Study reveals flaws in using social media to identify race and ethnicity for health research

May 27 2022

Credit: Pixabay/CC0 Public Domain

By using social media data, researchers can get insights into patients' experiences that are often overlooked or otherwise difficult to attain.

Social media also provides data more quickly than traditional epidemiological studies, which can take years to complete.

Trends

If researchers are able to identify key demographics of social media users they can help identify who is over or underrepresented and can help spot trends and gather information on the views and experiences of diverse groups.

While previous studies have looked at extracting or estimating features such as location, age, gender, language, occupation and class, this study is the first comprehensive review of the methods used to extract race or ethnicity.

The authors of the study, led by the University of York and published in the *Journal of Medical Internet Research*, say they have identified ethical concerns and doubts over the reliability of using Twitter to evaluate ethnicity or race from a user.

They say if the limitations are not addressed it could diminish the value of the information found.

The study, which looked at Twitter users in the U.S, found that researchers often rely on the information from the bios of Twitter users, as well as photos and the tweets themselves, often analyzing the language used and identifying any self-declarations.

Bias

Lead author, Dr. Su Golder from the University of York's Department of Health Sciences, says that "extracting race and ethnicity of Twitter users is particularly important to identify trends, experiences and

attitudes of racially and ethnically diverse populations."

"But we need to be sensitive to the methods used and mindful of bias. For example, when researchers look at photos they have an inclination to perceive their own race and using someone's name has disadvantages for women who often take their partner's name."

The study authors say another problem is the over-simplification of race into Black, White and Asian.

The researchers propose several new ways to improve identifying race or ethnicity from social media, including more representative research teams, a mixture of manual and computational methods and identifying self-declarations.

Disparities

Dr. Golder added that "the appeal of Twitter data is clear as it is one of the largest public facing social media platforms, with an ethnically diverse user base."

"However, the promising insights that can be derived from Twitter data are often limited by what is missing, specifically basic socio-demographic information of each user."

"In order to use social media and digital health research to address disparities, we need to know not only what is said on Twitter, but also who is saying what."

Provided by University of York

Citation: Study reveals flaws in using social media to identify race and ethnicity for health research (2022, May 27) retrieved 1 April 2023 from https://medicalxpress.com/news/2022-05-reveals-flaws-social-media-ethnicity.html