ISSN-P : 2521-2982
ISSN-E : 2707-4587

GLOBAL Political REVIEW
www.gprjournal.com
empowering humanity

VOL. XI ISSUE II, SPRING (JUNE-2026)

GPR

GLOBAL POLITICAL REVIEW

HEC-RECOGNIZED CATEGORY-Y

GLOBAL POLITICAL REVIEW (GPR)

HumanityPublications
sharing research
www.humapub.com
US | UK | Pakistan

DOI (Journal): 10.31703/gpr
DOI (Volume): 10.31703/gpr.2026(XI)
DOI (Issue): 10.31703/gpr.2026(XI-II)

Double-blind Peer-review Research Journal
www.gprjournal.com
© Global Political Review

Humanity Publications (HumaPub)
www.humapub.com
Doi: https://dx.doi.org/10.31703

Title: Generative AI and Autonomous Radicalization: Reshaping the future of Lone-Actor Terrorism

Abstract

The emergence of Generative artificial intelligence (AI) is rapidly transforming the digital ecosystems in which radicalization may occur easily. Earlier, search engines recommended limited information, mostly based on specific queries. AI has empowered human-machine interaction in a more personalized, continuous and adaptive way. This article examines the relationship between Generative AI, autonomous radicalization and lone-actor terrorism by integrating established theories, including Moghaddam's Staircase to Terrorism, Bandura's Social Learning Theory, and Kruglanski's 3N Model and proposes an Interactive Radicalization Interface (IRI). This model of radicalization provide a framework to understand grievances, AI-mediated interaction, narrative construction, reinforcement and ideological progression. The study argues that AI operate as a force multiplier of vulnerability but does not cause radicalization or terrorism. The transition from passive exposure to interactive reinforcement without AI governance has raised concerns for prevention and Counterterrorism strategies.

Keywords: Artificial Intelligence, Autonomous Radicalization, Generative AI, Interactive Radicalization Interface, Lone-Actor Terrorism, Online Radicalization, Terrorism

Authors:

Aisha Sarwar: Clinical Psychologist / Scholar, Terrorism and Political Violence, National Defense University (NDU), Islamabad, Pakistan.

Syed Kaleem Imam: (Corresponding Author)
Former Federal Secretary / IGP / UN Police Commissioner, Pakistan.
(Email: skimam98@hotmail.com)

Pages: 96-107
DOI:10.31703/gpr.2026(XI-II).08
DOI link: https://dx.doi.org/10.31703/gpr.2026(XI-II).08
Article link: https://gprjournal.com/article/generative-ai-and-autonomous-radicalization-reshaping-the-future-of-loneactor-terrorism
Full-text Link: https://gprjournal.com/article/generative-ai-and-autonomous-radicalization-reshaping-the-future-of-loneactor-terrorism
Pdf link: https://www.gprjournal.com/jadmin/Auther/31rvIolA2.pdf

Global Political Review

p-ISSN: 2521-2982 **e-ISSN:** 2707-4587
DOI (journal): 10.31703/gpr
Volume: XI (2026)
DOI (volume): 10.31703/gpr.2026(XI)
Issue: II Spring (June-2026)
DOI(Issue): 10.31703/gpr.2026(XI-II)

Home Page
www.gprjournal.com
Volume: XI (2026)
https://www.gprjournal.com/Current-issue
Issue: II-Spring (June-2026)
https://www.gprjournal.com/issue/11/2/2026
Scope
https://www.gprjournal.com/about-us/scope
Submission
https://humaglobe.com/index.php/gpr/submissions

Scan the QR to visit us

©2026 The Author(s). This work is licensed under a Creative Commons Attribution-Non-Commercial-No Derivatives 4.0 International License.

Humanity Publications (HumaPub)
www.humapub.com
Doi: https://dx.doi.org/10.31703

Citing this Article

Article Serial	08
Article Title	Generative AI and Autonomous Radicalization: Reshaping the future of Lone-Actor Terrorism
Authors	Aisha Sarwar Syed Kaleem Imam
DOI	10.31703/gpr.2026(XI-II).08
Pages	96–107
Year	2026
Volume	XI
Issue	II

Referencing & Citing Styles

APA	Sarwar, A., & Imam, S. K. (2026). Generative AI and Autonomous Radicalization: Reshaping the future of Lone-Actor Terrorism. *Global Political Review*, *11*(2), 96-107. https://doi.org/10.31703/gpr.2026(XI-II).08
CHICAGO	Sarwar, Aisha, and Syed Kaleem Imam. 2026. "Generative AI and Autonomous Radicalization: Reshaping the future of Lone-Actor Terrorism." *Global Political Review* 11 (2):96-107. doi: 10.31703/gpr.2026(XI-II).08.
HARVARD	SARWAR, A. & IMAM, S. K. 2026. Generative AI and Autonomous Radicalization: Reshaping the future of Lone-Actor Terrorism. *Global Political Review,* 11**,** 96-107.
MHRA	Sarwar, Aisha, and Syed Kaleem Imam. 2026. 'Generative AI and Autonomous Radicalization: Reshaping the future of Lone-Actor Terrorism', *Global Political Review*, 11: 96-107.
MLA	Sarwar, Aisha, and Syed Kaleem Imam. "Generative Ai and Autonomous Radicalization: Reshaping the Future of Lone-Actor Terrorism." *Global Political Review* 11.2 (2026): 96-107. Print.
OXFORD	Sarwar, Aisha and Imam, Syed Kaleem (2026), 'Generative AI and Autonomous Radicalization: Reshaping the future of Lone-Actor Terrorism', *Global Political Review,* 11 (2), 96-107.
TURABIAN	Sarwar, Aisha and Syed Kaleem Imam. "Generative Ai and Autonomous Radicalization: Reshaping the Future of Lone-Actor Terrorism." *Global Political Review* 11, no. 2 (2026): 96-107. https://dx.doi.org/10.31703/gpr.2026(XI-II).08.

©2026 The Author(s). This work is licensed under a Creative Commons Attribution-Non-Commercial-No Derivatives 4.0 International License.

Global Political Review
ISSN-P : 2521-2982 | DOI: https://dx.doi.org/10.31703/gpr.2026(XI-II).08 | ISSN-E : 2707-4587
www.gprjournal.com

Generative AI and Autonomous Radicalization: Reshaping the future of Lone-Actor Terrorism

Aisha Sarwar[1] **Syed Kaleem Imam** (Corresponding Author)[2]

[1] Clinical Psychologist / Scholar, Terrorism and Political Violence, National Defense University (NDU), Islamabad, Pakistan.

[2] Former Federal Secretary / IGP / UN Police Commissioner, Pakistan.
(Email: skimam98@hotmail.com)

Abstract

The emergence of Generative artificial intelligence (AI) is rapidly transforming the digital ecosystems in which radicalization may occur easily. Earlier, search engines recommended limited information, mostly based on specific queries. AI has empowered human-machine interaction in a more personalized, continuous and adaptive way. This article examines the relationship between Generative AI, autonomous radicalization and lone-actor terrorism by integrating established theories, including Moghaddam's Staircase to Terrorism, Bandura's Social Learning Theory, and Kruglanski's 3N Model and proposes an Interactive Radicalization Interface (IRI). This model of radicalization provide a framework to understand grievances, AI-mediated interaction, narrative construction, reinforcement and ideological progression. The study argues that AI operate as a force multiplier of vulnerability but does not cause radicalization or terrorism. The transition from passive exposure to interactive reinforcement without AI governance has raised concerns for prevention and Counterterrorism strategies.

Keywords: *Artificial Intelligence, Autonomous Radicalization, Generative AI, Interactive Radicalization Interface, Lone-Actor Terrorism, Online Radicalization, Terrorism*

Introduction

Terrorist movements are rapidly shifting to modern technologies for the same old longstanding purposes. Digitalization of Print, radio, television, and the internet / social media has changed the way propaganda was produced, distributed and consumed (United Nations Office of Counter-Terrorism [UNOCT], 2022). Generative artificial intelligence represents another major shift, different but much important: It interacts with the user while delivering the information (Conway, 2017; Weimann, 2016; Weimann, 2025).

Large language models and other generative systems are capable of producing and developing texts, images, audio and video in response to individual user. These capacities provide considerable benefits for education, health, research, governance and business. On the other hand, they raise concerns about misuse for propaganda, disinformation, manipulation and extremist communication (McGuffie & Newhouse, 2020). The expanding digital landscape has extended the access to conceptual material and the role social media play has increased its outreach regarding speed and networking potential. Generative AI could plausibly add other dimensions such as personalized and sustained interaction. (Weimann, 2025; Ganaie, 2026)

©2026 The Author(s). This work is licensed under a Creative Commons Attribution-Non-Commercial-No Derivatives 4.0 International License.

This key difference and nuanced approach are particularly relevant to lone-actor terrorism. Traditionally radicalization process often involved ideological mentors, extremist networks, recruiters and group interaction (Moghaddam, 2005; Horgan, 2014; Borum, 2011). At present, Individuals may develop extremist beliefs through combinations of grievance, ideological exposure while staying socially isolated and digital engaged. Autonomous radicalization therefore does not mean radicalization in complete isolation. It is a pathway without direct organizational recruitment or control while remaining influenced by a wider social and digital environment (Gill, 2015; Gill et al., 2014; Schuurman & Carthy, 2025).

Generative AI may reshape digital environment by letting users to ask questions, providing refined arguments and continually engage with tailored responses. Radicalization is a complex process included psychological, social, political and ideological factors shaped by human. The pertinent query therefore is not about AI "creates terrorists" but whether the AI-mediated communication can amplify the mechanisms which are already associated with radicalization (Conway, 2017; Gill et al., 2014; Weimann, 2025).

This article attempts to answer this question with the support of three already established perspectives: Moghaddam's Staircase to Terrorism, Bandura's Social Learning Theory and Kruglanski's 3N Model. Bringing the three together through a proposed Interactive Radicalization Interface (IRI). This proposed framework treats AI not as the source of radicalization but as a potential mediating environment connecting grievance, narrative formation, reinforcement and ideological progression. Generative AI may not alter the fundamental psychology underlying radicalization, but it may reshape the interactional architecture through which radicalization unfolds.

Autonomous Radicalization

Autonomous radicalization is a process in which an individual adopts extremist beliefs without direct exposure with a terrorist organization or sustained personal interaction with an extremist group (Conway, 2017; Von Behr et al., 2013). Radicalization doesn't occur in isolation rather external ideological ecosystems like political grievances online content, access to the chatbots and digital mediation shape the extremist attitudes and beliefs.

Earlier models generally assigned an important role to networks, mentors and organizations in reinforcing extremist attitudes and were not able to generate personalized responses. Digital platforms have reduced dependence on direct physical interaction by allowing individuals to encounter sociopolitical ethical or religious material independently. Generative AI potentially extends this development by introducing a more adaptive form of communication in which users can receive responses that are generated in relation to their questions. (Conway, 2017; Von Behr et al., 2013). A more specific explanation may therefore be organizationally autonomous but digitally facilitated radicalization.

Lone-Actor Terrorism

Lone actor terrorism represents an activity committed primarily by a single individual who is self-planning and executing without directions from the groups or organizations. These lone actors operating independently, but rarely are ideologically isolated (Gill, 2015). In such case consumption of extremist material is achieved through engagement with online communities while individuals draw inspiration from broader movements and narratives exposed through online sources (Spaaij, 2010; Schuurman & Carthy, 2025).

The term lone actor is preferable choice over to "lone wolf" because it avoids glamorizing terrorists' activities as guile or clever acts and better reflects contemporary academic terminology.

Lone actors are especially relevant because they are less dependent on identifiable organizational structures. Reduced dependability further complicates the conventional intelligence and counterterrorism approaches.

Problem Statement

This study has broadly examined existing online radicalization, extremist propaganda, social media and lone-actor terrorism. Emerging researches considers the security implications of generative AI. However, less consideration has been shown to the mechanisms through which AI-mediated interaction crisscross autonomous radicalization and lone-actor terrorism.

The real challenge is therefore not the extremist content exists online, but whether a new form of interactive technology will cause changes into how vulnerable individuals engage with narratives, grievances and ideological reinforcement.

Research Questions

Primary Question

How might Generative AI reshape pathways of autonomous radicalization and influence the future threat of lone-actor terrorism?

Secondary Questions

1. How does AI-mediated radicalization differ from earlier online ideological exposure? 2. How might repeated human-AI interaction have contributed existing autonomous radicalization pathways?

Research Objectives

1. To explore and examine the potential role of generative AI in autonomous radicalization and lone-actor terrorism.
2. To develop a framework explaining how AI-mediated interaction may influence already established radicalization processes.

Research Approach

This article draws upon established theoretical perspectives including radicalization, lone-actor terrorism, social learning and digital extremism with emerging concerns surrounding generative AI. The resulting model introduces AI mediated interaction as an additional dimension through which AI may influence the dynamics of radicalization and trajectory.

Rather than attempting to establish a direct causal relationship between AI use and terrorism, the study identifies mechanisms through which interactive AI environments may intersect with established radicalization processes. The IRI model is conceptual in its present form, and its primary purpose is to provide a theoretically grounded framework that can be empirically examined and validate through subsequent research in future.

Literature Review

Lone-Actor Terrorism and Digital Radicalization

Lone actor terrorism poses a distinctive challenge for security because individuals may pursue violent plans and undertake acts of violence without the structural communication typically associated with organized terrorist groups. (Borum, 2011).

Lone actors do not necessarily mean being socially or ideologically isolated, instead they rarely develop extremist beliefs in complete isolation. Evidence suggests that online exposure and interaction towards Personal grievances, social circumstances, ideological narratives can enhance the radicalization process and eventually violence may emerge (Gill et al., 2014; Schuurman & Carthy, 2025). Studies by Gill, Hamm and Spaaij, Conway, and others demonstrated the value of examining individual vulnerabilities, social factors and enabled ideological identification beyond boundaries of formal organizational affiliation.

The internet changed access. Social media changed reach and networking. Generative AI may change interaction. (Gill et al., 2014 Hamm & Spaaij, 2017; Conway, 2017; Weimann, 2025).

Schuurman and Carthy (2025) in a comparative study between group and lone actor terrorist and finds that lone actor terrorism frequently experiences individualized pathways to radicalization heavily relied on personal grievances, social isolation, and extensive online ideological exposure. They maintain the ideological connections through digital sources and may further accelerate through extremist engagement.

Miloshevska (2024) has explored the manipulative interface of generative artificial intelligence in her research work. The study indicated that AI can generate propaganda, manipulative communication, and active support depending on how radicalized individual adopt AI. Different factors may also involve like cost, accessibility and ideologically compatibility. The research further predict that AI is increasing the capabilities of individuals for future terrorist activities.

Generative AI and Radicalization

Research on AI and extremism increasingly considers automated content generation, personalization, synthetic media, propaganda and manipulation (Houser & Dong 2025). Baele et al. (2025) argued that generative systems can produce material rapidly and adapt responses to individual prompts and interests. The concern is therefore not simply an increase in the volume of information but the possibility of more responsive and individualized engagement. (Weimann, 2025; Ganaie, 2026).

Kunst et al. (2026) in recent research presented an architecture of radicalization by mapping psychological vulnerabilities interact with AI driven mechanisms across four stages.

This work highlights how Generative AI can reshape and accelerate established pathways of radicalization. when AI interacts with psychological vulnerabilities it can increase the radicalization process by reinforcing extremist beliefs and facilitating deeper engagement with extremist networks.

Houser and Dong (2025) in their systematic review, explained the convergence of AI and terrorism and identified a noticeable knowledge gap concerning terrorists' use of AI, while identifying propaganda and recruitment among areas where AI may augment terrorist activity. Weimann (2025) likewise examines how generative AI and large language models may enhance radicalization, recruitment and propaganda.

Yet there is no establish a causal relationship between AI and terrorism. Most individuals exposed to extremist material do not engage in terrorism. AI must instead be understood within a wider ecosystem involving grievance, identity, ideology, vulnerability and social relationships.

Research Gap

Three bodies of scholarship have largely developed alongside one another: research on radicalization and lone actors; research on internet and social-media-enabled extremism; and emerging work on generative AI and security.

A comparatively underdeveloped link is an explanation of how interactive AI mediation might connect these domains. Information can become interaction; interaction can produce reinforcement; and reinforcement can become individualized.

Kunst et al. (2026) address this emerging intersection through a four-stage AI architecture of radicalization that integrates AI, psychological risk factors and radicalization processes within these stages. This research provides a framework for understanding where AI may fit in with existing factors of radicalization.

The specific micro level interactional mechanism through which generative AI may continuously respond to, reinforce, and shape an individual's psychological and motivational processes remain less developed. The IRI model addresses this gap by conceptualizing AI-mediated interaction as a continuous mechanism.

IRI operates across six stages of radicalization and integrate Bandura's social learning principles, Kruglanski's 3N model, and Moghaddam's Staircase framework. Rather than treating generative AI simply as an environmental or psychological factor, IRI conceptualizes human to AI interaction as a potential mechanism through which existing radicalization processes may be reinforced, personalize, and progressively intensify. While remaining a conceptual framework, it requires future empirical testing.

Theoretical Framework

Radicalization is neither automatic nor linear. Most people experiencing grievance, injustice or ideological exposure never engage in terrorism. Theory should therefore not be used to predict terrorism mechanically but to understand processes that may increase or reduce the possibility of progression toward violence (Borum, 2011; McCauley & Moskalenko, 2017).

Modern researches revealed that online environments are more rapidly facilitating the processes that reinforce grievances and validating that ideology swiftly. These autonomous communication generators and innovative AI models made it much easier to facilitate the Individuals to step-up in the staircase. The present study has incorporated three classic models to explain the phenomenon of autonomous radicalization and lone-actor terrorism. These established theories not only help in understanding the process of autonomous radicalization but also explain why such behaviors occur and why only a limited number of individuals ultimately engage in terrorist activities. While these classical models successfully identify the pathways of radicalization, they provide limited explanation of how emerging Generative Artificial Intelligence may influence, facilitate, or reinforce autonomous radicalization in contemporary digital environments. The Present study incorporates these models into the suggested Interactive Radicalization Interface (IRI) in order to bridge this gap.

Kruglanski's 3N Model

The 3N framework emphasizes Needs, Narratives and Networks. Needs may arise from humiliation, exclusion, threatened identity or a search for significance. Narratives interpret these experiences and identify causes, enemies and possible responses. Networks reinforce narratives through belonging and social validation (Webber & Kruglanski, 2016).

Generative AI potentially intersects with these elements. Its prompt responses to the expressed needs, assist in constructing interpretations and simulate forms of narrative validation. However, it does not reproduce the influential human social networks.

Bandura's Social Learning Theory

Bandura's theory explains the learning process through observation, modelling, and reinforcement. When these principles are applied to radicalization, certain attitudes and behaviors may be strengthened through repeated exposure, perceived approval, and social feedback (Bandura, 1977).

Digital platforms provide support and facilitate such processes by exposing individuals to repeated content and social feedback. Generative AI adds an interactive dimension to communication, as its responses can be adapted according to user input. This may become particularly concerning when a vulnerable individual repeatedly seeks confirmation of existing beliefs rather than alternative interpretations.

Moghaddam's Staircase to Terrorism

Mughaddam (2005) describes the movement towards terrorism and presented a staircase model. In this model perceived injustice gradually move an individual from strong perception of blame to moral justifications and finally acting out the terrorism. It develops strong divisions between Us vs them and relatively a small no of individuals actually accepts violence as solution (Moghaddam, 2005). Although this theory developed before the emergence of AI the idea still offers an understanding the violent behavior. AI actually doesn't create the pathways but it can influence the movement of radicalized individual by how it is being used.

Lygre et al. (2011) in a critical review on Moghaddam's theory found that the model offers an important framework to explain radicalization but the transitions between its stages is not sufficiently supported.

Weimann (2025) in a recent study has identified areas like radicalization, recruitment, information gathering, planning, and propaganda in which generative AI/LLMs may enhance terrorist capabilities.

Then the question arises here is whether radicalization should be understood as a strictly linear sequence or personalized interaction might reduce friction between its steps by reinforcing grievances, attribution or ideological certainty.

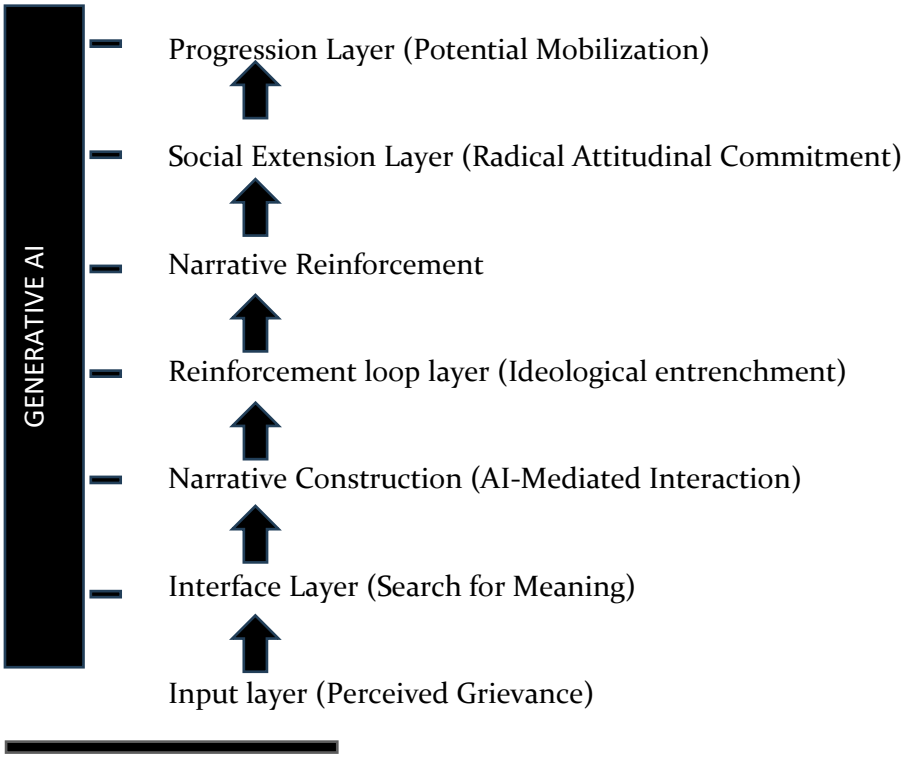
The Interactive Radicalization Interface

The Interactive Radicalization Interface (IRI) is the core conceptual contribution of this study, explaining how AI-mediated personalized continuous engagement interact with existing vulnerabilities. It also stablished radicalization mechanisms without treating AI itself as an autonomous radicalizing actor. This model has integrated three established theories of radicalization, learning and terrorism into a system-based architecture. Rather assuming a deterministic pathway, the model presents radicalization as dependent on specific conditions. AI-not only mediate the interaction but also reinforce and strengthen existing interpretations and vulnerabilities in certain individuals however, such reinforcements does not lead to ideological rigidity, radicalization, or violence in all cases.

The IRI model is composed of six interdependent multilayers.

Figure 1

Interactive Radicalization Interface (IRI)



Stage 1: Vulnerability and Grievance

The process begins with the individual, not the technology. Relevant factors may include perceived injustice, psychological needs, emotional distress, existing beliefs, cognitive biases and social isolation.

Stage 2: AI-Mediated Interaction

The individual enters an adaptive communication environment. Unlike a static webpage, a generative system responds to questions and subsequent prompts, producing iterative interaction.

Stage 3: Narrative Construction

Repeated interaction may assist users in interpreting experiences and organizing grievances into narratives. Importantly, the same technology may also expose users to alternative interpretations. Reinforcement is therefore possible, not inevitable.

Stage 4: Reinforcement Loop

Repeated interaction may strengthen existing interpretations through cognitive repetition, perceived validation and sustained engagement. The principal concern is the emergence of a low-

friction confirmation environment in which users repeatedly seek material consistent with existing assumptions.

Stage 5: Social Amplification

Online communities, influencers, algorithms and perceived consensus may extend AI-mediated interpretations into wider social environments. AI therefore interacts with rather than replaces human networks.

Stage 6: Possible Progression

Among a small subset of already vulnerable individuals, sustained reinforcement may contribute to ideological rigidity, sharper in-group/out-group distinctions, moral disengagement and possible normalization of violence.

IRI therefore does not propose another deterministic ladder to terrorism. It identifies an interaction layer through which existing psychological and social mechanisms may increasingly operate.

IRI model conceptualized the role of AI as a mediator throughout the layers, that may impact on individual experiences, evaluate shared information and their engagement with ideological narratives. This framework explains the role of Autonomous Intelligence in individual perceptions and pathways.

Discussion: From Exposure to Interaction

The principal transformation introduced by generative AI may be qualitative rather than merely quantitative. Digital environments surrounding extremism has progressively changed and have primarily expanded access to ideological information. Social media added recommendation, rapid dissemination of information and introduced greater opportunities for networking (Conway, 2017; Cinelli et al., 2021).

Generative AI by representing an innovative possibility of continuous personalized dialogue enables users to engage in an interactive session. This feature is of utmost importance as it clearly distinguished exposure from reinforcement. (Weimann, 2025). Exposure may make an idea accessible whereas, reinforcement make an existing interpretation appear increasingly consistent, more familiar and eventually persuasive through repetition. Accordingly, the potential relevance of Generative AI is significant because it not only helps generating entirely new extremist ideas but potentially enable users to organize, revisit and reinforce the ideas they already regarding meaningful (Borum, 2011; Bandura, 1977; Weimann, 2025; Ganaie, 2026).

Mechanisms such as personalization, synthetic social interaction, cognitive dependence, scalable persuasion and weaken informational boundaries potentially operates interconnectedly within such environments. These mechanisms should not be regarded as independent / isolated or sufficient causes of radicalization rather their effect remain embedded and operative within a broader psychological predispositions, social relationships and wider political and informational environment (Baele et al., 2025; Miloshevska, 2024).

The findings explain why AI can better conceptualized as a force multiplier rather than a root cause of radicalization. The potential drivers of radicalization including grievance, identity conflict, ideological predisposition, exclusion and political conditions are central to established approaches to understanding radicalization and existed well before emergence of generative AI (Borum, 2011; McCauley & Moskalenko, 2017; Moghaddam, 2005).

Technology changes the environment through which these factors may be interpreted and reinforced (Conway, 2017; Weimann, 2025; Ganaie, 2026).

The theoretical significance of IRI lies here. The 3N Model explains needs, narratives and networks; Social Learning Theory explains learning and reinforcement; the Staircase describes possible progression. IRI adds the interactive technological environment through which elements of all three may increasingly intersect.

Implications for Counterterrorism

Traditional counterterrorism has understandably concentrated on organizations, recruiters, financing, communication networks and operational planning. Lone actor terrorism poses a complex challenge to these conventional approaches of detection because limited organizational ties generate fewer observable indicators and fewer detectable warning signs.

Generative AI further complicate the problem by supporting increasingly individualized and dispersed ideological engagement, while synthetic material can be produced rapidly and at scale. Consequently, effective counterterrorism efforts, therefore cannot rely on traditional organizational monitoring or removing extremist content alone. (Houser & Dong, 2025; Weimann, 2025; Ganaie, 2026).

Yet the presence or the use of AI must not itself constitutes the grounds for suspicion, should not become a proxy of suspicion. Excessive and broad surveillance could compromise public privacy, freedom of expression and public trust without necessarily addressing the underlying causes and contributors of radicalization. The security objective should therefore focus on proportionate risk-sensitive prevention rather than technological suspicion.

Recommendations

First, Safety measures should be incorporated into the design and development of generative AI. This requires ongoing assessment, regular reviewing and evaluation of data to strengthen the safeguards against extremist manipulation and emerging risk of misuse.

Second, countering efforts should focus on creditable risk indicators rather than routine AI use. Counterterrorism departments should be more vigilant and distinguish legitimate digital interaction from credible signs that an individual may be moving toward violent intent.

Third, focus should be placed on strengthening AI and digital literacy so that individuals should be better equipped to understand personalization systems, identify synthetic content, misinformation as well as to recognize the ways how digital environments can reinforce pre-existing beliefs.

Fourth, Regulating AI systems alone is insufficient to address grievance, exclusion, identity conflict, political violence and other well-established drivers of radicalization. Efforts to prevent AI enabled risk should link technological safeguards with social interventions.

Fifth, encourage future research that should adopt an interdisciplinary approach. Greater collaboration across disciplines like AI, behavioral science, terrorism scholars and counterterrorism practitioners can help clarify the effects of sustained human-AI interaction and its effects on vulnerable individuals without compromising privacy and civil liberties.

A key priority for future research is to empirically test and evaluate the IRI model. Researchers can adopt Controlled research design, anonymized interaction data, comparative case analysis and longitudinal approaches to examine how AI-mediated interaction contributes to ideological rigidity and under what conditions progression toward violent extremism happens.

Limitations and Future Research

This study is presenting a conceptual framework to understand the underlying mechanism of violent activity. It does not claim to establish a causal relationship between generative AI and terrorist violence. The proposed IRI model is intended to outline or identify possible mechanisms through which AI mediated interaction may operate alongside existing radicalization processes rather than predicting particular individual behavior.

Therefore, further research is required to test the model empirically, including investigations that could identify both the risk and protective factors. Future studies could determine whether reinforcement through AI mediated interaction differs significantly from interaction with social media, online communities or human interlocutors.

The pace and the speed at which generative AI is rapidly evolving poses an important challenge as advances in technological capabilities, safeguards and patterns of use may change faster to outpaced the conventional research timelines. Consequently, the framework should remain open to empirical revision and should be regarded as dynamic model rather than a definitive explanation.

Conclusion

Generative AI is not responsible for originating grievances, discovering identity conflicts or identifying ideological struggles associated with violent extremism. The engagement with an AI system never transforms an individual into a terrorist. Radicalization is multifaceted complex context dependent process influenced by human vulnerabilities, relationships, narratives and social conditions.

Generative AI introduces a different form of interaction that is more personalized, echo user's view uncritically thereby, changing the architecture of interaction. Traditional online platforms allowed users to search for, consume and share information manually. Modern systems can respond, adapt and continue the conversation on their own. Vulnerable individuals who are already holding rigid ideological beliefs, may facilitate the consolidation for narrative reinforcement and cognitive entrenchment.

In this regard The Interactive Radicalization Interface captures the change by connecting individual vulnerability with AI-mediated interaction, narrative construction, reinforcement, social amplification and possible ideological progression. Rather than claiming an inevitable pathway toward terrorism. it only focuses on identifying mechanisms whose relationship and effect require systematic empirical investigation.

This distinction also requires necessary, a more precise and clearer conceptualization of autonomous radicalization, because autonomy does not mean complete isolation. Lone actors acting without organizational controls or direct commands from a formal hierarchy may remain immersed in digital narratives through online communities and interactive technologies can continue to shaping their beliefs.

Alone in organizational terms but digitally accompanied through online environment a lone actor is connected to numerous online narratives and communities, making the distinction between "lone "and "connected" difficult. The security situation has extended beyond easier access to extremist network to more challenging pathways in which information can now become interaction, interaction can become reinforcement, and reinforcement become individualized.

A next wave of lone actor terrorism may require nuanced understanding of the transition from traditional to modern AI, with particular attention to how digital interaction may contribute to harmful reinforcement without criminalizing technology or restricting freedom of expression.

References

Baele, S. J., Naserian, E., & Katz, G. (2025). Is AI-generated extremism credible? Experimental evidence from an expert survey. *Terrorism and Political Violence, 37*(8), 1060–1076. https://doi.org/10.1080/09546553.2024.2380089

Google Scholar Worldcat Fulltext

Bandura, A. (1977). *Social learning theory*. Prentice Hall.

Google Scholar Worldcat Fulltext

Borum, R. (2011). Radicalization into violent extremism II: A review of conceptual models and empirical research. *Journal of Strategic Security, 4*(4), 37–62.

Google Scholar Worldcat Fulltext

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences, 118*(9), e2023301118. https://doi.org/10.1073/pnas.2023301118

Google Scholar Worldcat Fulltext

Conway, M. (2017). Determining the role of the internet in violent extremism and terrorism: Six suggestions for progressing research. *Studies in Conflict & Terrorism, 40*(1), 77–98. https://doi.org/10.1080/1057610X.2016.1157408

Google Scholar Worldcat Fulltext

Ganaie, N. A. (2026). The role of artificial intelligence in radicalisation, recruitment and terrorist propaganda: Deconstructing violent extremism and reimagining counterterrorism in contemporary digital ecosystems. *Frontiers in Political Science, 7*, 1718396. https://doi.org/10.3389/fpos.2025.1718396

Google Scholar Worldcat Fulltext

Gill, P. (2015). *Lone-actor terrorists: A behavioural analysis*. Routledge. https://doi.org/10.4324/9781315766348

Google Scholar Worldcat Fulltext

Gill, P., Horgan, J., & Deckert, P. (2014). Bombing alone: Tracing motivations and antecedent behaviors of lone-actor terrorists. *Journal of Forensic Sciences, 59*(2), 425–435.

Google Scholar Worldcat Fulltext

Hamm, M. S., & Spaaij, R. (2017). *The age of lone wolf terrorism*. Columbia University Press.

Google Scholar Worldcat Fulltext

Houser, T., & Dong, B. (2025). The convergence of artificial intelligence and terrorism: A systematic review of the literature. *Studies in Conflict & Terrorism*. Advance online publication. https://doi.org/10.1080/1057610X.2025.2527608

Google Scholar Worldcat Fulltext

Kunst, J. R., Obaidi, M., Gollwitzer, A., Brandtzæg, P. B., Hinrichs, Y., Saini, N., & Schroeder, D. T. (2026). Intelligent systems, vulnerable minds: A framework for radicalization to violence in the age of AI. *Personality and Social Psychology Review, 30*(3), 395–426. https://doi.org/10.1177/10888683261430089

Google Scholar Worldcat Fulltext

Lygre, R. B., Eid, J., Larsson, G., & Ranstorp, M. (2011). Terrorism as a process: A critical review of Moghaddam's staircase to terrorism. *Scandinavian Journal of Psychology, 52*(6), 609–616. https://doi.org/10.1111/j.1467-9450.2011.00918.x

Google Scholar Worldcat Fulltext

McCauley, C., & Moskalenko, S. (2017). Understanding political radicalization: The two-pyramids model. *American Psychologist, 72*(3), 205–216.

Google Scholar Worldcat Fulltext

McGuffie, K., & Newhouse, A. (2020). *The radicalization risks of GPT-3 and advanced neural language models*. arXiv. https://doi.org/10.48550/arXiv.2009.06807

Google Scholar Worldcat Fulltext

Miloshevska, T. (2024). The potential exploitation of generative artificial intelligence by terrorists and violent extremists. *Annual of the Faculty of Philosophy in Skopje, 77*(1), 361–389. https://doi.org/10.37510/godzb02477361m

Google Scholar Worldcat Fulltext

Moghaddam, F. M. (2005). The staircase to terrorism: A psychological exploration. *American Psychologist, 60*(2), 161–169. https://doi.org/10.1037/0003-066X.60.2.161

Google Scholar Worldcat Fulltext

Schuurman, B., & Carthy, S. L. (2025). Who commits terrorism alone? Comparing the biographical backgrounds and radicalization dynamics of lone-actor and group-based terrorists. *Crime & Delinquency, 71*(6–7), 2092–2117. https://doi.org/10.1177/00111287231180126

Google Scholar Worldcat Fulltext

Spaaij, R. (2010). The enigma of lone wolf terrorism: An assessment. *Studies in Conflict & Terrorism, 33*(9), 854–870. https://doi.org/10.1080/1057610X.2010.501426

Google Scholar Worldcat Fulltext

United Nations Office of Counter-Terrorism. (2022). *Countering terrorist use of new technologies*. United Nations.

Google Scholar Worldcat Fulltext

Von Behr, I., Reding, A., Edwards, C., & Gribbon, L. (2013). *Radicalisation in the digital era: The use of the internet in 15 cases of terrorism and extremism*. RAND Europe.

Google Scholar Worldcat Fulltext

Webber, D., & Kruglanski, A. W. (2016). Psychological factors in radicalization: A "3 N" approach. In G. LaFree & J. D. Freilich (Eds.), *The handbook of the criminology of terrorism* (pp. 33–46). Wiley-Blackwell. https://doi.org/10.1002/9781118923986.ch2

Google Scholar Worldcat Fulltext

Weimann, G. (2016). *Terrorism in cyberspace: The next generation*. Woodrow Wilson Center Press.

Google Scholar Worldcat Fulltext

Weimann, G. (2025). Generative AI and terrorism. In P. Hacker (Ed.), *Oxford intersections: AI in society*. Oxford University Press. https://doi.org/10.1093/9780198945215.003.0100

Google Scholar Worldcat Fulltext