



Queensland University of Technology
Brisbane Australia

This may be the author's version of a work that was submitted/accepted for publication in the following source:

[Matamoros-Fernandez, Ariadna, Nelson, Lucinda, & Suzor, Nicolas \(2021\)](#)

QUT Digital Media Research Centre submission regarding 2021 report on combating glorification of Nazism, neo-Nazism and other practices.

This file was downloaded from: <https://eprints.qut.edu.au/216867/>

© Consult author(s) regarding copyright matters

This work is covered by copyright. Unless the document is being made available under a Creative Commons Licence, you must assume that re-use is limited to personal use and that permission from the copyright owner must be obtained for all other uses. If the document is available under a Creative Commons License (or other specified license) then refer to the Licence for details of permitted re-use. It is a condition of access that users recognise and abide by the legal requirements associated with these rights. If you believe that this work infringes copyright please provide details by email to qut.copyright@qut.edu.au

Notice: *Please note that this document may not be the Version of Record (i.e. published version) of the work. Author manuscript versions (as Submitted for peer review or as Accepted for publication after peer review) can be identified by an absence of publisher branding and/or typeset appearance. If there is any doubt, please refer to the published source.*

Special Rapporteur on contemporary forms of racism,
racial discrimination, xenophobia and related intolerance
By email: racism@ohchr.org

12 July 2021



QUT Digital Media Research Centre submission regarding 2021 report on combating glorification of Nazism, neo-Nazism and other practices

Prepared by Dr Ariadna Matamoros-Fernández, Lucinda Nelson and Prof Nicolas Suzor

We are researchers in QUT's Digital Media Research Centre and the QUT node of the ARC Centre of Excellence for Automated Decision-Making and Society. The DMRC is a global leader in digital humanities and social science research with a focus on communication, media, and the law. The ADM+S is a cross-disciplinary national research centre that supports the development of responsible, ethical and inclusive automated decision-making systems.

For more information about this submission, contact Lucinda Nelson:
l5.nelson@qut.edu.au.

Executive summary

We welcome the opportunity to contribute to this report. In general terms, our research focuses on the challenges of addressing the spread of harmful content online. One of the biggest challenges of content moderation is that platforms primarily focus on the most acute and blatant manifestations of racism, to the exclusion of more common forms of racism. Our research suggests that dealing with the normalisation of racism in everyday communication ought to be a key priority for digital platforms. We suggest that, to combat the glorification of Nazism and neo-Nazism, it is important to understand and address that:

1. Humour and play are used online to mask racist speech and conduct, incite hatred, and recruit members of extremist groups.
2. Extremists adopt other personas online to incite tension, spread disinformation, and dehumanise racial minorities.
3. Extremists hijack and appropriate keywords from social justice movements to manipulate search algorithms and promote racist content.
4. Bad actors use bots and automation to promote racist agendas across digital platforms.

These issues highlight the need for more research, advocacy, and experimentation to develop more sophisticated, cross-platform, contextually sensitive approaches to content moderation.

Use of humour

Extremist groups use humour online to discriminate, incite hatred,¹ and recruit new members.² Jokes and play are central elements of internet culture, and are promoted by the platform attention economy.³ Social media platforms struggle to identify and assess racist jokes,⁴ allowing extremist groups to use humour as a veil for hate, abuse

¹ Ben-David, A., & Matamoros-Fernández, A. (2016). Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication*, 10, 1167–1193; Matamoros-Fernández, A. (2018). Inciting anger through Facebook reactions in Belgium: The use of emoji and related vernacular expressions in racist discourse. *First Monday*, 23(9). <https://doi.org/10.5210/fm.v23i9.9405>.

² Fielitz, M., & Ahmed, R. (2021). *It's not funny anymore. Far-right extremists' use of humour*. 18. https://ec.europa.eu/home-affairs/sites/default/files/what-we-do/networks/radicalisation_awareness_network/ran-papers/docs/ran_ad-hoc_pap_fre_humor_20210215_en.pdf.

³ Matamoros-Fernández, A. (2020). "El Negro de WhatsApp" meme, digital blackface, and racism on social media. *First Monday*, 25(1). <https://doi.org/10.5210/fm.v25i12.10420>; Roberts, S. (2016). Commercial Content Moderation: Digital Laborers' Dirty Work. In S. U. Noble & B. Tynes (Eds.), *The Intersectional Internet: Race, Sex, Class and Culture Online*. Peter Lang Publishing. <http://ir.lib.uwo.ca/commpub/12>.

⁴ Roberts, S. (2016). Commercial Content Moderation: Digital Laborers' Dirty Work. In S. U. Noble & B. Tynes (Eds.), *The Intersectional Internet: Race, Sex, Class and Culture Online*. Peter Lang Publishing. <http://ir.lib.uwo.ca/commpub/12>.

and discrimination.⁵ Despite the ‘sticks and stones’ rhetoric, humour that punches down on historically marginalised groups causes real harm,⁶ and these new racist practices must be addressed to combat the glorification of extremist groups. Importantly, extremist groups constantly alter the way they communicate online to avoid content moderation, including by using coded languages, and digital platforms need to closely monitor and adapt to these changes.⁷

Use of automation

Extremist groups often use automation, including bot accounts, to push their agendas.⁸ To address this, platforms need to improve their efforts to identify and remove bot accounts, and monitor the way that their automated systems are used to promote extremist content. Hate and harassment campaigns are cross-platform in nature, and addressing the harmful use of automation will require a collaborative, multi-platform approach.⁹

Use of impersonation

Extremists use technological affordances on social media, including anonymity and pseudonymity, to cloak their identities and incite tension.¹⁰ For example, extremists sometimes pose as people from racial minorities, and weaponise these identities to stoke confrontation, spread disinformation, and incited hatred.¹¹ Combatting the glorification of extremism will require platforms to develop and adopt novel solutions to the use of impersonation.

⁵ Matamoros-Fernández, A. (2017). Platformed racism: The mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube. *Information, Communication & Society*, 20(6), 930–946. <https://doi.org/10.1080/1369118X.2017.1293130>.

⁶ Bell, M. C. (2021). John Stuart Mill’s Harm Principle and Free Speech: Expanding the Notion of Harm. *Utilitas*, 33(2), 162–179. <https://doi.org/10.1017/S0953820820000229>.

⁷ Sentropy Technologies. (2020, July 23). Flattening the Other Curve. *Medium*. <https://medium.com/sentropy/flattening-the-other-curve-e55b493a83a3>.

⁸ Marwick, A. E., & Lewis, R. (2017). *Media Manipulation and Disinformation Online*. Data&Society. https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf; Mueller, R. S. (2019). *The Mueller Report*. Musaicum Books.

⁹ Burgess, J., & Matamoros-Fernández, A. (2016). Mapping sociocultural controversies across digital media platforms: One week of #gamergate on Twitter, YouTube, and Tumblr. *Communication Research and Practice*, 2(1), 79–96. <https://doi.org/10.1080/22041451.2016.1155338>; Matamoros-Fernández, A. (2017). Platformed racism: The mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube. *Information, Communication & Society*, 20(6), 930–946. <https://doi.org/10.1080/1369118X.2017.1293130>;

Matamoros-Fernández, A., & Farkas, J. (2021). Racism, Hate Speech, and Social Media: A Systematic Review and Critique. *Television & New Media*, 22(2), 205–224. <https://doi.org/10.1177/1527476420982230>.

¹⁰ Farkas, J., Schou, J., & Neumayer, C. (2017). Cloaked Facebook pages: Exploring fake Islamist propaganda in social media. *New Media & Society*; Freelon, D., Bossetta, M., Wells, C., Lukito, J., Xia, Y., & Adams, K. (2020). Black Trolls Matter: Racial and Ideological Asymmetries in Social Media Disinformation. *Social Science Computer Review*. <https://doi.org/10.1177/0894439320914853>.

¹¹ Farkas, J., Schou, J., & Neumayer, C. (2017). Cloaked Facebook pages: Exploring fake Islamist propaganda in social media. *New Media & Society*; Freelon, D., Bossetta, M., Wells, C., Lukito, J., Xia, Y., & Adams, K. (2020). Black Trolls Matter: Racial and Ideological Asymmetries in Social Media Disinformation. *Social Science Computer Review*. <https://doi.org/10.1177/0894439320914853>.

Appropriating keywords from social justice movements

Extremists appropriate and hijack keywords from social justice movements to manipulate search algorithms and promote extremist content. They exploit ‘data voids’, which occur when there is insufficient content related to a particular search query for platforms to recommend, to push racist and misogynistic agendas.¹² For example, ‘bad actors’ use keywords related to newsworthy events such as terrorist attacks to publish Islamophobic content and make this content ‘get picked’ by platforms’ algorithms as ‘relevant’.¹³ They also use terms associated with racial justice organisations, such as Critical Race Theory, to spread racialised misinformation.¹⁴ To prevent the spread of extremist content, platforms will need to ensure that their ranking and recommendation algorithms are adapted to recognise and prevent this problematic hijacking of social justice keywords.¹⁵

¹² Golebiewski, M., & boyd, danah. (2018). *Data Voids: Where Missing Data Can Easily Be Exploited*. Data & Society Research Institute. https://datasociety.net/wp-content/uploads/2018/05/Data_Society_Data_Voids_Final_3.pdf.

¹³ Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to ‘ranking cultures’: Investigating the modulation of visibility in YouTube search results. *Convergence*, 24(1), 50–68. <https://doi.org/10.1177/1354856517736982>.

¹⁴ Collins-Dexter, B., & Donovan, J. (2021, March 11). How a racialized disinformation campaign ties itself to The 1619 Project. *Columbia Journalism Review*. <https://www.cjr.org/opinion/1776-keyword-squatting-right-wing-media.php>.

¹⁵ Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to ‘ranking cultures’: Investigating the modulation of visibility in YouTube search results. *Convergence*, 24(1), 50–68. <https://doi.org/10.1177/1354856517736982>.