Annotator Guidelines for HatemojiCheck

What is Hate?
The United Nations define hate speech as "speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor."

Hate speech has two key components:

- Hate involves expressing something negative, such as contempt, disparagement, derogation, demonization, dehumanising comparisons, expressions of harm or bias.

- Hate is directed against the identity of a group where identities can be understood as the social groups and affiliations that individuals belong to and are associated with. Protected identities include:

    – Religion

    – Race

    – Ethnicity

    – Gender

    – Sexuality and sexual preference

    – Immigration status

    – Nationality

    – Ableness / disability

    – Class

Put simply, hate speech is **abuse that is targeted at a protected group or at its members for being a part of that group**.

Annotation Task
Before starting the task please check you have the most recent version of your browser (e.g., Chrome, Firefox). This is required so the emoji render properly on your screen.

The task will be hosted on an individual survey link. You can stop the annotation work at any time, and when you re-open the link, you will automatically continue from where you left off.

You can also use the keyboard shortcuts 1,2,3,4 and 'Enter' to select and submit responses. If an item is too objectionable or disturbing for you to annotate, skip it by pressing '0'.