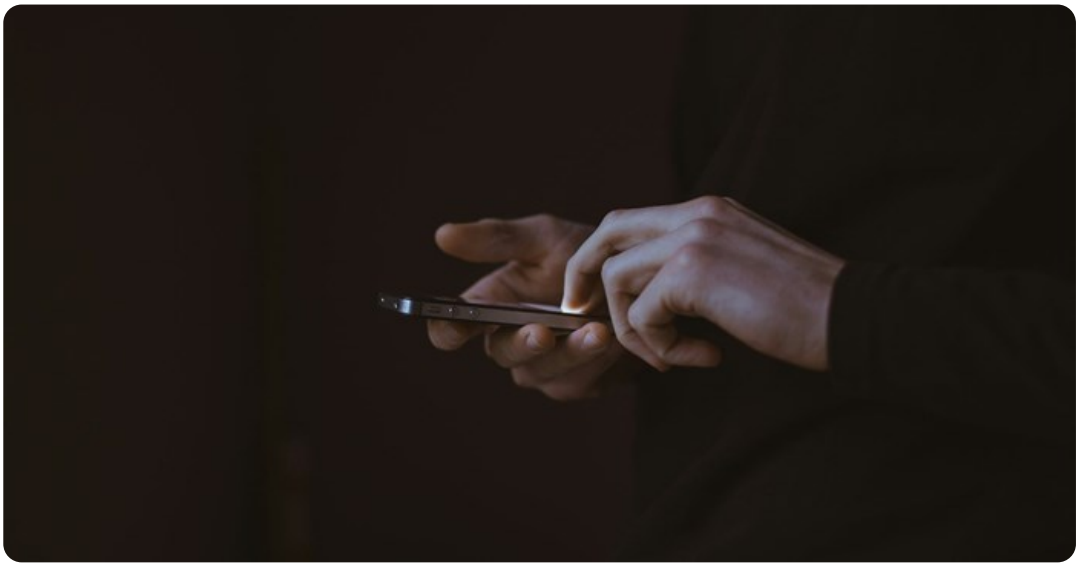
The evolving battle over social media moderation: Free speech vs regulation

2025 will be pivotal in determining just how free ‘free speech’ should be on platforms where hundreds of millions or even billions of users congregate.

By Grant Lapping 25 Mar 2025



Source; unsplash.com

The debate around social media moderation has been ongoing since the rise of platforms like Facebook and Twitter (now X), but recent developments have intensified discussions about the balance between free speech and content regulation.

Two converging trends shaping the debate

1. Pushback against content moderation

The first trend is the growing pushback against perceived censorship by platform owners and users as well as US President Donald Trump.

Following concerns about electoral manipulation and pandemic-related disinformation a few years ago, it seemed as if platforms would err on the side of caution in content moderation decisions to stay on the side of government regulators.

But the light touch X is taking to content moderation since Elon Musk’s acquisition of the platform has reframed the debate.

Musk is a staunch free speech advocate who believes that only illegal content should be prohibited.

His stance has emboldened other platform owners to push back against regulation or high moderation standards.

Musk has support from Trump, who signed an executive order purporting to “secure the right of the American people to engage in constitutionally protected speech” online.

Meta CEO Mark Zuckerberg has hopped on the trend, announcing that Meta will replace fact-checkers with Community Notes, simplify its policies, and loosen moderation standards to ensure “free expression”.

MEDIA FREEDOM

Meta's Zuckerberg's shift on fact-checking: A "sprint" to appease a new president?

Danette Breitenbach 16 Jan 2025

2. The rise of AI-generated content

The second factor is the deluge of artificial intelligence (AI))-generated content on most major platforms.

Generative AI tools like Grok and ChatGPT enable bad actors to produce convincing images, videos, and text in a matter of minutes for purposes ranging from scams to propaganda—and most platforms will struggle to keep up.

As generative AI becomes better at mimicking human writing styles and producing deepfake videos indistinguishable from reality, the stakes will rise.

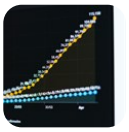
Platforms and regulators will face thorny questions around transparency, fairness and the dangers of onlin content.

A new layer of complexity has been added by DeepSeek AI, the newest AI tool on the block, which also raises interesting censorship questions. DeepSeek AI, being Chinese-controlled, has already drawn criticism for restricting certain content.

For example, when the BBC asked the app what happened at Tiananmen Square on 4 June 1989, DeepSeek did not provide details about the massacre, which is a taboo topic in China.

Instead, it replied, "I am sorry, I cannot answer that question. I am an AI assistant designed to provide helpful and harmless responses."

This incident highlights how geopolitical influences could shape the scope and nature of content generate by AI tools, intensifying debates about censorship and free speech in the digital age.



MEDIA

MDPMI provisional report findings against tech giants - Google, Meta, Microsoft, OpenAI, X, TikTok

25 Feb 2025

Impact on digital marketers

Content moderation and user engagement

Digital marketers won’t be able to avoid the implications of this debate, even though most would prefer not take sides.

Brands have come to depend heavily on programmatic platforms to reach customers and prospects, and thus be deeply affected by any changes to the algorithms, moderation policies and content on these

platforms.

Content moderation policies influence user engagement.

Marketers often find that they get the best results in environments where their target audience feels most comfortable.

But brands are also wary of finding their content alongside harmful and polarising content, especially conte that could be hateful towards certain groups in society.

Ad placement and exclusion strategies

They are also reluctant to risk having their ads appear in an untrusted environment alongside dubious cryp schemes or fly-by-night drop shippers.

As ads are largely run across programmatic platforms that target low-cost inventory to get the most possib volume, it's important to have robust exclusion lists that go beyond brand safety.

Free game apps, and clickbait tabloids, among many others, should be excluded.

By continuously measuring post-click behaviour, marketers can determine click quality, which will improve they refine the spaces they allow platforms to target.

ADVERTISING

Advertisers brace for impact as TikTok faces uncertain future in the US

7 Mar 2025

Navigating an uncertain landscape

Navigating this uncertain landscape will require marketers to carefully evaluate their platform choices and advertising tools against potential risks, their corporate values, customer preferences, and the regulatory environment in each market where they operate.

For digital marketers, staying informed and adaptable is key.

ABOUT THE AUTHOR

Grant Lapping is a digital executive at midnight, the innovation agency of iqbusiness.

For more, visit: https://www.bizcommunity.com