SJIF Impact Factor (2025): 8.688| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 10 | Issue: 10 | October 2025 - Peer Reviewed Journal

THE ROLE OF ARTIFICIAL INTELLIGENCE IN IDENTIFYING ONLINE RADICALIZATION PATHWAYS IN THE UNITED STATES: A SCOPING REVIEW

Aisha Mohammed Suleiman [a], Clement Aryee [b*]
[a] *University of Iowa, Iowa, USA, ORCID ID: 0009-0004-3996-0305*
[b] *Department of Sociology, Kwame Nkrumah University of Science and Technology, Ghana*

**Corresponding Author: Clement Aryee*

Article DOI: https://doi.org/10.36713/epra24594
DOI No: 10.36713/epra24594

ABSTRACT

Social media platforms have become powerful spaces for spreading extremist ideologies, with several violent attacks in the United States linked to online radicalization processes. Artificial Intelligence (AI) technologies, including machine learning and natural language processing, offer potential solutions for detecting extremist content and identifying radicalization pathways. This scoping review examines how AI has been utilized to identify online radicalization pathways in the U.S. context, synthesizing research across computer science, criminology, and security studies. A systematic search of electronic databases and grey literature sources was conducted following established scoping review methodology. The review reveals that AI applications in radicalization detection primarily focus on content classification, user behavior analysis, and network mapping, with techniques that range from traditional machine learning to advanced deep learning models. However, significant challenges remain, including definitional inconsistencies, dataset limitations, ethical concerns about surveillance and privacy, and the need to balance public safety with constitutional protections. This review identifies critical gaps in current research and highlights the need for interdisciplinary collaboration to develop effective, ethical AI-based approaches to counter online radicalization while preserving civil liberties.

KEYWORDS: *Artificial intelligence; Online Radicalization; United States; Radicalization pathways; Extremist; Machine learning; Natural language processing*

1. INTRODUCTION

Social media platforms have fundamentally transformed how extremist groups recruit members, disseminate ideologies, and coordinate activities (Irfan et al., 2025). In the United States, several high-profile violent attacks have demonstrated clear connections to online radicalization processes, including the 2019 El Paso Walmart shooting, the 2022 Buffalo mass shooting, and the January 6th Capitol attack (O'Connor et al., 2025; Zdjelar et al., 2025). Recently on September 10, 2025, the assassination of Charlie Kirk demonstrates how gaps in U.S. counter-radicalization strategies allow online extremism to escalate into physical violence. Despite extensive monitoring, extremist narratives continue to flood digital platforms, shaping grievances and providing ideological justification for political attacks. The above cases exposes the limitations of existing approaches and reinforce the need for artificial intelligence systems that are capable of mapping radicalization pathways and flagging high-risk online activity before it metamorphoses in acts of violence. Artificial Intelligence technologies, particularly machine learning and natural language processing, have emerged as promising tools for addressing online radicalization (Bouhlaoui, 2025; Kaur, 2025). These technologies can process vast amounts of digital content at scale, identify patterns in extremist communications, and track behavioral changes that may indicate progression along radicalization pathways (Antoliš, 2025; Hussain, 2025; Meena et al., 2025). For instance, current AI applications include automated content classification systems, social network analysis tools, and predictive models for identifying at-risk individuals (Aad & Hardey, 2025). In addition, AI detects extremist networks and identifies supporters through content analysis and network mapping (Suleiman, 2024).

However, the application of AI to radicalization detection presents significant challenges. Technical limitations include dataset biases, false positive rates, and algorithmic opacity that make it difficult to understand how these systems reach their conclusions (Matar, 2025). Also, ethical and legal concerns have centered on privacy rights, potential surveillance overreach, and the risk of censoring legitimate political discourses (Patil, 2025; Tariq, 2025). In the U.S. context, these concerns are particularly weak, given the existence of the First Amendment protections for free speech and the government's reliance on voluntary industry action rather than regulatory mandates.

SJIF Impact Factor (2025): 8.688| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 10 | Issue: 10 | October 2025 - Peer Reviewed Journal

Despite growing research interest across multiple disciplines, the evidence base remains scattered across several disciplines. Thus, this review surveys the use of AI to identify online radicalization pathways in the United States and highlights the technologies employed, the conceptual frameworks adopted, the ethical questions raised, and gaps in current literature.

2. METHODS

This scoping review followed the methodological framework outlined by Arksey & O'Malley, (2005) and the PRISMA Extension for Scoping Reviews (PRISMA-ScR) reporting guidelines. Electronic databases searched included PubMed, IEEE Xplore, ACM Digital Library, Web of Science, and Google Scholar. Grey literature sources included government reports from DHS, FBI, and DOJ. Search terms combined AI/technology concepts (artificial intelligence, machine learning, natural language processing), radicalization concepts (extremism, hate speech, terrorism), and U.S. context terms (social media, online, United States). Inclusion criteria encompassed studies describing AI use for detecting online radicalization in U.S. contexts, and published in English. Two reviewers independently screened titles, abstracts, and full texts using standardized criteria. Data extraction focused on the AI techniques used, definitions of radicalization, data sources, findings, and ethical considerations. A narrative synthesis approach was employed given the heterogeneous nature of the included studies.

3. RESULTS

This scoping review identified seven studies published between 2017 and 2025 that examine the role of artificial intelligence (AI) in identifying or shaping online radicalization pathways in the United States. The included studies span multiple domains like computer science, cybersecurity, criminology, political science, and security studies. This reflects the multidisciplinary nature of the field. **Table 1** summarizes the characteristics of the included studies, while **Table 2** synthesizes their findings into key thematic areas.

Table 1. Characteristics of Studies Included

Author/Year	Study Focus	Data Source	Methods	Key Outcomes
Al-Zewairi & Naymat, (2017)	Profiling Islamist radicals in the U.S. using PIRUS dataset	Profiles of Individual Radicalization in the U.S. (PIRUS), 1,473 individuals	Supervised ML: Deep Learning, Random Forest, Gradient Boosting, Naïve Bayes	High recall & precision in distinguishing Islamist radicals from others
Rajendran et al. (2022)	Detecting extremism on Twitter during the U.S. Capitol riot	Custom dataset of tweets from the Capitol riot	Deep Learning: Bi-LSTM, BERT, RoBERTa, DistilBERT	RoBERTa achieved 95% accuracy; effective classification of propaganda, recruitment, and radicalization
Jenkins et al. (2025)	Review of AI development in defense & national security	104 high-relevance studies (2022–2024)	Scoping literature review using MLII framework	Highlighted bias, adversarial risk, and black-box opacity; need for oversight
Burton (2023)	Explores how AI and algorithms contribute to radicalization, polarization, and violence	Conceptual/empirical review drawing on securitization theory	Critical review, qualitative analysis	AI is both a counter-radicalization tool and driver of polarization; calls for reconceptualization
Pemmasani (2023)	AI in national security: defense strategies, cyber threat intelligence, surveillance	Case studies (U.S., Israel, EU, China) + review literature	Descriptive review, applied examples	AI enhances real-time threat detection; ethical issues include bias, privacy, and autonomous weapons
Horowitz et al. (2018)	Assessment of AI's implications for U.S. security	Policy review + illustrative cases (DARPA, DHS, elections)	Policy analysis, scenarios	AI reshaping cybersecurity, propaganda, defense; double-edged potential
Dong et al., (2024)	Evaluated GPT-3.5 and GPT-4 for classifying extremist vs non-extremist posts in the U.S.	Far-left and far-right online posts	Large language models (LLMs) with prompt engineering	GPT-4 excelled in detecting far-right extremism; GPT-3.5 better for far-left; recall > precision; revealed bias and prompt sensitivity

SJIF Impact Factor (2025): 8.688| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 10 | Issue: 10 | October 2025 - Peer Reviewed Journal

Table 2. Thematic Synthesis of Results

Theme	Description	Supporting Studies
AI for Radicalization Detection	Use of ML/LLMs to classify extremist vs. non-extremist behavior online and offline	Al-Zewairi & Naymat (2017); Rajendran et al. (2022); Dong et al. (2024)
AI in Cybersecurity & Threat Intelligence	AI deployed for real-time detection of threats, anomaly detection, and cyber defense	Pemmasani (2023); Horowitz et al. (2018); Jenkins et al. (2025)
Bias, Ethics & Black-Box Risks	Risks associated with opacity, algorithmic bias, and ethical dilemmas in AI use	Jenkins et al. (2025); Burton (2023); Dong et al. (2024)
AI in Propaganda & Information Warfare	AI in spreading or countering disinformation and extremist narratives	Rajendran et al. (2022); Horowitz et al. (2018); Burton (2023)
AI in National Security & Defense	Strategic and operational integration of AI in U.S. defense and homeland security	Pemmasani (2023); Horowitz et al. (2018); Jenkins et al. (2025)
Need for Oversight and Reconceptualization	Calls for governance frameworks, ethical AI policies, and reframing securitization	Jenkins et al. (2025); Burton (2023)

4. DISCUSSION

This scoping review shows that artificial intelligence (AI) is playing an increasing role in identifying online radicalization pathways in the United States. Across the seven included studies, five main themes emerged: AI for radicalization detection, cybersecurity and threat intelligence, bias and black-box risks, propaganda and information warfare, and AI in national defense.

First, studies that apply AI to radicalization detection (Al-Zewairi & Naymat, 2017; Rajendran et al., 2022; Dong et al., 2024) demonstrate a strong technical capacity. Traditional supervised machine learning, deep learning architectures, and, more recently large language models (LLMs) all achieved promising results in classifying extremist versus non-extremist individuals and content. Nonetheless, a consistent challenge was the balance between recall and precision, with high recall increasing the risk of false positives and potential stigmatization.

Second, AI is central to cybersecurity and threat intelligence. Case studies and reviews (Pemmasani, 2023; Horowitz et al., 2018; Jenkins et al., 2025) highlighted how AI strengthens predictive analytics, anomaly detection, and situational awareness. However, these same studies stress that AI-driven systems remain vulnerable to adversarial manipulation, raising concerns about their reliability in high-stakes environments.

Third, bias, ethics, and black-box risks cut across the literature reviewed. For instance, several studies (Jenkins et al., 2025; Burton, 2023; Dong et al., 2024) have argued that the opacity of AI systems complicates accountability, while algorithmic biases may reinforce polarization or misclassify minority groups. This highlights the importance of clarity and transparency in AI deployments that touch on national security and civil liberties.

Fourth, the review revealed that AI is shaping the landscape of propaganda and information warfare. For instance, tools such as BERT-based classifiers (Rajendran et al., 2022) can identify extremist narratives. On the other hand, AI can also automate the production and dissemination of disinformation through bots, deepfakes, and tailored propaganda (Horowitz et al., 2018; Burton, 2023). Thus, this dual usage potential reinforces the need for governance frameworks that anticipate both beneficial and harmful applications.

Finally, the integration of AI in national security and defense contexts (Pemmasani, 2023; Horowitz et al., 2018) extends beyond counterterrorism into border control, surveillance, and military strategy. While such applications enhance operational readiness and situational awareness, they introduce ethical dilemmas around autonomous weapons, surveillance overreach, and democratic accountability. Thus, the current review reveals both the opportunities and risks of using AI to identify online radicalization pathways. Although technical advances have improved detection and monitoring, the broader implications on ethical, political, and strategic application demand equal attention.

SJIF Impact Factor (2025): 8.688| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 10 | Issue: 10 | October 2025 - Peer Reviewed Journal

5. CONCLUSION

This review mapped the growing but scattered literature on how AI is being used to understand and counter online radicalization pathways in the United States. The evidence suggests that AI can enhance the detection of radicalization, strengthen cybersecurity, and support national defense operations. At the same time, challenges related to bias, disinformation, black-box opacity, and the dual-use nature of AI highlight the urgent need for strategic monitoring and oversight.

AI operates simultaneously as a tool of prevention and an amplifier of radicalization dynamics. Accordingly, policymakers, practitioners, and scholars must strike a balance between technological advancement and ethical safeguards, transparency mechanisms, and effective regulatory oversight. Future work should not only refine AI models for accuracy but also integrate perspectives from law, security studies, and human rights to ensure their responsible and effective deployment.

6. LIMITATIONS

The relatively small pool of studies available reflects the early and fragmented nature of research on AI and online radicalization in the U.S. Additionally, the included studies varied significantly in design, data sources, and methodological rigor, which limited direct comparisons and synthesis. Additionally, most evidence came from experimental or conceptual work rather than large-scale, real-world applications, limiting the generalizability of the findings.

REFERENCES

1. *Aad, S. S., & Hardey, M. (2025). How GAI Works: Fundamentals. In After Generative AI (pp. 31-53). Emerald Publishing Limited.*
2. *Al-Zewairi, M., & Naymat, G. (2017). Spotting the Islamist Radical within: Religious Extremists Profiling in the United State. Procedia Computer Science, 113, 162-169.*
3. *Antoliš, K. (2025). ICT AND AI IN COMBATING TERRORISM. Public Security and Public Order, 37(1), 4-19.*
4. *Arksey, H., & O'malley, L. (2005). Scoping Studies: Towards a Methodological Framework. International Journal of Social Research Methodology, 8(1), 19-32.*
5. *Bouhlaoui, A. (2025). AI Agents for Counter-Extremism: Deployment Frameworks for Covert and Overt Digital Deradicalisation. Available at SSRN 5296073.*
6. *Burton, J. (2023). Algorithmic extremism? The securitization of artificial intelligence (AI) and its impact on radicalism, polarization and political violence. Technology in Society, 75, 102262.*
7. *Dong, B., Lee, J. R., Zhu, Z., & Srinivasan, B. (2024). Assessing large language models for online extremism research: identification, explanation, and new knowledge. arXiv preprint arXiv:2408.16749.*
8. *Horowitz, M. C., Allen, G. C., Saravalle, E., Cho, A., Frederick, K., & Scharre, P. (2022). Artificial Intelligence and International security.*
9. *Hussain, G. (2025). Artificial Intelligence (AI) and Radicalization. In Cyber Security in the Age of Artificial Intelligence and Autonomous Weapons (pp. 58-69). CRC Press.*
10. *Irfan, M., Almeshal, Z. A., & Anwar, M. (2025). Unleashing Transformative Potential of Artificial Intelligence (AI) in Countering Terrorism Online Radicalisation Extremism and Possible Recruitment.*
11. *Jenkins, R., Sullins, J. P., Kalu, O., Kamath, A., & Phumjam, K. (2025). Recent Insights in Responsible AI Development and Deployment in National Defense: A Review of Literature, 2022–2024. Journal of Military Ethics, 1-23.*
12. *Kaur, H. (2025). The Evolution of Terrorism in Digital Age: Cyber Jihad and Emerging Threats. International Journal of Multidisciplinary Education Research, 14(1), 3.*
13. *Matar, T. L. (2025). Mitigating the Threat of AI-Assisted Terrorism: Challenges and Counterterrorism Strategies (Doctoral dissertation, Master in International Relations, Strategy and Security, School of Social Science and Humanities, Neapolis University Pafos).*
14. *Meena, G., Raha, S., Selvakumar, P., Satyanarayana, P., & Vats, C. (2025). The Role of AI in Combatting Extremism and Radicalization on Social Media. In Ethical AI Solutions for Addressing Social Media Influence and Hate Speech (pp. 63-90). IGI Global Scientific Publishing.*
15. *O'Connor, D., Lewis Jr, R. H., Jacome Jr, T., Joseph, B., & Magnotti, L. J. (2025). Dissecting the Numbers: A Detailed Analysis of 23 Years of Mass Shootings across the United States. The American Journal of Surgery, 116516.*
16. *Patil, A. (2025). Ethical Considerations in AI Development: Safeguarding Human Rights and Privacy. Shodh Sagar Journal of Artificial Intelligence and Machine Learning, 2(1), 6-11.*
17. *Pemmasani, P. K. (2023). AI in National Security: Leveraging Machine Learning for Threat Intelligence and Response. The Computertech, 1-10.*
18. *Rajendran, A., Sahithi, V. S., Gupta, C., Yadav, M., Ahirrao, S., Kotecha, K., & Alhammad, S. M. (2022). Detecting Extremism on Twitter during US Capitol Riot Using Deep Learning Techniques. IEEE Access, 10, 133052-133077.*
19. *Suleiman, A. M. (2024). Enhancing the United States Counterterrorism Policy through Artificial Intelligence: A Comprehensive Analysis of Machine Learning Applications, Challenges, and Strategic Implications. International Journal of Scientific Research and Modern Technology, 3(5), 21–34.*

SJIF Impact Factor (2025): 8.688| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 10 | Issue: 10 | October 2025 - Peer Reviewed Journal

20. *Tariq, M. U. (2025). AI and the Ethical Crossroads: Navigating Privacy, Personhood, and Autonomy in the Digital Age. In Human Values, Ethics, and Dignity in the Age of Artificial Intelligence (pp. 49-72). IGI Global Scientific Publishing.*
21. *Zdjelar, I., Linning, S. J., Hart, M. B., & Davies, G. (2025). Using Crime Place Networks To Understand Terrorist Attacks: The Importance of in-Person and Online Crime-Involved Places. Journal of Policing, Intelligence and Counter Terrorism, 1-20.*