media + Society* 9, no.1 (2023): 1–12. https://doi.org/10.1177/20563051231157297.

Schuster, J., Franz, A., & Benlian, A. "What Makes Doxing Good or Bad? Exploring Bystanders' Appraisal and Responses to the Malicious Disclosure of Personal Information." In *Proceedings of the 57th Hawaii International Conference on System Sciences,* 2024, https://aisel.aisnet.org/hicss-57/cl/social_media/3.

Snyder, P., Doerfler, P., Kanich, C., & McCoy, D. "Fifteen minutes of unwanted fame: detecting and characterizing doxing." In *IMC '17: Proceedings of the 2017 Internet Measurement Conference,* 2017, https://doi.org/10.1145/3131365.3131385.

Strandell, J. "What is doxxing and how do you prevent it?" *Besedo,* May 27, 2024. https://besedo.com/blog/what-is-doxxing/.

Trottier, D. "Denunciation and Doxing: Towards a Conceptual Model of Digital Vigilantism." *Global Crime* 21, no. 3–4 (2019): 196–212. https://doi.org/10.1080/17440572.2019.1591952