**Instructions for Online Toxicity Annotation**

Attached to this email is a CSV with 1,000 tweets. We would like you to annotate whether you believe they are toxic. Please assign each tweet a value of '1' if you think it is toxic and '0' if you think it is not.

Toxicity is defined as: "*a rude, disrespectful, or unreasonable comment that is likely to make you leave a discussion*."

Please do not try to guess about how what proportion of the tweets should be considered toxic and do not make any assumptions about whether these tweets are representative of Twitter in general. Judge each tweet solely on its own merits.

If you have any questions, please reach out to the study authors.