

Robin Cameron, Gregory Stratton, and Anastasia Powell

## 30 Hate crime and networked hate

**Abstract:** The concept of hate crime is presented here as part of a wider structure of networked hate. Drawing on a digital criminological approach, this account seeks to highlight the wider structures that embolden and enable such crimes. These structures are defined by deep but less immediate and spectacular forms of harm that can be understood as a form of ‘slow harm.’ To avoid networked hate manifesting as hate crime we need to counter the slow violence that continues largely unaddressed in digital society.

**Keywords:** hate crime, violence, networks, digital, criminology

### Introduction

In July 2023, the social network Threads was launched, accompanied by a proliferation of hate across its communities. Even Mark Zuckerberg, the CEO of parent company Meta, was not safe. A reply to one of his early posts contained the ominous reply: “Hey Jew.” This was not a one-off post by this user, who made other racially hateful comments. Their first post on the new social media platform was: “It’s the Jews,” while their follow-up was a short video of notorious neo-Nazi Nick Fuentes, uttering extremely racist language. The user remains active at time of writing, despite a string of bigoted and hateful posts.

Similar patterns of hate were observable when notorious users from other platforms brought their networks of hateful infamy to Threads. White supremacist Roblox celebrity, Jon Mill, posted: “can I say the n word on this app or is that a big no no” (sic). His followers proceeded to spell out the racial slur, letter-by-letter, in the replies. Jack Posobiec, another white supremacist and progenitor of the ‘Pizzagate’ conspiracy, posted a series of hateful posts about LGBTQIA+ communities, most likely deliberate disinformation intended to test new moderation policies. Both Mill and Posobiec remain active on Twitter and Threads at time of writing, despite a string of bigoted and hateful posts.

These recent examples highlight that hateful online behavior is not the isolated actions of bigoted individuals. Rather, the rapid rise of targeted hate on Threads reflects the slow violence of accumulated networked hate, occurring through the operation and interaction of social, technical, and corporate networks. In a *digital society* various technological platforms are, on the one hand, mere tools for continuing longstanding forms of hate. They can, however, amplify, extend, and more deeply embed the harms of hate-based violence, discrimination, and exclusion; with very real, and potentially as yet unrealized consequences. In our conceptualization of Digital Criminology, *digital society* refers to an “understanding of the mutual and reciprocal shaping of technology and society” (Powell et al., 2018: 22). Digital criminology responds to this

by showing how the study of crime can move beyond ‘online/offline,’ ‘cyber/real,’ or ‘virtual/terrestrial’ dichotomies to recognize the “integration of the digital” in everyday life (Powell et al., 2018: 7).

In this chapter, we seek to apply a digital criminological approach to understanding both the harms of networked hate, and actions to address it. First, we elaborate the use and limitations of the conventional criminological focus on *hate crime* or *hate speech*, before supplementing the concept of *networked hate*. The normative effect of networked hate emboldens bigotry, and the cumulative harm of this hate is borne by vulnerable and marginalized people, constituting a ‘slow violence.’ Then, we discuss the implications of a digital criminology of networked hate that seeks to address legal, technical, and socio-cultural measures to reduce and prevent its harms (Wood, 2021). Finally, we discuss challenges that remain in the field and avenues for continued criminological research.

## Limitations of ‘hate crime’ in a digital society

There are many terms used across criminological and socio-legal literature addressing concerns with online hate. Terms such as hate crime, cyber hate, hate speech, and hateful extremism have all been widely used for example, as well as the identification of specific issues including gender-based hate or online misogyny, cyber racism, homophobic and/or transphobic aggression, and disability-based hate. Though there is not the scope here to consider in-depth the implications of these different terms, we do find it useful to consider how criminology has traditionally engaged with online ‘hate crime,’ before making a case for a broadening of focus towards the harms of networked hate (Powell et al., 2018).

Though legislation varies in different countries globally, in the simplest of terms, a hate crime is a crime that is motivated by a bias against people or groups with specific characteristics that are protected by law. This often involves both a crime under a legislated crimes act, and protected characteristics under human rights or anti-discrimination legislation. Many countries have human rights and/or anti-discrimination legislation that identifies protected characteristics such as race, religion, sexual orientation, gender identity, and disability. Though the criminal element may often be a physical crime, such as homicide, assault, threats and/or vandalism, some countries have legislated specific hate crimes that are equally applicable to digitally mediated harm.

In some Australian states, for example, it is a specific criminal offense to knowingly display a Nazi symbol in public without a reasonable excuse, and there are specific laws against publicly threatening or inciting violence on grounds of race, religion, sexual orientation, gender identity or intersex or HIV/AIDS status (e.g., Crimes Act 1900 [NSW]). In this context, ‘publicly’ can include any form of communication such as speaking, writing, broadcasting and communicating through social media or other electronic methods to the public (see Social Media by Twigt). Additionally, many jurisdictions have specific civil law anti-discrimination protections against hate speech, which

may not constitute criminal offenses, but rather provide a mechanism for redress (such as through an apology, compensation, or other reparations).

There are, however, several limitations with hate crime, and even hate speech, as concepts for a criminological approach to the harms that emerge from the proliferation of hateful discourse in online spaces. Specifically, under these concepts the onus often remains on individuals to seek recourse when they have been specifically targeted (such as in order to report a hate crime), or to have specifically experienced a harmful incident which can form the basis of a claim for redress (such as hateful speech that resulted in lost employment, and/or psychological injury). Additionally, there are many challenges and tensions in policing and law enforcement of hate crime legislation across jurisdictions globally (Clough, 2015). For example, many vulnerable groups are not protected under current legislative frameworks (see Vulnerability by Ranchordas and Beck). Gender, for instance, may be a protected characteristic in specific anti-discrimination legislation, but not a protected characteristic under specific laws against publicly threatening or inciting violence (D'Souza et al., 2018). Additionally, protections from hate speech often come into conflict with free speech protections. Attempts to address hate at the level of individual speech acts are likely to be perceived by those it restricts as an encroachment on free speech (Powell et al., 2018, 124–126). Not only are restrictions on speech likely to further inflame such bigoted impulses, they fail to address the broader networks that sustain and embolden those views.

## The ‘slow violence’ of networked hate

While online communications cannot be understood as the *cause* of networked hate, neither are social media and other digital platforms free of influence in co-producing and amplifying racism, misogyny, and other forms of bigotry. In other words, digital technologies are not mere tools of human action and interaction, nor are humans merely responsive to the direction of such technologies; networked hate is, like so many other practices in a digital society, inherently *technosocial*. The very human and social problem of discriminatory ideology and behaviors, combines with the very technological problem of networked communications that enable hateful content to spread with a virality and a perceived legitimacy that draws more humans in to engage with it. Research continues to show that hateful content spreads more widely and rapidly than more neutral or positive content (Daniels, 2018; Esko and Koulu, 2022). But this is no accident. The virality of networked hate is arguably driven by several key factors, such as a broader mainstreaming of extremism (Miller-Idriss, 2018), the function of algorithmic content delivery (Daniels, 2018), and the profit imperatives of technology companies themselves (Powell et al., 2018). Together these technosocial processes result in a form of ‘slow violence’ (Nixon, 2011; see also Brydolf-Horwitz, 2022) that, we suggest, is incrementally and insidiously causing harm. However, it does not attract the same criminological or legal recognition that arises from hate-based attacks or homicide.

Together the mainstreaming of extremism, algorithmic content delivery, and profit imperatives have had the effect of amplifying the social harms of inequality, discrimination, and hate. Yet, the severity and scale of these harms is arguably under-recognized within criminology; that is, until they erupt into criminal incidents of physical harm such as violent protests, vandalism, assaults, or fatalities. It is here that Rob Nixon's (2011) concept of slow violence could prove useful in drawing greater scholarly, and indeed public, attention to the harms of networked hate. As Nixon (2011: 2) describes: "Violence is customarily conceived as an event or action that is immediate in time, explosive and spectacular in space, and as erupting into instant sensational visibility." Slow violence, by contrast, "occurs gradually and out of sight, a violence of delayed destruction that is dispersed across time and space, an attritional violence that is typically not viewed as violence at all" (Nixon, 2011: 2). Applying this concept to the specific examples of online harms such as harassment and bullying, Brydolf-Horwitz (2022) further elaborates, highlighting the ways in which online abuse frequently fails to attract the same legal and societal outrage and action compared with physical forms of abuse.

The slow violence of networked hate has uneven impacts; and its harms are both cumulative and collective in nature. The *unevenness* of networked hate is evident in the disproportionate harms experienced by minorities and marginalized groups; such that racial, sexuality, gender identity, and disability minorities are among those most often targeted (Powell et al., 2018). These same groups already bear the brunt of discrimination in settings such as education, the workplace, healthcare, and civic participation, as well as being disproportionately targeted by physical violence and abuse (Powell et al., 2018). The *cumulativeness* of networked hate meanwhile, refers to the psychological toll that builds with each daily or regular insult, microaggression, act of exclusion, denial of personhood or identity, and that might be further interspersed with more intensive incidents such as a barrage of hateful communications, threats, blackmail, or abuse. Finally, by the *collective* nature of the harms of networked hate, we mean to highlight that marginalized and at-risk groups can still be harmed by the effects of amplified online hate and/or hate-based disinformation, even without experiencing online hate directly themselves. When hate-based discourse, misinformation, insults, and aggression towards minorities proliferate online, they reinforce and embolden these views and behaviors in other domains, such as education, workplaces, and healthcare referred to above; in effect, widening the discrimination and exclusion gap.

## **Preventing hate crime by addressing networked hate**

As the discussion in this chapter has identified, networked hate is an inherently technosocial problem. Addressing it is, in part, hampered by the lack of urgency that comes with forms of slow violence. Among the key implications of a digital criminological ap-

proach to networked hate, is a call for research and policy leadership that recognizes the slow violence of networked hate and advocates for action to address these serious, yet not always criminal, harms. Slow violence is likely where harms arising from inequality are present. Towards this goal, integrating an awareness of, and concern for, how compounding forms of inequality perpetuate slow violence should be integrated into criminological research agendas to build greater empirical understanding how it can exacerbate the impact of crime against protected and vulnerable people. A further implication of the understanding of networked hate is that we cannot simply code our way out of this problem. In other words, given the technosocial nature of networked hate in our digital society; it will take a combination of regulatory, legal, technical, and socio-cultural solutions to address these harms.

There is arguably a role for criminologists to contribute to regulatory and legal reforms that may be needed, but also to work collaboratively outside of government agencies and with technology providers to inform technical solutions as well as community programs. Given the global scope and economic power of tech companies, national laws alone will not alone create the required change. Attempts to create transnational normative change can be seen in an initiative such as the 'Christchurch Call' in 2019, led by then Prime Minister of New Zealand Jacinda Arden as a response to the far-right terror attacks (MFAT, 2019). This led to 120 government, online service providers, and global civil society organizations signing onto a voluntary commitment to a co-ordinated approach to reduce violent online extremism (Christchurch Call, 2019). In the realm of regulatory and legal reform, there is a wide field of cybercrime research which acknowledges the challenges of criminal approaches to crime that occurs via global networks (Clough, 2015; Holt and Bossler, 2020). Policing resources, jurisdictional barriers, and incongruous legislation are among some of these key challenges that would apply just as readily to any legal reforms addressing networked hate.

Nonetheless, there are some promising examples, and the European Union is the clear leader in the implementation of laws and regulations that affect the behavior of large tech companies. This chapter opened with examples of networks of coordinated hate speech on Threads, however, this is not evident in the member states of the EU because at the time of writing Threads is not available there. This is because Threads is not yet compliant with regulatory laws passed by the EU, causing Meta to delay its European launch. In late-2022 the EU passed the Digital Services Act that places an obligation on online service providers to ensure that "what is illegal offline, is also illegal online" (European Council, 2021). This is the fourth in a suite of EU laws, following on from the General Data Protection Regulation, the Digital Markets Act, and the Data Governance Act intended to encourage safe, sustainable, transparent, and fair business practices by tech companies.

Meanwhile, in the Australian context the eSafety Commissioner, an independent authority, has been established by government to receive complaints of online harm. In responding to these complaints, the commissioner has been legislatively empowered to direct technology providers to remove content and impose financial penalties if take-down orders are not enacted (eSafety Commissioner, 2021). However, after writing

to Twitter or 'X' with concerns about online hate and disinformation in 2022, the Office of the eSafety Commissioner has confirmed that there are no longer any Australian-based representatives to respond to requests to remove harmful material and online hate (Evans, 2023). They now face potential fines of up to \$A700,000 per day due to non-compliance with eSafety Commissioner requests (Doran, 2023). Enquiries to the press office email have been met with a 'poop emoji' reply from Twitter or 'X' (Treisman, 2023).

Technical solutions should seek to build in inclusion and safety by design. For example, *inclusion by design* refers to the intentional and proactive incorporation of diversity and accessibility considerations into the planning, development, and design of devices, services, applications, or systems (e.g., Holmes, 2020). Further, *safety by design* calls attention to identifying and mitigating potential risks and vulnerabilities during the early stages of design, rather than addressing them as an afterthought or reactively after a product or system has been developed. Together, these efforts seek to ensure that technologies are accessible, usable, and proactively address the safety needs or risks for a wide range of individuals, including those with a diversity of backgrounds. In applying these processes to the problem of networked hate, technologists could proactively consider empowering at-risk users with more control over curation and filtering of the content they receive as well as effective tools for blocking, reporting, and removal of online hate. However, it is also vital to consider mechanisms for proactively disrupting the harmful actions of users engaged in online hate. For example, through continuously improving artificial intelligence (AI) (see Artificial Intelligence by Van Brakel) algorithms to identify and remove hate speech, offensive content, and harmful behavior in real-time (Windisch et al., 2022), as well as identifying and deplatforming leaders and key influencers in hate networks (Thomas and Wahedi, 2023).

Finally, addressing networked hate also requires socio-cultural solutions aimed at tackling the root causes of hate within society through strategies that promote understanding, empathy, and inclusivity. These approaches seek to change attitudes, behaviors, and social norms to create a more equitable and harmonious society. Specific examples include fostering inclusive online communities through clear and enforced conduct standards; prevention campaigns and resources for countering hate and discrimination; supporting parents with resources for promoting digital safety and inclusion conversations at home; and digital citizenship education in schools (see, e.g., Citron and Norton, 2011).

## Conclusion

Networked hate encompasses a profound but less 'immediate' and 'spectacular' form of harm than traditional hate crime. It is, we suggest, a form of slow violence that substantially shifts social norms of equality and inclusion to those of discrimination and exclusion (Nixon, 2011; see also Brydolf-Horwitz, 2022). Unfortunately, as criminologists, we have tended to focus our attention almost exclusively on explosive events,

such as when online hate transgresses into physical assault, homicide, and/or terrorism. On this basis we can say that:

- digital criminology acknowledges that the technosocial nature of our digital society requires attentiveness to the dispersed, cumulative, and often hidden forms of networked hate.
- to minimize the potential for a portion of these harms to convert to more traditionally recognizable violence, we need to counter the insidious slow violence that is growing largely unabated. Historical harms, and ongoing practices of hate-based discrimination, harassment, and violence cannot be coded away.
- meaningfully addressing the harms of networked hate will require sustained intervention into the profit-driven development of social technologies.

The seriousness of hate crimes and wider slow violence of networked hate requires a whole of society response, encompassing legal, technical, and social solutions. In addition to formal justice responses, public and private organizations need to ensure their technology use actively constructs an era of equality by design as a means of crime prevention.

## Suggested reading

Brydolf-Horwitz, R. (2022). Embodied and entangled: Slow violence and harm via digital technologies. *Environment and Planning C: Politics and Space*, 40(2), 391–408.

Powell, A., Stratton, G., & Cameron, R. (2018). *Digital Criminology: Crime and Justice in Digital Society*. London: Routledge.

Wood, M. A. (2021). Rethinking how technologies harm. *The British Journal of Criminology*, 61(3), 627–647.

## References

Brydolf-Horwitz, R. (2022). Embodied and entangled: Slow violence and harm via digital technologies. *Environment and Planning C: Politics and Space*, 40(2), 391–408.

Christchurch Call. (2019). *To Eliminate Terrorist & Violent Extremist Content Online*. Christchurch Call (accessed 17 August 2023). <https://www.christchurchcall.com/>.

Citron, D. K., & Norton, H. (2011). Intermediaries and hate speech: Fostering digital citizenship for our information age. *Boston University Law Review*, 91, 1435–1484.

Clough, J. (2015). *Principles of Cybercrime*. Cambridge: Cambridge University Press.

Daniels, J. (2018). The algorithmic rise of the “alt-right”. *Contexts*, 17(1), 60–65.

Doran, M. (2023). *Australia issues Elon Musk's Twitter with a 'please explain' notice over surge in online hate*. ABC News (accessed 17 August 2023). <https://www.abc.net.au/news/2023-06-22/cph-e-safety-commissioner-compels-twitter-to-produce-documents/102505546>.

D'Souza, T., Griffin, L., Shackleton, N., & Walt, D. (2018). Harming women with words: The failure of Australian law to prohibit gendered hate speech. *The University of New South Wales Law Journal*, 41(3), 939–976.

eSafety Commissioner. (2021). *Additional Government funding to keep Australians safer online*. eSafety Commissioner (accessed 17 August 2023). <https://www.esafety.gov.au/newsroom/media-releases/additional-government-funding-keep-australians-safer-online>.

Esko, T., & Koulu, R. (2022). Rethinking research on social harms in an algorithmic context. *Justice, Power and Resistance*, 5(3), 307 – 313.

European Council. (2021). *What is illegal offline should be illegal online: Council agrees on position on the Digital Services Act*. European Council (accessed 17 August 2023). <https://www.consilium.europa.eu/en/press/press-releases/2021/11/25/what-is-illegal-offline-should-be-illegal-online-council-agrees-on-position-on-the-digital-services-act/>.

Evans, J. (2023). *Twitter has axed Australian team the eSafety regulator contacted to report child abuse material*. ABC News (accessed 17 August 2023). <https://www.abc.net.au/news/2023-02-20/australian-twitter-staff-dealing-with-child-abuse-material-axed/101998040>.

Holmes, K. (2020). *Mismatch: How Inclusion Shapes Design*. Cambridge, MA: MIT Press.

Holt, T. J., & Bossler, A. M. (eds.) (2020). *The Palgrave Handbook of International Cybercrime and Cyberdeviance*. London: Palgrave Macmillan.

MFAT. (2019). *Christchurch Call*. Ministry of Foreign Affairs & Trade website, accessed 17 August 2023. <https://www.mfat.govt.nz/en/peace-rights-and-security/international-security/christchurch-call/>.

Miller-Idriss, C. (2018). *The Extreme Gone Mainstream: Commercialization and Far Right Youth Culture in Germany*. Princeton, NJ: Princeton University Press.

Nixon, R. (2011). *Slow Violence and the Environmentalism of the Poor*. Cambridge, MA: Harvard University Press.

Powell, A., Stratton, G., & Cameron, R. (2018). *Digital Criminology: Crime and Justice in Digital Society*. London: Routledge.

Thomas, D. R., & Wahedi, L. A. (2023). Disrupting hate: The effect of deplatforming hate organizations on their online audience. *Proceedings of the National Academy of Sciences*, 120(24), e2214080120.

Treisman, R. 2023. *Got a question for Twitter's press team? The answer will be a poop emoji*. NPR (accessed 17 August 2023). <https://www.npr.org/2023/03/20/1164654551/twitter-poop-emoji-elon-musk>.

Windisch, S., Wiedlitzka, S., OlaghHERE, A., & Jenaway, E. (2022). Online interventions for reducing hate speech and cyberhate: A systematic review. *Campbell Systematic Reviews*, 18(2), e1243.

Wood, M. A. (2021). Rethinking how technologies harm. *The British Journal of Criminology*, 61(3), 627 – 647.