

Written evidence submitted by Hossein Dabbagh, Northeastern University London & Oxford University

Algorithmic Echoes: Social Media, Polarisation and the Surge of Extremism in British Democracy

Personal Overview

I am an Assistant Professor in Applied Ethics at Northeastern University London and a Philosophy Tutor at Oxford University. My research, which focuses on applied ethics and public policy, intersects with issues related to digital technology, AI, and their impact on democracy and social justice. Through my teaching and scholarly work, I aim to illuminate the ethical challenges posed by emerging technologies and advocate for responsible innovation that upholds democratic values and human rights. I place a special emphasis on the responsibilities of technologists and policymakers in safeguarding democracy. By blending philosophical insights with practical ethics, I critically assess the ways in which technology influences political discourse, social dynamics, and individual behaviour.

Reason for Submission

The reason for my submission to the JCNSS inquiry into ‘Defending Democracy’ stems from my concern over the erosion of democratic norms facilitated by AI-driven social media algorithms. As an academic with expertise in applied ethics, public policy and philosophy, I have observed firsthand the capacity of these technologies to manipulate public opinion, exacerbate social and political polarisation, and empower extremist ideologies.

The phenomena of rising nationalism, xenophobia, Islamophobia, and antisemitism in the UK and beyond are particularly troubling, as they threaten the foundational principles of democracy and social cohesion. My research and teaching have increasingly focused on these issues, seeking to understand the mechanisms by which digital technologies can undermine democratic processes and to explore ethical frameworks for mitigating these effects.

Through this submission, I aim to offer a nuanced ethical analysis of the challenges posed by AI and social media to UK democracy. I seek to contribute my expertise to inform the committee’s understanding of these complex issues and to suggest actionable recommendations that prioritise ethical considerations in the defence of democracy. My goal is to engage in a productive dialogue with policymakers, technologists, and fellow academics to develop strategies that strengthen democratic resilience in the face of digital threats.

Executive Summary

The rapid advancement of Artificial Intelligence (AI) has brought about significant benefits for society, enhancing efficiencies in various sectors. However, deploying AI within social media algorithms profoundly challenges democratic values and social cohesion. This report delves into the political and social ramifications of AI-driven social media platforms, particularly focusing on their role in facilitating political and social polarisation. Such polarisation is a contributing factor to the resurgence of right-wing extremism in the UK, marked by increased nationalism, xenophobia, Islamophobia, and antisemitism. These developments pose a considerable threat to the fabric of democracy and the long-term stability of democratic institutions.

This report emphasises the need for proactive measures to mitigate the harmful effects of AI-driven polarisation on democracy. It is essential for policymakers, social media companies, and civil society to collaborate in addressing these challenges, ensuring that the digital landscape serves to strengthen rather than undermine democratic principles and social cohesion.

Introduction

AI-driven algorithms, fundamental to the operation of major social media platforms, play a pivotal role in content curation and dissemination. While designed to enhance user engagement, these algorithms often inadvertently amplify divisive content, fostering echo chambers and reinforcing extreme viewpoints. This dynamic has significant implications for political discourse, potentially exacerbating societal divides and facilitating the spread of right-wing extremism.

The Mechanisms of Polarisation

Social media algorithms are designed to maximise user engagement by curating content that most likely resonates with individual preferences and biases. This design principle creates 'echo chambers' or 'filter bubbles', where users are predominantly exposed to viewpoints that align with their own, while opposing perspectives are minimised or excluded altogether. Such environments facilitate polarisation by reinforcing existing beliefs and amplifying partisan content, making users more susceptible to radicalisation. The mechanism thrives on the human tendency to seek confirmation of pre-existing beliefs, a phenomenon known as 'confirmation bias'. As these algorithms better predict and influence user behaviour, the divide between differing viewpoints widens, making consensus and mutual understanding increasingly difficult to achieve (Pariser, 2011). As Sunstein (2017) highlights, these echo chambers can diminish shared experiences and facts necessary for the democratic process, reducing the opportunity for compromise and understanding between differing political views.

Impact on Right-wing Extremism

The polarisation facilitated by AI-driven content curation has been linked to the rise of right-wing extremism in the UK. Platforms have been criticised for their role in disseminating extremist propaganda, including nationalist, xenophobic, Islamophobic, and antisemitic content. Such content can radicalise individuals, promoting ideologies antithetical to democratic values and social harmony (Bartlett & Miller, 2010). For example, Munn (2019) discusses how right-wing groups exploit social media to spread their message, often using sophisticated propaganda techniques that exploit algorithmic biases towards sensational and divisive content.

Case Studies and Evidence

Several case studies underscore the relationship between social media polarisation and the rise of right-wing extremism. As a notorious case, the Brexit campaign in the UK saw widespread use of social media to spread misinformation and xenophobic rhetoric, contributing to heightened societal divisions (Bastos & Mercea, 2019). Similarly, the role of social media in the 2016 US Presidential election highlighted how platforms can be used to amplify divisive content and misinformation, influencing public opinion and political outcomes. Another example is the rise of far-right groups in the UK, such as Britain First,

which have used social media to disseminate Islamophobic content, resulting in increased offline activities and demonstrations (Allen, 2020).

Research studies, such as those by Müller and Schwarz (2021), have empirically linked social media usage with an increase in hate crimes, suggesting that online polarisation can have direct, harmful consequences in the real world. These case studies underscore the real-world impact of online polarisation and the facilitation of extremist ideologies through social media.

Social Media's Role in Boosting Nationalism and Xenophobia

Social media's algorithmic prioritisation of emotionally charged content can exacerbate nationalist and xenophobic sentiments. Content that invokes fear, anger, or moral outrage is more likely to be engaged with and shared, creating a feedback loop that amplifies these us-versus-them narratives and emotions. Nationalist groups often exploit these dynamics to spread messages that frame immigrants and minorities as threats to societal well-being, thereby stoking xenophobic attitudes. This mechanism spreads harmful stereotypes and undermines the social fabric by fostering distrust and animosity among different community groups (Marwick & Lewis, 2017).

The Consequences for Democracy

The long-term consequences of social media-fuelled polarisation and extremism for democracy are profound and multifaceted. Polarisation erodes the common ground necessary for democratic discourse and governance, making it increasingly difficult to reach consensus on critical issues. This fragmentation undermines the essence of popular sovereignty and political equality, as polarisation limits the ability of citizens to engage in meaningful dialogue and make informed decisions collectively. As a result, the very ideals of democratic tradition, which include mutual respect and the pursuit of the common good, are at risk, potentially signalling the onset of a 'post-democracy' era where the symbols of democracy persist, but its core practices and values are hollowed out (Crouch, 2000).

The rise of right-wing extremism further exacerbates these challenges, threatening the principles of inclusivity, pluralism, tolerance, and diversity that underpin democratic societies. This can lead to an increase in hate crimes and social unrest, fracturing communities and deepening societal divides. Moreover, undermining rational debate by emotionally charged and misleading content distorts the democratic process, influencing elections and policy decisions based on misinformation rather than informed debate. Such dynamics facilitate the spread of populist narratives that reject the complexity of democratic governance in favour of simplistic solutions to complex problems, further eroding the quality of public discourse.

In the long term, these dynamics can weaken democratic institutions and erode public trust in the democratic system. When citizens lose faith in the ability of democratic processes to address their concerns and resolve societal issues effectively, the stability and integrity of democratic governance are at risk. This loss of trust can lead to disengagement from democratic processes, a decline in voter turnout, and an increased susceptibility to authoritarian appeals, posing a significant threat to the fabric of democratic societies.

Therefore, the implications of social media-fuelled polarisation and extremism are far-reaching, potentially leading to a fundamental transformation of the democratic landscape. The erosion of shared experiences, facts, and rational debate essential for a functioning democracy underscores the urgent need for strategies to rebuild the common ground, promote

informed and respectful public discourse, and safeguard the principles of democracy for future generations.

Policy Recommendations

1. **Regulatory Intervention:** Implement regulations to ensure transparency in social media algorithms and to hold platforms accountable for content that promotes extremism, ensuring a balance between free expression and the prevention of harm (Gillespie, 2018).
2. **Algorithmic Adjustments:** Encourage social media companies to adjust algorithms to promote a diversity of viewpoints and reduce the spread of divisive content. This involves refining content recommendation systems to avoid amplifying extremist narratives and increasing the visibility of counter-narratives and factual information.
3. **Digital Literacy and AI Ethics Education:** Invest in comprehensive digital literacy programs that include AI ethics education. Such programs should enhance critical reasoning skills, enabling users to critically assess and understand the algorithms that curate their online content. By fostering a more informed and discerning online populace, individuals can better navigate the complexities of digital spaces and resist the influence of polarising content (Floridi & Cowls, 2019; Dabbagh et al., 2024).
4. **Research and Monitoring:** Support ongoing research into the impact of AI on democracy and monitor the spread of extremist content online. This should include studies on the efficacy of different interventions to combat online radicalisation and the development of tools to detect and counteract emerging threats to democratic discourse (Benkler et al., 2018).
5. **Promoting Critical Engagement:** Develop initiatives to encourage critical engagement with information sources. This includes supporting platforms that facilitate discussions across political divides and creating spaces where constructive dialogue can occur, thus working against the segregation fostered by echo chambers.

Conclusion

The intersection of AI, social media, and democracy presents complex challenges. While AI has the potential to enhance democratic engagement, its current deployment within social media algorithms poses significant risks, notably by exacerbating political and social polarisation and facilitating the spread of right-wing extremism. Addressing these challenges requires a multifaceted approach, combining regulatory measures, technological adjustments, and educational initiatives to safeguard democratic values and institutions.

11 March 2024

References:

- Allen, C. (2014), Britain First: The 'Frontline Resistance' to the Islamification of Britain. *The Political Quarterly*, 85: 354-361.
- Bartlett, J., & Miller, C. (2010). The Power of Unreason: Conspiracy Theories, Extremism and Counter-Terrorism. Demos.
- Bastos, M. T., & Mercea, D. (2019). The Brexit Botnet and User-Generated Hyperpartisan News. *Social Science Computer Review*, 37 (1), 38-54.
- Benkler, Y., Faris, R., & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press.

- Crouch, C. (2000). *Coping with Post-Democracy*, Fabian Society.
- Dabbagh H., Earp B. D., Porsdam Mann S., Plozza M., Salloch S., Savulescu J. (2024). AI Ethics Should be Mandatory for Schoolchildren, *AI and Ethics*, forthcoming.
- Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review* 1 (1).
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Press.
- Marwick, A., & Lewis, R. (2017). Media Manipulation and Disinformation Online. Data & Society Research Institute.
- Müller, K., & Schwarz, C. (2021). Fanning the Flames of Hate: Social Media and Hate Crime. *Journal of the European Economic Association*, Vol. 19, Issue 4, 2131–2167.
- Munn, L. (2019). Alt-Right Pipeline: Individual Journeys to Extremism Online. *First Monday*, Vol. 24, No. 6.
- Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Books.
- Sunstein, C. R. (2017). *Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.