

Annotation Guidelines for Labeling Emotion in Multilingual Tweets

Version 1.0

LT3 Technical Report – LT3 23-12

Pranaydeep Singh, Aaron Maladry and Els Lefever

LT3 – Language and Translation Technology Team
Department of Translation, Interpreting and Communication
Ghent University

All rights reserved.

LT3, Faculty of Arts, Humanities and Law, Ghent University, Belgium.

December 16, 2023

ANNOTATION GUIDELINES

EXALT : Emotion for X-linguAL Tweets

Pranaydeep Singh, Aaron Maladry, Els Lefever

1. Introduction

In this project, we aim to understand how emotions are expressed differently in various languages on social media. In this initial study we work with English, Dutch and Slovene, annotating tweets from the year 2017. We also introduce a detailed layer to assist us in our goal of language-independent emotion understanding, namely *Trigger Words*. These are words or spans that help an annotator determine the correct emotion in the tweet. The annotations will be done using the [Inception](#) toolkit.

For each tweet, we aim to have an *Emotion Tag*, followed by a set of *Trigger Words*. For English we aim to have 5000 annotated tweets, and for Dutch and Slovene, 2000 annotated tweets. This is excluding the *Discard* category since those tweets will be removed from the final dataset. For any further questions please contact Pranay (pranaydeep.singh@ugent.be) or Aaron (aaron.maladry@ugent.be).

2. Emotion Categories

For this project we will be working with 5 basic emotion categories: Joy, Love, Anger, Sadness and Fear. Inherited from [An emotional mess! Deciding on a framework for building a Dutch emotion-annotated corpus](#). A 6th Neutral category is added to indicate tweets where no emotions are expressed. And finally, a Discard category is added to indicate other issues, listed in more detail below. The goal is to focus on the user's perspective, so the emotion indicated should represent the feelings the 'tweeter' attempted to convey and not just based on Emoji's or words listed in the text. Below you can find helpful related terms and examples to distinct between the emotions. In case of confusion or uncertainty, you should make sure at least 1-2 of the words in the related emotions fit into the tweet's expression.

Anger

Related emotions: *rage, outrage, fury, wrath, hostility, ferocity, bitterness, hate, loathing, scorn, spite, vengefulness, dislike, resentment, disappointment, dismay, displeasure, disgust, revulsion, contempt, frustration, exasperation, irritation, aggravation, agitation, annoyance, grouchiness, grumpiness*

Example:

@someuser Get. Him. Out. I don't care what policies he pursues, he's a madman and we need to get. him. out.

Fear

Related emotions: *alarm, shock, fright, horror, terror, panic, hysteria, mortification, nervousness, anxiety, tenseness, uneasiness, apprehension, worry, distress, dread*

Example:

However much this scares me.. it's obviously nothing to how much it must scare you. 10 years ago is crazzzzzy @someuser @someuser <http://someurl>

Joy

Related emotions: *amusement, bliss, cheerfulness, gaiety, glee, jolliness, joviality, delight, enjoyment, gladness, happiness, jubilation, elation, satisfaction, ecstasy, euphoria, contentment, pleasure, enthusiasm, zeal, zest, excitement, thrill, exhilaration, optimism, eagerness, hope, pride, triumph, relief*

Example:

@someuser Thank you for using the term constitutional republic and not "our democracy " wish our "leaders" would do the same. Well done!

Love

Related emotions: *adoration, affection, fondness, attraction, caring, tenderness, compassion, sentimentality, enthrallment, rapture, longing, lust, arousal, desire, passion, infatuation, sympathy*

Example:

@someuser Thank you mate lots of tracks ready for the new year so I'll keep you posted!

Sadness

Related emotions: *depression, despair, hopelessness, gloom, glumness, unhappiness, grief, sorrow, woe, misery, melancholy, pity, sympathy*

Example:

From my perspective most people find me heavily captivating with this kind of magnetism til they suddenly don't and I'm last weeks philosophy...I'll write a book about how charm has an expiration date.

Neutral

If none of the above options are adequately represented, then the tweet must be indicated as Neutral. Types of Neutral tweets can be:

1. Stating a fact
2. Presenting an opinion without any indication of emotion
3. Customer service communication asking questions
4. Tweets from Bots

And many more such categories.

Note: Beware that tweets with issues or ones you are unsure about are not Neutral. Check the Discard section below to make sure tweets with issues are indicated correctly into that category.

Examples:

Finished with The Wishing Spell, by Chris Colfer <http://someurl> (Emotionally Neutral)

@someuser @someuser Shine had already gone onto his edge to make the turn. Him removing his hand was to make himself big and eliminate options for the player with the puck, and it looks like he tried to brace himself and the goalie when the goalie made contact. (Presenting an argument/fact without emotion)

California city official is gunned down while vacationing in Mexico <http://someurl> #FoxNews (Presenting a NEWS update)

Discard-General

These are tweets that we would not want to use in our analysis. So, one must be careful about indicating too many tweets as 'Discard', since these would not be useful for us.

Tweets to be indicated in this category could have some of the following issues:

1. If you feel like there is an Emotion present but none of the labels adequately fits the text
2. If more than one emotion category fits perfectly and there is no justification for indicating only one of them
3. Wrong language or doesn't make sense
4. Refers to an attached image or requires an image to understand
5. If you are unsure because the tweet requires some additional context or information of a topic/conversation that you are not aware about. Eg: Politics in a specific country, etc.

Note: There can be certain instances where you are conflicted between Neutral or Discard. In these cases, it's always better to pick Discard rather than have a wrongly annotated Neutral tweet in the final version.

Examples:

*@someuser Thats a very large 0. I suspect youre right. But wait... Is a large 0 *more* or *less* evidence than just plain 0? (Missing context)*

When I first saw Michael B. Jordan in The Assistants I thought he was cute. Then he went through a meh and ugly phase for some years and now hes a daddy. (Multiple emotions)

Discard-Irony

These are also tweets which will be discarded from the final set because of Ironic or Sarcastic sentiments.

Example:

I just love walking home from school in the rain (Ironic)

3. Emotion Triggers

Finally, if you have indicated one of the emotion categories other than Discard or Neutral. You should mark the *trigger words*. Trigger words are the words that help you make your decision in deciding the expressed emotions. These would ideally but not exclusively be words that indicate the positive/negative sentiment expressed.

A simple way to check if you've annotated the correct trigger words would be to think if your annotated emotion category would be the same if the trigger words were not present in the tweet. Of course, it might often be the case that the whole tweet or a large chunk of it is the trigger for the represented emotion and in that case, it is okay to not tag any trigger words in particular.

Examples:

@someuser Thank you mate ___ lots of tracks ready for the new year so I'll keep you posted!

@someuser Get. Him. Out. I don't care what policies he pursues, he's a madman and we need to get. him. out.

From my perspective most people find me heavily captivating with this kind of magnetism til they suddenly don't and I'm last weeks philosophy...I'll write a book about how charm has an expiration date. (No trigger words)

4. Additional conventions and ambiguous tweets

While the annotation guidelines are meant to be as straightforward as possible, some tweets can still be open to interpretation because multiple rules apply. In this case we try to follow a convention to make sure the annotations are as consistent as possible.

1. Empathetic Automatic or Official Communication

From one perspective, you could say this should be labeled as a 'Neutral' tweet because it contains "Customer service communication asking a question". Still, it contains quite some emotion and the poster is very sympathetic and compassionate towards the customer. For this type of tweets, where the emotion is very present and strongly expressed (i.e. you are certain about the emotion), you can still annotate this as "Love".

Such as this tweet, for example:

@someuser Hi Arne, how annoying for you. Can you send me your customer details and your contact number? Then I will send this to my colleagues from business support so that they can contact you about this. Valerie

2. Emotional irony

In ironic tweets, which should receive the discard_irony label, you can sometimes clearly tell if the user is enjoying the situation and making an ironic joke for fun or if the intent is malicious. If this intent is clear and the annotator has no doubts about the underlying emotion, you can add the emotion label (Anger or Joy) **in addition** to the discard_irony label. Note that this is the only time where you should use two emotion labels.

This is an example of an angry ironic tweet:

*(original) @someuser Wat fijn [NEWLINE] Hier altijd extreem vroeg wakker
(translated) @someuser How nice [NEWLINE] Always awake here extremely early*

3. Trolling

Occasionally, you will find people trolling on social media, possibly enjoying themselves but mostly intending to insult or degrade someone. Recognizing this malicious intent, we would not discard these tweets but annotate them as Anger because of the frustration. We illustrate this with the following example:

@someuser WAAIIITTT!!! Your whole page is the same thing . You're either really old and bore

5. Love or Joy

These two emotions are quite similar and may cause some confusion because both tap into positivity and supportiveness. In general, we would say love is more personal towards a specific person (linked to the related emotions Adoration, Tenderness, Caring and Affection). If you would consider something "cute" then the best emotion is probably Love. Joy, on the other hand, is more enthusiastic

and energetic with the related emotions being *gaiety, cheerfulness, enthusiasm, optimism and excitement*.

Joy:

@someuser Haha, exactly that. And, indeed, its always good to see a sign of life from you on Twitter!

Love:

@someuser I had a friend with depression and he preferred to be alone when it got really dark. I would text: Though I respect you want to be alone today, know that you are not actually alone. I am here for you, and I can be beside you in 45 minutes if you change your mind.

While this should help you with most doubtful cases, there is still a subjective element to it. Some people may consider a text more enthusiastic than others, which may cause one annotator to label a tweet as Joy while another annotator would use Love. This is not a problem, as long as you can support your case. Very enthusiastic loving wishes to a special person, for example, could be either, as shown here:

Have fun together!

6. Discard or Neutral or other Emotion

Often, you may consider annotating a tweet as Neutral because there is no clear dominant emotion. Still, you may sense some hints of multiple emotions and imagine some situations where a tweet could be slightly angry or sad. Identifying how much emotion a tweet can contain before it is neutral will remain slightly subjective. Therefore, if you notice that you have to guess or assume too much context, we prefer to Discard the tweet.

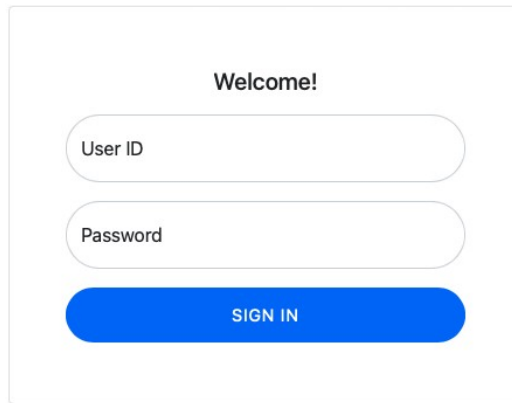
Note: if you are certain about the emotion in a tweet and believe you can justify the emotion with a trigger word or common sense, you can still annotate the emotion.

Finally, we would like to remind you to check the related emotions when in doubt. Usually this will help you make a decision. If you would still have any doubts, feel free to reach out and send us a message. The tweet you're not certain about may help us clarify some of the guidelines. Some additional examples for Dutch can be found in this document:

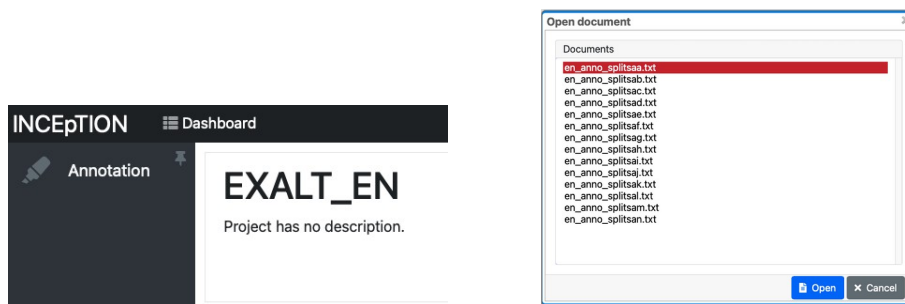
<https://docs.google.com/document/d/13BFYKxp8So0CION4kASRLWFaZTxFta8sdhUSI54EFpw/edit?usp=sharing>

5. Annotating using Inception

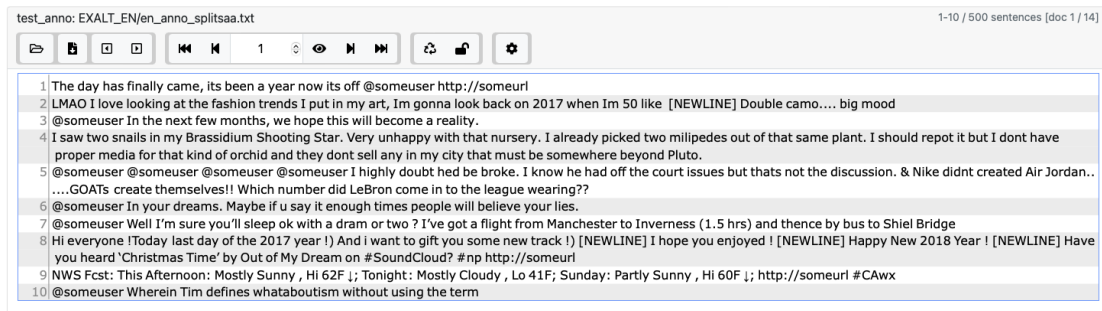
- Navigate to inception.lt3.ugent.be/login.html and login with your credentials:



- Click on the Annotation button in the Left-side Menu Bar, and then select a document to work with.

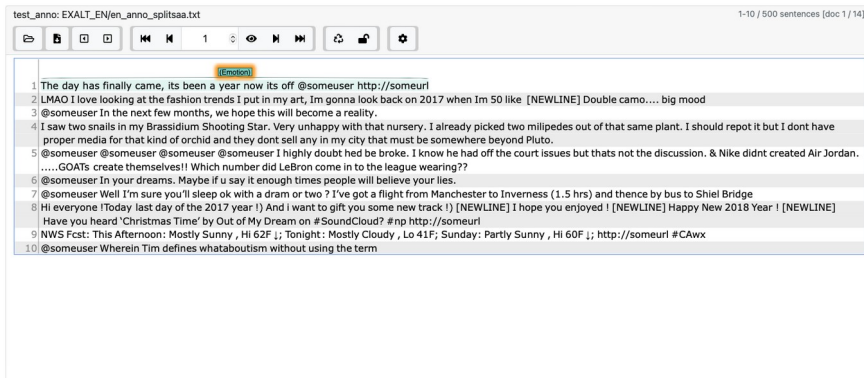


- You will now be on the main annotation screen. The screen shows 10 tweets at a time, and you can navigate to the next or previous 10 tweets using the buttons in the toolbar above the tweets.

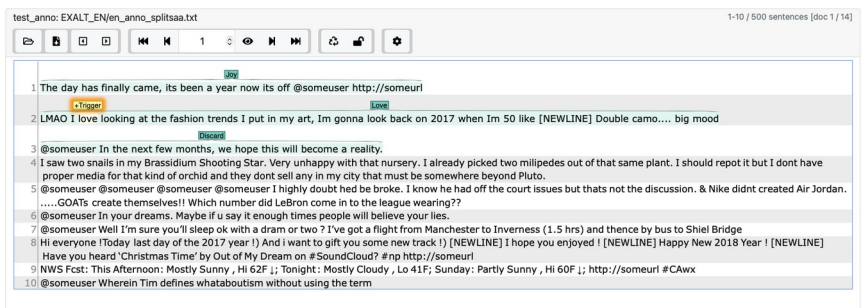


- To annotate your first tweet, first select the Emotion layer from the right-hand side annotation menu. Then select any character on the tweet, which will activate the Emotion layer menu. Select an Emotion from the drop-down menu on the right.





- Next select the Trigger layer from the right hand menu, and specifically select the words you want to mark as a Trigger Word. This will activate the Trigger layer menu. Select 'Yes' in the Trigger layer menu.



- Repeat the same process for all 10 tweets on a page, and then navigate to the next page. Inception automatically saves your annotations so there is no need to save anything separately.
- To delete an incorrectly annotated tag simply click on the annotation to reactivate the Layer menu, then use the red Delete button on the top right.

