Terrorist content lurks all over the internet—regulating only six major platforms won't be nearly enough

March 20 2024, by Marten Risius and Stan Karanasios

Credit: Unsplash/CC0 Public Domain

Australia's eSafety commissioner has sent legal notices to Google, Meta, Telegram, WhatsApp, Reddit and X (formerly Twitter) asking them to show what they're doing to protect Australians from online extremism. The six companies have 49 days to respond.

The notice comes at a time when governments are increasingly cracking down on major tech companies to address online harms like child sexual abuse material or bullying.

Combating online extremism presents unique challenges different from other content moderation problems. Regulators wanting to establish effective and meaningful change must take into account what research has shown us about extremism and terrorism.

Extremists are everywhere

Online extremism and terrorism have been pressing concerns for some time. A stand-out example was the 2019 Christchurch terrorist attack on two mosques in Aotearoa New Zealand, which was live streamed on Facebook. It led to the "Christchurch Call" to action, aimed at countering extremism through collaborations between countries and tech companies.

But despite such efforts, extremists still use online platforms for networking and coordination, recruitment and radicalization, knowledge transfer, financing and mobilization to action.

In fact, extremists use the same online infrastructure as everyday users: marketplaces, dating platforms, gaming sites, music streaming sites and social networks. Therefore, all regulation to counter extremism needs to consider the rights of regular users, as well.

The rise of 'swarmcasting'

Tech companies have responded with initiatives like the Global Internet Forum to Counter Terrorism. It shares information on terrorist online content among its members (such as Facebook, Microsoft, YouTube, X and others) so they can take it down on their platforms. These approaches aim to automatically identify and remove terrorist or extremist content.

However, a moderation policy focused on individual pieces of content on individual platforms fails to capture much of what's out there.

Terrorist groups commonly use a "swarmcasting" multiplatform approach, leveraging 700 platforms or more to distribute their content.

Swarmcasting involves using "beacons" on major platforms such as Facebook, Twitter and Telegram to direct people to locations with terrorist material. This beacon can be a hyperlink to a blog post on a website like Wordpress or Tumblr that then contains further links to the content, perhaps hosted on Google Drive, JustPaste.It, BitChute and other places where users can download it.

So, while extremist content may be flagged and removed from social media, it remains accessible online thanks to swarmcasting.

Putting up filters isn't enough

The process of identifying and removing extremist content is far from simple. For example, at a recent US Supreme Court hearing over internet regulations, a lawyer argued platforms could moderate terrorist content by simply removing anything that mentioned "al Qaeda".

However, internationally recognized terrorist organizations, their members and supporters do not solely distribute policy-violating extremist content. Some may be discussing non-terrorist activities, such as those who engage in humanitarian efforts.

Other times their content is borderline (awful but lawful), such as misogynistic dog whistles, or even "hidden" in a different format, such as memes.

Accordingly, platforms can't always cite policy violations and are compelled to use other methods to counter such content. They report using various content moderation techniques such as redirecting users, pre-bunking misinformation, promoting counterspeech and offering warnings, or implementing shadow bans. Despite these efforts, online extremism continues to persist.

What is extremism, anyway?

All these problems are further compounded by the fact we lack a commonly accepted definition for terrorism or extremism. All definitions currently in place are contentious.

Academics attempt to seek clarity by using relativistic definitions, such as

"extremism itself is context-dependent in the sense that it is an inherently relative term that describes a deviation from something that is (more) 'ordinary', 'mainstream' or 'normal'."

However, what is something we can accept as a universal normal? Democracy is not the global norm, nor are equal rights. Not even our understanding of central tenets of human rights is globally established.

What should regulators do, then?

As the eSafety commissioner attempts to shed light on how major platforms counter terrorism, we offer several recommendations for the commissioner to consider.

1. Extremists rely on more than just the major platforms to disseminate information. This highlights the importance of expanding the current inquiries beyond just the major tech players.

2. Regulators need to consider the differences between platforms that resist compliance, those that comply halfheartedly, and those that struggle to comply, such as small content storage providers. Each type of platform requires different regulatory approaches or assistance.

3. Future regulations should encourage platforms to transparently collaborate with academia. The global research community is well positioned to address these challenges, such as by developing actionable definitions of extremism and novel countermeasures.

Provided by The Conversation