

# Volunteers for dataset labeling

The dataset is comprised of three possible classes:

- Hateful
- Benign (Not Hateful)
- Not a meme

## Let's start with NOT A MEME

An image is not a meme when:

- It is not comprehensible for an English speaker (contains text in Russian, Spanish .. etc). Please note that a meme could contain a Russian text for example. We still need to consider it if its target is the English speaking audience.



- The image does not contain a text:



## A HATEFUL MEME

A hateful meme could be any of the following:

- The meme references something that is hateful and could be found offensive for some people. For example:



This references the holocaust

- It attacks specific people in a hateful manner



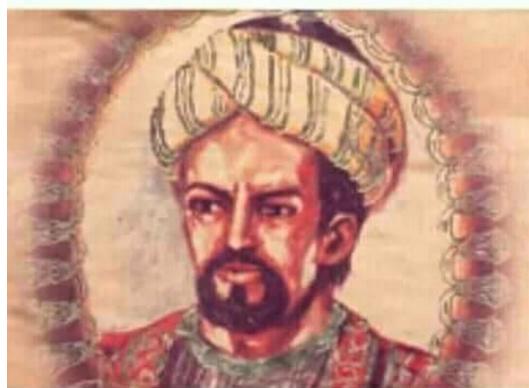
This one attacks LGBT individuals

- You can try to find a way that you can get offended by this meme and if you feel that there is a group of people that could be offended by this meme you can define it as hateful.
- 

## BENIGN MEMES

Any meme that has English text and does not attack anyone in a hateful manner is benign. For instance, calling someone stupid in a funny manner can be labeled as benign.

A wise Arabic poet once said:



أَلِقْتِي مِنْ هُسْ بِلْ عَنْ هُقْ



## HOW TO USE THE APP

Volunteer id field: you need to write your id (it could be anything just use it as a login to conserve your progress state). After you write your id you need to press enter either on the keyboard or mobile keypad. (Andriod or web app);

