

# Guidelines for Sentiment Annotation

**Purpose:** The purpose of this annotation task is to categorise the use of profanity in a specific context into one of three sentiment categories: Positive, Neutral or Negative. For this task, ‘profanity’ is defined as follows:

*“a swear word, also known as a curse word, profanity, or expletive, is a word or phrase considered offensive vulgar or taboo in a particular culture or context.”*

This may include explicit language (e.g. fuck, shit, etc.), slurs/taboo terms (e.g. the n-word, faggot, etc), sexual references (e.g. pussy, dick, etc.), or milder exclamations (e.g. hell, crap, etc.). An additional fourth annotation category may be used for words that were recognised as swear words but in fact are not (see definitions and examples below). For ease of annotation, the following tags should be used to annotate each line in its respective category:

0 - N/A

1 - Positive

2 - Neutral

3 - Negative

## Definitions & Examples

- **N/A:** Some words might have been recognised as profanity by the program, but are either actually just the use of the word as a noun without implying anything profane, or are in fact not profanity. For example:

*“yeah he do the walk of life  
here comes Johnny HELL tell you the story  
hand me down my walkin shoes”*

The word ‘hell’ is recognised as a swear word, however, due to the removing of punctuation in the preprocessing of the data, this was actually ‘he’ll’ (he will) and so, is not actually a swear word. Additionally, in a different context, the word ‘hell’ might just be used as meaning ‘a place or state of misery and torment’, rather than being used to indicate something profane.

- **Positive:** Contexts that express an uplifting, supporting or favourable emotional state.

For example: *“anyway  
its bad BITCH oclock yeah its thick thirty  
ive been through a lot but im still flirty okay”*

Here, the word ‘bitch’ (specifically in combination with ‘bad’) is used positively as a term of empowerment and self-confidence.

- **Neutral:** Contexts that do not display a clear positive or negative sentiment. Profanities in these contexts may be used as intensifiers or in adjectival form. If the annotator is unsure and the context is ambiguous, the lines should be noted as neutral.

For example:       “feasible rhymes that you can vision and feel  
                          conducted and formed this is a HELL of a concept  
                          we make it hype and you want to step”

In the phrase ‘a hell of a concept’, the use of ‘hell’ means that the concept being referred to is particularly impressive, remarkable, or powerful. It's used to emphasise the strength or quality of the concept being discussed, not as an expression of positive or negative opinion about the concept itself.

- **Negative:** Contexts that express dissatisfaction, disapproval, anger or overall negative sentiment.

For example:       “this shit wicked to get their names buzzin  
                          some NIGGAS just gotta go lay in a tomb  
                          and media thirsty for clicks”

The use of the n-word here refers to the idea that that some individuals engage in unethical behaviour to gain fame or notoriety, and is thus used as criticism or disapproval.