

Oral Submission of Karen White
Twitter International Company
Oireachtas Committee on Justice and Equality
9 October 2019

I would like to thank the Committee for your invitation to Twitter to participate in today's session. My name is Karen White, I am Director of Public Policy for Twitter in Europe. I am joined by my colleague Ronan Costello, Public Policy Manager for Twitter in Europe.

Twitter is an open, public service. Our singular mission is to serve the public conversation. We serve our global audience by focusing on the needs of the people who use our service, and we put them first in every step we take.

Twitter is committed to improving the collective health, openness, and civility of public conversation on our platform. Our success is built and measured by how we help encourage more healthy debate, conversations, and critical thinking. Conversely, abuse, malicious automation, and manipulation detracts from it.

Controlling the Twitter Experience

We provide people on Twitter with a range of tools so that they can control their experience and manage the types of content and accounts they see. These range from being able to keep your account private, to blocking, muting or reporting other individuals on the service.

We also give people control over what they see in search results through safe search mode which, enabled by default, excludes potentially sensitive content from the search results, such as spam, adult content, and the accounts an individual has muted or blocked.

Online Abuse and Hateful Conduct

Twitter strives to provide an environment where people can feel free to express themselves and we recognise that if people experience abuse on Twitter, it can jeopardise their ability to do this.

Our policies outline that an individual using the platform is not permitted to promote violence against or directly attack or threaten other people on the basis of a range of protected categories. You may not engage in abusive behavior which is an attempt to harass, intimidate, or silence someone else's voice.

We do not allow individuals to use hateful images or symbols in their profile image or profile header. Individuals using the platform are not allowed to use their username, display name, or profile bio to engage in abusive behaviour, such as targeted harassment or expressing hate towards a person, group, or protected category.

Under this policy, we take action against behaviour that targets individuals or an entire protected category with hateful conduct.

Twitter's Efforts to Combat Self-Harm and Suicide

After we assess a report of self-harm or suicide, Twitter will contact the reported user and let them know that someone who cares about them identified that they might be at risk. We will provide the reported user with available online and offline resources and encourage them to seek help. In response to certain keyword searches relating to these issues, we direct individuals using Twitter's search function to online prevention resources. This is available in Ireland, where we have partnered with the Samaritans and direct individuals to their website and support services.

Twitter's Policies on Non-Consensual Nudity

Twitter does not allow individuals on the service to post or share intimate photos or videos of someone which were produced or distributed without their consent. This is sometimes referred to as revenge porn and this content poses serious safety and security risks. We inform our users that sharing explicit sexual images or videos of someone online without their consent is a severe violation of their privacy and the Twitter Rules. .

Results of Our Efforts

People who don't feel safe on Twitter shouldn't be burdened to report abuse to us. Earlier this year we made it a priority to take a proactive approach to abuse. Today, by using technology, 38 percent of abusive content that's enforced is surfaced proactively for human review.

We have made meaningful progress in creating a healthier service. Since we announced our focus on improving the health of the conversation occurring on Twitter, we have seen:

- A 16 percent year-over-year decrease in reports from people about interactions with other users allegedly abusing them on Twitter.
- A 45 percent increase in the number of account suspensions for those users who attempted to create new accounts following their original suspension.

- Over the first quarter of 2019 this amounted to over 100,000 account suspensions for these "reoffenders".
- 3 times more abusive accounts suspended within 24 hours after a report.
- 2.5 times more private information removed with a new, easier reporting process.

Appeals

We have always provided an appeals mechanism for individuals who may have had enforcement action taken against their accounts but earlier this year we launched a new feature that allows people to appeal right within the Twitter app. This change has meant that we've been able to respond to people 60% faster than before.

Working with Law Enforcement

We have well-established relationships with law enforcement agencies, including An Garda Síochána. We have continuous global coverage to address reports from law enforcement around the world and have a dedicated online portal to swiftly handle law enforcement requests. We provide regular trainings on Twitter's policies and procedures for working with law enforcement which are also publicly available on our website entitled "Guidelines for Law Enforcement".

Legislation on Harassment, Harmful Communications and Related Offences

The Committee has asked for our recommendations on proposed legislative changes. It is important that the legislation is as consistent as possible with existing legal frameworks to avoid uncertainties and discrepancies. This is particularly pertinent with reference to Articles 12 to 15 of the e-Commerce Directive and Article 25 of the Directive on combating the sexual abuse and sexual exploitation of children and child pornography.

The effectiveness of any legislative solution relies on it being proportionate, technically feasible, and flexible, particularly given the diversity of companies within the digital ecosystem. In this context, the Committee will need to consider and assess how different legal and illegal harms manifest themselves across different platforms and varied jurisdictions.

We all share the objective of protecting our systems of due process and our commitment to freedom of expression. Preserving these tenets in regulatory proposals can be achieved by collectively ensuring there is clarity on the obligations of all stakeholders, thereby avoiding an outcome whereby companies could overreach and erroneously remove content that should otherwise be kept online. A clearly defined scope will assist Twitter and others.

In order to ensure people can continue to express themselves freely and safely on Twitter, we know we must continue investing further in proactive technology, safety tools, as well as developing policies which keep pace with the changing contours of public conversation.

We stand ready to work with the Committee as we continue to explore options to ensure that all people are protected from online harassment and harmful communications.

Thank you.