# Research finds online hate speech is often hidden in plain sight

*The study found that Afghan nationals were repeatedly targeted by derogatory comments on Instagram despite the platform utilising AI-powered detection tools*

CAMBRIDGE, UNITED KINGDOM, February 7, 2023 /EINPresswire.com/ -- New research indicates the scale of online hate speech may be worse than previously feared due to the implicit ways that some content is phrased.

The study, led by Dr Vahid Parvaresh of Anglia Ruskin University (ARU), found that the majority of posts which meet the description of hateful content do not contain overtly derogatory words or expressions. Instead, hateful sentiments are frequently conveyed through covert or modified wording.

a.r.u.

This inventiveness exhibited in the subtle uses of language to express hate online is in many cases believed to be a conscious act designed to lower or remove the risk of detection, in the same vein as the much more blatant practice of inserting intentional typos into disparaging comments.

The findings, published in the Journal of Pragmatics, are informed by instances of online hate speech directed at Afghan nationals in the comment sections of 58 Instagram posts between 2019 and 2021.

Each post was specifically selected for containing a news item about Afghanistan or the Afghan people, who have experienced one of the world's most extreme displacements in recent times.

Meta, Instagram's parent company, says it uses artificial intelligence to detect the spread of

online hate, with unclear cases referred to review teams. On its platforms hate speech is defined as a "direct attack against people – rather than concepts or institutions – on the basis of what we call protected characteristics: race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease".

Despite this, the study contains numerous examples of comments which, while free of markedly hateful expressions, are unarguably discriminatory against the person being targeted. Such posts, translated from their original language, include:

"If you are that unhappy, why don't you pack up your stuff and just go back? Every day there are more of you. Thanks to you, there is not even a single house available in the villages and in the outskirts of the towns. You make a lot of money and then keep nagging. If you are so unhappy, stay in your country and make it better."

And:

"With so much tragedy and war, how come Afghans have not died out yet? ! !"

The study's author, Dr Vahid Parvaresh, Associate Professor of Linguistics at Anglia Ruskin University (ARU), said:

"Due to the availability and spread of computer-mediated communication, instances of hate speech have been on the rise. This is continuing despite varying levels of intervention from social media companies and governments around the world.

"It is especially true for more innovative and creative ways of expressing hate. It seems that these covert methods are allowing haters to avoid detection online and a primary motivation behind my research is to raise awareness of this fact.

"In order to tackle hate speech, more attention needs to be paid—by all stakeholders—to covert and implied ways of expressing hate."

David Styles
Anglia Ruskin University
david.styles@aru.ac.uk

This press release can be viewed online at: https://www.einpresswire.com/article/615612771