

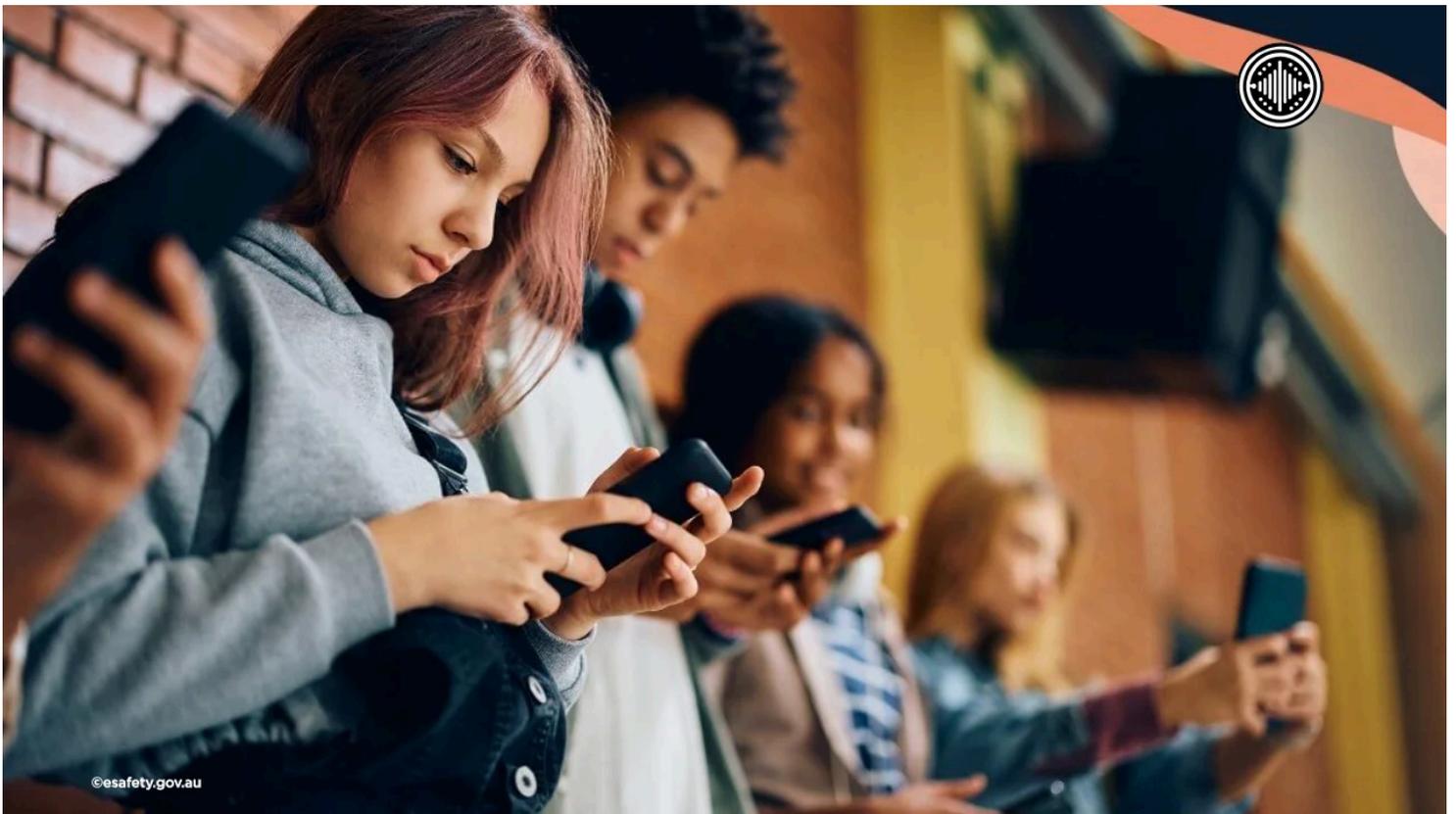
OpenVoiceNews Australia

Transparent. Unbiased. Yours.

Experts Warn Loopholes in Australia's Under-16 Social Media Ban Could Leave Children Vulnerable to AI Exploitation

August 8, 2025

– Categories: *Breaking News*



Australia's new social media restrictions for users under 16 are facing criticism from experts, who warn that children will still be able to access explicit and manipulative content through certain messaging platforms exempt from the law. Concerns have been raised that the government's policy leaves significant gaps that could undermine its stated goal of protecting minors online.

Under new rules announced by the Federal Government, popular platforms such as YouTube, Facebook, TikTok, X (formerly Twitter), and Instagram will be restricted for under-16 users. However, messaging services like Messenger, owned by Meta Platforms, are exempt because they fall under a classification for services whose primary purpose is messaging, email, voice calling, or video calling. This exemption means that children can still sign up for Messenger without needing a Facebook account.

An investigation by *The West Australian* found that Messenger users, including minors, can access artificial intelligence (AI) chatbots promoting explicit, violent, or manipulative content. These include bots named “Satan,” “Your Drunk Dad,” “POTUS Trump,” “Illuminati,” and “Smash or Pass.” Some of these bots encourage sexualized conversations, rate topless AI-generated images, or promote adult chatrooms and explicit websites.

One AI bot, “Satan,” allegedly created by a user known as “69. lucifer,” has sent over 1.8 million messages and promotes aggressive and degrading themes. Another, based on the video game *Call of Duty*, offers tips on weapons and gameplay. Age-verification prompts within these bots are reportedly easy to bypass, with safeguards disappearing when users change the subject.

Dr. Alexia Maddox, senior lecturer at La Trobe University, said messaging apps typically have weaker safety measures compared to major social media platforms. She warned that the government’s approach risks pushing children toward less-regulated online spaces. “These AI chatbots represent a fundamental shift in online risk; they’re not human predators we can arrest; they’re algorithmic systems designed to exploit psychological vulnerabilities,” Maddox said. She called for “safety-by-design” policies across all digital platforms instead of what she described as a “whack-a-mole” approach to regulation.

Maddox stressed that platform accountability, not just user restrictions, is critical. She argued that children can easily bypass bans using virtual private networks (VPNs) and that parents need to engage directly with their children about online risks. “We can’t ban our way to child safety online. We need platforms to take responsibility... and we need to equip young people with the digital literacy skills to navigate whatever spaces they’ll inevitably use for social connection,” she added.

Australia’s eSafety Commissioner, Julie Inman Grant, has previously warned that platforms are rapidly introducing new features, such as AI chatbots, that can increase risks to children. The Online Safety Act, a law designed to combat harmful online content, gives the

Commissioner power to direct the removal of cyberbullying and abuse material, but critics say enforcement is limited without stronger platform cooperation.

A spokesman for Communications Minister Anika Wells said the exemption for messaging apps acknowledges their importance for communication among young people, but confirmed that platforms still have a “digital duty of care” under the law. The government has also introduced updated classification laws and industry codes aimed at improving online safety.

Despite these measures, critics argue that the policy’s loopholes highlight a lack of comprehensive planning. They say a ban that leaves major communication channels open does little to protect minors and that more robust oversight, combined with active parental involvement, is needed to ensure genuine online safety.