Two mass shootings show cracks in social media content moderation

Article

The news: Social media platforms have struggled to stop the spread of upsetting footage and misinformation surrounding the recent mass shootings in Buffalo, New York, and Ulvade, Texas, prompting criticisms and concerns about their ability to curb harmful content.

Shooting videos spread online: Footage of the May 14 Buffalo supermarket shooting, which was broadcast by the gunman on **Twitch**, rapidly spread across social media platforms

throughout the day.

- The Twitch broadcast of the shooting was viewed by at least 22 people and taken down less than two minutes after beginning, per a statement from the company, but the video and the shooter's manifesto, posted on **Facebook**, had already made its way to other platforms where it was viewed by millions.

- Videos of the attack could be found on Twitter and Facebook in the hours and days after the event, per Vice News.

- This wasn't the first time social media platforms—including Twitch and Facebook specifically — were used to livestream mass murders, prompting criticism over their lack of preparedness. The gunman in the 2019 Christchurch incident in New Zealand broadcast it on Facebook and directly referenced a popular **YouTuber**, and Twitch was used to circulate another 2019 shooting in Germany.

- In the days after the Buffalo shooting, both New York and New Jersey launched investigations into Twitch and social gaming app **Discord**.

Ulvade conspiracy theories: This week's mass shooting at an elementary school in Ulvade, coming just days after the Buffalo incident, reignited concerns about social media's role in the aftermath of violent events after misinformation about the shooter spread online.

- A conspiracy theory that began on the alt-right messaging board **4chan** falsely accusing a transgender woman of being the shooter quickly made its way to sites like **Reddit** and **Twitter**, where they were shared by conservative commentators and even US House Rep. **Paul Gosar** of Arizona, per CNBC.

- Twitter released a statement saying it would require removal of all posts "that share misleading claims about the identity of the perpetrator with the intent to incite fear or spread fearful stereotypes."

- But at the time of this writing, posts and photos falsely identifying the shooter could easily be found on both Facebook and Twitter, nearly 24 hours after the incident.

Attitudes Toward Whether Social Media Companies' AI Can Find False Information as Effectively as Humans According to US Adults, Nov 2021

% of respondents

AI does a better job	19%
AI does about the same job	25%
AI does a worse job	22%
Not sure	32%

Note: ages 18+; responses of "no answer" not shown; numbers may not add up to 100% due to rounding
Source: Pew Research Center, "AI and Human Enhancement: Americans' Openness Is Tempered by a Range of Concerns," March 17, 2022

274424   InsiderIntelligence.com

Consequences for platforms: Platforms have routinely struggled to contain harmful content after violent events, causing brand safety concerns and hurting their already-low consumer sentiment during a period of intense competition over digital advertising dollars.

- There have been some steps taken to stop the spread of misinformation and upsetting videos, but the two recent shootings show that there are still gaps when it comes to moderating posts—and especially live video—about violent incidents.

- In 2019, platforms and tech companies including **Microsoft**, Facebook, Twitter, and others contributed to a joint effort called **The Global Internet Forum to Counter Terrorism** that helps platforms identify and shut down harmful content quickly. Twitch credited the program for its ability to quickly take down the Buffalo video.