

An open letter to Roblox on safety and ID verification

metatablecat

Foreword

I've been using Roblox for 15 years, the way this platform has evolved over the years has not only been amazing but also quite surprising, for something that started as a physics simulator with Lego-like blocks, to see it evolve into a fully fledged game engine that can compete the likes of Unity and Unreal, is amazing work.

The platform has also come under scrutiny for its safety practises, and I have seen this first hand, the moderation here needs to be improved but thats not what this letter is for.

For disclosure, I am ID verified on Roblox, so that I can have access to paid assets, I doubt this update will affect me that much.

Who this letter is for

This letter is addressed to Roblox's safety team, and relates to how ID verification actually is not this shining silver bullet you think it is, I hope that this letter may change your opinion on this.

If you are not Roblox, and instead someone else reading this document, I ask that you share it so that it gets more attention. I am not on social media because I think it is not good for society, but if you are, feel free to share this around.

1. ID Verification is unsafe for user privacy

I would not trust any private company with my face scan, let alone my ID. We have already seen two big leaks^{[1][2]} where Tea and Discord were both not properly deleting user information, the former not even bothering to secure the S3 bucket containing said IDs, the latter just not deleting them. Given that Roblox has also had its own issues with their moderation panel being socially engineered, why should I trust Roblox to be any different here and actually delete the IDs after they have been used for their legitimate reason.

2. Age estimation is easy to bypass and has tons of false flags

We all know about fake IDs right? How its used to get into bars much younger than you should. It is going to happen to Roblox as well. You could have all the systems in the world to detect fake IDs, but false positives mean some people will slip through the cracks with them. On age estimation, some age estimators think I am 13, I am almost definitely not but this is the problem with using AI to guess someone's age. What will happen to other

older users on Roblox who look a lot younger than they actually are?

TikTok^[3] banned KreekCraft because it thought he was a lot younger than he actually was. I really hope you do not do this here because your moderation panel will be flooded with appeals if such, and it is quite well known that Roblox's moderation panel seems to just reject every appeal it gets.

3. Focus on actual problems, do not take away social features

Condos are still a problem. There are still users trying to get people off-platform through other means. I do not see how taking away social features to non-verified users will change anything here. Discord exists, and I am sure many predators would be fine spinning up a Studio instance with an experience that violates Roblox's Terms of Service while talking inappropriately on Discord, effectively bypassing the whole purpose of locking away chat features.

I do believe under 13 users should not be able to chat on Roblox, it was wild that they could to begin with. But age estimation as a new way to prevent that will not change anything. Predators are slimy, they will find bypasses.

The site has a filter, it may not be perfect, but any filter is better than none. Discord on the contrary, does not have a filter, and its moderation practices are a lot worse than Roblox.

Note on Legal Compliance

I am aware that some countries have introduced digital safety laws, such as the United Kingdom's Online Safety Act, or Australia's act of the same name. These acts are insanely overreaching and should be pushed back against at any opportunity given. I do not expect Roblox to fight against this legislation given it would be years of legal battles, but this should only be rolled out to countries where it is actually required, and not the platform as a whole.

Conclusion

In conclusion, I do not see how age estimation will change anything, when sites like Discord exist, and I can think of other ways myself that this could be bypassed by predators.

Sources

1. *Tea app hacked: 13,000 photos leaked after 4chan call to action, 2025, NBC* - <https://www.nbcnews.com/tech/social-media/tea-app-hacked-13000-photos-leaked-4chan-call-action-rcna221139>
2. *ID photos of 70,000 users may have been leaked, Discord says, 2025, BBC* - <https://www.bbc.co.uk/news/articles/c8jmzd972leo>.

3. *KreekCraft almost loses TikTok account after app mistakenly says hes under 13, Dexerto, 2025* - <https://www.dexerto.com/tiktok/kreekcraft-almost-loses-tiktok-account-after-ai-says-hes-under-13-3236652/>