

Annotation Guidelines

Dota 2 Message Classification for Auto-Moderation Purposes

Richard Motorgeanu - motorger - 400436012
Jacquelyn Pohl - pohlj - 400455088
Nicholas Gyorgypal - gyorgypn - 400446785

February 2, 2025

Content Warning: This document deals with the categorization of toxic content and messages found online that contain profane, vulgar, or offensive content. Proceed with caution.

1 Overview

A concise overview of the annotator's job

2 Annotation Guidelines

The set of labels that can be applied with descriptions of their meaning (for each item in a numeric scale, or for each class)

Unambiguous rules or criteria that should be followed in order to label the data (what is the thinking process you should go through to label the data)

2.1 Examples

At least 1 example of each label along with an explanation for why that is the correct label (and if helpful, why the other labels would not apply to that instance)

3 Contact Information

- Richard Motorgeanu - motorger@mcmaster.ca - Can contact on Teams as well.
- Jacquelyn Pohl - pohlj@mcmaster.ca - Can contact on Teams as well.
- Nicholas Gyorgypal - gyorgypn@mcmaster.ca - Can contact on Teams as well.