DCU
Ollscoil Chathair
Bhaile Átha Cliath
Dublin City University

Plain Language - Job Description

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 gain foundational knowledge of computer programming and 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 play a key role in developing authentic test cases aimed at evaluating current automated systems designed to detect harmful or abusive online content. Beyond test case creation, the assistant will critically analyse and adapt cyberbullying scenarios from academic literature to ensure their relevance and applicability to modern social media platforms such as TikTok, Instagram, Snapchat, YouTube, Discord, and WhatsApp. Furthermore, the junior research assistant will acquire foundational skills in Python programming and prompt engineering, equipping them with essential competencies for thriving in the AI-driven landscape. This experience will not only enhance their technical abilities but also provide valuable insights into the ethical and social implications of AI in online environments.
Please note that this work may involve exposure to some potentially offensive content; however, these examples will be textual in nature, benign, and similar to what one might typically encounter on social media.

Benefits of participating?

Your involvement directly contributes to the creation of tools that effectively address harmful online content, playing a crucial role in enhancing user safety on social media platforms. By participating, you'll help shape improved guidelines and strategies for protecting users from online harassment. Additionally, you'll gain valuable skills in areas like AI, equipping you to navigate and excel in the rapidly evolving digital landscape. This experience not only empowers you with technical expertise but also places you at the forefront of efforts to create a safer, more responsible online environment.

[1] Formerly Twitter

DCU
Ollscoil Chathair
Bhaile Átha Cliath
Dublin City University

Plain Language Statement

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. Students are free to withdraw from the study at any time without incurring any penalties or academic consequences. This study has received full ethics approval from the DCU Ethics Committee, and all necessary support is in place to ensure the well-being of participants. A list of available resources can be accessed at any time via the provided website, and participants will also receive a copy of these resources. The researcher will be available throughout the study to ensure participants' well-being and provide any necessary support.

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.