RSIS COMMENTARY

17 July 2025

The Use of Artificial Intelligence in Countering Online Radicalisation in Indonesia

Raneeta Mutiara

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

Ponder the Improbable since 1996

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

RSIS COMMENTARY

www.rsis.edu.sg

No. 157 – 17 July 2025

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.

The Use of Artificial Intelligence in Countering Online Radicalisation in Indonesia

By Raneeta Mutiara

SYNOPSIS

New and emerging Artificial Intelligence tools could help to counter the threat of online radicalisation and violent extremism in Indonesia. Based on interviews with practitioners in the field, the author argues that more investment and training are needed to equip counter-terrorism teams with these new skills. A national centre to act as a secure hub for such efforts, along with a localised Large Learning Model attuned to the domestic context and culture, are among the proposed ideas.

COMMENTARY

Digitalisation of the activities of Islamic State in Iraq and Syria (ISIS) has been a longstanding issue in Southeast Asia. In recent years, the nature of this threat has become more widespread and complex. In countries like Indonesia, where radicalisation is primarily offline, online platforms still play a role in spreading extremist ideas and maintaining ideological networks. The phenomenon of online radicalisation can erode social cohesion, highlighting the need for strategic measures to counter its destabilising impact.

Indonesia has made several attempts to combat online radicalisation. The National Counter Terrorism Agency of Indonesia (BNPT) initiated the *Duta Damai Dunia Maya* campaign to counter harmful content on the Internet. Other online initiatives, such as *BincangSyariah* and *Islamidotco*, have also been promoting Islamic literacy, moderating religious interpretations, and correcting misleading narratives.

Nevertheless, Indonesia still encounters online radicalisation cases. In July 2024, Indonesia's elite counterterrorism unit, Densus 88, detained a 19-year-old student who had expressed allegiance to ISIS through social media and was believed to be planning attacks on religious sites before he was caught.

The swift progress of Artificial Intelligence (AI), especially in areas of machine learning (ML) and natural language processing (NLP), presents both opportunities and challenges in combating online radicalisation in Indonesia. AI, generally defined as machines mimicking human intelligence, enables systems to recognise patterns, analyse content, and produce outputs in text, images, and videos. Within this AI landscape, ML allows models to enhance themselves through data, while NLP, as a specific ML application, deals with understanding and generating human language. These advancements provide possibilities for creating early detection systems, content moderation tools, and sentiment analysis tools that can spot and counter extremist messages online.

For the research, the author conducted interviews with fifteen experts across different fields, including law enforcement officers, academics, representatives from civil society organisations (CSOs), and employees of AI start-ups in Indonesia. The qualitative data collected from this process have been analysed through thematic analysis, and the preliminary findings reveal that AI can indeed complement the conventional CVE (countering violent extremism) methods in the country, albeit not without challenges.

Potential Roles of AI in Countering Violent Extremism

Most respondents interviewed believe that AI can enhance CVE efforts in Indonesia because it can increase the efficiency of the CVE apparatus and aid in social media monitoring. Integrating AI into automated systems significantly improves the functionality of digital tools used for counter-radicalisation efforts.

AI-driven tools such as NLP chatbots, built on Large Language Models (LLMs), can aid CVE initiatives by providing personalised, responsive interactions on campaign platforms and public websites. LLMs excel at detecting abstract themes, sentiments, and contextual subtleties, which can indicate potential radicalisation, like endorsing violence and using divisive language. They can also be trained to redirect users to relevant educational content or support services. While these bots will not change beliefs instantly, they can serve as effective early intervention tools, particularly among vulnerable individuals.

Even though AI does not "predict" human behaviour in a deterministic way, ML algorithms can recognise recurring patterns and anomalies based on past and behavioural data. For instance, clusters of posts with growing engagement in violent ideologies or shifts from passive viewing to active sharing may indicate the early phases of online radicalisation. These signals can act as risk indicators, helping authorities prioritise monitoring, investigate more deeply, or intervene sooner.

The Challenges: Institutional and Societal Barriers

Respondents have identified persistent barriers to the integration of AI into Indonesia’s CVE strategies, in particular, limited financial resources allocated to digital infrastructure and low public engagement with technology-based counter-radicalisation tools. These challenges are not exclusive to Indonesia, and their effects are especially severe in the CVE context, where slow adoption of cutting-edge technologies could allow extremist narratives to proliferate unchecked online.

Although the BNPT has sub-directorates for counterpropaganda and intelligence gathering, it has not used AI to its full potential in Indonesia. One reason for this is the lack of expertise in interpreting AI-generated insights. This skills gap stems from limited government investment in applied AI research and development concerning public safety and security. This has resulted in AI-illiterate officers incapable of maximising the use of AI for CVE. Consequently, the roles related to AI are often handled by third parties, which are mostly private companies.

Additionally, the concept of radicalisation is not widely understood by Indonesian society. The process of radicalisation, particularly its online and psychological aspects, remains largely unfamiliar to the general public. On the other hand, in a culture characterised by passive information consumption, people often perceive AI as the “expert”, thus taking all AI-generated information uncritically without the need for deeper analysis. In this environment, a concerning irony emerges: while government agencies find it challenging to leverage AI for CVE, the public might become more susceptible to extremist narratives amplified by AI.

These complications stem from the government’s failure to provide adequate education and training in CVE and AI to the public. As a result, there is limited public awareness of the promising roles of AI in countering online radicalisation. Most respondents believe this has adversely affected public trust in the government when it comes to using AI in CVE, creating legal frictions against its implementation. A debate has sparked over whether AI-assisted social media monitoring violates Indonesia’s Personal Data Protection (PDP) Law No. 27 of 2022.

Recommendations

To address these challenges effectively, Indonesia should consider building a national AI centre for public security. Such a centre would not only serve as a secure, centralised data hub for analysing extremist content and trends but also act as a national training and certification platform for CVE practitioners, law enforcement officers, and digital analysts. This could address the current AI-illiteracy and overreliance on third-party contractors.

To complement this institutional effort, Indonesia might explore creating its own LLMs specifically designed for CVE. While international LLMs are trained on extensive multilingual datasets, they often lack the cultural understanding, linguistic sensitivity, and contextual relevance necessary to identify and interpret radical content in Indonesian languages, dialects, or local metaphors. Building a localised LLM helps the law enforcers focus on relevant data.

Training models on linguistically and culturally specific content also reduces false positives, improves accuracy on thematic identification, and better detects radical narratives unique to the Indonesian context. More importantly, if the development process is transparent and involves civil society stakeholders, it can restore public confidence by showing responsible use, protecting privacy, and complying with Indonesia’s legal requirements, such as the PDP law.

Raneeta Mutiara is an Associate Faculty member in the Public Safety and Security (PSS) Programme at the Singapore University of Social Sciences (SUSS). Her research interests include countering violent extremism in Southeast Asia and examining the roles of women in public safety and security.

S. Rajaratnam School of International Studies, NTU Singapore
Block S4, Level B3, 50 Nanyang Avenue, Singapore 639798