LEXARA
Journal of Law and Legal Studies
Lexara: Journal of Law and Legal Studies
Vol. 1, No. 1, 2026
E-ISSN: XXXX-XXXX
punggawajournal.com/lexara

From Propaganda to Synthetic Operations: Generative AI and the Transformation of Cyberterrorism

Nadiah Khaeriah Kadir[1]
[1] Fakultas Hukum, Universitas Hasanuddin, Indonesia
Correspondence: nadiahkhaeriah@unhas.ac.id

Article Info

Article history:
Received July 10th, 2026
Revised July 25th, 2026
Accepted July 29th, 2026

Keyword:
Cyberterrorism; Generative AI; Propaganda; Synthetic Operations.

ABSTRACT

Existing studies largely treated generative AI in cyberterrorism as a tool for synthetic propaganda, leaving unclear when its outputs began to support operational preparation and how existing law could respond. This study examined the movement from AI-assisted propaganda toward synthetic operations and assessed its legal implications. It used normative legal research with conceptual and statutory approaches, supported by scholarly literature, open-source documents, and European Union law. The findings indicated that documented terrorist use remained concentrated in content production, translation, adaptation, and reuse. Generative AI nevertheless acquired operational relevance when outputs could be applied with limited correction and connected to later tasks in preparing or carrying out terrorist activity. Open evidence did not establish autonomous or systematic AI-driven terrorist workflows; synthetic operations therefore remained an analytical threshold rather than a settled practice. The synthetic origin of content also did not determine legal responsibility. Criminal attribution continued to depend on human intent, use, and connection to prohibited conduct, while regulatory duties varied according to the role and capacity of model providers, applications, hosting services, and online platforms. Technology-neutral rules remained relevant, although proof, attribution, and content-focused regulation became more difficult when assistance was distributed across services and occurred before public dissemination.

INTRODUCTION

Digital technology was long treated in cyberterrorism scholarship mainly as a means of distributing propaganda. Online networks also allowed communication to continue when organizations were under physical pressure. Generative artificial intelligence has gradually altered that position because digital systems no longer merely carry messages already formulated by human actors. Generative models can read source material, compose outputs, and prepare responses that are used again at a later stage of the same workflow. AI is beginning to function as a layer of production and information processing rather than simply as a channel. A change at one stage may affect subsequent work without removing human control. This development does not mean that every use of AI by extremist groups automatically constitutes cyberterrorism. The term is used here for digital activities that provide material support to terrorist purposes or conduct, including cyber-enabled terrorism, while excluding every radical expression found online (Kadir, 2025a). This boundary matters because synthetic propaganda and cyber operations do not carry the same legal weight. The word "from" in the title likewise does not describe a linear replacement. Propaganda remains the most visible use, while the functions of AI are beginning to extend into other areas (Winter et al., 2020; Macdonald et al., 2022).

AI use is still often judged through the final artifact. Deepfakes and automated content make the technology visible, but the origin or visual quality of a piece of material does not explain its role in a sequence of conduct. A sophisticated video may remain symbolic, while a simple text tailored to a particular recipient may be more useful in maintaining communication. Artifact-centered analysis is therefore too narrow. It does not show who reviewed the output, how much correction remained, or whether the result was reused in later work. Operational value depends on context, the user's competence, and the resources available. A simple output may be immediately useful to an actor who can verify it, whereas a more elaborate answer may remain general information. The relevant question is therefore not how synthetic a piece of content appears, but what work it actually performs before and after publication (GIFCT, 2025; Gregory, 2022; Vaccari & Chadwick, 2020).

The open evidence is not equal in weight. Monitoring reports have actor validity because they show that terrorist or violent extremist environments have used AI, mainly to produce and adapt communication materials (Europol, 2025). Tech Against Terrorism has reported thousands of AI-assisted items in such spaces. That number does not represent an equal number of operations or actors: one channel may generate many items, and the origin of each item cannot always be established (Tech Against Terrorism, 2023). Technical studies proceed from another direction. The two bodies of evidence answer different questions. Some benchmarks show that language-model agents can complete parts of cybersecurity tasks when they receive sufficient information and tool access. Such studies offer capability validity, but they do not establish terrorist adoption. Laboratory testing shows that a function may be performed. Direct-use reports may provide stronger attribution but little operational depth. This distinction is often lost, allowing an experimental capability to be read as an established practice (Conway & Macdonald, 2021).

Synthetic operations refer here to terrorist activity in which AI assistance does not end with the production of a single artifact but enters a workflow and reduces the work needed to produce a later action. The threshold lies in actionability and workflow integration. An output must be transferable to another stage with limited revision and then used within a broader process. Full autonomy is unnecessary. Human actors continue to determine aims and important decisions, while AI assists particular parts of the work. Under this definition, an AI-generated poster is not necessarily a synthetic operation. A modest form of assistance may matter more when it is repeatedly used and connected to other tasks.

The literature has not adequately explained when this transition occurs. Propaganda research generally documents growth in the amount or quality of content but rarely asks whether the output can actually be used at an operational stage. Technical studies measure model capabilities without a sufficient basis for connecting them to patterns of terrorist use. Governance research still concentrates on visible material found on platforms, even though AI assistance may occur in private exchanges or document processing that is difficult to observe and rarely leaves a public trace. Technical possibility can therefore be mistaken for an event, while gradual change may be overlooked because it does not resemble a readily identifiable autonomous attack (Unlu & Yilmaz, 2025).

The urgency of this inquiry lies in a growing mismatch: terrorist use of generative AI is increasingly visible in communication, while its significance for operational preparation and legal responsibility remains uncertain (Baele et al., 2025; Herath & Whittaker, 2023). Treating every synthetic artifact as operational support would make cyberterrorism too broad, yet limiting analysis to content production may overlook assistance that reduces work across a terrorist workflow. This study therefore aims to identify when AI-assisted propaganda begins to acquire operational significance, determine whether verifiable evidence supports the emergence of synthetic operations, and assess how existing European Union law addresses criminal attribution and provider duties in that setting. Direct-use evidence is kept separate from demonstrations of technical capability and governance materials. The article does not presume that a technical capability proves terrorist adoption or that a new technology creates a legal vacuum; its central claim is that AI has so far altered the division of labor more clearly than it has displaced human control.

RESEARCH METHODS

This study is normative legal research using statutory and conceptual approaches. The statutory approach examines legal rules concerning terrorism, artificial intelligence, digital services, and the responsibilities of platform operators. The conceptual approach is used to examine cyberterrorism,

synthetic propaganda, and synthetic operations, including the boundary between AI used to generate content and AI that begins to support operational activity. The units of analysis are legal norms, legal concepts, and documented uses of generative AI in terrorism-related activity drawn from open sources (Negara, 2023).

Research materials were collected through library research. Primary legal materials consist of relevant legislation and international instruments. Secondary materials include books, scholarly articles, and research on cyberterrorism, terrorism, artificial intelligence, and platform governance. Official counterterrorism reports, policy documents, and technical studies of AI capabilities are used as supporting materials. Legal sources were selected according to relevance and continuing validity. Reports on AI use and technical research were concentrated on publications issued between January 2023 and July 2026, when generative AI services became widely available. A source was used when its full text was accessible and its informational basis or method could be examined (van Gestel, 2023).

The materials were analyzed qualitatively. Legal norms were interpreted, positions in the literature were compared, and both were related to documented uses of AI. The analysis distinguishes observed terrorist use from capabilities demonstrated only through research or testing. It also distinguishes AI used for propaganda from AI that begins to assist operational activity. The findings are used to assess whether existing legal concepts and rules can adequately explain and respond to these developments. The conclusions are confined to uses traceable through open sources and do not measure their prevalence or causal effect on the number of terrorist attacks (Taekema, 2021).

RESULTS AND DISCUSSION

1. From Synthetic Content to Adaptive Propaganda Practices

Propaganda produced with AI does not automatically become cyberterrorism. It acquires legal relevance only when it has a sufficiently clear relationship with terrorist activity, whether through direct encouragement of violence, recruitment, or support for preparatory conduct. This distinction matters because institutional reports often employ the broader category of violent extremist content, which is not identical to terrorist propaganda in law. A poster circulating in a sympathizer channel may reveal ideological approval without proving organizational involvement or an operational purpose. Conversely, a simple message may carry greater weight when it is designed to maintain contact with a prospective recruit or is reused in later communication. Assessment cannot stop at radical content or the synthetic origin of the material. Its relationship to the actor, the audience, and the function for which it is used must also be examined. AI primarily changes the production and adjustment of material; it does not by itself convert all propaganda into operational support. Legal classification continues to depend on human conduct and a concrete relationship between the material, the actor, and the violence being supported (KhosraviNik & Amer, 2022).

Open documentation since 2023 has most often recorded AI use in the production of propaganda, speech transcription, and translation (Ganaie, 2026). Tech Against Terrorism identified posters likely generated by image models, automated transcriptions of messages by Islamic State leaders, and guides to AI services in supportive environments. It also archived more than 5,000 AI-assisted items. That figure cannot be read as 5,000 operations or 5,000 actors. A single channel can produce many items, while the origin and organizational connection of each one cannot always be established. The evidence therefore indicates the breadth of experimentation rather than settled organizational adoption (Tech Against Terrorism, 2023).

The production of variants is the most readily observable change. A base text can be rewritten for another style or platform with far less time and effort, allowing one message to circulate in several forms. Variation, however, is not the same as adaptation. Adaptation requires some knowledge of the audience or its earlier responses, whereas most reports establish only that new versions can be produced (Kadir, 2025b). Europol recorded greater use of generative AI for propaganda and hate speech, particularly in far-right extremist environments, but did not show that each item was built from a specific recipient profile (Europol, 2025). More content also says little about persuasion. Rapid production can yield awkward or socially misplaced material, while even technically polished output may fail when the source lacks credibility. Trust still depends on social relations and the standing of the messenger (Macdonald & Lorenzo-Dus, 2021). Open evidence therefore does not support the stronger claim that AI-generated propaganda is more effective in recruitment or radicalization (Nieweglowska et al., 2023).

Translation offers a more concrete advantage. Automated transcription can shorten the initial work needed before a message is rendered into another language, while generative models can produce a more readable version. This function matters to networks that rely on sympathizers to transcribe and translate material. Yet fluent translation does not make a message legitimate or convincing to a new audience. Religious terminology, for example, can acquire different meanings across communities, and wording that is grammatically correct may remain socially misplaced. JCAT identified translation, text-to-speech, and voice cloning as uses that warrant attention, but it did not establish that cultural barriers had disappeared. AI can reduce language work; the legitimacy of a message still depends on human actors and the network that circulates it (JCAT, 2024).

Recreating material can also complicate moderation, although it does not necessarily make content undetectable. Minor alterations may still be recognized through perceptual hashing or other classification systems. Greater difficulty arises when a text is substantially rewritten, an image is generated anew, or its relationship with earlier material can be understood only through a wider social context. Platforms must then determine whether the new item still carries the same prohibited message. Tech Against Terrorism has described variant recycling, while Europol has emphasized the speed with which material can be edited and redistributed. These findings support the possibility of a heavier review burden, not the conclusion that moderation has failed. The legal question concerns function and context rather than file similarity alone (Gorwa et al., 2020; He et al., 2024).

Chatbots may reduce the work needed to answer repeated questions, but sustained terrorist use of such systems has not been widely documented. GIFCT treats the prospect mainly as an emerging risk, and the ability to maintain a conversation should not be confused with evidence of lasting changes in belief or behavior. Radicalization still depends on social and personal conditions that cannot be reduced to interaction with a model (GIFCT, 2025; Kunst et al., 2026). The stronger finding is narrower: production, translation, reuse, and some forms of targeted communication have become easier. That does not yet amount to an adaptive influence infrastructure. A single poster, translation, or chatbot reply is insufficient to constitute a synthetic operation. The boundary begins to shift only when an AI result is reused to interpret information, prepare a decision, or assist another task beyond publication. At that point, analysis must follow the sequence of work rather than the amount of propaganda produced.

The analytical boundary developed in this section is summarized in Table 1.

Table 1 Analytical Boundary Between Synthetic Propaganda and Synthetic Operations

Analytical dimension	Synthetic propaganda	Synthetic operations
Primary function	Creates, translates, adapts, or redistributes persuasive material.	Supports a later task connected to terrorist preparation or execution.
Use of the output	The result usually remains a communicative artifact.	The result is transferred into another stage of work.
Analytical threshold	Content production or variation alone.	Actionability combined with workflow integration.
Current evidence	The most clearly documented form of terrorist use since 2023.	Fragmented and emerging; no systematic autonomous workflow has been established.
Main legal question	Meaning, context, audience, and dissemination of the message.	Human intent, concrete use, and connection to prohibited conduct.

2. Synthetic Operations: The Boundary between Technical Assistance and Operational Support

Highly detailed information does not necessarily assist the preparation of terrorism. Its value can be assessed only after considering how much work remains before an AI output can be used. A document summary, an explanation of terminology, or technical advice may do no more than help a user understand an issue. Assistance changes character when the result is tailored to a purpose, can be applied with limited revision, and is actually placed within conduct directed toward preparation or execution. This distinction is important in normative legal analysis. Publicly available information does

not become prohibited assistance merely because it is sought by a person with extremist views. A connection with conduct must be established. Knowledge of the user's purpose may strengthen that connection, particularly when an output is highly specific and later used in a particular plan. These conditions will not always occur together, so legal assessment cannot be built from the content of the answer alone. Synthetic operations are not treated here as an established terrorist practice. The term is used as an analytical threshold for identifying situations in which AI no longer merely provides information but begins to save work required at a preparatory stage (Gaudette et al., 2022).

JCAT reported that foreign terrorist organizations and their supporters had shown interest in generative AI, published guidance, and held training activities. Up to April 2024, the clearest use remained content production and dissemination. Programming assistance, instructional chatbots, and cyberattacks were described as possible uses for planning or training, not as a pattern of operations already demonstrated at scale (JCAT, 2024). Those terms cannot be flattened into one category. Interest indicates a direction of attention, while training reflects an effort to acquire capability. Neither establishes that AI output was used in preparing a specific attack.

Technical research shows that language-model agents can complete parts of cybersecurity tasks, but success is highly sensitive to testing conditions. Fang et al. found sharply different results depending on whether the agent received a CVE description (Fang et al., 2024), while CVE-Bench reported lower success on web-application vulnerabilities in a more constrained environment (Zhu et al., 2025). The percentages are not directly comparable because the tasks, tools, success criteria, and available assistance differ; both studies were also preprints when used here. Their value is therefore limited but useful. Initial information and tool access shape performance, and humans still have to recognize erroneous output (Ferrag et al., 2025). Laboratory evidence shows what a system may do under specified conditions, whereas counterterrorism reporting shows what has actually been observed among actors (Ibrar et al., 2025). Combining the two without restraint turns technical possibility into an event. Operational support must remain tied to a traceable terrorist use rather than inferred from the sophistication of the model alone.

Model error also affects the value of assistance. A user who understands the task is more likely to recognize and correct a weak answer, whereas a user without prior knowledge may accept advice that cannot be applied. The sources examined do not establish who obtains the greatest benefit. Conceptually, however, the value of AI must be compared with the checking work that remains with the human user. Speed is of little consequence if every step must be tested again. AI may save time at one point and add work at another. An operational advantage is plausible only when the saving survives the need for correction and additional data (Brynjolfsson et al., 2025; Ji et al., 2023; Noy & Zhang, 2023).

The CT-AI Benchmark tested 27 models with almost 2,500 prompts drawn from terrorist and violent extremist misuse cases. Approximately one third of the responses were rated as providing useful uplift beyond information readily available through ordinary web search. That rating reflects the benchmark's criteria; it is not evidence that the material was used in an operation. Some responses that began with a refusal still contained information assessed as usable. The test, however, was limited to single-turn prompts in English and did not reproduce long conversations, multilingual use, or connections to external tools (Tech Against Terrorism, 2026).

One AI answer is rarely enough to establish a synthetic operation. Greater significance arises when a result is used at a later stage and the link between those stages can be demonstrated. A model may help a person understand a document; after human review, that result may be used to prepare other material. Open evidence through July 2026 still shows isolated uses more often than complete terrorist workflows. Integration must therefore be proved rather than assumed. Human decision-making also remains central. A model does not independently form a terrorist purpose in the legal sense, even when parts of the work are delegated or accelerated. Responsibility may nevertheless extend beyond the immediate user because providers and platforms possess different degrees of knowledge and control. Those differences matter for attribution, but they do not remove the need to prove intent and a concrete relationship between assistance and conduct (Puranam, 2021).

Open evidence currently shows fragmented operational support rather than a synthetic workflow systematically used by terrorist organizations. Synthetic operations remain useful as a conceptual boundary when applied strictly: an AI output must be usable with limited revision, placed within another sequence of work, and demonstrably related to the preparation or execution of terrorism.

Technical capability alone does not meet that threshold. Information seeking unconnected to a particular plan cannot be treated as an operational contribution. The legal issue is not whether an artifact was made with AI, but what human actors did with the result and how closely its function related to prohibited conduct. This boundary keeps the analysis focused on acts rather than identity or belief alone.

3. Governing Synthetic Operations in the European Union: Criminal Attribution, Provider Duties, and the Limits of Content Regulation

The fact that material was produced with AI does not determine who may be held responsible for its use. European Union law is used here as the selected normative framework rather than as a description of global law because it contains criminal-law rules on terrorism, a special regime for terrorist content online, digital-service duties, and AI regulation that can be read together. This choice also keeps the limits of the analysis visible: its doctrinal conclusions concern the structure of EU law, while any relevance to other jurisdictions remains conceptual. The four instruments do not regulate the same matter. Directive (EU) 2017/541 addresses terrorist offences and related conduct; Regulation (EU) 2021/784 governs the removal of terrorist content disseminated to the public through hosting services. The Digital Services Act establishes due-diligence duties for intermediary services, while the AI Act imposes specified duties on AI operators. Read together, these instruments expose a problem that is missed when analysis merely searches statutes for the words generative AI. An AI result may move from a private exchange into preparation and later reappear as different content or conduct. Law must assess the role of that result in a sequence of acts, not only the way the material was made (Fernández-Llorca et al., 2025).

Directive 2017/541 remains centered on human conduct and fault. The offences in Articles 5 to 8 contain different elements, and the use of AI does not collapse those distinctions. A technical answer becomes relevant to criminal liability only when it can be connected to regulated conduct and a terrorist purpose that can be proved. The recitals also direct attention to context and resulting danger when public provocation is assessed. Radical or detailed content is therefore insufficient; the prosecution must still establish how the assistance was requested and used (Lowe, 2022). The absence of an offence that expressly names generative AI does not itself create a legal vacuum. Technology-neutral rules may still apply when the relevant elements are satisfied. Difficulty arises when the link between an AI result and a particular plan cannot be proved, or when a provider of general capability is treated as if it shared the user's knowledge. A new prohibition should respond to such a concrete gap rather than to the novelty of the tool alone (Watkin, 2023).

System records, watermarks, or metadata may help identify the service used, but technical evidence will not always identify the responsible person. An AI result may be edited by another actor or passed through an application that does not preserve the initial context. Even when technical origin can be established, that evidence does not reveal who selected the result, understood its purpose, or decided to use it. Criminal attribution requires a relationship between the person, the relevant mental element, and concrete use. Knowledge and control are useful analytical tools, although they are not a single legal test that summarizes the whole of EU law. Knowledge may indicate whether a person understood the purpose of the assistance. Control helps identify who could direct or stop the use. Technical position remains relevant but is not enough by itself to prove criminal fault. A model provider may control safety design without knowing a particular request, while the user may understand the purpose without controlling the underlying system. An application provider between them may see more context but have limited authority over the base model. This difference explains why a user's criminal liability must not be confused with the regulatory duties of service providers. They arise from different legal bases and standards of proof (Dathathri et al., 2024; Teichmann, 2026).

Regulation 2021/784, the Digital Services Act, and the AI Act operate at different layers. Regulation 2021/784 addresses terrorist content disseminated to the public through hosting services, while the DSA supplies a horizontal framework for intermediary services and prohibits general monitoring obligations. Its systemic-risk duties for very large platforms and search engines do not turn every private use of AI into removable content. This matters because an AI system may be used to read or prepare material without producing anything that enters the public-content regime (Zornetta & Pohland, 2022; Fasel & Weerts, 2024). The AI Act approaches the problem through provider duties. Articles 53 and 55 impose documentation, evaluation, risk-mitigation, incident-reporting, and cybersecurity obligations for relevant general-purpose AI models. These preventive duties do not make

a provider criminally responsible for every terrorist misuse; they allocate obligations according to role and risk. Criminal participation still requires a separate assessment of knowledge and conduct (Schuett, 2024).

Lawful use also limits intervention. Article 1(3) of Regulation 2021/784 excludes material disseminated for educational, journalistic, artistic, research, or counterterrorism purposes from the definition of terrorist content once its actual purpose is assessed. Recital 40 of Directive 2017/541 likewise keeps radical or controversial views in public debate outside public provocation and protects legitimate research from being treated as terrorist training. Sensitive keywords or technical requests are therefore an inadequate sole basis for restriction. Human review remains a reasonable safeguard when a decision may terminate access, disclose user data, or trigger legal proceedings (Sullivan, 2025). Taken together, the EU framework does not yet demonstrate a need for a separate offence of synthetic operations. The harder problems are evidentiary: proving how an AI result was used, separating criminal fault from regulatory duties, and identifying which party knew the context or could intervene. Legal change should be tied to those concrete gaps rather than to technological novelty itself.

CONCLUSION

Generative AI has not transformed cyberterrorism into activity that operates without human control. The clearest change remains in the production, translation, adjustment, and reuse of propaganda. Its added value lies in reducing the time and labor required to prepare material, not in evidence that AI-generated propaganda is more effective in recruitment or violence. Movement toward operational support begins when AI results are used to interpret information, prepare material, or assist later work. Open evidence through July 2026 does not yet establish an autonomous and systematic AI-based terrorist workflow. Synthetic operations are therefore better understood as an analytical threshold for identifying situations in which AI assistance no longer ends with information but becomes connected to the preparation or execution of terrorism.

The fact that material was made with AI is likewise insufficient to determine legal responsibility. Assessment must examine how the result was used, its connection with prohibited conduct, and the parties that understood the purpose or exercised control over the process. Technology-neutral law remains capable of application, although proof becomes more difficult when assistance is distributed among users, model providers, applications, and platforms. Content regulation remains important for published propaganda but cannot reach every use of AI before material circulates. Intervention must also preserve the distinction between research, lawful expression, and real assistance to terrorism. The decisive measure is not model sophistication or the synthetic origin of content, but the function of AI output within a sequence of conduct and its proximity to the preparation or execution of terrorism.

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

Brynjolfsson, E., Li, D., & Raymond, L. R. (2025). Generative AI at work. *The Quarterly Journal of Economics, 140*(2), 889–942. https://doi.org/10.1093/qje/qjae044

Conway, M., & Macdonald, S. (2021). Introduction to the special issue: Extremism and terrorism online—Widening the research base. *Studies in Conflict & Terrorism*, 1–7. https://doi.org/10.1080/1057610X.2020.1866730

Dathathri, S., See, A., Ghaisas, S., Huang, P.-S., McAdam, R., Welbl, J., Bachani, V., Kaskasoli, A., Stanforth, R., Matejovicova, T., Hayes, J., Vyas, N., Al Merey, M., Brown-Cohen, J., Bunel, R., Balle, B., Cemgil, T., Ahmed, Z., Stacpoole, K., … Kohli, P. (2024). Scalable watermarking for identifying large language model outputs. *Nature, 634*(8035), 818–823. https://doi.org/10.1038/s41586-024-08025-4

Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 on combating terrorism and replacing Council Framework Decision 2002/475/JHA and amending Council Decision 2005/671/JHA. (2017). *Official Journal of the European Union, L 88*, 6–21. https://eur-lex.europa.eu/eli/dir/2017/541/oj/eng

Europol. (2025). *European Union terrorism situation and trend report 2025 (EU TE-SAT)*. Publications Office of the European Union. https://www.europol.europa.eu/cms/sites/default/files/documents/EU_TE-SAT_2025.pdf

Fang, R., Bindu, R., Gupta, A., & Kang, D. (2024). *LLM agents can autonomously exploit one-day vulnerabilities*. arXiv. https://doi.org/10.48550/arXiv.2404.08144

Fasel, M., & Weerts, S. (2024). Can Facebook's community standards keep up with legal certainty? Content moderation governance under the pressure of the Digital Services Act. *Policy & Internet, 16*, 588–606. https://doi.org/10.1002/poi3.391

Fernández-Llorca, D., Gómez, E., Sánchez, I., & Mazzini, G. (2025). An interdisciplinary account of the terminological choices by EU policymakers ahead of the final agreement on the AI Act: AI system, general purpose AI system, foundation model, and generative AI. *Artificial Intelligence and Law, 33*, 875–888. https://doi.org/10.1007/s10506-024-09412-y

Ferrag, M. A., Alwahedi, F., Battah, A., Cherif, B., Mechri, A., Tihanyi, N., Bisztray, T., & Debbah, M. (2025). Generative AI in cybersecurity: A comprehensive review of LLM applications and vulnerabilities. *Internet of Things and Cyber-Physical Systems, 5*, 1–46. https://doi.org/10.1016/j.iotcps.2025.01.001

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

Gaudette, T., Scrivens, R., & Venkatesh, V. (2022). The role of the Internet in facilitating violent extremism: Insights from former right-wing extremists. *Terrorism and Political Violence, 34*(7), 1339–1356. https://doi.org/10.1080/09546553.2020.1784147

Global Internet Forum to Counter Terrorism. (2025). *Artificial intelligence: Threats, opportunities, and policy frameworks for countering violent non-state actors*. https://gifct.org/wp-content/uploads/2025/04/GIFCT-25WG-0425-AI_Report-Web-1.1.pdf

Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society, 7*(1), 1–15. https://doi.org/10.1177/2053951719897945

Gregory, S. (2022). Deepfakes, misinformation and disinformation and authenticity infrastructure responses: Impacts on frontline witnessing, distant witnessing, and civic journalism. *Journalism, 23*(3), 708–729. https://doi.org/10.1177/14648849211060644

He, Q., Hong, Y., & Raghu, T. S. (2024). Platform governance with algorithm-based content moderation: An empirical study on Reddit. *Information Systems Research, 36*(2), 1078–1095. https://doi.org/10.1287/isre.2021.0036

Herath, C., & Whittaker, J. (2023). Online radicalisation: Moving beyond a simple dichotomy. *Terrorism and Political Violence, 35*(5), 1027–1048. https://doi.org/10.1080/09546553.2021.1998008

Ibrar, W., Mahmood, D., Al-Shamayleh, A. S., Ahmed, G., Alharthi, S. Z., & Akhunzada, A. (2025). Generative AI: A double-edged sword in the cyber threat landscape. *Artificial Intelligence Review, 58*, 285. https://doi.org/10.1007/s10462-025-11285-9

Joint Counterterrorism Assessment Team. (2024). *Violent extremists' use of generative artificial intelligence*. National Counterterrorism Center. https://www.dni.gov/files/NCTC/documents/jcat/firstresponderstoolbox/151s_First_Responders_Toolbox-Violent_Extremists_Use_of_Generative_Artificial_Intelligence.pdf

Ji, Z., Lee, N., Frieske, R. M., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), Article 248. https://doi.org/10.1145/3571730

Kadir, Z. K. (2025a). Kejahatan Berbasis Identitas Digital: Menggagas Kebijakan Kriminal untuk Dunia Metaverse. *Jurnal Litigasi Amsir*, *12*(2), 124–137.

Kadir, Z. K. (2025b). Kriminologi Reaksioner: Sayap Kanan, Moralitas, dan Kriminalisasi Budaya Pop dalam Era Post-Truth. *Jurnal Intelek Insan Cendikia*, *2*(5), 10108–10121.

KhosraviNik, M., & Amer, M. (2022). Social media and terrorism discourse: The Islamic State's (IS) social media discursive content and practices. *Critical Discourse Studies, 19*(2), 124–143. https://doi.org/10.1080/17405904.2020.1835684

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

Lowe, D. (2022). Far-right extremism: Is it legitimate freedom of expression, hate crime, or terrorism? *Terrorism and Political Violence, 34*(7), 1433–1453. https://doi.org/10.1080/09546553.2020.1789111

Macdonald, S., & Lorenzo-Dus, N. (2021). Visual jihad: Constructing the "good Muslim" in online jihadist magazines. *Studies in Conflict & Terrorism, 44*(5), 363–386. https://doi.org/10.1080/1057610X.2018.1559508

Macdonald, S., Jarvis, L., & Lavis, S. M. (2022). Cyberterrorism today? Findings from a follow-on survey of researchers. *Studies in Conflict & Terrorism, 45*(8), 727–752. https://doi.org/10.1080/1057610X.2019.1696444

Negara, T. A. S. (2023). Normative legal research in Indonesia: Its originis and approaches. *Audito Comparative Law Journal (ACLJ), 4*(1), 1–9. https://doi.org/10.22219/aclj.v4i1.24855

Nieweglowska, M., Stellato, C., & Sloman, S. A. (2023). Deepfakes: Vehicles for radicalization, not persuasion. *Current Directions in Psychological Science, 32*(3), 236–241. https://doi.org/10.1177/09637214231161321

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187–192. https://doi.org/10.1126/science.adh2586

Puranam, P. (2021). Human–AI collaborative decision-making as an organization design problem. *Journal of Organization Design, 10*, 75–80. https://doi.org/10.1007/s41469-021-00095-2

Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online. (2021). *Official Journal of the European Union, L 172*, 79–109. https://eur-lex.europa.eu/eli/reg/2021/784/oj/eng

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC (Digital Services Act). (2022). *Official Journal of the European Union, L 277*, 1–102. https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending specified Union legislation (Artificial Intelligence Act). (2024). *Official Journal of the European Union, 2024/1689*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

Schuett, J. (2024). Risk management in the Artificial Intelligence Act. *European Journal of Risk Regulation, 15*(2), 367–385. https://doi.org/10.1017/err.2023.1

Sullivan, G. (2025). Algorithmic governance of 'terrorism' and 'violent extremism' online. *London Review of International Law, 13*(1), 47–75. https://doi.org/10.1093/lril/lraf005

Taekema, S. (2021). Methodologies of rule of law research: Why legal philosophy needs empirical and doctrinal scholarship. *Law and Philosophy, 40*, 33–66. https://doi.org/10.1007/s10982-020-09388-1

Tech Against Terrorism. (2023, November 8). *Early terrorist adoption of generative AI*. https://techagainstterrorism.org/news/early-terrorist-adoption-of-generative-ai

Tech Against Terrorism. (2026, July 2). *AI terrorism blind spot: First benchmark built to measure it finds frontier models give attackers usable help*. https://techagainstterrorism.org/news/press-release-ai-terrorism-benchmark

Teichmann, F. (2026). General-purpose AI under the EU AI Act: A conceptual allocation of duties across the value chain. *SCRIPTed: A Journal of Law, Technology & Society, 23*(1). https://doi.org/10.2218/scrip.12300

Unlu, A., & Yilmaz, K. (2025). Online terrorism studies: Analysis of the literature. *Studies in Conflict & Terrorism, 48*(9), 1032–1055. https://doi.org/10.1080/1057610X.2022.2122108

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society, 6*(1), 1–13. https://doi.org/10.1177/2056305120903408

van Gestel, R. (2023). Quality, methodology, and politics in doctrinal legal scholarship. *Law and Method, 2023*(jan–mar), 1–24. https://doi.org/10.5553/REM/.000070

Watkin, A.-L. (2023). Countering online terrorist content: A social regulation approach. *Policy & Internet, 15*(4), 528–543. https://doi.org/10.1002/poi3.373

Winter, C., Neumann, P., Meleagrou-Hitchens, A., Ranstorp, M., Vidino, L., & Fürst, J. (2020). Online extremism: Research trends in internet activism, radicalization, and counter-strategies. *International Journal of Conflict and Violence, 14*, 1–20. https://doi.org/10.4119/ijcv-3809

Zhu, Y., Kellermann, A., Bowman, D., Li, P., Gupta, A., Danda, A., Fang, R., Jensen, C., Ihli, E., Benn, J., Geronimo, J., Dhir, A., Rao, S., Yu, K., Stone, T., & Kang, D. (2025). *CVE-Bench: A benchmark for AI agents' ability to exploit real-world web application vulnerabilities*. arXiv. https://doi.org/10.48550/arXiv.2503.17332

Zornetta, A., & Pohland, I. (2022). Legal and technical trade-offs in the content moderation of terrorist live-streaming. *International Journal of Law and Information Technology, 30*(3), 302–320. https://doi.org/10.1093/ijlit/eaac020