# Annotation guidelines

## Emotion detection in Dutch tweets during the COVID-19 pandemic

**The goal of the annotation task**

The annotation goal is to annotate Dutch tweets with emotion tags. The annotated data will be used during the course to train machine learning models and then test them on automatically predicted emotions. The tweets that will be used for annotation are tweeted during the COVID-19 pandemic period from February 1 till April 11. Besides annotating the data for machine learning, it is also interesting to quantify emotions in the pandemic tweets.

**What to annotate**

Tweeters often use tweets to express emotions towards people, events, decisions, and so forth. Sometimes emotions are not explicitly expressed in tweets, but they still can be deduced to some extent. Additionally, a tweet might trigger emotions in the reader that might not be explicitly intended by the tweet. The annotation task is to identify emotions conveyed or triggered by the tweets (even if they are not explicitly stated) and mark the words triggering the emotions.[1]

**Emotion tags (etags)**

We employ 1+9 tags for emotions. Nine of the tags are named after the eight basic emotions proposed by Plutchik, plus one etag for *Caring*. The **Other** etag is used for the emotions covered by these nine named etags.

**Anger** (annoyance, frustration, rage)
**Anticipation** (expectation, interest, vigilance)
**Caring** (affection, empathy, friendliness, love)
**Disgust** (disinterest, dislike, hate, indifference)
**Fear** (anxiety, apprehension, concern, panic, terror, worry)
**Joy** (calmness, ecstasy, happiness, serenity)
**Sadness** (disappointment, grief, guilt, pensiveness, pessimism, regret, sorrow)
**Surprise** (amazement, confusion, distraction, indecision)
**Trust** (acceptance, admiration, confidence, hopefulness, like, optimism)

The etags account for more general classes of emotions than their names suggest. For example, *dislike*, *hate*, and *indifference* all fall in the class of **Disgust**. To help better interpret each emotion class, in the annotation tool every etag is followed by an incomplete list of sub-emotions. By expanding the etag ⊞ you can see definitions of some of the sub-emotions (which non-native speakers of English might find useful).

In case you have difficulties to classify an emotion X of a tweet with the provided nine named etags, use the **Other** etag, but also provide the name of the emotion X as a note in brat. While giving the emotion name, try to use one of the emotions found on this page[2]. Note that

[1] ~~All emojis and emoticons in tweets will be masked with ▉ symbol. This allows annotators to focus on the linguistic content of tweets to identify emotions. Also, the links to the original tweets are removed to encourage the annotators to consider only the content of the tweets and to prevent them from reading the original unmasked tweets and their contexts.~~

[2] Before providing the emotion name as a note, check (search with Ctrl+F) whether it is mentioned in the provided page.

**Other** can be something else than *irony*, for example, *distrust*, *greed*, *envy*, *jealousy,* etc. After watching the video tutorial, many annotators interpreted **Other** only as *irony*. Provide a lower-cased noun word for emotion X, e.g., *happiness* is prefered over *happy*. Use the note field of the **Other** etag only for an emotion name and don't provide additional explanations for it.

**Confidence level for etags**

Provide a **confidence level** (high or low) for each etag. By default, a confidence level is set to High. Note that confidence level is not about the intensity of emotion, but about how confident you are that there is a particular emotion in a tweet. For example, you can be highly confident that Joy is mildly/weakly expressed in a tweet.[3]

**Main etag**

For each tweet with a high-confidence (⇧) emotion, mark one of the ⇧-etag as main which expresses the main emotion of the tweet. The main status of an etag can be chosen from the confidence values. The main etag is marked with ★. All spans of the main emotion should be marked with ★. In other words, ★ means ⇧and additionally marks the main etag of the tweet.[4]

**Selecting spans for etags**

The etags are applicable to spans, where spans are contentful strings containing no white space.[5] The same etag can be applied to one or more spans (to apply an etag to several spans use Add Frag. in the annotation window of brat), and the same span can trigger more than one emotions.

**Which spans to select?** Spans are emotion indicators. Select spans in a top-down fashion. First, understand a tweet and identify emotions it carries. Then, find those minimal spans that are triggering the emotions. The emotional trigger spans are such that when these words are ignored/hidden, the twitter stops carrying (or hardly carries) the emotion.[6] For example, I don't like you expresses/triggers Disgust, and this is done by the spans *don't like* because they trigger the emotion together. Depending on their contribution to emotions, #hashtags and @handles can also act as emotional triggers (during span selection, # and @ symbols must be excluded). Remember that spans can get more than one etag if they trigger more than one emotion.

**Annotating a tweet**

If a tweet carries or triggers no emotion, then the tweet shouldn't contain any span marked with an etag. In the end, used etags have to reflect the overall emotion carried by the tweet. Each tweet will get possibly multiple labels based on the used etags. A tweet gets a label 'no emotion' if no etag occurs in its annotation. Feel free to provide concise and clear notes via brat's annotation window.

---

[3] The confidence level will be used to check whether keeping only high-confidence etags lead to high inter-annotator agreement.

[4] Marking the main etag helps to frame the emotion detection as a classification task (one class per tweet) instead of labelling (possibly multiple labels per tweet).

[5] When selecting spans don't include punctuation marks if possible. For example, trailing punctuations must be omitted (e.g., happy. ) but word-internal symbols should be a part of a span (e.g., self-confidence and can't ).

[6] Also, think about an emotional trigger span as an explanation for the emotion.

Use only the content of the tweet when annotating it. For example, don't follow URLs mentioned in the tweet.

**Samples**

To give you a better intuition about the annotation, the following link has several tweets annotated based on this guideline:
https://naturallogic.pro/brat/git/index.xhtml#/Anno4ML/covid-19/demo

**Checklist for finalizing annotation**

- If a tweet is expressing an emotion, then there should be at least one span with the corresponding etag.
- Each etag has either a confidence level (⇧ or ⇩) or the main status ★, which are depicted with symbols preceding the etag.
- Spans shouldn't include white space and # or @ in cases of hashtags and handles.
- **Other** etag must have the name of a specific emotion as a note.

**Remember, this guideline provides an incomplete interpretation of the etags and no guidelines can be complete about the emotion annotation. Therefore, it is normal to have difficulties while annotating certain tweets. Identifying pre-defined emotions in tweets is a notoriously difficult task. Use your intuition and interpretation of a tweet while annotating emotions. It is important to be consistent with your annotations.**

# Ignore the rest

## Examples:

http://ccc.inaoep.mx/~villasen/bib/Identifying%20Expression%20of%20Emotion%20in%20Text.pdf \

## References:

https://github.com/collab-uniba/EmotionDatasetMSR18/blob/master/Emotion%20Annotation%20Guidelines.pdf
http://groups.inf.ed.ac.uk/ami/corpus/Guidelines/EmotionAnnotationManual-v1.0.pdf
Sentiment, emotion, purpose, and style in electoral tweets
Identifying Expressions of Emotion in Text

Ekman:
https://www.aclweb.org/anthology/P19-1096.pdf
https://www.aclweb.org/anthology/S07-1013.pdf
https://dl.acm.org/doi/pdf/10.1145/1363686.1364052
https://www.aclweb.org/anthology/W18-6206.pdf
http://ccc.inaoep.mx/~villasen/bib/Identifying%20Expression%20of%20Emotion%20in%20Text.pdf mixed + no
https://www.aclweb.org/anthology/H05-1073.pdf Su+, Su-
https://www.aclweb.org/anthology/S12-1033.pdf

Plutchik:
https://www.aclweb.org/anthology/W10-0204.pdf
https://www.aclweb.org/anthology/S18-1001.pdf +love + optimism + pessimism
https://www.aclweb.org/anthology/L18-1199.pdf