RSIS COMMENTARY

28 January 2026

Countering Online Radicalisation in Indonesia – Policy Needs to Keep Pace with Changes

Nauval El Ghifari

PONDER THE IMPROBABLE

RSiS S. RAJARATNAM SCHOOL OF INTERNATIONAL STUDIES
Nanyang Technological University, Singapore

RSiS 30

RSiS S. RAJARATNAM SCHOOL OF INTERNATIONAL STUDIES
Nanyang Technological University, Singapore

RSIS COMMENTARY

www.rsis.edu.sg

No. 016 – 28 January 2026

RSIS Commentary is a platform to provide timely and, where appropriate, policy-relevant commentary and analysis of topical and contemporary issues. The authors' views are their own and do not represent the official position of the S. Rajaratnam School of International Studies (RSIS), NTU. These commentaries may be reproduced with prior permission from RSIS and due credit to the author(s) and RSIS. Please email to Editor RSIS Commentary at RSISPublications@ntu.edu.sg.

Countering Online Radicalisation in Indonesia – Policy Needs to Keep Pace with Changes

By Nauval El Ghifari

SYNOPSIS

Recent incidents involving an Indonesian youth in Jordan exposed to ISIS content, and the bomb attack on Jakarta's State Senior High School 72, reveal that youth radicalisation in Indonesia is increasingly occurring through mainstream digital platforms rather than closed online spaces. There is a need for policy to keep pace with these evolving circumstances.

COMMENTARY

Online extremism today no longer resembles a structured enemy with clear organisational boundaries or identifiable ideological markers. Instead, it increasingly operates as a "ghost" within digital spaces. It is anonymous, fragmented, and difficult to trace, thriving in environments where attribution is weak and identities are easily concealed. This "ghost-like" nature allows extremist narratives to circulate across platforms with minimal disruption, often escaping early scrutiny by both authorities and platform moderation systems.

At the same time, online extremism operates like a "poisonous chameleon". Instead of positioning itself against mainstream culture, it embeds itself within it. Extremist narratives increasingly mimic popular digital forms such as memes, viral sounds, popular music, humour, and visual aesthetics that resonate with younger audiences. By adopting familiar cultural cues, these narratives become harder to distinguish from ordinary content. They circulate without triggering immediate suspicion, normalising extremist ideas through repetition and emotional resonance rather than explicit ideological instruction.

This dual character poses a fundamental challenge for counter-radicalisation efforts. If online extremism is invisible, like a “ghost”, and adaptive, like a “poisonous chameleon”, it cannot be effectively addressed through countermeasures designed for static and easily identifiable threats. Yet many existing responses continue to rely on detection models that assume identifiable patterns, stable narratives, and predictable forms of expression in online extremism. This mismatch raises a critical question. Why do counter-radicalisation strategies remain largely reactive when the threat they seek to counter is fluid, anonymous, and constantly changing shape?

Why Current Countermeasures Are Misaligned

Current counter-radicalisation approaches remain poorly aligned with the evolving nature of online extremism. Government responses, including in Indonesia, continue to prioritise blunt takedowns after incidents have already occurred. These measures are largely reactive by design, focusing on content removal rather than early-stage intervention. While takedowns may limit short-term exposure, they do little to address how extremist narratives emerge, adapt, and embed themselves within everyday digital culture.

Social media platforms, meanwhile, rely heavily on AI-driven moderation systems built around pattern recognition and rule-based enforcement. These systems assume that extremist content can be identified through stable markers such as keywords, symbols, or recurring visual cues. However, this assumption increasingly fails when extremist narratives survive precisely by altering their appearance. When content is anonymous and constantly shifting, attribution becomes weak, and detection models struggle to keep pace. Algorithmic governance without cultural literacy creates blind spots, particularly when extremist narratives draw on local references, humour, or visual styles that appear benign to external reviewers.

As a result, counter-radicalisation efforts often end up chasing a threat that has already changed its form. By the time harmful content is identified and removed, new variations have already surfaced elsewhere. This reactive cycle reinforces enforcement over prevention and leaves policymakers perpetually one step behind an adversary that thrives on adaptation.

This gap points to a structural weakness in how platforms govern harmful content. If extremist narratives no longer appear in fixed and recognisable forms, moderation cannot rely solely on automated detection. Social media companies, therefore, need to invest not only in more sophisticated AI, but in institutionalised human judgement.

One practical step is to create dedicated “trust and safety” units tasked with developing a continuously updated “narrative hub”, a glossary that maps how extremist ideas are articulated, reframed, and culturally embedded across digital spaces. Such a system would track not only explicit keywords, but symbolic references, visual tropes, humour, music, and aesthetic cues through which extremist narratives circulate.

Closing this gap demands more than technical capacity alone. It calls for cultural and linguistic literacy among professionals in these units, particularly an understanding of local youth culture and online subcultures, so platforms can anticipate narrative shifts rather than merely react to them. Without this, algorithmic governance will remain structurally blind to the very forms of extremism that are now most effective.

When Policy Design Lags Behind Youth Digital Reality

Youth occupy a paradoxical position within online radicalisation dynamics. They are the primary audience of digital cultural spaces and, increasingly, the primary targets of extremist narratives. Yet within policy design, youth are rarely treated as active stakeholders. More often, they are framed as passive recipients of messaging or as risk groups in need of protection, rather than as contributors capable of shaping preventive strategies.

This blind spot persists despite growing recognition at the global level. The Youth, Peace, and Security (YPS) agenda has emphasised the importance of youth participation in peacebuilding and conflict prevention. However, its application in digital counter-radicalisation remains limited. In practice, youth inclusion is often reduced to consultation exercises or awareness campaigns that do not meaningfully influence how policies are designed or implemented.

A significant generational gap further compounds the problem. Policymakers and security institutions do not inhabit the same digital ecosystems as younger users. The platforms, cultural references, and modes of expression that shape youth’s online experience are often poorly understood by decision-makers.

The gap between policy language and lived digital reality continues to widen. Whereas official counter-narratives depend on formal messaging and institutional authority, extremist narratives resonate by drawing on belonging, identity, and shared cultural codes.

A strategy that excludes youth perspectives is therefore structurally incapable of understanding how online radicalisation works. Without insight into how narratives resonate within youth culture, prevention efforts remain detached from the environments where radicalisation takes root.

As Indonesia prepares its National Action Plan on Preventing and Countering Violent Extremism (RAN PE), 2025-2029, this gap should not be repeated. The National Counter Terrorism Agency (BNPT) needs to treat youth not merely as target audiences, but as policy partners. This requires structured collaboration with young digital practitioners, social media companies, and civil society organisations working on youth-focused Preventing and Countering Violent Extremism (P/CVE) work, so that preventive policy reflects how radicalisation unfolds online, not how institutions assume it does.

Conclusion

Online extremism is no longer a fixed or easily identifiable threat. It operates invisibly like a “ghost”, enabled by anonymity, and adapts like a “poisonous chameleon” by embedding itself within mainstream digital culture. As long as counter-radicalisation strategies remain reactive and youth remain marginal to policy design, prevention efforts will continue to chase a threat that has already shifted form.

Addressing online radicalisation, therefore, requires more than enforcement and takedowns. It demands a preventive approach grounded in cultural literacy, anticipatory governance, and meaningful youth participation. Without a shift from reactive enforcement to culturally informed prevention, counter-radicalisation policy will remain structurally misaligned and permanently behind the threat it aims to manage.

Nauval El Ghifari is a Master of Science student in Strategic Studies at the S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University (NTU), Singapore. He is an alumnus of Young Leaders for the Online Prevention and Countering of Violent Extremism (PCVE) in Southeast Asia organised by the United Nations Office of Counter-Terrorism (UNOCT).

Please share this publication with your friends. They can subscribe to RSIS publications by scanning the QR Code below.