MANOHAR PARRIKAR INSTITUTE FOR DEFENCE STUDIES AND ANALYSES
मनोहर पर्रिकर रक्षा अध्ययन एवं विश्लेषण संस्थान

MP-IDSA
Issue Brief

AI and Extremist Propaganda: An Assessment

Saman Ayesha Kidwai

March 25, 2026

Summary

AI has rapidly accelerated the transformation of the global violent extremist landscape by acting as a force multiplier in the manufacturing and dissemination of extremist propaganda. This presents a broader set of challenges for states and reinforces the need for technologically grounded counter-violent extremist frameworks.

Introduction

Artificial Intelligence (AI), primarily through rapidly emerging and multifaceted platforms, is transforming the pace, scope, nature and intensity of information dissemination globally. At the same time, it carries with it major ramifications for the global violent extremist landscape. Non-state actors have increasingly used AI tools to disseminate propaganda and cultivate echo chambers, thereby legitimising their extremist worldviews. This has amplified the stakes for state actors in the era defined by strategic storytelling and narrative warfare.

AI tools and AI-enabled content have frequently been misused by non-state actors, with a platform being equally pivotal in furthering extremist worldviews and countering them. As a result, this has underscored the multifaceted nature of this broader domain. This is particularly concerning because even AI tools that promote free trial versions, let alone those that offer subscriptions with far-reaching uses, can, if utilised strategically, dramatically redefine the next phase of global violent extremism.

Furthermore, this has reinforced the compounding need to implement policy reforms to counter the exploitative uses of AI. Such challenges have widened due to geopolitical and geo-economic polarisation, despite the growing need to accommodate existing realities. Yet, despite AI's visible impact on the broader info-tech ecosystem, its profound security implications and potential to be repurposed as a counter-violent extremist tool remain underexamined.

Shift from a Centralised to a Decentralised Propaganda Machinery

One of the major transformative shifts in the global violent extremist landscape has been that a hierarchical media wing within a centralised violent extremist organisation is no longer the sole authority spearheading the propaganda ecosystem. Instead, the manufacturing and dissemination of extremist propaganda has become increasingly decentralised. Here, the role of micro cells, digital communities, lone wolves, and online recruiters has become increasingly prominent. Youth, in particular, have rapidly become among the primary consumers of gamified and memefied propaganda, especially post-COVID-19 outbreak.

According to the report jointly released by the United Nations Interregional Crime and Justice Research Institute and the United Nations Office on Counter-Terrorism in November 2025, online radicalisation, particularly among young

gamers, has become a focal point of concern. As per the findings, it has been observed that:

> *Globally, terrorists and violent extremist actors have exploited gaming platforms… Tactics range from building violently racist and xenophobic games to recreating attacks on religious sites, directly targeting gamers for recruitment via in-game chats, and taking advantage of misogynistic, racist, and insular elements of some gaming communities to radicalise members to violence. Since 2019, over nine major attacks have involved the use of gaming platforms or services.*[1]

This has been the case as violent extremist groups, including White Supremacists or Islamists, including ISIS sympathisers or recruits, have migrated to the digital sphere to remain relevant amid increased physical crackdowns on organisations, battlefield defeats, or impositions of sanctions resulting in the curbing of their terror financing operations. These trends have been on the rise in an increasingly interconnected society with widespread digital penetration. It has broadly been observed that extremist groups, lone wolves, and online communities have largely benefitted from wide-ranging issues persisting globally, including digital illiteracy and the inability of online participants on gaming or social media platforms to detect disinformation and misinformation, pushing many on the path of individual or collective radicalisation.

Digitalisation and Youth Radicalisation

The youth constitute one of the fastest-growing recruitment pools within the extremist ecosystem, particularly in Europe and North America. For example, as per the Global Terrorism Index Report (2026), "Children and adolescents accounted for 42 per cent of all terror-related incidents in Europe and North America, a threefold increase since 2021."[2] Similar trends are evident in Sub-Saharan Africa, an area regionally considered the most volatile due to terror-related incidents. This highlights that youth-based radicalisation has been rising globally.

Notably, youth-centric radicalisation has historical roots. The current landscape has raised the stakes for counter-terror authorities. This is due to the unprecedented pace and reach of indoctrination and recruitment that is underway, driven by a worldwide mobilisation towards the use of digital platforms and low-cost technology. It has been observed that the average radicalisation timeline has contracted dramatically: from 18 months in 2005 to 13 months in 2016. Today,

[1] **"Level Up: Gaming and Violent Extremism in Africa"**, Report, United Nations Interregional Crime and Justice Research Institute and the United Nations Office on Counter-Terrorism, 18 November 2025.

[2] **"Global Terrorism Index 2026"**, Report, Institute for Economics & Peace, March 2026.

radicalisation can occur within a matter of weeks.[3] Such trends must be closely watched by Indian policymakers, given the expanding digital footprint of its rapidly expanding youth demographic, even in remote corners.

One of the more interesting findings revealed that extremist recruiters online mapped their recruitment activities as per a strategic timeline, for example, in the United Kingdom (UK). That is, they primarily seek to manipulate younger audiences by establishing emotional connections at times when they would otherwise lack consistent interaction and support systems, particularly during the holiday season.[4] This is when teenagers have been observed spending their time behind screens rather than engaging in social interactions that would allow an adult to notice trends and deviations from established patterns and behaviours.

This signals a concerning trend which could be replicated in other geographical spaces where political polarisation is often accompanied by socio-economic disparities and festering psycho-social grievances, which create fertile ground for recruitment by extremist outfits across the vast ideological spectrum. One must remember that extremist campaigns, generally speaking, frequently focus on legitimate socio-economic concerns, such as corruption, housing and employment crises, among other issues, to indoctrinate and recruit new followers before following that up with virulent rhetoric against specific communities or entities, which is normalised over the course of time as part of their extremist worldview.

This has the potential to undermine the established social contract and institutional frameworks that facilitate governance and lawmaking, particularly in democratic countries and penetrable digital ecosystems. Furthermore, AI apps such as Z.ai and Pixazo.ai, as evidenced by the snapshots below, have the potential to draft campaign outlines and visual aids, respectively, that resonate with the targeted audience using specific prompts, and could be disseminated across chat platforms like Telegram or Discord, which extremist sympathisers often favour.

AI-Generated Propaganda Samples Based on Targeted Prompts

The following examples illustrate how AI tools and platforms can rapidly produce tailored propaganda for targeted audiences, based on prompts from subversive elements. The Z.ai platform was used to draft a sample campaign script, while

[3] Ibid.

[4] Libby Brooks, **"Far-right Extremists Using Games Platforms to Radicalise Teenagers, Report Warns"**, *The Guardian*, 31 July 2025.

Pixazo.ai was used to draft a visual infographic based on the guidelines compiled by Z.ai.

Figure a. Z.ai (19 March 2026)

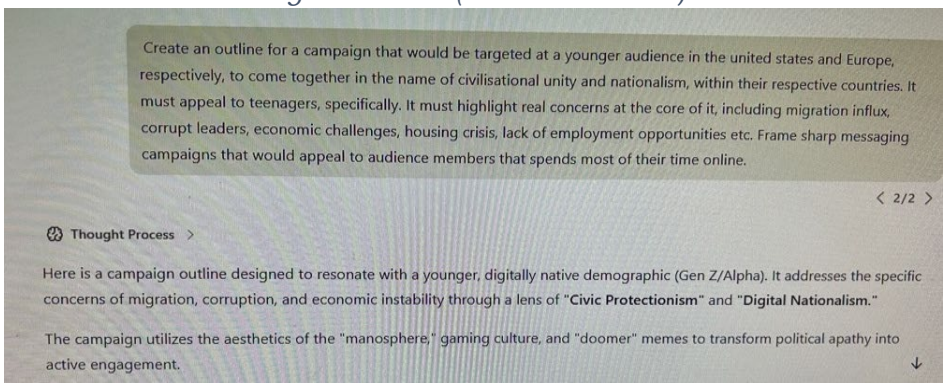


Figure b. Z.ai (19 March 2026)

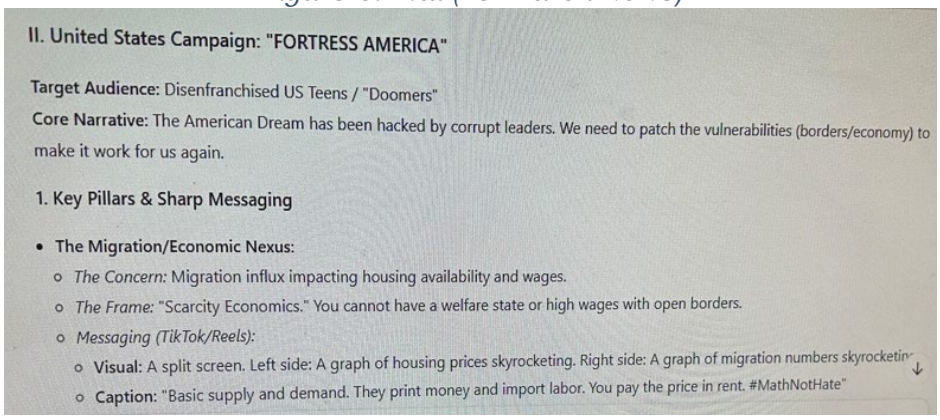


Figure c. Z.ai (19 March 2026)

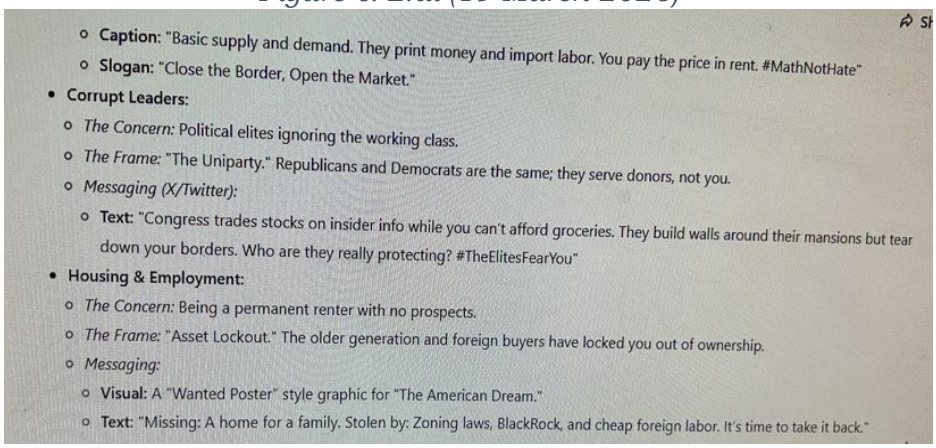


Figure d. Z.ai (19 March 2026)

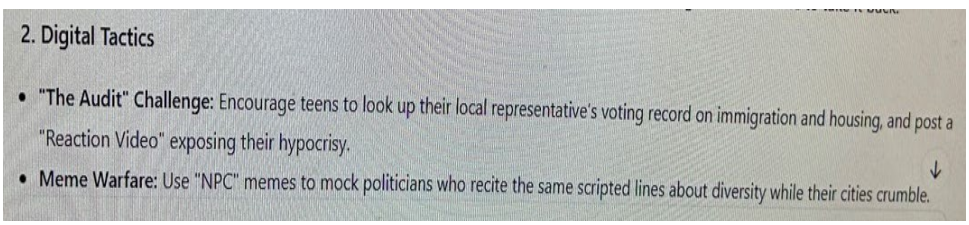


Figure e. Pixazo.ai (19 March 2026)

Assessing Post-Organisational Extremism

The evolving violent extremist landscape underscores the need to move beyond organisation-centric metrics while determining the threat matrix. This is because AI has begun to act as a force multiplier in the dissemination of extremist propaganda. This has largely been recognised as central to the extremism toolkit among a new generation of recruits and self-proclaimed foot soldiers. They, having been inspired by a specific extremist worldview, have, without any formal association, carried out stabbing rampages, mass shootings, or used personal vehicles to inflict physical violence at public gatherings in the name of violent extremist groups.

As a result, the role of AI-generated propaganda has proved fundamental in ensuring the ideological resilience of violent extremist movements and terror narratives despite organisational degradation, leadership decapitation, and territorial losses. It has served as a consistent lynchpin, ensuring the reinforcement and survival of violent radical ideas in top-down and bottom-up formats.

What we see today are the manifestations of the post-organisational nature and increasing decentralisation of extremism. With them, the media and strategies for retaining the relevance of specific extremist worldviews have also transformed, largely due to the convergence of technical and digital domains.

On the one hand, these developments are, arguably, emerging due to the widespread digitalisation and decentralisation of access to low-cost technology and weapons. On

the other hand, the prevalence of lone wolves and anonymous digital leaders in committing or planning attacks, bound transnationally through digital threads and shared values, must be treated as another crucial factor contributing to the ideological resilience of violent extremism globally. Today, an extremist ideology can survive independently of an organisation or a centralised command-and-control structure.

Expanding Spectrum of AI-Enabled Threats

AI's role has gradually expanded beyond content generation to include the amplification of extremist worldviews through deepfakes and systematic, synthetic propaganda in easily shareable formats on mainstream platforms such as Pinterest and Instagram, while maintaining anonymity.[5] This considerably lowers the overall threshold while engaging with extremist narratives.

Simply put, behind the digital vanguard of their screens, the manufacturers and consumers of extremist propaganda can easily exchange radical worldviews and expressions in their physical manifestations, with minimal social, legal, or punitive costs.

It is also important to underscore how deepfake platforms such as HeyGen, Reface and Viggle, among others, provide limited access to high-quality deepfake videos, accessible even for beginners. As of 20 March 2026, HeyGen, according to figures advertised on its website, had generated 113,457,133 deepfake videos and 87,580,981 avatars.

India's Counter-Terror Framework and the Way Forward

These developments within the AI domain have become particularly significant for India's national interests.

On 23 February 2026, India unveiled its first comprehensive national counter-terrorism policy and strategy, PRAHAAR, in light of emerging security threats.[6] As the first pivotal national counter-terror framework introduced by the Indian government, it is positioned to spearhead a cohesive approach towards mitigating and neutralising emerging security threats. Its focus signals a growing recognition and inclusion of pathways to address challenges arising from non-conventional security threats to India, including unregulated crypto wallets, the cyber domain,

[5] Mariam Shah, **"The Digital Weaponry of Radicalisation: AI and the Recruitment Nexus"**, Insights, Global Network on Extremism & Technology, 4 July 2024.

[6] **"National Counter-Terrorism Policy & Strategy"**, Ministry of Home Affairs, Government of India, 23 February 2026.

drones, the dark web, and exploitative uses of instant messaging platforms and social media.

Nevertheless, despite its comprehensive outlook, certain aspects remain that must be addressed to enable a more robust counter-response to the evolving security landscape. The next phase of the national counter-terror framework must place greater emphasis on more strategically addressing the security challenges posed by a youth-driven, decentralised and digitalised extremist landscape that AI is constantly redefining. At the same time, it must retain the focus on the more physical manifestations of terrorism and organisational degradation of proscribed outfits.

In that regard, some of the key reforms that require inclusion in subsequent legislation and strategic frameworks must centre around the following:

a. Provisions targeted at AI-driven propaganda specifically and capable of accommodating meteoric technological advancements over time.

b. A specialised nodal agency to coordinate real-time efforts of multiple actors involved in drafting and executing a new-age counter-violent extremism framework grounded in existing realities and transformations within the digital and technological domains.

c. The introduction of more stringent regulatory mechanisms for mainstream and fringe social media and AI-driven platforms.

d. Digital literacy initiatives, mental health assistance and early interventions for youths, particularly teenagers.

e. Capacity-building initiatives for content moderators, policymakers, intelligence officials and law enforcement agencies to effectively navigate the overwhelming volume of extremist content and assess threat levels in real time.

f. Budgets specially devoted to promoting training in the use of AI tools, OSINT analysis, and countering disinformation/misinformation in real time.

g. Frequent and periodic mental health assessments and assistance, including financial support, provided to those encountering harmful content online, specifically content moderators, to ensure they can effectively contribute to counter-violent extremism strategies without falling into the trap of radicalisation themselves. These interventions would prove critical in ensuring that national security remains largely uncompromised.

AI-Generated Counter-Violent Extremist Video Sample Based on Targeted Prompts

Despite the evident use and misuse of AI-enabled content by non-state actors, educators working in tandem with state agencies can collectively stage targeted interventions to contain the spillover of radicalisation and digital indoctrination. Vibrant campaigns and sharp messaging introduced within elementary school ecosystems can lay the foundations of a more inclusive, stable and cohesive society, wherein the younger generation, especially Gen-Z, can be prevented from crossing the threshold of violent extremism and radicalisation.

For this purpose, the following samples have been generated using ChatGPT and HeyGen on 23 March 2026 to provide evidentiary support. These samples demonstrate how AI can play a vital role in staging counter-violent extremism interventions. The video has been generated in response to the targeted prompt; please refer to Figures 'f' and 'g'. Additionally, the poster, i.e., Figure 'i ', has been generated in response to the targeted prompts; please refer to Figures 'j' and 'k'.

Figure f. ChatGPT (23 March 2026)

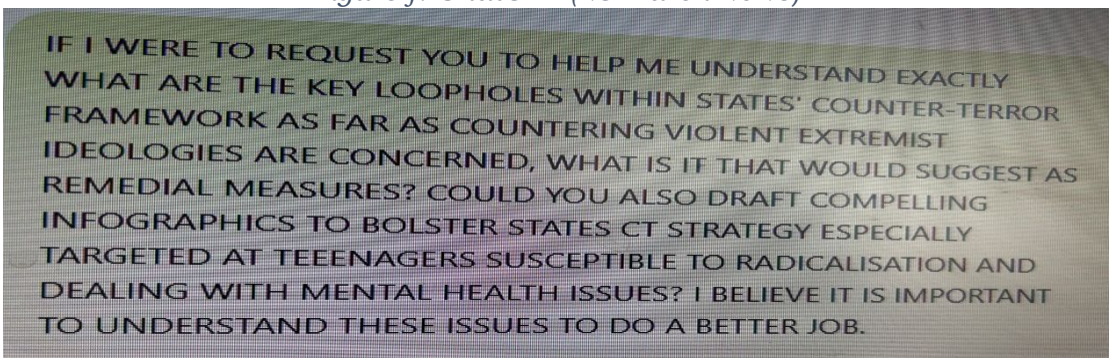


Figure g. ChatGPT (23 March 2026)

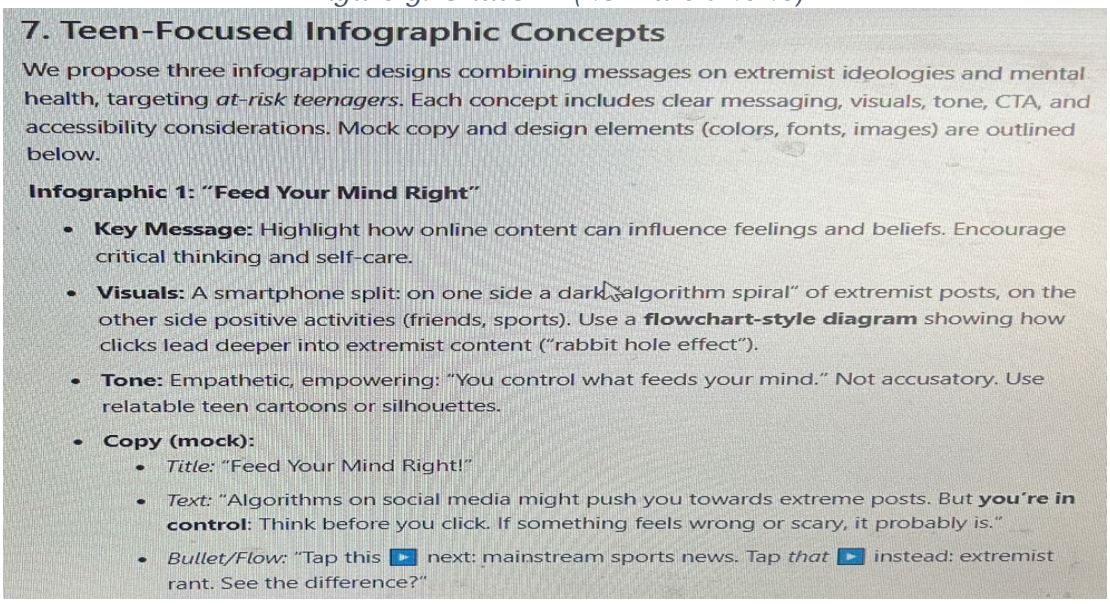


Figure h. ChatGPT (23 March 2026)

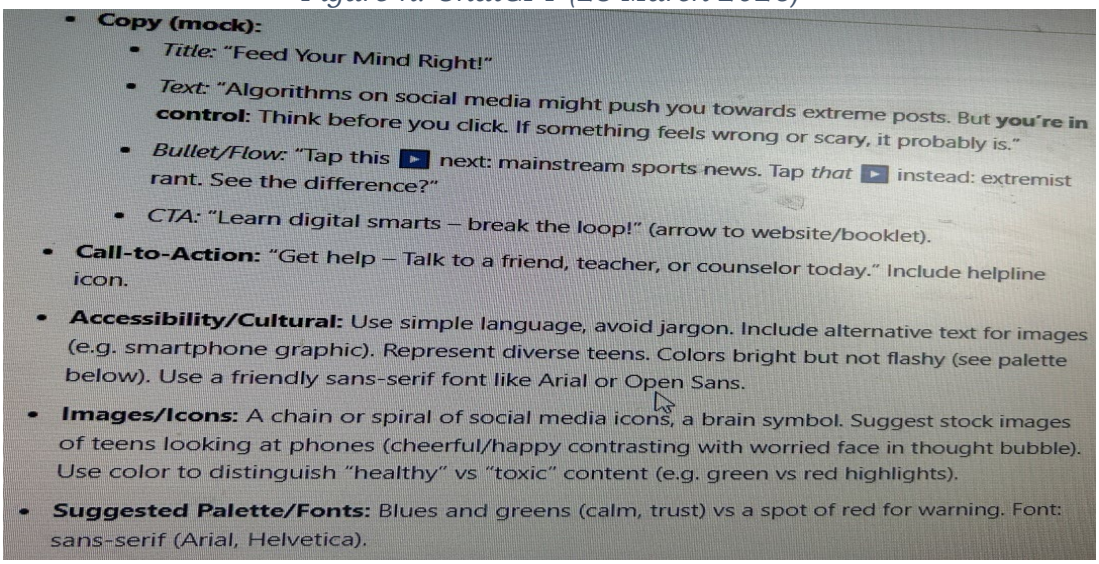


Figure i. HeyGen (23 March 2026)

Figure j. ChatGPT (23 March 2026)

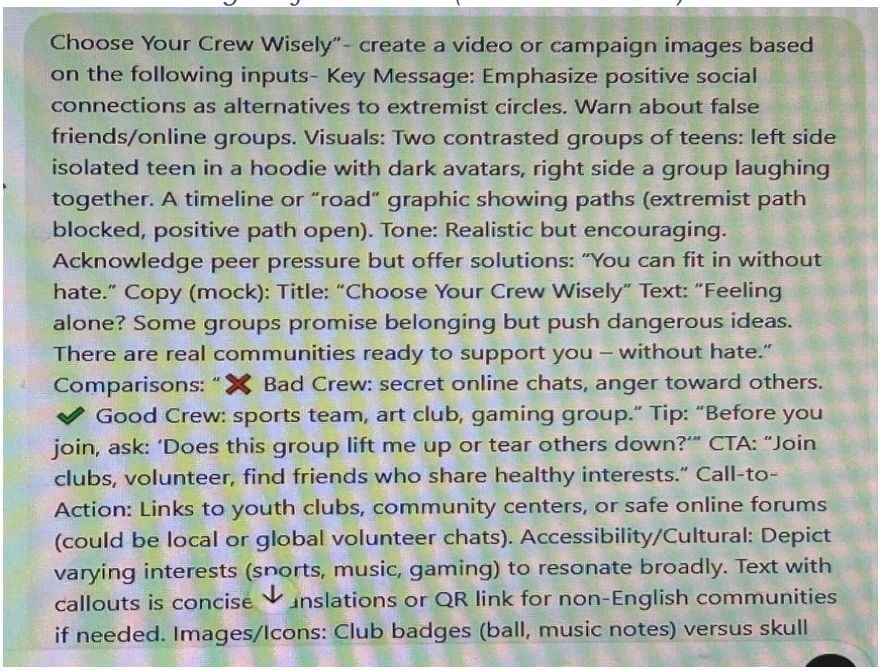
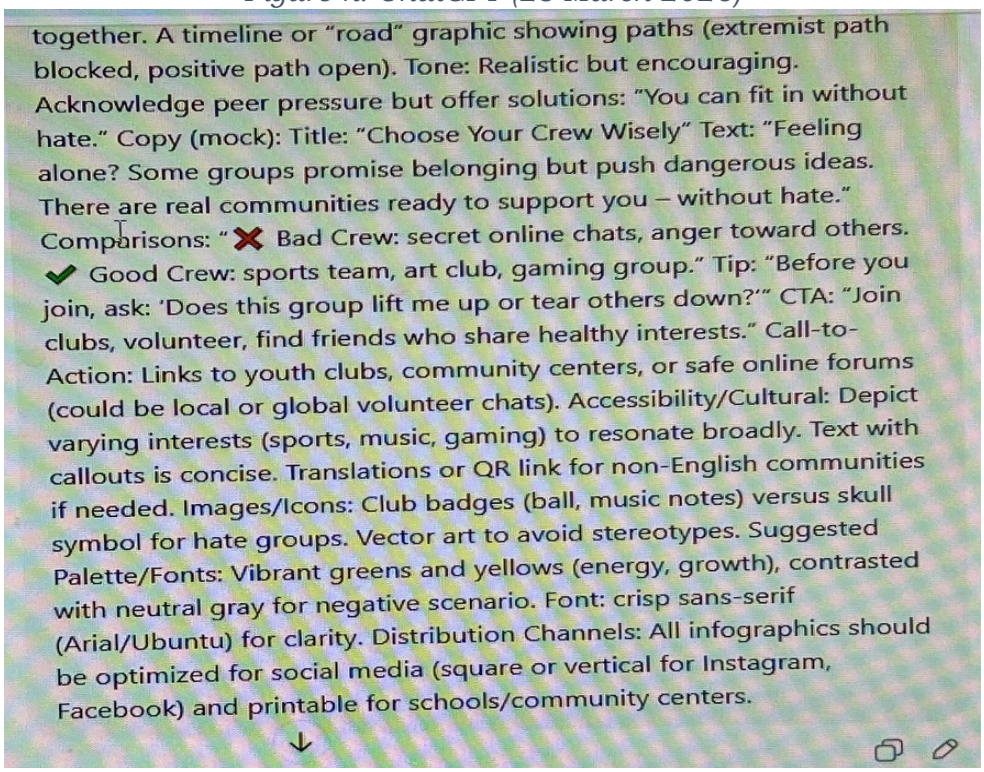


Figure k. ChatGPT (23 March 2026)



together. A timeline or "road" graphic showing paths (extremist path blocked, positive path open). Tone: Realistic but encouraging. Acknowledge peer pressure but offer solutions: "You can fit in without hate." Copy (mock): Title: "Choose Your Crew Wisely" Text: "Feeling alone? Some groups promise belonging but push dangerous ideas. There are real communities ready to support you – without hate." Comparisons: "✗ Bad Crew: secret online chats, anger toward others. ✓ Good Crew: sports team, art club, gaming group." Tip: "Before you join, ask: 'Does this group lift me up or tear others down?'" CTA: "Join clubs, volunteer, find friends who share healthy interests." Call-to-Action: Links to youth clubs, community centers, or safe online forums (could be local or global volunteer chats). Accessibility/Cultural: Depict varying interests (sports, music, gaming) to resonate broadly. Text with callouts is concise. Translations or QR link for non-English communities if needed. Images/Icons: Club badges (ball, music notes) versus skull symbol for hate groups. Vector art to avoid stereotypes. Suggested Palette/Fonts: Vibrant greens and yellows (energy, growth), contrasted with neutral gray for negative scenario. Font: crisp sans-serif (Arial/Ubuntu) for clarity. Distribution Channels: All infographics should be optimized for social media (square or vertical for Instagram, Facebook) and printable for schools/community centers.

Figure l. ChatGPT (23 March 2026)

Conclusion

The assessment of the evolving extremist landscape reveals a stark reality—AI has effectively transformed the extremist propaganda machinery and the tools used to disseminate it for the foreseeable future. Additionally, the broader landscape has become far more decentralised, digitalised and mutable. This, in turn, has resulted in a domino effect, i.e., it has contributed towards making violent extremist worldviews more ideologically durable, as they are no longer reliant on centralised command-and-control structures within a parent terror organisation for their survival.

Such trends create a non-negotiable requirement for countries, including India, to integrate AI, youth-centric propaganda, and threat matrices into their broader counter-terror frameworks and national security strategies. Finally, the most fundamental challenge for all states lies in mitigating and redirecting the digital ecosystems that create conditions conducive to the proliferation of extremist narratives and propaganda.

As the security landscape evolves, the competition between state and non-state actors over narratives and digital ecosystems will likely intensify. Therefore, the outcome will ultimately hinge on which actor can more effectively influence and shape contested digital narratives and technology-driven information ecosystems.

About the Author

Ms. Saman Ayesha Kidwai is Associate Fellow at the Manohar Parrikar Institute for Defence Studies and Analyses, New Delhi.

Manohar Parrikar Institute for Defence Studies and Analyses is a non-partisan, autonomous body dedicated to objective research and policy relevant studies on all aspects of defence and security. Its mission is to promote national and international security through the generation and dissemination of knowledge on defence and security-related issues.

Manohar Parrikar Institute for Defence Studies and Analyses
1, Development Enclave, Rao Tula Ram Marg
New Delhi 110 010 India
T +91-11-2671 7983 F +91-11-2615 4191
www.idsa.in
Twitter @IDSAIndia
www.facebook.com/ManoharParrikarInstituteforDefenceStudiesAnalyses