Technology Demo. Annotation Tool for Tweets

Gunther Jikeli, Deepika Awasthi, Daniel Miehling, Denizhan Pak

Indiana University, Bloomington IN 47405, USA
{gjikeli, dawasthi, damieh, denpak@iu.edu}@iu.edu

Manual annotation is the most time-consuming part in building an annotated corpus that can be used as a gold standard for machine learning programs. Annotation tools accelerate the annotation process and can improve accuracy. We needed a tool that visualizes each tweet of our samples and a form that allows annotators to quickly answer questions about the tweet.

There are many annotation tools available, mostly for text annotation but increasingly also for the annotation of images. (Neves and Ševa, 2019) However, none of them shows the post in their natural environment with a few clicks and provides a form that annotators can click through easily.

Our "Annotation Portal" is designed to facilitate the annotation and to reduce annotation errors. It can be used remotely by any registered users who can then export their annotations in CSV format and use them for their own analysis. The Annotation Portal is accessible at https://annotationportal.com. The Portal is a website built with Django Framework. The server is hosted on Indiana University's Jetstream supercomputers. New annotators register at the welcome page with a

username, password, and email address. They choose the annotation type with different annotation forms related to different themes, such as antisemitism, racism, or anti-Muslim hatred. The forms can easily be adjusted by adding or changing questions. Annotators then select a sample from the sample sets we uploaded previously. This generates an annotated sample associated with the annotator with the default selections for each tweet in the selected sample.

The Annotation Portal links to live tweets of predefined samples by referencing their ID number. It does not use deleted or suspended tweets so that it is in compliance with Twitter requirements. Changes in the annotation for a tweet are registered when the annotator hits the next button. The annotator can go back to each tweet and change the selections they made at any time. When logging in again after a break the annotator will be reverted to the tweet where they left when they logged out.

If annotators are unsure of the meaning of the tweet they can click on the tweet which then opens a new tab in the browser and reverts to the tweet on the Twitter page where annotators can see the entire thread, replies, likes, and comments. Annotators can export their annotation at any time in a CSV file. We have tested the Portal with lay annotators and professional annotators and produced a labeled dataset of 4,137 tweets (labelling them antisemitic or not antisemitic, among other annotations).

We will present the frontend and backend of the Portal.