

Tagging Instructions

Our model is a supervised learning model. That means we have to provide it with pre-tagged data *that we're all sure about* in order for it to learn properly. As such, the team is going to work together to tag a large amount of training data.

I've found a free and seemingly fully-featured solution [here](#). (That's my fork, since I had to make some changes to get it to work).

Step 0: Setup

There's a good bit of setup required to get our tagging software to work. The good news: you only have to do it once. Even better, this is going to make your tagging so much faster and easier.

Repo

Start by cloning [this repo](#).

Docker

You'll first need to install [Docker](#). It's software that allows us to set up "containers", which are basically mini-VMs. That means we won't have to worry about incompatibility on our computers, and you'll (hopefully) never have to deal with issues because of your specific device.

Step 1: Running the App

To run the app, simply navigate to the repo you cloned above and run `docker-compose up`. That should spin up both of the containers you need.

Step 2: Tagging:

You'll be assigned some number of text files. Upload them to the site and you'll be able to tag them.

Start by adding categories for "SKILL" and "ATTRIBUTE". All caps, to make sure we're all writing the exact same thing.

Then, you can tag by selecting the category you want, then highlighting the word that matches that category. Switch between categories as needed. Once you've confirmed that everything is tagged properly, click "Save".

You'll be given an alert at the very end telling you that you've gone through all sections. Once you're there, click the "export" button.