YouTube changes community translation system to contain trolls

YouTube has changed its rules around community contributions after the feature was misused for a harassment campaign. Under the new rules, contributed translations won't be published to YouTube until the channel owner has manually approved them, making it harder for trolls to sneak abusive or self-promoting translations onto a popular channel.

Community translations were envisioned as a way for channels with active fanbases to crowdsource translation work. Filipino fans might be happy to translate their favorite YouTuber's titles and descriptions into their native language, thus making the channel more accessible to international audiences without any additional work from YouTube or the channel owner. Under the old system, creators could review individual contributions and had the option of disabling community contributions for specific videos or groups of videos, but most saw little downside to simply leaving the option enabled and letting translations populate without review.

> *Thanks for your patience — Based on the feedback we've heard, we are introducing some changes to Community Contributions. Moving forward, creators that have turned on this feature will need to manually review their Community Contributions and check for spam before publishing.*
>
> *— TeamYouTube (@TeamYouTube)* ***August 30, 2019***

YouTube seems to have been aware of the possibility for abuse

within that system from the beginning. In a Creator Academy post from 2017, YouTube explicitly warns creators about the importance of reviewing contributions. “While most community contributors are there to help, abuse can happen,” the post warns. “Regular review can keep things like spam links and offensive language out of your translated content.”

In this case, the problem seems to have been targeted harassment rather than spam or swearing. After running a video on bad behavior in PewdiePie’s translations, a YouTuber named JT found himself targeted by a large-scale harassment campaign, largely waged through the translations for PewDiePie’s videos. As JT recounted in a video a week before the change, the harassing titles quickly spread to other channels, often crowding out accurate translations.

JT took the fight to YouTube’s @teamyoutube account on Twitter, which recommended that he simply flag the offending translations for review.

“So you want us to report every title of every video?” JT replied. “Why not just roll out an actual fix… like having to approve them first?” A few days later, YouTube seems to have seen the logic of that request.

In a triumphal video posted September 1, JT shared news of the policy change and thanked his fans for keeping the pressure on. “Honestly, I thought YouTube was going to add a system where they check it for you,” he says in the video. “But at least there is a fix, and I’m not longer involved in this situation.”

https://www.theverge.com/2019/9/2/20844910/youtube-translation-changes-jt-pewdiepie-harassment