**DCU**
Ollscoil Chathair
Bhaile Átha Cliath
Dublin City University

## About this research study

Online bullying and other online harms manifest in diverse forms on social media apps like TikTok, Snapchat, Instagram, YouTube, X[1]. While few online safety tools claim to assert superior moderation of harmful content, there lacks a high quality benchmark for evaluating online bullying prevention tools. Through this week-long activity, the participant will aid in developing real-life test cases to evaluate such tools.

## Taking part in this research means?

In this study, the junior research assistant will aid in generating authentic test cases to evaluate existing automated detection systems for hate speech or offensive content in a more realistic manner. The participant will also adapt and assess realistic cyberbullying scenarios to suit present-day social media platforms like TikTok, Instagram, Snapchat, YouTube, Discord, WhatsApp.

## Benefits of participating?

Your involvement aids in the development of tools to effectively respond to mean or hurtful online content. By participating, you contribute to a better understanding of creating improved guidelines for ensuring user safety on social media apps.

## Voluntary Participation

Please note you can withdraw from the research activity at any point of time. Participation in the research activity is entirely voluntary. Your voluntary contribution will help independent researchers design and develop tools that can help tackle online harmful content.

## Who are we?

We are researchers working on the project Data-driven Toolkit to Combat Cyberbullying among Teens. Our project includes a group of experts, Kanishk Verma (PhD Student), Dr Brian Davis at Dublin City University, Dr Tijana Milosevic at University College Dublin, and Dr Rebecca Umbach at Google. This project is supported by the Irish Research Council and Google under the grant number ESPSG/2021/161.

We hope you'll consider participating in our study to help make the online world a better place for everyone !

---

[1] Formerly Twitter