

Artificial Intelligence and Countering Violent Extremism: A Primer

By Marie Schroeter



Global Network
on Extremism & Technology

Radicalisation can take place offline as well as online.
To what extent the internet plays a role remains contested.

Manipulated search engines and recommendation systems can contribute to counter-radicalisation by pointing to moderate content.

A balanced recommendation system would counter malicious narratives with opposing content or spread information about projects and contact points on the prevention and countering of violent extremism.

Tackling disinformation and manipulated content online lacks automated solutions.

Superhuman AI is not going to 'sound an alarm' if individual radicalises online.

1. There is not enough data to feed an algorithm with definite information on radicalisation and when a radicalised individual turns to violence.

2. Predicting the behaviour of individuals would require clearly assignable data on individuals, which would give away every aspect of privacy and potentially result in surveillance on an unprecedented scale.

Natural language processing (NLP) can help to translate minority languages for better content moderation and support content moderation of niche websites in the long run.

Improved NLP can help to translate content into languages in which large numbers of experienced and trained moderators operate.

This report looked into the ability of artificial intelligence (AI) applications to contribute to countering radicalisation.