RESEARCH PAPERS

AI-Generated Propaganda and Synthetic Media in Terrorist Recruitment

Author: Bhavya Chaturvedi • **Published:** July 05, 2026 • **Source:** Centre for New Age Warfare Studies



EXECUTIVE BRIEF & KEY FINDINGS

The convergence of Artificial intelligence and synthetic media has created a new threat vector for terrorist organisations transforming the process of radicalisation, recruitment and mobilization. This article examines how AI generated propaganda including deepfakes, voice cloning and generative AI systems is reshaping extremist communication. With a focused lens on Indian context it analyses experimentation of AI enabled tools the advantages these technologies offer over traditional recruitment the challenges these pose to counter-radicalisation frameworks, and the limitations of existing content moderation mechanism. The article concludes by proposing regulatory, technological and strategic measures to mitigate the risks posed by AI-generated propaganda and synthetic media in terrorist recruitment ecosystems.

1. Introduction

With the emergence of the modern nation-state, the ability to shape narratives and mobilise support has become central to political legitimacy and organisational survival. Traditionally the state relied on print media, pamphlets and official channels to legitimize governance and public opinion. Similarly, non-state actors, including religious movements and

extremist organisations, employed interpersonal networks and printed material to recruit supporters and disseminate ideological narratives. Although these methods were often effective, their reach remained geographically limited. The advent of the internet and digital communication technologies fundamentally transformed this landscape, enabling information to travel across borders almost instantaneously. While these developments facilitated political mobilisation and social movements, they also provided extremist organizations with unprecedented opportunities to disseminate propaganda, recruit followers and radicalise vulnerable populations.

Recent advances in Artificial Intelligence have amplified these trends along with success in medical and technological fields, impacting millions, guiding them to the right track, Extremist organisations increasingly exploit AI to expand support networks. Generative AI and synthetic media are one such tool that is actively employed by the terrorist organisations to increase their recruitment drive to challenge the status quo whenever they can. These tools can now generate text, images, audio, and video content in multiple languages and formats. It enables the terrorist organisations to tailor content for specific audiences, create realistic synthetic media, and automate the system of recruitment and radicalisation. Thus, AI is increasingly viewed as a force multiplier for terrorist information operations.

2. Traditional to AI technological evolution

In the early 2000s, various terrorist fronts like al-Qaeda established a digital presence by creating websites and funding online campaigns to garner support from the masses, although their ability to tailor content remained limited, as manual intervention was always needed to make their content multilingual to increase its scalability, thus one message was catering to all the diverse population. With the emergence of AI it has significantly enhanced three characteristics namely tailored content, scalability and plausible deniability for these organisations these features exploit psychological vulnerabilities by understanding their vulnerabilities in linguistic profiles and social system thus enhancing the persuasive appeal of extremist narratives.

Generative AI systems, particularly Large Language Models (LLMs) such as GPT, can generate human-like text, images, audio, and video content with minimal user input. According to the International Center for Counter Terrorism (ICCT,2024) Gen AI can support terrorist organizations across multiple functions including propaganda production, recruitment, translation, operational planning, cyber activities and fund raising thereby expanding both the scale and efficiency of extremist operations. These technologies have legitimate applications in education, healthcare, governance, and scientific research. However, they can also be exploited by extremist actors to create persuasive propaganda, fabricate narratives, and automate recruitment efforts. Unlike traditional AI systems that primarily analyse existing data, generative AI can create entirely new content, making it a powerful tool for information manipulation. This capability significantly enhances the efficiency of terrorist propaganda operations and allows extremist organisations to engage target audiences through highly personalised messaging (Warrick, 2022).

The growing sophistication of synthetic media has increased the difficulty of distinguishing authentic content from manipulated information. The Gen AI and synthetic media are manufacturing consent, as Gramsci considered that the organic intellectuals/capitalists creates the common sense which actually favours them, in the same way through the use of deepfake videos and audios audiences may be misled and become susceptible to propaganda. Chesny and Citron (2019) termed it as a liar's dividend- a phenomenon where due to the abundance of false and illegitimate information sometimes the authentic information by the legitimate authority is also disregarded. The National Investigation Agency in its report (2023) claimed that LeT, JeM, ISIS-k has been using WhatsApp and other social media platforms creating deepfakes for the circulation of their propaganda and indoctrinating the youths of Maharashtra, Kerala, and Jammu and Kashmir. In 2018 a mere whatsapp forward which was morphed and doctored led to lynching of innocent tribals and migrant workers (Banaji et al., 2019). The realism of fabrication has been amplified through use of

AI, which is sometimes impossible to detect without the professional forensic eye.

3. Experimentation with AI and synthetic media

Berger and Morgan(2015) observed that ISIS was among the first organisation to systematically exploit social media ecosystems for recruitment and mobilisation. Terrorist recruitment and propaganda are increasingly facilitated with these tech innovations although these operations are conducted covertly but academic researchers have quoted that use of LLM's and AI chat bots are working 24*7, analysts at Global network on extremism and technology claims that Telegram application is used to circulate doctored AI videos of battlefield victories, speeches of the top leadership using Midjourney and Stable Diffusion to create a sense of victory and momentum amongst supporters(Bloom and Moskalenko,2023). Winter (2015) described ISIS's propaganda architecture as "Virtual Caliphate" highlighting its ability to sustain recruitment, ideological dissemination and supporter of mobilization across geographical boundaries through digital platforms. These images, videos and scripts create a sense of events that never occurred in reality. Similar incidents have occurred in west Asia where doctored pictures of children and women were used to incite violence and create uncertainty in the social structure of the society. Hamas affiliated groups using gen AI are uploading pictures of IDF soldiers wearing diapers to undermine the Israeli Army leading to spread of misinformation in Gaza.

The domestic politics is also is also getting affected by it as in Europe and USA extremists nationalists are operating on approximately 40 telegrams groups which daily share the propaganda driven material on the groups which was earlier humanly impossible to sustain (institute of strategic dialogue, 2023).This is the industrialisation of extremist content production and extremist propaganda content. *Institute for digital forensics and research lab* has ascertained that using apps like Eleven labs Voice cloning is used to mimic the Muslim clerics to facilitate recruitment and indoctrinate Muslim youth to join Jihadist causes using moderation blind distribution mechanism for which social media regulations are largely ineffective for example the online gaming platforms. It is claimed that Pakistan sponsored agents cloning to be Indian citizens practicing in dissemination of Anti India narrative in Jammu and Kashmir under its "bleed India with Thousands cut military doctrine" of ISI, its termed as persona farming where narrative is formed outside the nation but prima facie it appears to be within the nation.

4. Advantages over traditional methods

4.1. Cost effectiveness and accessibility

The classic example of Dabiq and Rumiyah magazine required full scale professional videographers and writers still its reach was limited to a particular set of nationalities but with the AI a person with the basic familiarity with generative AI tools can create professional like content in hours which suggests that even with less financial liabilities a professional content can be created catering to the needs of diversified nationalities.

4.2. Linguistic and cultural penetration

In India there is a vast cultural and linguistic vibrance earlier content usually circulated around in Arabic which becomes difficult for the Tamil, Telugu, Marathi, etc to comprehend while in current scenario by the simple use of GPT and a prompt high end animated videos cloning the voice of the religious leader are created which generates false legitimacy and thus creates broader audience reach and engagement.

4.3. Emotional and psychological poignancy

Presence of daily updates of status on social media creates a particular pattern for the users AI can easily understand it and can mark the potential targets for recruitment as it’s majorly driven by the emotional connection. The content generated is deduced at the micro level to adjust the user behaviour and create more emotional resonance with the recruiters. Extremist recruitment functions by exploiting the fault lines of the social system like grievances, distrust and aloofness (Sageman,2004).

4.4. Operational swiftness

With the advent of Virtual borderless world the speed of news has come at the par of thought. Within seconds after any event we are witnessing law and order problems as we saw in the Delhi 2020 riots. Reaction speed towards a narrative is quicker than forming a counter narrative measure by the authorities. AI could intensify the speed of operations hence the impacts of it are disastrous without proper regulatory measures over the content.

4.5. Plausible deniability

Through the use of onion routers it becomes difficult for the agencies to trace the source of the content thus no one group or human can be blamed. Intelligence gathering for the incitement of extremism is restricted thus maintaining plausible deniability. Behind this veil the generation and distribution of content is carried out as large scale operations.

5. AI-Enabled Intensification of Recruitment

Before AI, large volumes of writing were the basis for propaganda comprehension, which was limited, thus leading to lower conversion in recruitment, making it a wide-mouth, low-yield enterprise. AI has funnelled down to scouting particular users with the same psychological understanding, increasing recruitment by targeting a narrower audience and improving conversion.

The two Pyramid models of radicalisation by McCauley and Moskalenko (2017) bifurcated between opinion radicalisation, i.e., holding of extreme beliefs, and Action Radicalisation, i.e., committing violent acts. Majorly individuals falls under the opinion radicalisation pyramid as by consuming the propaganda material they form an extremist orientation wherein they perceive other community as threat they talk about it openly in kin network but does not act or commit violence while in the Action radicalisation pyramid people have the same set of orientations while they are convinced that extremist and violent action is the only way forward for them to protect their interest and ideals as we saw in the Nuh riots in haryana in 2023 where a simple whatsapp generated narrative led to a clash between the two prominent groups of the region. Ai system can easily target the users falling at the apex of the Action pyramid model with higher rates of conversion in the recruitment.

The foreign intelligence agencies and proxy actors have historically attempted to exploit social and political grievances to influence vulnerable populations in what is known as the At Risk Populations comprising of the vulnerable youths in the regions of Western UP, Mewat region Kerala and the peripheral urban areas of Mumbai, having high rates of unemployment, poor education outcomes, and histories of communal violence (Basit, 2015; IPCS Research, 2016). In the regions of North east India and Kashmir Telegram groups are operated by extremist organisations like UNLF, KNF, TRF and alike using propaganda strategies, Firstly by creating a sense of a community with its members, secondly by demonizing and defaming the enemy Audience and lastly by creating propaganda narrative for the Universal audience by making Us vs Them and victimhood narrative.

The AI is consistently being used to generate gender specific propaganda increasingly aimed at women and girls through tailored narratives and synthetic media used to target by making patterned content by cloning the female figurehead voice, generating content and inciting by visual oppression of women, glorifying the militant lifestyle. Some individuals may be drawn towards narratives of supporting their brother fighters and having children to further the process of radicalisation and violence. Karim Pakzad, of the French Institute of International and Strategic Relations, said some young women had “an almost romantic idea of war and warriors. While young girls are picturising idealised portrayals of life in the Terrorist centers. They are being contacted by strangers on social media and are indoctrinated by being given the marriage narrative, where they are promised to move to an utopian society.

6. Challenges to counter radicalisation efforts

Multiple challenges surround the counter-radicalisation process, which is led by the State, civil society, and other platforms. These include:

Verification Challenges: Through the use of the LLM’s and AI chatbots, the video, audio, texts, and images are created of the war zones, false victimhood is claimed and fake speeches of the religious leaders are circulated, so it becomes crucial for the counter-radicalisation agencies to primarily debunk the myths of the content, which becomes resource-intensive. The burden of proof shifts towards the State; the counter-narrative must disprove what appears to be authentic.

Accelerating-attribution gap: Counter-terrorism agencies, including NIA, the Indian Cyber Crime Coordination Center, and state Cyber units, follow a bureaucratic protocol in conducting their operations, while the extremist deepfake video is circulated to several people and places, by the time agencies crack down, it may already have been disseminated across multiple platforms.

Psychological impacts of deepfakes: Lewandosky et al. (2012) pointed out the “continued influence effect” of misinformation, suggesting that even after debunking of the deepfakes, the initial claims are rarely nullified. Deepfake videos of prominent personalities promoting radicalisation serve as propaganda and are used by people to support their arguments. The latter faces allegations of inciting violence and its deep psychological impacts . Exposure to such videos brings strong emotional reactions like shock, anger, loyalty, which persist even after their refutation and erosion of social cohesion.

Reactive Nature of Existing Counter-Radicalisation Frameworks: The State De-radicalisation Structure, like the Multi-agency Centre, de-radicalisation cells of the state governments working under the aegis of MHA, are more predominantly reactive than pre-emptive due to the scalability of AI. Although these institutions are trying to keep up with technological developments, the conceptual reorientation of the users by consuming content based on their algorithm is to be studied in order to anticipate extremist propaganda.

7. Effectiveness of current detection and content moderation mechanism

Deepfake detection has become an asymmetric technological contest in which advances in content generation are increasingly outpacing improvements in detection capabilities, the generation of the content is becoming easier and stealthy while detection of propaganda content and upgradation in such technologies are becoming difficult in other words generative capacity is out pacing the detection capacity.

Deepfake detection tools including Meta deepfake detection challenge (DFDC) Microsoft Video Authenticator and open source system brings 80-90 % dataset accuracy in detection under laboratory conditions (Dolhansky et.al.,2020). Just

as Anti-virus software protects from the malware that detects the fake content in the propaganda tools. But the real problem lies in analysis as these softwares works on the Western-based datasets; they can easily decode fake narratives for the western populations but generally fail when it comes to other nationalities.

While in the other social media apps whenever propaganda is circulated like whatsapp due to its image/ video compression abnormalities created by AI can be ignored Platform content moderation operates on hash-matching and classifier-based systems. PhotoDNA and its video equivalents are effective for previously identified content but are structurally blind to novel synthetic media, which by definition does not match any prior hash. AI-generated extremist content, produced fresh for each deployment, evades hash-based systems entirely.

India's legal framework for content moderation centres on the Information Technology Act, 2000 (amended 2008) and the IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 creates compliances for the online platforms its uses and privacy of the data but lack of statutory definition for AI and synthetic media creates interpretive ambiguity which is exploited by these terror networks. Various state level digital initiatives along with Broadcast seva portal and PIB fact check identify hundreds of manipulated videos and images but even a modest resource adversary through the LLM’s can generate thousands of narrative media on a daily basis.

Academic literature assesses current mechanisms as providing a necessary but deeply insufficient baseline. Indian content moderation mechanism faces three deficits:

- **Regulatory fragmentation:** Every Ministry in India addresses different sections of governance hence there is no base rule which can be applied everywhere leading to poor coordination and slow responses.
- **Lack of Indigenous technology:** Unavailability of advanced AI detection tech like Natural language processing and computer vision softwares which understands the patterns, language, and culture for enhanced detection and content moderation.
- **Limited transparency requirements:** Social media companies are not legally compliant to identify and segregate AI and nonAI videos making it difficult for the authorities to distinguish between authentic and synthetic information.

8. Regulatory and technological measures to reduce future risks

To challenge the extremist propaganda, a layered response is needed involving interagency cooperation, technological investment, international cooperation, and a broad legal structure.

8.1. Legal and regulatory reforms

India needs a stronger, broad-based IT regulation with clear legal definitions of synthetic media and AI-generated content. The proposed Digital India act replacing IT act 2000 should include clear provisions for the regulation of deepfakes AI generated videos and synthetic audios. Whenever any AI video is uploaded it should become compliant that the producer has to clearly label AI-generated content in the case of non compliance penalties should be imposed for inciting violence, misinformation and for the support of terrorist recruitment. India can learn from the international development on legal Compliances for AI and synthetic media like European Union AI act (2024) and US defiance act (2024) both of which seeks to improve transparency and accountability ,it also involves provisions like: a)users are to be alerted when consuming AI videos, b)deepfake must be clearly labeled c) transparency disclosures of synthetic text, audio, video or images d)Citizens can sue creators for non consensual synthetic images e) establishment of legal liability even when the content is completely synthetic.

In addition, the Ministry of Electronics and Information Technology should adopt a content provenance system which uses digital signature and metadata to analyse whether the data is authentic or synthetic. Coalition for content provenance and authenticity can help verify content provenance before widespread dissemination. This can in turn enhance the efficiency of Indian counter radicalisation agencies as they can verify the information and make the content information more transparent.

8.2. Indigenous technological investments

India is investing heavily on generating AI detection tools based on south asian datasets under the aegis of India AI which is an independent business division and national initiative under MEITY. Institutions like centers *for artificial intelligence and robotics* along with IIT's are developing tools like *Sakhya multi agent, AI vishleshak* detecting deepfake videos. MHA has also created a synthetic media threat assessment cell to monitor emerging threats related to deepfakes, AI generated propaganda and online radicalisation. This specialised unit could analyse new trends and patterns based on indigenous LLM's on radicalisation and extremism with the intelligence agencies. A grievance appellate committee is also established which provides a platform to address complaints regarding harmful and misleading information.

8.3. International Cooperation

India should leverage its membership in the international forums such Global Counter Terrorism forum, FATF, and QUAD to make binding international rules on synthetic media provenance. Bilateral international agreements with USA, France, Britain and Israel could be made as these countries have formed strict content moderation policies and can share real time synthetic media signatures and extremist content libraries. At the multilateral level India should work closely ith *UN's Counter terrorism committee executive directorate* to ensure that AI generated propaganda, deepfakes and synthetic media are included in international counter-terrorism discussions based on UNSC resolution 2354(2017) which focuses the importance of countering the terrorist propaganda in order to reduce radicalisation and extremism.

8.4. Need of a Counter narrative ecosystem

Suppression and removal of the content is not enough. India needs to build a robust ADI detection ecosystem to diminish the False consciousness effects. A comprehensive effort is needed from all the sections of civil society, educational institutions, religious leaders and scholars to create messages that strengthen social cohesion thus creating social harmony, inclusion and national unity. These counter narratives should address issues like Unemployment, social exclusion and marginalisation. Contemporary initiatives like Saajha program in Uttar pradesh and Aman ki Asha demonstrate how community engagement can act as a tools for community resilience and de-radicalisation

9. Conclusion

The combination of AI and synthetic media is not a future threat now. The convergence of artificial intelligence and synthetic media is no longer a hypothetical future threat; it is an emerging security challenge with significant implication on terrorist recruitment radicalisation and information warfare. The mass content generation and funnel strategy to reach people even at the micro level spreading of the propaganda and rise in recruitment sitting in the hands of developers in a small room can be produced by a small number of actors with limited resources and technical expertise, while being completely anonymous to the world.

With India's large cultural, religious, linguistic and ethnic diversity it stands at the centre of challenges offered by AI and synthetic media creating security challenges. Extremists groups use them to spread propaganda and target specific

vulnerable individuals and can influence public opinion at a significant level. The growth in the AI detection and content moderation systems need to be further accelerated by investing on indigenous detection tools, strengthening legal frameworks, improving intelligence sharing mechanisms and working with international partners to counter threat vectors. India possesses such a policy framework which can address such issues with great intensity.

Fundamentally the challenge is not simply technological but strategic as it is a part of hybrid warfare which aims at distorting a country's viable resources and increasing societal vulnerabilities to hostile influence operations. The future of counter- terrorism lies on how efficiently India tackles the AI and synthetic media pressures and focusing more on building resilience and develop effective mechanisms to counter extremist propaganda before in influences vulnerable populations. If proactive measures are not taken then a time could come where AI can reach a sophistication level that makes reactive responses increasingly irrelevant.

References:

1. Banaji, Shakuntala, Ram Bhat, Anirudh Agarwal, Nirmalya Passanha, and Sadhana Pravin. *WhatsApp Vigilantes: An Exploration of Citizen Reception and Circulation of WhatsApp Misinformation Linked to Mob Violence in India*. London: London School of Economics and Political Science, 2019.
2. Basit, Abdul. *Radicalisation of Indian Muslims: Concerns and Realities*. RSIS Working Paper No. 294. Singapore: S. Rajaratnam School of International Studies, 2015.
3. Bloom, Mia. *Bombshell: The Many Faces of Women Terrorists*. London: Hurst & Company, 2011.
4. Bloom, Mia, and Sophia Moskalenko. *From Caliphate to Algorithm: Islamic State's Evolving Digital Propaganda Ecosystem*. London: Global Network on Extremism and Technology (GNET), 2023.
5. Chesney, Robert, and Danielle Keats Citron. “Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security.” *California Law Review* 107, no. 6 (2019): 1753–1820.
6. Dolhansky, Brian, Joanna Bitton, Ben Pflaum, Jikuo Lu, Russ Howes, Menglin Wang, and Cristian Canton Ferrer. “The Deepfake Detection Challenge (DFDC) Dataset.” *arXiv Preprint* arXiv:2006.07397 (2020).
7. European Union. *Regulation (EU) 2024/1689 of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act).* Brussels: Official Journal of the European Union, 2024.
8. Institute for Strategic Dialogue (ISD). *AI-Generated Extremism: How Generative Tools Are Reshaping Far-Right Propaganda*. London: ISD Global, 2023.
9. Institute of Peace and Conflict Studies (IPCS). *Mapping Radicalisation in India: Social Media, Education and Identity*. Research Paper No. 6. New Delhi: IPCS, 2016.
10. Lewandowsky, Stephan, Ullrich K. H. Ecker, Colleen M. Seifert, Norbert Schwarz, and John Cook. “Misinformation and Its Correction: Continued Influence and Successful Debiasing.” *Psychological Science in the Public Interest* 13, no. 3 (2012): 106–131.
11. McCauley, Clark, and Sophia Moskalenko. “Understanding Political Radicalization: The Two-Pyramids Model.” *American Psychologist* 72, no. 3 (2017): 205–216.
12. National Investigation Agency. *Annual Report 2023*. New Delhi: Ministry of Home Affairs, Government of India, 2023.
13. Sageman, Marc. *Understanding Terror Networks*. Philadelphia: University of Pennsylvania Press, 2004.
14. Statista Research Department. “WhatsApp Users in India 2024.” Statista, 2024.
15. Warrick, Joby. *The ISIS Files: Propaganda, Caliphate-Building, and the War for Minds*. Washington, D.C.: Brookings Institution Press, 2022.

16. Berger, J. M., and Jonathon Morgan. *The ISIS Twitter Census: Defining and Describing the Population of ISIS Supporters on Twitter*. Washington, D.C.: Brookings Institution, 2015.
17. Winter, Charlie. *Documenting the Virtual Caliphate*. London: Quilliam Foundation, 2015.
18. International Centre for Counter-Terrorism (ICCT). *The Exploitation of Generative AI by Terrorist Groups*. The Hague: International Centre for Counter-Terrorism, 2024.
19. Jacoby, Tim. "The Islamic State's Use of the Qur'an in its Magazines, Dabiq and Rumiyah." Discourse & Society 35, no. 3 (2024). https://doi.org/10.1177/09579265231208660.
20. Nair, Anjali. "Extremist Propaganda on Social Media: Impact, Challenges and Countermeasures." Observer Research Foundation, October 4, 2024. https://www.orfonline.org/research/extremist-propaganda-on-social-media-impact-challenges-and-countermeasures.
21. Sherwood, Harriet, Sandra Laville, Kim Willsher, Ben Knight, Maddy French, and Lauren Gambino. "Schoolgirl Jihadis: The Female Islamists Leaving Home to Join ISIS Fighters." The Guardian, September 29, 2014. https://www.theguardian.com/world/2014/sep/29/schoolgirl-jihadis-female-islamists-leaving-home-join-isis-iraq-syria.
22. United States Senate. *Disrupt Explicit Forged Images and Non-Consensual Edits Act of 2024 (DEFIANCE Act of 2024).* Washington, DC: United States Senate, 2024. https://www.durbin.senate.gov/imo/media/doc/defiance_act_of_2024.pdf.
23. Mathias, Lloyd, and Harsh Lailer. "Deepfake Challenge: India Needs Provenance, Not Just Platform Moderation." *Deccan Herald*, January 19, 2026. https://www.deccanherald.com/opinion/deepfake-challenge-india-needs-provenance-not-just-platform-moderation-3866943.
24. Ministry of Electronics and Information Technology, Government of India. "IndiaAI Mission Launches AIKosha and AI Compute Portal to Democratise Access to AI Resources." Press Information Bureau, September 19, 2025. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2175698.
25. United Nations Security Council. Resolution 2354 (2017): Implementation of the Comprehensive International Framework to Counter Terrorist Narratives. S/RES/2354 (2017). New York: United Nations, May 24, 2017.

The views expressed are those of the authors and do not represent the views of CNAWS.

Image Source: ChatGPT