International Journal of Advanced Engineering and Management Research

Vol. 11, No. 02; 2026

ISSN: 2456-3676

AI-enabled Behavioral Threat Detection: Integrating Machine Learning with Cyberpsychological Indicators of Targeted Violence

Dr. Robb Shawe (Lead Author)
Dr. Robert W. Clark (Co-Author)
Capitol Technology University, Department of Sustainability, Department of Cyber-Psychology, 11301 Springfield Road, Laurel, MD 20708, USA

doi.org/10.51505/ijaemr.2026.11216 URL: http://dx.doi.org/10.51505/ijaemr.2026.11216

Received: Mar 09, 2026 Accepted: Mar 17, 2026 Online Published: Mar 24, 2026

Abstract
The increasing prevalence of digital communication has created unprecedented opportunities for identifying behavioral indicators associated with targeted violence. While traditional threat assessment frameworks rely heavily on observable behaviors and investigative intelligence, emerging advances in artificial intelligence (AI) and machine learning offer new capabilities for analyzing large volumes of digital behavioral data. This article explores integrating cyberpsychological indicators with machine learning to enhance early detection of potential threats in online environments. Drawing upon interdisciplinary literature from behavioral threat assessment, cyberpsychology, and computational intelligence, the article proposes a conceptual framework for AI-enabled behavioral threat detection. The framework illustrates how machine learning models may assist analysts in identifying patterns of grievance formation, digital identity reinforcement, and behavioral leakage within online discourse. The analysis highlights both the operational potential and the ethical considerations associated with AI-supported threat-detection systems. The article concludes by outlining future research directions and policy implications for integrating computational analytics into behavioral threat assessment practices.

Keywords: behavioral threat assessment, artificial intelligence, machine learning, cyberpsychology, targeted violence, digital behavioral analysis

1. Introduction

The rapid expansion of digital communication platforms has fundamentally altered the ways individuals express grievances, construct identities, and communicate intent. In recent years, researchers and practitioners have increasingly recognized that individuals who engage in targeted violence frequently exhibit observable behavioral indicators prior to an attack (Meloy et al., 2012; Silver et al., 2018). These indicators often manifest not only in offline environments but also across digital platforms where individuals communicate with peers, participate in ideological communities, and publicly express personal grievances.

Vol. 11, No. 02; 2026

ISSN: 2456-3676

Traditional threat assessment methodologies emphasize behavioral observation, investigative intelligence, and multidisciplinary threat assessment teams (Calhoun & Weston, 2015; Fein & Vossekuil, 1998). However, the scale and velocity of online communication present significant challenges for human analysts attempting to monitor large volumes of digital information. Consequently, emerging research has explored the potential role of artificial intelligence (AI) and machine learning technologies in identifying behavioral signals within digital environments (Ferrara et al., 2016; Olteanu et al., 2018).

Machine learning systems are increasingly capable of detecting linguistic patterns, sentiment indicators, and behavioral signals within large-scale datasets through automated pattern recognition and predictive modeling techniques (Jordan & Mitchell, 2015; Ferrara et al., 2016). When integrated with established behavioral threat assessment frameworks, these technologies may provide analysts with additional tools for identifying concerning behavioral trajectories prior to violent incidents.

This article forms part of a broader program of research examining digitally mediated behavioral threats and the evolving analytical frameworks required to identify, interpret, and responsibly govern them. The research program investigates the progression of online behavioral indicators, the mechanisms through which grievance expression may escalate into behavioral leakage, and the role of digital intelligence methods and artificial intelligence–enabled analytics in supporting early threat detection. Complementing these analytical dimensions, the program also explores the ethical and governance considerations necessary to ensure that emerging detection capabilities are implemented responsibly and in alignment with democratic norms and civil liberties. Collectively, this body of work seeks to advance an integrated interdisciplinary framework for understanding and managing digitally mediated threat environments.

Moreover, this article explores the integration of machine learning techniques with cyberpsychological indicators of targeted violence. Specifically, it proposes a conceptual framework for AI-driven behavioral threat detection that synthesizes insights from behavioral threat assessment, cyberpsychology, and computational social sciences. The study expands upon existing foundational work examining cyberpsychological indicators of targeted violence, pathways of behavioral leakage, and social media intelligence frameworks for early risk identification. Collectively, this research promotes an interdisciplinary approach that consolidates behavioral threat assessment, cyberpsychology, and computational analytics to enhance the detection of emerging threats within digital environments.

Furthermore, the article contributes to the scholarly literature by developing an interdisciplinary conceptual framework that integrates behavioral threat assessment, cyberpsychology, and machine-learning analytics to improve the early detection of digital behavioral indicators associated with targeted violence. This publication is part of an ongoing research initiative that explores digitally mediated behavioral threats, including cyberpsychological indicators,

Vol. 11, No. 02; 2026

ISSN: 2456-3676

pathways of behavioral leakage, AI-enabled threat detection systems, and ethical governance frameworks.

2. Literature Review

Behavioral Threat Assessment

Behavioral threat assessment frameworks have emerged as one of the most widely accepted approaches for preventing targeted violence. Rather than attempting to profile potential offenders based on demographic characteristics, modern threat assessment models focus on identifying concerning behaviors and contextual indicators that may signal a progression toward violence (Borum et al., 1999; Meloy & O'Toole, 2011).

Research examining pre-incident behaviors has consistently demonstrated that individuals who commit targeted violence frequently communicate their intentions or grievances prior to an attack. This phenomenon, commonly referred to as behavioral leakage, represents a critical component of contemporary threat assessment practices (Meloy et al., 2012; Silver et al., 2018).

Cyberpsychology and Digital Behavior

The emergence of online communication environments has introduced new dimensions to the study of behavioral indicators associated with violence. Cyberpsychology research suggests that digital environments may facilitate increased emotional expression, identity experimentation, and social reinforcement processes (Suler, 2004; Weimann, 2016). These dynamics may influence how individuals communicate grievances and interact with ideological communities.
Online disinhibition effects, combined with algorithmically driven information environments, may amplify expressions of grievance and hostility. In some cases, digital communities can reinforce narratives that legitimize violence or normalize extreme viewpoints (Conway, 2017; Koehler, 2014).

Artificial Intelligence and Digital Threat Detection

Artificial intelligence technologies have increasingly been applied to analyze large-scale patterns in digital communication. Machine learning techniques such as natural language processing, probabilistic topic modeling, and sentiment analysis allow researchers to detect linguistic indicators and emerging behavioral patterns across large datasets (Blei, 2012; Ferrara et al., 2016). These computational approaches enable analysts to identify thematic clusters, ideological narratives, and shifts in sentiment that may signal escalating grievance expression or behavioral leakage within online discourse.

While these technologies cannot predict violent behavior with certainty, they may help analysts identify patterns that warrant further investigation. When combined with expert interpretation and contextual analysis, machine learning tools may enhance situational awareness within threat assessment processes.

Vol. 11, No. 02; 2026

ISSN: 2456-3676

3. Conceptual Framework

The conceptual framework proposed in this article integrates cyberpsychological indicators with machine-learning analysis to support behavioral threat detection in digital environments. The framework conceptualizes threat development as a progression across three interconnected stages.

First, individuals may experience grievance formation, characterized by perceived injustices, personal failures, or ideological conflicts. These grievances often become visible through online expressions of frustration, resentment, or hostility.

Second, digital identity reinforcement may occur as individuals seek validation or support within online communities. Through repeated interaction with ideologically aligned networks, individuals may reinforce narratives that legitimize their grievances or encourage adversarial worldviews.

Third, behavioral leakage may emerge when individuals communicate violent fantasies, threats, or operational intentions within digital environments. These expressions may appear in the form of direct threats, symbolic language, or references to violent actors and events.

Machine learning systems may assist analysts in identifying linguistic patterns, sentiment shifts, and behavioral signals associated with these stages. By analyzing large volumes of digital communication data, computational models may highlight patterns that warrant further review by trained threat assessment professionals.

4. Methodological Orientation

This article adopts a conceptual and interdisciplinary methodological orientation. Rather than conducting empirical data collection, the study synthesizes insights from behavioral threat assessment research, cyberpsychology literature, and computational social science.
Conceptual analysis enables researchers to integrate theoretical perspectives from multiple disciplines to develop new frameworks for understanding complex phenomena (Rocque & Duwe, 2018). In the context of targeted violence prevention, interdisciplinary synthesis is particularly valuable because behavioral indicators often emerge across psychological, social, and technological domains.

Interdisciplinary synthesis approaches are particularly valuable when examining digitally mediated behavioral phenomena that span psychological, technological, and social domains.

5. Analysis and Discussion

Integrating machine learning technologies with behavioral threat assessment frameworks presents several potential advantages. Computational systems may assist analysts in detecting patterns across large digital datasets that would otherwise be difficult to analyze manually. For

Vol. 11, No. 02; 2026

ISSN: 2456-3676

example, natural language processing algorithms may identify linguistic indicators associated with grievance expression, hostility, or ideological alignment.

However, the application of AI technologies to threat detection also raises important ethical and operational considerations, particularly regarding data privacy, algorithmic bias, and the governance of large-scale behavioral data systems (Boyd & Crawford, 2012). False positives, algorithmic bias, and privacy concerns must be carefully addressed when developing AI-supported monitoring systems. Human oversight remains essential to ensure that computational findings are interpreted within appropriate contextual frameworks.

Ultimately, AI technologies should be viewed as analytical support tools rather than autonomous decision-making systems. When integrated responsibly within multidisciplinary threat assessment processes, machine learning may enhance analysts' ability to identify concerning behavioral trajectories within digital environments.

Scholars examining big data governance emphasize that algorithmic decision-support systems must operate within transparent accountability frameworks to prevent misuse, amplification of bias, or unintended violations of civil liberties (Boyd & Crawford, 2012).

6. Policy and Operational Implications

The integration of AI technologies into behavioral threat assessment practices has significant implications for policymakers and security practitioners alike. Government agencies and security organizations may consider investing in analytical platforms that combine computational analysis with expert interpretation.

Operationally, multidisciplinary threat assessment teams may benefit from incorporating data science expertise into their analytic processes. Collaboration between behavioral analysts, cybersecurity professionals, and machine learning specialists may improve the effectiveness of digital threat detection initiatives.

At the policy level, clear governance frameworks are necessary to ensure that AI-supported threat detection systems operate within legal and ethical boundaries. Transparency, accountability, and oversight mechanisms are critical for maintaining public trust while leveraging technological capabilities to prevent violence.

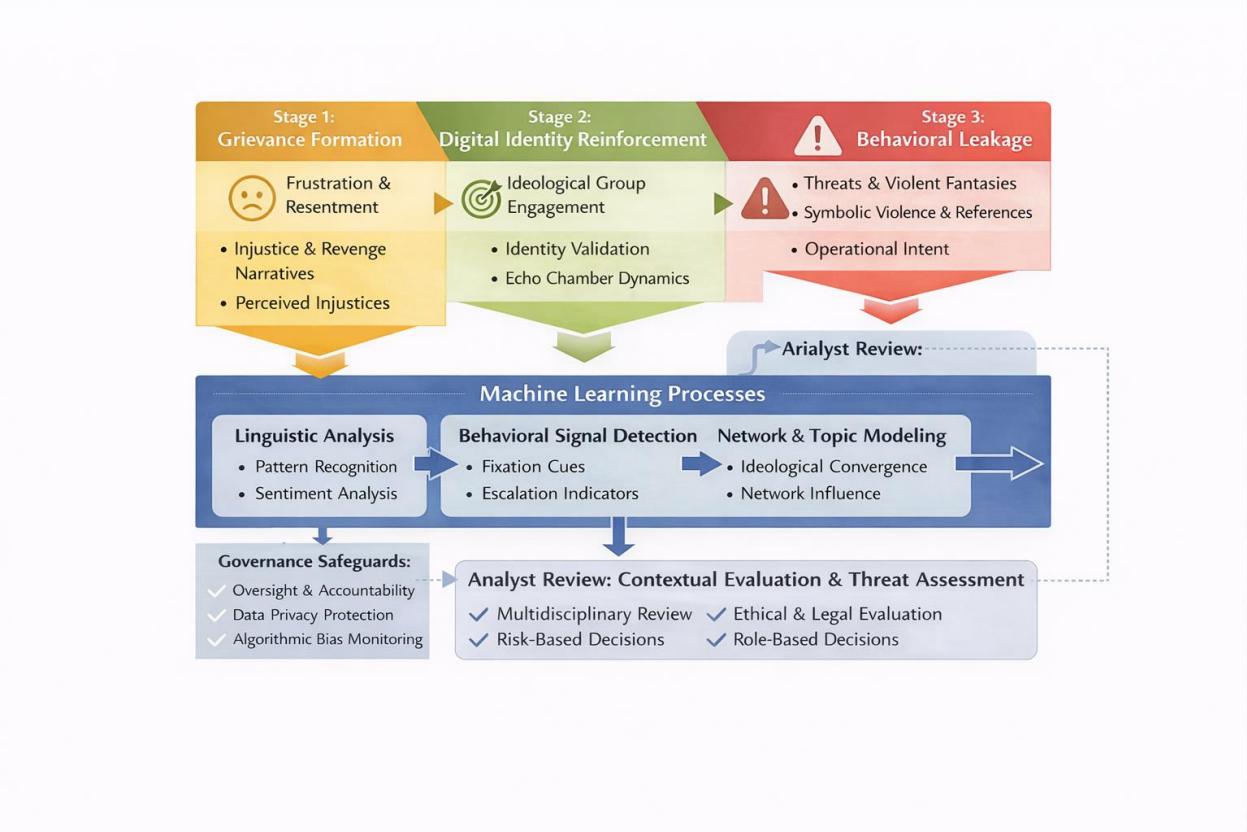
To extend the operationalization of cyberpsychological indicators through computational analytics and demonstrate how machine learning can augment structured threat assessment processes, the following conceptual model illustrates the integration of AI-enabled analysis with behavioral threat detection frameworks. The conceptual model presented in Figure 1 illustrates how artificial intelligence–enabled analytical processes can assist in identifying behavioral indicators associated with grievance formation, digital identity reinforcement, and behavioral leakage within digital environments.

Vol. 11, No. 02; 2026

ISSN: 2456-3676

Figure 1

AI-Enabled Behavioral Threat Detection Model



Note. The model illustrates how machine learning processes can assist analysts in detecting cyberpsychological indicators associated with grievance formation, identity reinforcement, and behavioral leakage. Analyst review ensures contextual interpretation, ethical evaluation, and risk-informed threat assessment.

Taken together, the model demonstrates that AI-enabled analytical systems can enhance the detection of cyberpsychological threat indicators by identifying large-scale behavioral patterns, but their effectiveness depends on integration with structured threat assessment frameworks, human analytic oversight, and governance safeguards that ensure ethical interpretation and responsible decision-making.

7. Limitations and Future Research

This article presents a conceptual framework rather than an empirically validated model. Future empirical validation of AI-enabled behavioral threat detection models will require carefully governed datasets, interdisciplinary collaboration between behavioral scientists and computational researchers, and robust ethical oversight mechanisms. Consequently, the proposed

Vol. 11, No. 02; 2026

ISSN: 2456-3676

integration of cyberpsychological indicators and machine learning techniques requires empirical testing. Future research may examine how computational models perform on real-world datasets containing digital behavioral indicators associated with targeted violence.
Additionally, interdisciplinary collaboration between behavioral scientists, computer scientists, and threat assessment practitioners will be essential for refining AI-supported analytical methods. Further research may explore how emerging technologies, such as deep learning and network analysis, can enhance behavioral threat-detection capabilities.

8. Conclusion

The increasing digitization of social interaction has transformed the landscape of behavioral threat detection. Individuals who engage in targeted violence frequently communicate grievances, identities, and intentions within digital environments prior to an attack. Integrating insights from cyberpsychology with advances in artificial intelligence may offer new opportunities to identify these behavioral indicators at earlier stages.

The conceptual framework proposed in this article illustrates how machine learning technologies may support analysts in identifying patterns associated with grievance formation, digital identity reinforcement, and behavioral leakage. While technological tools cannot replace expert judgment, they may enhance situational awareness within multidisciplinary threat assessment processes.

Continued interdisciplinary research will be essential for ensuring that AI-enabled behavioral threat detection systems are both effective and ethically responsible. Collectively, this research series establishes a foundational interdisciplinary framework for integrating cyberpsychological dynamics, behavioral leakage patterns, and computational analytics to enhance the early detection of targeted violence in digitally mediated environments.

Authorship Statement

Dr. Robb Shawe and Dr. Robert W. Clark jointly conceptualized the research framework and analytical orientation of this manuscript. Dr. Clark contributed practitioner expertise derived from federal law enforcement leadership and public safety governance experience. Dr. Shawe contributed a scholarly synthesis across cyberpsychology, critical infrastructure protection, and public safety governance, including the integration of literature and manuscript preparation. Both authors reviewed and approved the final manuscript.

Author Note and Copyright Statement

Dr. Robb Shawe, PhD, is a scholar-practitioner, U.S. Navy veteran, and Chief Executive Officer of New York Security Consulting Professionals L.L.C. His research focuses on critical infrastructure protection, cybersecurity governance, and resilience systems. Dr. Shawe served as the lead author and was primarily responsible for conceptualization, development of the analytical framework, and manuscript preparation.

Vol. 11, No. 02; 2026

ISSN: 2456-3676

Dr. Robert W. Clark, PhD, currently serves as Deputy Mayor of Public Safety for the City of Los Angeles and previously served as an Assistant Special Agent in Charge with the Federal Bureau of Investigation. His research focuses on mass violence prevention, behavioral threat assessment, and the intersection of cyberpsychology and public safety intelligence. Dr. Clark contributed to the literature synthesis, conceptual refinement, and collaborative development of the research framework. Both authors participated in reviewing and approving the final version of the manuscript.

The authors declare no conflicts of interest related to this research. The manuscript represents original scholarly work and is not currently under consideration by another publication outlet. Copyright will remain with the authors unless transferred to a publisher upon acceptance for publication.

Copyright Notice

References

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77–84.

Borum, R., Fein, R., Vossekuil, B., & Berglund, J. (1999). Threat assessment: Defining an approach for evaluating risk of targeted violence. *Behavioral Sciences & the Law, 17*(3), 323–337.

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society, 15*(5), 662–679.

Calhoun, F. S., & Weston, S. W. (2015). *Threat assessment and management strategies: Identifying the howlers and hunters.* CRC Press.

Clark, R. W. (2025). *Mass shootings in America: A law enforcement response* (Unpublished doctoral dissertation). Capitol Technology University.

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.

Fein, R. A., & Vossekuil, B. (1998). Preventing assassination: The Secret Service exceptional case study project. *Annals of the American Academy of Political and Social Science, 576*(1), 62–74.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96–104.

Fox, J. A., & Levin, J. (2015). *Extreme killing: Understanding serial and mass murder* (3rd ed.). Sage Publications.

Vol. 11, No. 02; 2026

ISSN: 2456-3676

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260.

Koehler, D. (2014). The radical online: Individual radicalization processes and the role of the internet. *Journal for Deradicalization, 1*, 116–134.

Meloy, J. R., & O'Toole, M. E. (2011). The concept of leakage in threat assessment. *Behavioral Sciences & the Law, 29*(4), 513–527.

Meloy, J. R., Hoffmann, J., Guldimann, A., & James, D. (2012). The role of warning behaviors in threat assessment: An exploration and suggested typology. *Behavioral Sciences & the Law, 30*(3), 256–279.

Olteanu, A., Castillo, C., Diaz, F., & Kıcıman, E. (2018). Social data: Biases, methodological pitfalls, and ethical boundaries. *Frontiers in Big Data, 1*, 13.

Rocque, M., & Duwe, G. (2018). The patterns and correlates of mass murder in the United States. *Homicide Studies, 22*(2), 131–145.

Shawe, R. (2025). *Exploring smart grid technologies' impact on sustainable energy management in New York State* (Publication No. 32284053) [Doctoral dissertation, Capitol Technology University]. ProQuest Dissertations & Theses Global.

Silver, J., Simons, A., & Craun, S. (2018). *A study of the pre-attack behaviors of active shooters in the United States between 2000 and 2013.* Federal Bureau of Investigation.

Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior, 7*(3), 321–326.

Weimann, G. (2016). *Terrorism in cyberspace: The next generation.* Columbia University Press.