Center on Law, Ethics and National Security

Essay Series
Number 19
July 25, 2023

Not Just Words:
Grappling with the Doxing of Civilians in War

By Riley Flewelling

DUKE LAW

NOT JUST WORDS: GRAPPLING WITH THE DOXING OF CIVILIANS IN WAR

RILEY FLEWELLING[1]

ABSTRACT

As warfare evolves, so must the law of armed conflict. Modern war has spilled off the battlefield. States leverage sanctions in the hopes of weakening their opponents or the opponents of their allies, and information operations seek to change the "truth" of what happens on the ground. This paper reckons with another way war might expand out of its traditional box: the use and abuse of civilian data through doxing. Doxing is the process of publishing private information about an individual or organization, often as a form of punishment. Doxing is often thought of in reference to political scandals or corporate leaks, but it could be used in a wartime setting as well. As civilians share greater and more personal swaths of information online, and certain states develop wide-reaching control of the internet, doxing could be used to draw targets on the backs of particular groups. This phenomenon appears to have happened in countries like Ukraine and Myanmar. This paper explores existing frameworks in the law of armed conflict which could be used for the emerging threat of doxing. This paper ultimately argues that incremental steps forwards in the ways the law of armed conflict has handled propaganda may provide the most fitting solution.

[1] Duke University School of Law, J.D./L.L.M. expected May 2024. A tremendous thank you to Major General Charles J. Dunlap, Jr. USAF (Ret.) for his guidance and support, and to the LENS Center for furthering student opportunities to engage with these topics.

INTRODUCTION

Imagine the following situation. Country X experiences an insurgency led by a socially progressive faction of its society. Country X has purchased surveillance technology from China, and with the China's help, has established a far-reaching ability to monitor its citizens' private online habits and control information flows. Country X commands a large cyber force dedicated to surveillance and control activities, a force which falls under the control of the military. The dominant culture in Country X is socially conservative and has strong homophobic and patriarchal leanings.

As the insurgency begins to pick up momentum, Country X engages in a doxing campaign on a max scale. The country's cyber force publishes lists of civilians who have indicated sympathy for the insurgent movement on social media, including pornographic photos of the women on the list and whether the listed individuals have accounts on gay dating apps. Alongside these lists, the cyber force provides the home addresses of the listed individuals and engages in a massive information campaign encouraging loyal citizens to take matters into their own hands and take vengeance on the insurgent-sympathizers. The cyber force's involvement is not public; they disguise themselves as patriotic hackers and concerned citizens.

In the following days, people are harmed. Reports abound of those killed, injured, and raped by individuals taking up the call of the "patriotic hackers." In some cases, the only insurgent "support" the targeted individuals had offered was "liking" a prominent rebel leader's Facebook post.

This situation is intentionally dramatic, but it fails to read entirely like science fiction. Digital technologies have led to waves of good in conflict situations,[2] but these benefits are matched by new concerns and threats. Around the world, the interaction between conflict, government surveillance, and the swaths of data people disclose online creates reason for concern. This paper focuses on one such reason for concern: the phenomenon of doxing. The Merriam-Webster dictionary defines doxing as "publicly identify[ing] or publish[ing] private information about [someone] especially as a form of punishment or revenge." [3] Doxing is a form of targeting individuals by acquiring their personal data, often through hacking, and releasing it to the public.[4] Revenge porn, or the non-

[2] *See* Int'l Comm. Red Cross, *Testimonies: How Humanitarian Technologies Impact the Lives of Affected Populations*, *in Digital Technologies and War*, 102 INT'L REV. RED CROSS 23, 23–24 (2020) (offering an example of family members reconnecting years after being separated from one another post-conflict, through websites which feature photo galleries of people looking for lost family members).

[3] *Dox*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/dox (last visited Apr. 11, 2023). Doxing is often spelled "doxxing" as well, and the sources this paper cites use both forms. This paper will use "doxing."

[4] Rob Lever, *What is Doxxing?*, U.S. NEWS (Dec. 16, 2021), https://www.usnews.com/360-reviews/privacy/what-is-doxxing.

consensual release of intimate photos (both real and fabricated), also falls under this umbrella.[5]

The law of armed conflict ("LOAC)" has not yet grappled with doxing. Attacking civilians violates the LOAC principle of distinction,[6] but most online harms to civilians are not considered "attacks" because they do not "cause injury or death to persons or damage or destruction to objects."[7] Targeting and abusing data does not fit easily into this definition.[8] Similarly, the criminalization of speech acts offer an unclear framework for dealing with doxing. Propaganda which constitutes incitement to genocide or instigation of crimes against humanity has been punished under LOAC,[9] but it is unclear how far the logic behind these offenses extends. There are also myriad other categories of conduct which could prevent the doxing of civilians, including the prohibition against humiliating and degrading treatment, the responsibility of states to protect civilians, and the prohibition against terrorizing civilians, but none of them offer a perfect fit.[10]

The problem of doxing grapples with a world in which individualized online targeting of civilians can lead directly to physical harm. This premise mirrors discussions in domestic criminal law[11] and contemporary conversations surrounding domestic terrorism.[12] Threats online do not always remain there, and the unprecedented abilities of governments to access and weaponize civilian data propels LOAC towards adaptation.

Importantly, how to reckon with the doxing of civilians is not the only question in the intersecting world of war, cyber, and surveillance. This paper does not explore the important question of when a civilian should be considered a direct combatant based on their online actions, nor does it make claims about broader use of information warfare by states. Importantly, this paper is limited to *jus in bello* framing and does not touch on state use of doxing against individuals in times where war is absent. Finally, this paper does not tackle human rights debates on the balance between doxing and free speech, or the right to privacy online.

The rest of the paper proceeds as follows. Part I explores a few instances of doxing connected with armed conflict, before explaining state capacity and

---

[5] *See* Daisy Schofield, *'He Found Out Where I Live' - Sex Workers Are Getting Doxxed by Clients*, VICE (Jan. 4, 2021, 6:00 AM), https://www.vice.com/en/article/v7m38y/doxxing-.

[6] Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I) art. 48, Jun. 8, 1977, 1125 U.N.T.S. 3 [hereinafter Protocol I].

[7] GARY D. SOLIS, THE LAW OF ARMED CONFLICT 248 (3d ed. 2022) (citing TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER OPERATIONS, r. 92, 92.5 (Michael N. Schmitt ed., 2017) [hereinafter TALLINN 2.0]).

[8] *See infra* Part II.B.

[9] *See infra* Part II.C.

[10] *See infra* Part II.D.

[11] *See* Hannah Mery, *The Dangers of Doxing and Swatting: Why Texas Should Criminalize These Malicious Forms of Cyberharassment*, 52 ST. MARY'S L. J. 905 (2021) (calling for criminalization of online threats using the example of "Gamergate," a phenomenon where prominent female videogamers were threatened with rape and death online alongside facing physical security concerns).

[12] *See generally* ALEXANDRA T. EVANS & HEATHER J. WILLIAMS, RAND, HOW EXTREMISM OPERATES ONLINE (2022) (explaining how domestic extremist groups use the internet, including how they amplify their ideologies and incentivize others to act based solely on online messaging).

potential incentives to engage in this behavior. Part II touches on key LOAC frameworks, exploring the definition of "attack" in the cyber context, investigating jurisprudence regarding the criminalization of speech, and evaluating other characterizations of doxing in current LOAC protections. Finally, Part III explores the best path forward for punishing doxing as a war crime and notes limiting principles. Ultimately, this paper concludes that criminalizing doxing as a form of instigation to commit war crimes is the best fitting solution, but that individual criminal liability can likely only attach when the instigated crimes occur.

I. STATE USE OF DOXING

The incidence of states using doxing is small, but not nonexistent. This Part explores how doxing has been used in armed conflicts in Ukraine and Myanmar, before diving into state incentives and capacity to engage in doxing campaigns.

A. Ukraine, Russia, and Myanmar

1. Doxing has occurred in the war between Russia and Ukraine. Ukraine offers one example of doxing leading to physical violence. The Ukrainian website 'Peacemaker' appears to have been harassing and doxing individuals it views to be anti-Ukrainian since Russia's invasion of Crimea in 2014.[13] The site, reportedly established by "a member of Ukraine's interior ministry"[14] publishes the personal information of those it calls "enem[ies] of Ukraine."[15] In 2015, a journalist and a former politician who had publicly expressed "pro-Russia" views were both shot days after Peacemaker published their home addresses.[16] Both killings reportedly happened at or near the individual's homes.[17]

The website remains active in the current war between Russia and Ukraine.[18] When listed individuals are killed, the word "liquidated" is put over their

[13] Elise Thomas, *Project Nemesis, Doxxing and the New Frontier of Informational Warfare*, INST. FOR STRATEGIC DIALOGUE: DIGITAL DISPATCHES (June 23, 2022), https://www.isdglobal.org/digital_dispatches/project-nemesis-and-the-new-frontiers-of-informational-warfare/.

[14] *Id. See also*, U.S. DEP'T OF STATE, 2020 COUNTRY REPORTS ON HUMAN RIGHTS PRACTICES: UKRAINE, 32 (2020), https://www.state.gov/reports/2020-country-reports-on-human-rights-practices/ukraine/ ("The [database] . . . reportedly maintained close ties to the country's security services").

[15] @DFRLab, Atlantic Council's Digital Forensic Research Lab, *UN Calls for Investigation of Ukrainian Digital Blacklist*, MEDIUM (Sep. 20, 2017), https://medium.com/dfrlab/un-calls-for-investigation-of-ukrainian-digital-blacklist-14fec836753f.

[16] Manisur Mirovalev, *Peacemaker: The Ukrainian Website Shaming Pro-Russia Voices*, AL JAZEERA (Aug. 27, 2019), https://www.aljazeera.com/features/2019/8/27/peacemaker-the-ukrainian-website-shaming-pro-russia-voices.

[17] Reuters Staff, *Ukrainian Journalist with Pro-Russian Views Shot Dead in Kiev*, REUTERS (Apr. 16, 2015, 7:33AM), https://www.reuters.com/article/ukraine-crisis-crime/ukrainian-journalist-with-pro-russian-views-shot-dead-in-kiev-idINKBN0N718720150416.

[18] Thomas, *supra* note 13.

photos on the website.[19] While it is unclear if the Ukrainian government exercises active control over Peacemaker, the published information could be connected to civilian deaths, at least in the case of the 2015 killings.[20] The European Parliament called for the blog to be shut down in 2021,[21] but there has been pushback to characterizations of the blog as a "kill list," as well as claims that Russian propaganda efforts have created that characterization.[22]

Similar websites have emerged out of Russia, including "Project Nemesis," a website and Telegram channel publishing the personal information of "Ukrainian Nazis" and "those who help them."[23] Posts related to Project Nemesis often include calls for violence, commonly featuring homophobic rhetoric and encouraging attacks against LGBTQ individuals.[24] However, this effort appears to be largely targeted at combatants, such as Ukrainian military members and security services.[25] It is unclear if direct violence against any civilians has occurred as a result of these efforts, and similarly unclear if the Russian government has any control over those doing the doxing.[26] Overall, both Peacemaker and Project Nemesis offer examples of what doxing might look like in international armed conflicts.

2. Protesters in Myanmar are frequent victims of doxing. Individualized online harm is also salient in Myanmar.[27] Myanmar's military, the Tatmadaw, took power in a coup in 2021 after opposing the results of the country's democratic election.[28] The state jailed journalists and opposition leaders, and peaceful protests against the military grew violent in a matter of weeks.[29] The country is now split between the Tatmadaw and a resistance government, the National Unity Government ("NUG"), which claims to control half the country and commands

[19] *See* Mark Trevelyan, *Russian Ex-Submarine Officer on Ukraine Blacklist Gunned Down*, REUTERS (July 11, 2023, 1:59 PM), https://www.reuters.com/world/europe/russian-military-official-ukraine-blacklist-shot-dead-morning-run-2023-07-11/.
[20] *See supra* notes 16–17 and accompanying text (noting that the journalist and politician were killed near their homes after their addresses were published on Peacemaker).
[21] Freedom House, *Ukraine*, *in* FREEDOM ON THE NET 2022, https://freedomhouse.org/country/ukraine/freedom-net/2022.
[22] *See* Vitaly Portnikov, *The "Peacemaker" Phenomenon*, RADIOLIBERTY RUSSIA (June 9, 2018), https://www.svoboda.org/a/29275087.html (Offering an interview with Anton Gerashchenko, a Ukrainian internal affairs minister, where he said "[t]his whole dirty campaign against the "Peacemaker" project was started and developed by Russian propagandists and special services. They called on the world community, found collaborators in Ukraine who shouted that "Peacemaker" was bad"). The interview appears in a translated form.
[23] Thomas, *supra* note 13.
[24] *Id.*
[25] *Id.* For discussion of LOAC and the doxing of combatants, *see* Eric Jensen & Sean Watts, *Ukraine Symposium, Doxing Enemy Soldiers and the Law of War*, ARTS. WAR, LIEBER INST. (Oct. 31, 2022), https://lieber.westpoint.edu/doxing-enemy-soldiers-law-of-war/.
[26] *Id.*
[27] Rodlyn-mae Banting, *Burmese Women Protesting the Military Coup are Having Their Sex Tapes Leaked*, JEZEBEL (Feb. 7, 2023), https://jezebel.com/burmese-women-protesting-the-military-coup-are-having-t-1850083991.
[28] *Id.*
[29] *Id.*

armed forces.[30] The conflict in Myanmar constitutes a non-international armed conflict.[31]

While the Tatmadaw appears to be harming civilians in numerous ways,[32] doxing through revenge porn is a new weapon in the conflict.[33] Female dissidents and sympathizers of the opposition government have had explicit videos and photos of themselves published online, alongside their names and addresses.[34] These posts are largely on Telegram, a messaging platform with accountability problems.[35] There are a few general patterns to the posts. Sometimes they involve explicit photos and videos, other times they just publish women's names and call that they be punished.[36] The photos and videos are sometimes doctored or faked.[37] Men have been targeted too, but the messaging is different, usually describing the dissidents as terrorists without including sexualized content.[38] Intermittently, the police arrest the doxed individuals.[39] While the use of doxing to identify protesters is not new,[40] the use of revenge porn accompanied by direct calls for violence seems to be novel.

One woman who was arrested after organizing peaceful protests against the junta experienced doxing after she was released from prison.[41] Users online called for her to be killed and raped, saying things such as "[a]fter every [sic] has f**ked her, deliver her verdict."[42] Comments often include dehumanizing and racist

[30] *Id.*
[31] Human Rights Council, Situation of Human Rights in Myanmar Since 1 February 2021, U. N. Doc. A/HRC/49/72, at 1 (Mar. 15, 2022) ("Myanmar is caught in a downward spiral of violence characterized by . . . several non-international armed conflicts.").
[32] *See Myanmar: Abuses Mount Since Military Coup*, HUMAN RIGHTS WATCH (Jan. 12, 2023, 12:00 AM), https://www.hrw.org/news/2023/01/12/myanmar-abuses-mount-military-coup ("the security forces have been implicated in mass killings, arbitrary arrests and detention, torture, sexual violence, and attacks on civilians in conflict areas.").
[33] Banting, *supra* note 27.
[34] Pallabi Munsi, *They Released a Sex Video to Shame and Silence Her*, CNN (Feb. 8, 2023), https://www.cnn.com/2023/02/07/asia/myanmar-military-sexual-images-doxxing-telegram-as-equals-intl-cmd/index.html.
[35] *See* Veronika Velch, *Telegram: A Growing Social Media Refuge, for Good and Ill*, JUST SEC. (Feb. 26, 2021), https://www.justsecurity.org/74947/telegram-a-growing-social-media-refuge-for-good-and-ill/ (describing how the messaging platform is used to organize both pro-democracy protesters and extremists who encourage violence).
[36] Munsi, *supra* note 34.
[37] Banting, *supra* note 27 (discussing one user who posted "doctored pornographic images purporting to be female opposition figures"); Amara Thiha, *Revenge Porn Has Become a Political Weapon in Myanmar*, THE DIPLOMAT (Aug. 9, 2021), https://thediplomat.com/2021/08/revenge-porn-has-become-a-political-weapon-in-myanmar (discussing intimate photos "both real and faked").
[38] *Id.*
[39] Munsi, *supra* note 34.
[40] *See,* @DFRLab, Atlantic Council's Digital Forensic Research Lab, *Telegram Channels Used to Doxx and Report Hong Kong Protesters to Chinese Authorities*, MEDIUM (Sep. 25, 2019), https://medium.com/dfrlab/telegram-channels-used-to-doxx-and-report-hong-kong-protesters-to-chinese-authorities-91bed151f345. *See also*, Jaclun Peiser, *Internet Detectives are Identifying Scores of Pro-Trump Rioters at the Capitol*, WASH. POST (Jan. 8, 2021, 6:54 AM), https://www.washingtonpost.com/nation/2021/01/08/capitol-rioters-fired-doxed-online/.
[41] Munsi, *supra* note 34.
[42] *Id.*

language.[43] The woman in question hid in a safe house, stating that "military supporters and religious extremists started keeping watch in the neighborhoods [she] was likely to be in."[44] While the doxing campaigns have not been attributed directly to the state, some of the accounts appear to be working directly with the military to dox women and arrest them.[45] There is indication that pro-democracy individuals have used the same approach against pro-junta women,[46] but a CNN study estimates that 90% of the posts target democracy advocates.[47] Regardless of whether the state is actually behind these campaigns, they offer an example of what state-sponsored doxing could look like in a non-international armed conflict.

The events in both Ukraine and Myanmar suggest that doxing could be weaponized by state actors and other armed groups. In Ukraine, civilians are not the only targets, and in Myanmar, it's unclear that the government is actively posting any of this information itself. However, it is easy to imagine a world in which civilians are the sole targets, and a state cyber unit is the true source of the posts, or even just an amplifying force. The unit need not even be particularly advanced; apps exist which would allow them to produce content such as revenge porn in seconds.[48] The heavy use of explicit content targeting women in Myanmar also showcases the impact cultural norms can have on the nature of targeting. This tactic could be especially effective in states where female promiscuity is heavily policed, such as in Uganda, where female victims of revenge porn are punished by the country's anti-pornography statutes.[49]

In both case studies, doxing has resulted in reductions in civilian safety, and in the case of Ukraine, actual harm. The next Section discusses factors which make it possible that this phenomenon will increase in future conflicts.

B. State Capacity and Incentives

1. States have broad abilities to access civilian data. TikTok offers a clear example of civilian data being collected by a private company in massive quantities. TikTok gathers users' names, ages, phone numbers, emails, approximate locations,

[43] Al Jazeera Staff, *Myanmar Women Target of Online Abuse by Pro-Military Social Media*, AL JAZEERA (Jan. 26, 2023), https://www.aljazeera.com/news/2023/1/26/myanmar-women-target-of-online-abuse-by-pro-military-social-media.
[44] Munsi, *supra* note 34.
[45] *Id.*
[46] Thiha, *supra* note 37.
[47] Munsi, *supra* note 34.
[48] *See* James Vincent, *New AI Deepfake App Creates Nude Images of Women in Seconds*, THE VERGE (June 27, 2019), https://www.theverge.com/2019/6/27/18760896 (offering an example of one such application).
[49] *See generally*, Twasiima Patricia Bigirwa, *Twice Shamed: The Use of Uganda's Anti-Pornography Act to Turn ~~Revenge Pornography~~ Non-Consensual Image Distribution Victims into Villains*, 22 GEO. J. GENDER & L. 565. *See also*, Abdi Latif Dahir, *'We Will Hunt You': Ugandans Flee Ahead of Harsh Anti-Gay Law*, N.Y. TIMES (Apr. 20, 2023), https://www.nytimes.com/2023/04/20/world/africa/uganda-anti-gay-bill-lgbtq.html (offering an example of a situation where doxing with a homophobic bent could also be effective).

and IP addresses,[50] but its algorithms assume much more about a user's preferences beyond those basic facts.[51] Data collected on social media in general, much of which users share willingly, can identify factors which define aspects of an individual's identity, such as sexual orientation and political leanings.[52]

Further, the actor doing the data collection might be the state itself. While those in charge of TikTok claim they never share data with the Chinese government, the US government does not appear to believe them.[53] TikTok aside, states can purchase massive swaths of data whenever they would like,[54] or they can gather it themselves. For example, China tracks people constantly, irrespective of their online activity.[55] The state expanded its abilities to monitor people even further during the pandemic.[56] China has also exported this technology to other countries around the world, often through the state's Belt and Road Initiative,[57] suggesting that capacity to collect and hold extensive records about civilian behavior will only continue to proliferate.

2. Whether legal or not, states have incentives to dox. During armed conflicts, threatening to expose the private data of protesting citizens could serve as a deterrent even in its most mild, and arguably legal,[58] forms. Protesters could fear losing their jobs or being shunned by their communities.[59] This fear could be enough to deflate support for insurgencies or resistance governments in a civil war.

---

[50] Christian Hetrick, *Terms of Misuse?: Breaking Down the Data TikTok Collects on Its U.S. Users*, DOT.LA (July 19, 2022), https://dot.la/what-data-does-tiktok-collect-2657689460.html.
[51] *See* Ben Smith, *How TikTok Reads Your Mind*, N.Y. TIMES (Dec. 5, 2021), https://www.nytimes.com/2021/12/05/business/media/tiktok-algorithm.html (explaining how the algorithm monitors the time you spend engaging with content, alongside other factors to predict and steer your preferences on everything from "socialism or Excel tips or sex, conservative politics or a specific celebrity").
[52] Rebecca J. Rosen, *Armed With Facebook 'Likes' Alone, Researchers Can Tell Your Race, Gender, and Sexual Orientation*, ATLANTIC (Mar. 12, 2013), https://www.theatlantic.com/technology/archive/2013/03/armed-with-facebook-likes-alone-researchers-can-tell-your-race-gender-and-sexual-orientation/273963/.
[53] *See* Bobby Allyn, *Congress Grills TikTok's CEO About Security of User Data*, NPR (Mar. 23, 2023), https://www.npr.org/2023/03/23/1165699549/congress-grills-tiktoks-ceo-about-security-of-user-data ("Deceptive, evasive, unconvincing – these are . . . the ways lawmakers in Washington today described the CEO of TikTok.").
[54] *See* Charles J. Dunlap, *The Hyper-Personalization of War*, GEO. J. INT'L AFFS. 108, 110 (2014) (citations omitted) (noting the phenomenon of data profile brokers). *See also*, Laura Hecht-Felella, *Federal Agencies Are Secretly Buying Consumer Data*, BRENNAN CENT. JUST. (Apr. 16, 2021), https://www.brennancenter.org/our-work/analysis-opinion/federal-agencies-are-secretly-buying-consumer-data (noting the phenomenon of U.S. state agencies purchasing private data).
[55] Xiao Qiang, *The Road to Digital Unfreedom*, 30 J. DEMOCRACY 53, 56–57 (2019).
[56] Zeyi Yang, *The Chinese Surveillance State Proves That the Idea of Privacy is More "Malleable" Than You'd Expect*, MIT TECH. REV. (Oct. 10, 2022), https://www.technologyreview.com/2022/10/10/1060982.
[57] Sheena Chesnut Greitens, *Dealing with Demand for China's Global Surveillance Exports*, BROOKINGS 5 (2020). One of China's law enforcement technologies, its "Safe City" program, has been adopted in countries such as Malta, Serbia, and Turkey *id.*
[58] *See infra* Part II.C.
[59] *See* David Lee, *Trump Supporters Lose Jobs and Businesses After Participation in Capitol Riot*, COURTHOUSE NEWS SERV. (Jan. 10, 2021), https://www.courthousenews.com/trump-supporters-

Adding calls for violence against the civilians being doxed likely only increases the strength of the deterrent, as well as actively assists the suppression of particular groups. Calls for ethnic and gender-based violence have frequently characterized modern conflict. The legality of propaganda campaigns which incentivize violence will be discussed in Part II.C, but history shows that armed groups sometimes use public messaging as a part of their war fighting efforts. Rwanda serves as an example, where media was used to incentivize ethnic killing.[60] Further, intimidation through threatened or actual sexual violence, an umbrella under which revenge porn falls (if uneasily), has historically been used in war.[61]

As there are state incentives to encourage violence against civilians with a nexus to war, the country examples merely underscore that this can now be done in a more individualized and rapid manner on the internet. Combined with the proliferating mass surveillance and data collection powers of modern states, countries around the world could possess both incentives and capacity to engage in mass doxing campaigns with potentially violent consequences.

II. IMPLICATIONS OF DOXING: LEGAL SOLUTIONS?

A. General Concepts

Four elements are essential for conduct to constitute a war crime. The event at issue must occur during an armed conflict; the offense "must be incorporated in an applicable criminal or prosecutorial code;" there must be a nexus between the conflict and the charged act; and "if the prosecuting tribunal is a law of war military commission . . . the charged offense must be an internationally recognized violation of the laws and customs of war."[62]

The harm this paper discusses has not been directly addressed by LOAC. On the one hand, doxing of civilians seems to invoke the core concept of distinction, or the basic rule that combatants must not make civilians the objects of attacks.[63] Making civilians the object of an attack is never acceptable and is internationally treated as a war crime.[64] However, LOAC has been hesitant to expand the definition of what constitutes an "attack" in the cyber realm.[65]

---

lose-jobs-and-businesses-after-participation-in-capitol-riot/ (offering an example of doxed protesters who lost jobs or faced other consequences).

[60] Mathias Ruzindana, *The Challenges of Understanding Kinyarwanda Key Terms Used to Instigate the 1994 Genocide in Rwanda*, *in* PROPAGANDA, WAR CRIMES TRIALS, AND INTERNATIONAL LAW: FROM SPEAKERS' CORNER TO WAR Crimes 145, 145 (Pedrag Dojčinoivić ed., 2012).

[61] *See, e.g.*, Prosecutor v. Kunarac, Appeals Judgment, Nos. IT-96-23 & IT-96-23/1-A, ¶ 195 (Int'l Crim. Trib. for the Former Yugoslavia June 12, 2022) (establishing rape as a war crime under customary international law).

[62] SOLIS, *supra* note 7 at 248.

[63] Protocol I, *supra* note 6, art. 48 (explaining the "[b]asic rule" in terms of civilian protections).

[64] *See* Rome Statute of the International Criminal Court art. 8, July 17, 1998, 2187 U.N.T.S. 90 [hereinafter Rome Statute] (showing the international criminalization of targeting civilians).

[65] *See* INT'L COMM. RED CROSS, THE PRINCIPLE OF DISTINCTION 2–3 (2023), https://www.icrc.org/sites/default/files/wysiwyg/war-and-law/03_distinction-0.pdf (discussing the current narrow view of the definition of attack and differing State views).

Approaching the problem from another angle, doxing could be considered a speech act aimed to incite others to violence, conduct which has been criminalized in the context of genocide and criminalized as instigation regarding crimes against humanity. Finally, arguments can be made that doxing falls into various descriptions of behavior already considered to be war crimes by the international community.

This Part explores each of these legal avenues and their fit with the described behavior to determine if customary international law would support the inclusion of doxing as a war crime. This paper does not have room to explore the nuance of how doxing might be treated differently in international armed conflicts ("IACs") and non-international armed conflicts ("NIACs") or the respective geographic constraints on application under either classification. This paper assumes that the protections offered to civilians do not meaningfully differ between the two, as customary international law protections have often filled in the "gaps" in treaty based NIAC protections.[66]

B. What is an "Attack" in the Cyber Context?

Civilians can proportionally be injured as collateral damage in attacks directed at military objectives, but they cannot be the objects of attacks themselves.[67] If doxing was considered an "attack" it would clearly be a war crime when directed against civilians. Under conventional understandings, an "attack" is defined under LOAC as an "[act] of violence against the adversary, whether in offense or in [defense]."[68] This definition feels straightforward in the physical, kinetic context,[69] but it does not easily map onto the modern world of warfare where cyber impacts all rules of the game.

1. The prevailing understanding of attack in cyberspace is limited. The Tallinn Manual is a research effort compiled by legal experts from around the world and provides a general overview of international law applicable to cyber operations.[70] The Tallinn Manual 2.0 was published in 2017,[71] and the Tallinn Manual 3.0 is expected to come out in 2026.[72] Tallinn 2.0 defines a cyberattack as

[66] *See* Jean-Marie Henckaerts, *Study on Customary International Law*, 87 INT'L REV. RED CROSS 175, 198–212 (2005) (offering a study of the customary international law protections which apply in both IACs and NIACs).
[67] SOLIS, *supra* note 7, at 228.
[68] Protocol I *supra* note 6, art. 49(1).
[69] *See contra*, Geoffrey S. Corn, *Beyond Human Shielding: Civilian Risk Exploitation and Indirect Civilian Targeting*, 96 INT'L LAW STUDS. 118 (2020) (offering an example of where this definition does not feel as straightforward in the conventional context by discussing human shielding).
[70] Eric Talbot Jensen, *The Tallinn Manual 2.0: Highlights and Insights*, 48 GEO. J. INT'L L. 735, 738.
[71] TALLINN 2.0, *supra* note 7.
[72] *See* Charlie Dunlap, *International Law and Cyber Ops: Q & A With Mike Schmitt About the Status of Tallinn 3.0*, LAWFIRE (Oct. 3, 2021), https://sites.duke.edu/lawfire/2021/10/03/international-law-and-cyber-ops-q-a-with-mike-schmitt-about-the-status-of-tallinn-3-0/ ("[Mike Schmitt] still anticipate[s] completion by 2025, with a Cambridge University Press release date in early 2026.")

a "cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects."[73] Tallinn 2.0 couches this understanding of attack in its violent effects, which are not limited to damage to the cyber system itself and could include "any reasonably foreseeable consequential damage, destruction, injury, or death."[74] But this reasonable foreseeability is limited; Tallinn 2.0 notes that "brief . . . interruption of non-essential cyber services, clearly do *not* qualify as cyberattacks."[75]

2. Doxing does not easily fit into this definition. There are two steps a state would be taking if it adopted a large-scale doxing campaign. Step one involves collecting or stealing civilian data. Step two involves publishing it online.

States legally collect civilian data all the time.[76] Acquiring the same data through hacking might be a crime under domestic law, but it likely would not constitute an attack under the definition described above, as "[n]on-violent operations, such as psychological cyber operations and cyber espionage, do not qualify as attacks."[77] Operations against data, which Tallinn 2.0 describes as a "non-physical entity," could still be considered attacks based on the foreseeable impacts on individuals or objects.[78] However, the majority position did not include cyber operations aimed at civilian datasets in the definition of attack.[79] There is indication that civilian data sets could be protected under other aspects of LOAC,[80] but it is likely that hacking and stealing civilian personal information would not currently be considered an attack.[81]

Regarding step two of the doxing process, or the publication of data, ignoring for now calls for violence attached to the data, it is similarly unlikely this would be considered an attack. Merely posting information online does not inherently lead to physical harms, and "psychological operations such as dropping leaflets or making propaganda broadcasts" are not considered to violate the principle of distinction.[82] Further, operations which "merely cause inconvenience or irritation to the civilian population" are not considered attacks.[83] While having personal data published online may feel like more than mere inconvenience to the

[73] TALLINN 2.0 *supra* note 7, at r. 92.
[74] *Id.* at para. 5. Note, this includes effects equivalent to "chemical, biological, or radiological attacks" which don't involve *kinetic* violent force but are considered attacks "as a matter of law." *Id.* at para. 3.
[75] TALLINN 2.0, *supra* note 7, at r. 71, para. 8.
[76] *See supra* Part I.B.1.
[77] TALLINN 2.0, *supra* note 7, at r. 92, para. 2.
[78] *Id.* at para. 6.
[79] *See id.* at r. 100, para. 7 (noting that only a minority of experts believed that operations against civilian data sets should also be contemplated under the definition of attack).
[80] *See id.* (explaining that a minority position believed alterations or deletions of civilian datasets should be protected because the focus should be on "the severity of the operation's consequences, not the nature of the harm").
[81] *But see id.*, at r. 94, para. 5 (showing that there might be a difference when it comes to *altering* civilian data, stating that, "[f]or instance, consider the case of a cyber operation intended to harm a particular individual by manipulating her medical information stored in a hospital's database. She would be the object of attack . . .").
[82] *Id.* at r. 93, para. 5.
[83] *Id.* at r. 92, para. 14.

affected individual, the publication does not directly cause injury or death to the individual. While Tallinn 2.0 indicates willingness to extend the effects of attacks to include "severe mental suffering that [is] tantamount to injury,"[84] it is unclear if mere publication of the data will lead to these effects.

While there is evidence of doxing being connected to violence, as seen in the Ukrainian example,[85] these effects may be too attenuated from the act of publishing the data. Ignoring the explicit calls for violence, the event itself must be "reasonably expected to cause injury or death to persons."[86] The bare bones elements of a third party seeing the information online and then acting on it are unlikely to fall within reasonable expectations, even if the information is accompanied by a call for violence. The actions of the third party which cause the injury or death might make publication a de minimis cause, which does not reach the relevant threshold of harm.[87] Thus, the act of publication also would not be an attack.

3. Doxing is also not clearly implicated in an expanded definition of attack. Many have argued that the definition of attack should be expanded in the cyber context considering the prevalence with which cyber elements dictate the conduct of modern warfare. The International Committee of the Red Cross ("ICRC") has expressed concern at the current definition of attack and noted it would advocate for a broader understanding.[88] The ICRC believes the "foreseeable direct and indirect (or reverberating) effects" should also be definitively included.[89] For example, this would include "the death of patients in intensive care units caused by a cyber operation on an electricity network that results in cutting off a hospital's electricity supply."[90] The ICRC would also expand protections over civilian datasets because of the importance they hold in modern life, but it remains unlikely that pure theft of data would be included in the expansion they advocate for.[91]

Arguments for expanded definitions often arise in the context of information operations, or "any coordinated or individual deployment of digital resources for cognitive purposes to change or reinforce attitudes or [behaviors] of the targeted audience."[92] Some have argued that cyber misinformation operations

---

[84] *See id.* at r. 92, para. 8 ("[T]he International Group of Experts agreed that it is, in light of the law of armed conflict's underlying humanitarian purposes, reasonable to extend the definition to serious illness and severe mental suffering that are tantamount to injury.") This conversation connects to the prohibition against spreading terror amongst the civilian population, which will be discussed in Part II.D.3.

[85] *See supra* Part I.A.1.

[86] TALLINN 2.0, *supra* note 7, at r. 92.

[87] *Id.* at para. 4.

[88] *International Humanitarian Law and Cyber Operations During Armed Conflicts*, 102 INT'L REV. RED CROSS 481, 489 (2021).

[89] *Id.*

[90] *Id.*

[91] *See id.* at 490. If a hacker altered the data and the alteration caused foreseeable effects, this might be included. *Id.*

[92] Dapo Akande, Antonio Coco, Talita de Souza Dias, Duncan B. Hollis, James C. O'Brien, & Tsvetelina van Betham, *Oxford Statement on International Law Protections in Cyberspace*, JUST SEC. (June 2, 2021), https://www.justsecurity.org/76742).

could rise to the level of a use of force in the *jus ad bellum* context based on the harm they cause,[93] and an argument can be made that this logic should similarly apply to the treatment of information operations during armed conflicts.[94] Others note that while the element of direct causation is difficult, information operations could be brought into the fold by expanding towards an understanding of "indirect attacks" which violate distinction.[95] The concept of indirect attacks has been applied to behavior like using human shields, where a defending state uses the attacking force as an agent to attack civilians.[96] In the context of doxing, this liability could be paralleled by stating that one armed party publishes information to cause third party agents to attack civilians. However, this linkage is shaky in the doxing context, particularly considering that the information is usually posted in public online forums.

Thus, doxing does not easily fit into either current or proposed understandings of what constitutes an attack in the cyber context. Still, Tallinn 2.0 does concede that cyber operations could "amount to war crimes and thus give rise to individual criminal responsibility under international law,"[97] suggesting that the conduct could still be penalized if another basis in customary international law was determined. Notably, Tallinn 2.0 notes that "posting online exhortations to continue the slaughter of civilians of a particular religious group during an armed conflict could amount to abetting if said exhortations were likely to be effective."[98] While this paper will not have the room to explore aiding and abetting liability, the mere inclusion of "exhortation" in this example suggests that speech crimes are a worthy next area for analysis.

C. Criminalizing Speech: Propaganda, Incitement, and Instigation

Speech "targeted" at civilians is similarly not considered an attack, but the conduct can constitute a war crime. Propaganda is an ambiguous term; Black's Law Dictionary defines it as "[t]he systematic dissemination of doctrine, rumor, or selected information to promote or injure a particular doctrine, view, or cause."[99] Publishing civilians' information online, whether falsified or not, is about communicating information to impact the perceptions of a target audience. Propaganda can be used as an umbrella term for speech acts designed to incite others to action, including violence, and doxing fits underneath this umbrella. This Section evaluates doxing in light of other criminalized forms of speech, as a form of propaganda to commit war crimes.

---

[93] Marko Milanovic & Michael N. Schmitt, *Cyber Attacks and Cyber (Mis)information Operations During a Pandemic*, 11 J. NAT'L SEC L. POL'Y 247, 269 (2020).
[94] Henning Lahmann, *Protecting the Global Information Space in Times of Armed Conflict*, 102 INT'L REV. RED CROSS 1227, 1241.
[95] Eian Katz, *Liar's War: Protecting Civilians from Disinformation During Armed Conflict*, 102 INT'L REV RED CROSS 659, 670 (2020) (citing Corn, *supra* note 69).
[96] Corn, *supra* note 69 at 129.
[97] TALLINN 2.0, *supra* note 7, at r. 84.
[98] *Id.* at para. 15.
[99] *Propaganda*, *Black's Law Dictionary* (11 ed. 2019).

Propaganda to commit war crimes is not explicitly incorporated in any relevant criminal codes, nor has it been recognized as customary international law. These are two necessary components for conduct to be considered a war crime.[100] However, key international tribunals are permeated with discussions regarding propaganda. While no statute of any international criminal tribunal includes propaganda as an offense itself,[101] propaganda has established the basis to prosecute individuals for incitement to genocide and instigation to crimes against humanity.[102] This Section proceeds by tracing the jurisprudence on the issue and its applicability to doxing.

1. The International Military Tribunal offers early guidance. During World War II, the Nazis used mass propaganda campaigns to bolster support for the war effort, Jewish suppression, and claims of German national and ethic superiority.[103] At the International Military Tribunal ("IMT"), the temporary international court created after the war, prosecutors often described these campaigns in the charges they levied against specific Nazis.[104] These charges were based on Article 6 of the Charter of the International Military Tribunal ("The Charter)", which gave the Tribunal jurisdiction over "[c]rimes against peace," "[w]ar crimes," and "[c]rimes against humanity."[105]

Prosecutors used propaganda as evidence in many of the trials, but two individuals, Julius Streicher and Hans Fritsche, were charged "exclusively on the basis of their speech and the influence they wielded over… the media."[106] Streicher advocated for the extermination of Jewish people through various publications.[107] Because he engaged in this conduct with knowledge of the ongoing Holocaust, the IMT found him guilty of incitement as a crime against humanity, even though incitement was not listed in the Charter.[108] The IMT ruled that this incitement was "persecution … in connection with war crimes . . . [which] constitutes a crime against humanity."[109] Thus, Streicher's conviction represents an early example of individual criminal responsibility based solely on propaganda generation.

Hans Fritzsche's charges of crimes against humanity, war crimes, and crimes against peace were also based solely on his speech acts.[110] Fritzsche was a Nazi government official who hosted his own radio show and was involved with

---

[100] *See* SOLIS, *supra* note 7, at 248.
[101] Michael G. Kearney, *Propaganda in the Jurisprudence of the ICTY*, *in* PROPAGANDA, WAR CRIMES TRIALS, AND INTERNATIONAL LAW: FROM SPEAKERS' CORNER TO WAR Crimes 231, 234 (Pedrag Dojčinoivić ed., 2012).
[102] Propaganda has also popped up in trials regarding aiding and abetting and joint criminal enterprises, but this paper will not have the time to dive into this side of the scholarship.
[103] MICHAEL G. KEARNEY, THE PROHIBITION OF PROPAGANDA FOR WAR IN INTERNATIONAL LAW 34–35 (2007).
[104] *Id.*
[105] *Id.* (citing Agreement for the Prosecution and Punishment of the Major War Criminals of the International Military Tribunal, art. 6, Aug. 8, 1945, 82 U.N.T.S. 280 [hereinafter The Charter]).
[106] KEARNEY, *supra* note 103, at 40.
[107] *Id.* at 41.
[108] *Id.*
[109] *Id.* at 41–42.
[110] *Id.* at 42.

daily newspapers.[111] However, the court acquitted him of crimes against humanity and war crimes charges because he did not "urge persecution or extermination of Jews" and it was unclear if he knew of the active extermination efforts.[112] Without this knowledge, the IMT did not consider Fritzsche's speech a direct enough cause of the underlying crimes, and thus not incitement.[113] Regarding crimes against peace, the IMT noted that Fritsche's aims were to "arouse popular sentiment in support of Hitler and the war effort," conduct which the IMT was not prepared to establish individual criminal liability for.[114] Notably, Major General I. T. Nikitchencko indicated in dissent that he would have expanded the theory of causation, claiming that German Fascism would not have been able to engage in as many crimes *without* propaganda efforts such as Fritzsche's.[115]

Ultimately, Streicher was the only individual convicted for propaganda efforts alone, suggesting that individual criminal responsibility required both explicit calls for the underlying crimes and knowledge that they were ongoing. When the Geneva Convention Relative to the Protection of Civilian Persons in Time of War ("Geneva IV") was drafted, neither incitement nor propaganda were included.[116] However, Streicher's conviction at the IMT did appear to have some influence, as "direct and public incitement [to genocide]" was included as a distinct punishable act in the Genocide Convention in 1948.[117]

A few of the principles underwriting Streicher's conviction seem immediately applicable to the doxing context. Similar to the difference between Streicher's conviction and Fritzsche's acquittal, it would likely be necessary that the doxing was done with the purpose of putting individuals in danger, rather than inflaming peoples' political passions in favor of one group over another. Similarly, the knowledge requirement of awareness that that harm was actively occurring could be an important component in an online forum, where posters could claim that they believed their content was harmless. However, the sparseness of these standards warrants further investigation of the jurisprudence.

2. The ICTR developed standards for incitement to genocide. Speech acts were raised again during the International Criminal Tribunal for Rwanda ("ICTR") [118] and the International Tribunal for the former Yugoslavia ("ICTY").[119] During the armed conflict in Rwanda, massive propaganda campaigns encouraged the

[111] *Id.*
[112] *Id.* at 43.
[113] *Id.*
[114] *Id.*
[115] Kearney, *supra* note 101, at 232. Nikitchencko was the judge from the Soviet Union. *Id.*
[116] *See* Geneva Convention Relative to the Protection of Civilian Persons in Time of War, Aug. 12, 1949, 75 U.N.T.S 287 [hereinafter Geneva IV].
[117] Convention on the Prevention and Punishment of the Crime of Genocide, art. 3, Dec. 9, 1948, 78 U.N.T.S 277.
[118] Statute of the International Criminal Tribunal for Rwanda art. 2, Nov. 8, 1994, S.C. Res. 995 [hereinafter ICTR Statute].
[119] Statute of the International Criminal Tribunal for the Former Yugoslavia, May 25, 1993, S.C. Res. 827 [hereinafter ICTY Statute].

genocide of an estimated over 500,000 Tutsi people.[120] In the armed conflict over the former Yugoslavia, Serbian leaders used propaganda to play on Serbian fears and encourage support for nationalist policies.[121] Each tribunal established key holdings on the use of propaganda in the "incitement to crimes of an international dimension."[122]

Propaganda holdings from the ICTR revolve around incitement to genocide. Like that of the ICTY,[123] the statute itself criminalized "direct and public incitement to commit genocide."[124] Genocide is the only offense in either statute in which incitement is listed as a basis of individual criminal liability, mirroring the gravity of Streicher being the sole individual prosecuted in the IMT specifically for his propaganda activities because he advocated for extermination.[125]

At the ICTR, incitement was considered "direct and public" if "the persons for whom the message was intended immediately grasped the implications thereof."[126] Jean-Paul Akayesu was found guilty of direct and public incitement to genocide based on his propaganda speech acts.[127] In one act determined to be direct and public, Akayesu encouraging a crowd to eliminate the "sole enemy" and then read aloud the names of individual Tutsi people, knowing that they would immediately be targets for violence.[128] In *Prosecutor v. Nahimana*[129] the ICTR clarified that incitement to genocide required "more than a vague or indirect suggestion," noting that hate speech alone would not meet the standard.[130] In *Nahimana*, use of "mass media," through the radio and tabloids, was also considered public and direct, as the defendants' radio station commonly discussed contempt for Tutsis and encouraged Hutus to exterminate them.[131] Thus, an incitement conviction called for explicit advocacy of activities akin to genocide.

Regarding causation, the Trial Chamber in *Nahimana* held that there did not need to be a "specific causation requirement linking the expression at issue with the demonstration of a direct effect" to satisfy the charge of incitement to genocide, citing *Streicher*.[132] Nahimana and his co-defendants were charged with incitement for activity which had occurred for years leading up to the genocide itself.[133] This relaxed standard for causation mirrored the lack of a requirement that the incitement

---

[120] *Numbers*, *Leave None to Tell the Story: Genocide in Rwanda*, HUM. RTS. WATCH, https://www.hrw.org/reports/1999/rwanda/Geno1-3-04.htm.
[121] Kearney, *supra* note 101, at 234 (citing Prosecutor v. Tadić, Case. No. IT-94-1, ¶ 130, (Int'l Crim. Trib. for the Former Yugoslavia May 7, 1997)).
[122] KEARNEY, *supra* note 103, at 222 (quoting Prosecutor v. Akeyasu, Case No. ICTR-96-4-T, Trial Judgment, ¶ 550 (Sept. 2, 1998)).
[123] ICTY Statute, *supra* note 119, art. 4.
[124] ICTR Statute, *supra* note 118, art. 2
[125] *See supra* notes 117 and accompanying text.
[126] *Akeyasu*, *supra* note 122, ¶ 558.
[127] *Id.* ¶ 550.
[128] KEARNEY, *supra* note 103, at 225.
[129] Case No. ICTR 99-52-A, Appeals Judgment (Int'l Crim. Trib. for the Former Yugoslavia Nov. 28, 2007)
[130] *Id.* ¶ 692.
[131] KEARNEY, *supra* note 103, at 223, 228.
[132] *Nahimana*, ¶ 1007.
[133] KEARNEY, *supra* note 103, at 222–23.

be "successful," as the incitement itself constituted an inchoate offense.[134] Both of these standards emphasized the danger of the speech acts themselves, underscoring that direct and public advocacy for such a grave crime was dangerous enough to warrant individual criminal liability.

Overall, convictions for incitement seemed to require language which directly advocated for conduct akin to genocide in a public manner. It is unclear if this requirement would be the same in the context of doxing, where information is often communicated through euphemism. The lack of a strict causation standard for incitement demonstrates the weight with which the international community treats the threat of genocide, and this standard was notably not extended to prosecution for crimes against humanity in other forums. It is likely this standard would need to be stricter in the doxing context, particularly as the connection between those who post on online forums and those who read the posts might be particularly attenuated.

3. The ICTY offers requirements for instigation. In the ICTY, like the IMT, propaganda was often used as evidence of liability for other crimes rather than conduct which constituted an individual criminal act.[135] The first trial of the Tribunal, *Prosecutor v. Tadić*,[136] established the prevalence of propaganda throughout the conflict, noting Serbian leaders' constant reliance on it.[137] The ICTY's Charter did not mention propaganda and incitement was only listed in reference to genocide, as in the ICTR.[138] However, instigation for any of the described crimes was listed as a basis of liability for individual criminal responsibility.[139] Propaganda ultimately held the most weight when it came to these charges of instigation.

Initially, in the cases which relied on propaganda, the court remained unwilling to elevate propaganda to an act which independently established liability for instigation. For example, Dario Kordić was convicted of the crime of persecution, a crime against humanity, after the prosecution included his propaganda activities in their arguments regarding instigation.[140] The Trial Court was not willing to say that his hateful propaganda independently rose to the level of instigation to a crime against humanity, noting that international case law had only previously criminalized speech acts as incitement to murder and extermination in *Streicher*, and genocide in cases in the ICTR.[141]

However, a few years later, Radoslav Brdjanin was convicted of instigating the crimes against humanity of deportation and forcible transfer based partially on

[134] *Nahimana*, *supra* note, ¶ 1017.
[135] Kearney, *supra* note 101, at 236.
[136] *Supra* note 121.
[137] Kearney, supra note 101, at 234.
[138] ICTY Statute, *supra* note 119, art. 2–5.
[139] *Id.* art. 6.
[140] Kearney, *supra* note, 101, at 234–35.
[141] Prosecutor v. Kordić, Case No. IT-95-14/2-T, Trial Judgment, ¶ 209 n.272 (Int'l Crim. Trib. for the Former Yugoslavia Feb. 26, 2001).

his propaganda activities.[142] The Trial Chamber noted that this was the "only reasonable conclusion" which could be drawn from decisions made after Brdjanin's "unambiguous public statements" where he "[called] upon the non-Serb population to leave . . . stating that only a small percentage of non-Serbs would be allowed to stay."[143] However, the Court noted that Brdjanin's political role, efforts to coordinate the crimes, and knowledge that "force and fear" would be required to complete the deportation and forcible transfer efforts further demonstrated that he "intended to induce" the crimes.[144]

In one of the last judgements of the ICTY, the Tribunal made an important move regarding speech acts and convicted Vojislav Šešelj of instigation to persecution based on one public speech in Hrtkovci, Serbia.[145] While the Trial Court was not satisfied that Šešelj's speech was a direct enough cause for the persecutory acts which followed it, the Court of Appeals noted that there was a "striking parallel between his inflammatory words and the acts subsequently perpetrated."[146] They ultimately determined instigation was made out because of the "specificity of Šešelj's words; the short time lapse; . . . regular threats to non-Serbians remaining in the village . . . [by] individuals who had attended the rally; Šešelj's visible influence over the crowd and the fact that the criminal acts corresponded closely to the content of the speech."[147]

In contrast, for some of Šešelj's other speeches, the Court of Appeals held that lapses in time affected the element of causation, and it wasn't clear that his statements "substantially contribute[d] to the commission of the specific crimes."[148] The Court also indicated that a personal relationship between the instigator and the perpetrator was required, and that the Hrtkovci speech was distinct in the personal hold Šešelj seemed to exercise over the crowd.[149]

Overall, this conviction was significant because Šešelj's single speech was equated with the gravity of instigating a crime against humanity. The propaganda stood out on its own, rather than supporting instigation in a more general way.

From the *Kordić*, *Brdjanin*, and *Šešelj* trials, the standard for instigation to crimes against humanity appears to have a stricter causality requirement than that of incitement to genocide under the ICTR. This is reasonable, particularly as the underlying crime must normally occur for liability for instigation to incur.[150]

---

[142] Prosecutor v. Brdjanin, Case No. IT-99-36, Trial Judgment, ¶¶ 574–77 (Int'l Crim. Trib. for the Former Yugoslavia Sep. 1, 2004).
[143] *Id.* ¶ 574.
[144] *Id.* ¶ 575.
[145] Wibke K. Timmerman, *International Speech Crimes: Šešelj Judgment*, *in* PROPAGANDA AND INTERNATIONAL CRIMINAL LAW: FROM COGNITION TO CRIMINALITY 105, 110 (Pedrag Dojčinoivić ed., 2nd ed. 2020).
[146] *Id.* (quoting Prosecutor v. Seslji, MICT-16-99-A, Appeals Judgment, ¶ 154 (Int'l Crim. Trib. for the Former Yugoslavia Apr. 11, 2018)).
[147] Timmerman, *supra* note 145, at 110.
[148] *Šešelj*, ¶ 132.
[149] Timmerman, *supra* note 145, at 119.
[150] *See* Partv V: Inchoate and Preparatory Acts, *in* MODES OF LIABILITY IN INTERNATIONAL CRIMINAL LAW, 337 (Jérôme de Hemptinne, Robert Roth & Elies van Sliedregt eds. 2019) ("[I]nstigation can only be sanctioned if it substantially contributed to the actual commission of an international crime.").

However, there still appear to be requirements that the speech be direct and "unambiguous," as in *Brdjanin*,[151] not mere hate speech, as in *Kordić*.[152] The causality requirement may be linked to time lapses and the "personal" connection between the instigator and perpetrator, as in *Šešelj*.[153] These standards seem to be more directly applicable to the context of doxing, as does the requirement that the underlying crime occur. This applicability will be explored more in Part III.

D. Other Treaty Based and Statutory Options

Beyond the cyber context and criminalization of speech acts, there are few other ways doxing could be incorporated into current LOAC frameworks. Looking to Geneva IV, Protocol I and the Rome Statute, the general harms brought about by state use of doxing might already fall under existing crimes. This Section quickly explores these statutory paths forward.

1. "Humiliating and degrading treatment" might not reach far enough. First, the Rome Statute lists "[c]omitting outrages upon personal dignity, in particular humiliating and degrading treatment" as a war crime in both IACs[154] and NIACs,[155] and Article 3 of Geneva IV (commonly called "Common Article 3" because it exists in all four Geneva Conventions in some form) and Protocol I both prohibit "outrages upon personal dignity, in particular humiliating and degrading treatment."[156] The ICRC defines "humiliating or degrading treatment" as "acts which cause real and serious humiliation or a serious outrage upon human dignity, and whose intensity is such that any reasonable person would feel outraged."[157] Significantly, the definition does not involve a physical component, and different terms are used for more physical abuses like torture.[158]

As examples of degrading treatment, the ICRC lists "forced public nudity" and "enduring the constant fear of being subjected to physical, mental or sexual violence."[159] These examples are offered in the context of commentary on the Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field ("Geneva I"),[160] but it is unlikely the ICRC would interpret the same phrase drastically differently in Geneva IV. Both examples have parallels in doxing; for instance, the use of revenge porn could be considered a form

[151] *Brdjanin*, *supra* note 142, ¶ 574.
[152] *Kordić*, *supra* note 141, ¶ 209.
[153] Timmerman, *supra* note 145, at 110.
[154] Rome Statute, *supra* note 64, art. 8(2)(b)(xxi).
[155] *Id.* art. 8(2)(c)(ii).
[156] Geneva IV, *supra* note 116, art. 3; Protocol I, *supra* note 6, art. 75(2)(b).
[157] *Torture and Other Forms of Ill-Treatment: The Definitions Used by the ICRC*, INT'L COM. RED CROSS (Jan. 1, 2016), https://www.icrc.org/en/document/torture-and-other-forms-ill-treatment-definitions-used-icrc.
[158] *Id.*
[159] Commentary of 2016, Article 3 – Conflicts not of an international character, Geneva Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field, Geneva, 12 August 1949, para 672. https://ihl-databases.icrc.org/en/ihl-treaties/gci-1949/article-3/commentary/2016#_Toc465169912.
[160] *Id.*

of "public nudity." While nudity online may cause a less immediate mental impact to the victim, the enduring quality of intimate images shared online could extend the length of the active harm. Further, as seen in Myanmar, doxing campaigns have pushed some victims into hiding out of constant fear of violence.[161]

However, there may be a necessary control element to these provisions. Protocol I forbids this conduct against those "in the power" of a state; and Geneva IV defines protected people as those who are "in the hands of" a state (of which they are not nationals).[162] Similarly, in the cyber context, the Tallinn Manual only notes protection against humiliating and degrading treatment in reference to detained parties.[163] While some have argued that the object and purpose of these provisions warrants a broader understanding in the cyber context,[164] it's unclear if these protections would currently cover situations where a victim is not actively in combatant control.

2. State responsibility to protect civilians is complicated in the cyber context. Protocol I calls on states to provide civilians "general protection against dangers arising from military operations" under Article 51[165] demands that "constant care . . . be taken to spare the civilian population, civilians and civilian objects" in offensive operations under Article 57,[166] and that the state take "necessary precautions to protect the civilian population . . . under their control against the dangers resulting from military operations" in defensive preparation under Article 58.[167]

Some have argued that these provisions should carry more weight in modern discussion of cyber operations, and that they should limit targeting civilians in instances that don't qualify as attacks.[168] The Tallinn Manual's Rule 114 interprets Protocol I Article 57 by stating that "[d]uring hostilities involving cyber operations, constant care shall be taken to spare the civilian population."[169] The Tallinn Manual also states that Rule 114 "supplements" the obligations of distinction and the rule of proportionality in the cyber context.[170] Overall, doxing would clearly violate the principle of "constant care" as military commanders would not be "avoid[ing] any unnecessary effects" of cyber operations on civilians.[171] Increasing the insecurity of individual civilians and making their personal information accessible to those who might wish them harm is likely an unnecessary effect.

---

[161] *See supra* notes 43–44 and accompanying text.
[162] Lahmann, *supra* note 94, at 1237.
[163] TALLINN 2.0, *supra* note 7, at r. 135.
[164] *See* Lahmann, *supra* note 94, at 1237 ("[C]onsidering the object and purpose of the obligation and the fact that the digital transformation has vastly expanded the possibilities to negatively impact a civilian person's dignity, an expansive interpretation that encompasses such conduct might be justifiable.").
[165] Protocol I, *supra* note 6, art. 51.
[166] *Id.* art. 57.
[167] *Id.* art. 58(c).
[168] Cordula Droege, *Get Off My Cloud: Cyber Warfare, International Humanitarian Law, and the Protection of Civilians*, 94 INT'L REV. RED CROSS 533, PAGE (June 1, 2012) (citation omitted).
[169] TALLINN 2.0, *supra* note 7, at r. 114.
[170] *Id.* at r. 114(3).
[171] *Id.* at r. 114(4).

Regarding defenders, they could be called upon to help mitigate the effects of doxing aimed at civilians by exercising control over their information spaces and removing content once it has been posted to the extent feasible. However, this likely strays beyond the responsibilities detailed in Article 58, which has traditionally been applied to preemptively removing civilians and civilian objects from the vicinity of military objectives.[172] Applying Article 58 to doxing in the cyber environment is different, as civilians themselves are the targets of these operations. Beyond blocking certain types of content from being posted, a move that would not be feasible in messaging forums like Telegram, there are few preemptive actions a state could take.

Further, beyond the general complexities of adapting each of these provisions to the unique orientation of doxing, a failure on the part of the state to protect civilians does not carry with it individual criminal responsibility.[173]

3. The prohibition against terrorizing civilians is a tricky fit. Beyond the general calls of Protocol I to protect the civilian population, Protocol I more specifically notes that "[a]cts or threats of violence the primary purpose of which is to spread terror among the civilian population are prohibited."[174] Article 33 of Geneva IV similarly states that "all measures… of terrorism are prohibited."[175] The ICRC believes this prohibition extends to operations carried out through "cyber means or methods of warfare"[176] and a minority of experts while drafting the Tallinn Manual believed in a "customary norm" which prohibits cyber operations "intended (whether the primary purpose or not) to [terrorize] the civilian population."[177]

State doxing campaigns invoke aspects of attacks meant to spread terror. Through doxing, a state, or an armed group, could engender the fear of being targeted online in an entire population. It could be used to punish and subdue dissenters, such as the Myanmar protesters. Terror spread through doxing campaigns could alter the way civilians live their day to day lives.

Similarly, the prohibition against terror has precedent in international jurisprudence. In the ICTY, Stanislav Galić was found guilty of the crime of terror as a violation of LOAC by instigating a campaign under which civilians were continuously shelled and sniped.[178] To establish the offense of terror, the Trial Chamber cited travaux preparatoires "to the 1974-77 Diplomatic Conference held

---

[172] *See* Eric Jenson, *Ukraine and the Defender's Obligations*, ARTS. OF WAR, LIEBER INST. WEST POINT (Mar. 2, 2022), https://lieber.westpoint.edu/ukraine-defenders-obligations/ ("If the military headquarters in Kyiv is near a residential district, the government should, to the maximum extent possible, relocate those civilians.").

[173] *See* Aural Sari, *Urban Warfare: The Obligations of Defenders*, LAWFARE (Jan. 24, 2019), (https://www.lawfareblog.com/urban-warfare-obligations-defenders) (noting that "[c]urrently, violations of Article 58 do not attract individual criminal liability").

[174] Protocol I *supra* note 6, art. 51(2).

[175] Geneva IV, *supra* note 116, art. 33.

[176] *International Humanitarian Law and Cyber Operations During Armed Conflicts*, *supra* note 88, at 486.

[177] TALLINN 2.0, *supra* note 7, at r. 98(7).

[178] Prosecutor v. Galić, Case No. IT-98-29-A, Appeals Judgment, ¶¶ 105–09, (Int'l Crim. Trib. for the Former Yugoslavia Nov. 30, 2006).

under the auspices of the ICRC."[179] In these documents, it was clear that many states believed propaganda could be a means of spreading terror among civilians.[180] As doxing could be considered a form of propaganda, this could lend support to viewing doxing as a form of terror which causes continuous fear of harm.

However, the prohibitions against terror all refer to conduct which constitutes an attack under LOAC or threat thereof.[181] In *Galić*, the conduct in question easily met this requirement, as it was sniper action.[182] As it is unclear if doxing constitutes an attack,[183] framing the crime as one of terror might be an unsuccessful effort. Overall, none of the potential treaty based or statutory paths to pull doxing into LOAC offer a clear means of success.

III. A VIABLE PATH FORWARD

A. Embracing a Theory of Instigation

Classifying doxing as a form of instigation to commit war crimes is likely the most workable path forward. The conduct does not currently fit into international understandings of "attacks" in the cyber realm,[184] and no other statutorily listed or treaty based war crimes offer a winning fit.[185] Thus, to pull doxing into the world of individual accountability, advocacy for a theory of instigation similar to that adopted in forums like the ICTY offers more solid ground. While pursuing this approach would bring with it a host of difficulties, the nature of war continues to change, and establishing new conceptions of war crimes as related to the cyber realm will be necessary to address current and future harms.

1. The instigation standards of the ICTY could be applied to war crimes. First, while individuals were only prosecuted in the ICTY for instigation to crimes against humanity, the potential nexus between doxing and armed conflict supports extending the theory to war crimes. Myanmar and Ukraine both provide examples of doxing being used during active wars, seemingly for efforts related to the conflict.[186] As in Myanmar, doxing could be used to promote direct violence against dissenting movements.[187] It could also be used to encourage and provoke ethnic or identity-based violence, paralleling the "mass media" strategy used by the *Nahimana* defendants.[188] Proving that an armed group engaged in doxing for the purpose of aiding or accomplishing a military objective would be crucial to

[179] Kearney, *supra* note 101, at 237.
[180] *Id.* at 237. The US was the sole dissenter. *Id.*
[181] Lahmann, *supra* note 94, at 1238 (noting that "the communicative act in question must either amount to an attack within the meaning of IHL or a threat thereof" while discussing making disinformation a form of terror).
[182] *Galić*, *supra* note 178, ¶ 105.
[183] *See supra* notes 74–86 and accompanying text.
[184] *See supra* Part II.A.
[185] *See supra* Part II.D.
[186] *See supra* Part I.A.
[187] *See supra* Part I.A.2.
[188] *See supra* note 131 and accompanying text.

establishing the underlying conduct as a war crime. Still, the rising ways in which online forums could be leveraged to control and harm civilians in war supports expanding the instigation standard beyond crimes against humanity.

The holdings of the ICTY regarding instigation could reasonably be adopted to the context of doxing. At the outset, criminalizing doxing would likely require "unambiguous" speech, as in *Brdjanin*.[189] This would likely mean that publishing civilians' personal information without encouraging harm against them would not rise to the level of unambiguous speech, even in the context of revenge porn. There would need to be some level of encouragement or direction of action alongside the provision of information. This encouragement could feasibly come through the post itself, as in examples from Myanmar,[190] or in the way the publishing forum portrays itself. For example, if the Peacemaker blog in Ukraine did portray itself as something more akin to a "hit list," then the doxed information on the website could arguably fall into the realm of "unambiguous" instigating speech. Overall, the requirement of unambiguous language would likely be necessary to differentiate between doxing which might cause a mere "inconvenience" to civilians,[191] and that which presents a true threat of physical harm.

Further, similar to the standard in *Šešelj*, it would be prudent to require some form of personal connection between the instigator and the perpetrator, which would assist in establishing causation. The ICTY was only willing to hold Šešelj accountable for instigating crimes against humanity for *one* of his speeches because he seemed to have a personal hold over that particular crowd.[192] The court cited Šešelj's specific language, the short time lapse between his speech and action taken, the actions taken by others who attended his speech, and his "visible influence" over the crowd.[193] Passing information through the internet inherently seems less personal than speaking directly to a crowd, particularly if the doxed information was posted in public forums or passed through messaging platforms far beyond where the information was initially posted.

Perpetrators' online engagement with the posted content could help to establish a personal relationship, and shorter timelines between the content being posted and the crimes being committed could also speak to a more personal relationship. A perpetrator's re-posting or referencing of the doxed content and any parallels between the specific language used in the initial posts and the action ultimately taken could also signify a tighter connection. Evidence that the perpetrator relied on the information exposed through doxing, such as details of the victim's location, would also support causation. It would further be prudent to include the requirement from *Streicher*, that the instigator had knowledge war crimes of the type they were instigating were actually happening.[194] Ultimately, it would be required to prove that the individual's action was not a de minimis cause of the war crime.

---

[189] *See Brdjanin*, *supra* note 142, ¶574.
[190] *See supra* note 40 and accompanying text.
[191] *See* TALLINN 2.0, *supra* note 6, r. 92 para. 14 (indicating that cyber operations which merely inconvenience civilians are not treated as worthy of criminalization in war).
[192] Timmerman, *supra* note 145, at 110.
[193] *Id.*
[194] *See supra* notes 106–109 and accompanying text.

Finally, unlike incitement to genocide, liability for instigation to war crimes through doxing would also require that the underlying crimes actually occur. While pursuing an inchoate standard like that of incitement to genocide would lead to the criminalization of a broader swath of doxing conduct, past tribunals have hesitated to expand incitement for good reason.[195] The position of genocide is unique in international law, as seen in the way it has been removed from the category of "crimes against humanity" to a specific category itself.[196] Criminalization of genocide's inchoate elements should thus remain unique from a deterrence perspective. While all harms to civilians during war are tragic, individual instances of harm should not be equated with the overall harm and societal danger of genocide.

2. There are drawbacks to this theory, but it is worth pursing nonetheless: There are many critiques which could be raised to adopting a theory of instigation. Importantly, attribution is very difficult in the cyber context, and could be even more difficult in the context of social media and double encrypted messaging platforms.[197] Even if an armed group was engaging in this conduct at a massive scale, it could be difficult to determine exactly which instances of doxing the group was responsible for. However, identifying the difficulty of prosecuting these crimes does not mean criminalization should not occur.

Still, one could feasibly argue that expanding the criminalization of speech in conflict could be a "slippery slope," in which free speech could ultimately be a victim. Requiring that the underlying crime occur for liability to establish cuts against this critique and prevents the over-broadening of the range of criminalized activity. Yet, the necessary narrowing of the range of conduct covered might be *too* narrow for some, who may believe that instigators should be punished even when their attempts at instigation are unsuccessful. However, this argument runs into the same logic which supports pursuing a standard of instigation over that used for incitement to genocide.

Ultimately, none of these critiques fully undercut the utility of pulling doxing into LOAC. Like all legal doctrines, it would ultimately need to be adapted to the facts at hand and widened and narrowed as needed. However, as the facts of current armed conflicts show that doxing will likely continue to be a persistent phenomenon, the international community should hasten to begin this process of adaption.

---

[195] *See supra* notes 110–115 and accompanying text (noting that the IMT was not willing to prosecute Fritzsche for his propaganda activities because they did not aim to incite extermination)
[196] *See* Convention on the Prevention and Punishment of the Crime of Genocide, *supra* note ##, pmbl. para. 2, art. 1 ("Recognizing that at all periods of history genocide has inflicted great losses on humanity . . . "[t]he Contracting Parties confirm that genocide, whether committed in time of peace or in time of war, is a crime under international law which they undertake to prevent and to punish.").
[197] *See* Zak Doffman, *Yes, Telegram Really is 'Dangerous' For You*, FORBES (Apr. 22, 2021, 6:00 AM), https://www.forbes.com/sites/zakdoffman/2021/04/22/forget-whatsapp-new-telegram-warning-for-millions-of-windows-10-users/ (describing some attribution difficulties on a platform like Telegram).

CONCLUSION

As long as the human race has existed, people have developed new ways to harm one another. Increasing use of cyber tactics, operations, and strategies are testing the bounds of LOAC to respond to modern conflict. The phenomenon of doxing is an issue for many existing legal frameworks, domestic and international alike. But, as seen in Ukraine and Myanmar, its potential use in armed conflict is especially concerning, particularly given the legal gray area in which the conduct currently falls. The distance between information shared online and physical harm continues to grow smaller, and states are poised to capitalize on the shrinkage during conflict. However, rushing to criminalize the conduct in the broadest means possible is not a prudent approach, as such a jump could lessen the unique weight of a crime like incitement to genocide. Pursuing a theory of individual criminal liability through instigation to commit war crimes offers LOAC a means of adapting to the new threat of doxing, while remaining rooted in past treatment of propaganda during war. Pushing LOAC to expand incrementally while retaining strong connections to its historic core must be the international community's united goal in the wars to come, as the legal questions imposed by modern conflict will continue to demand answers.