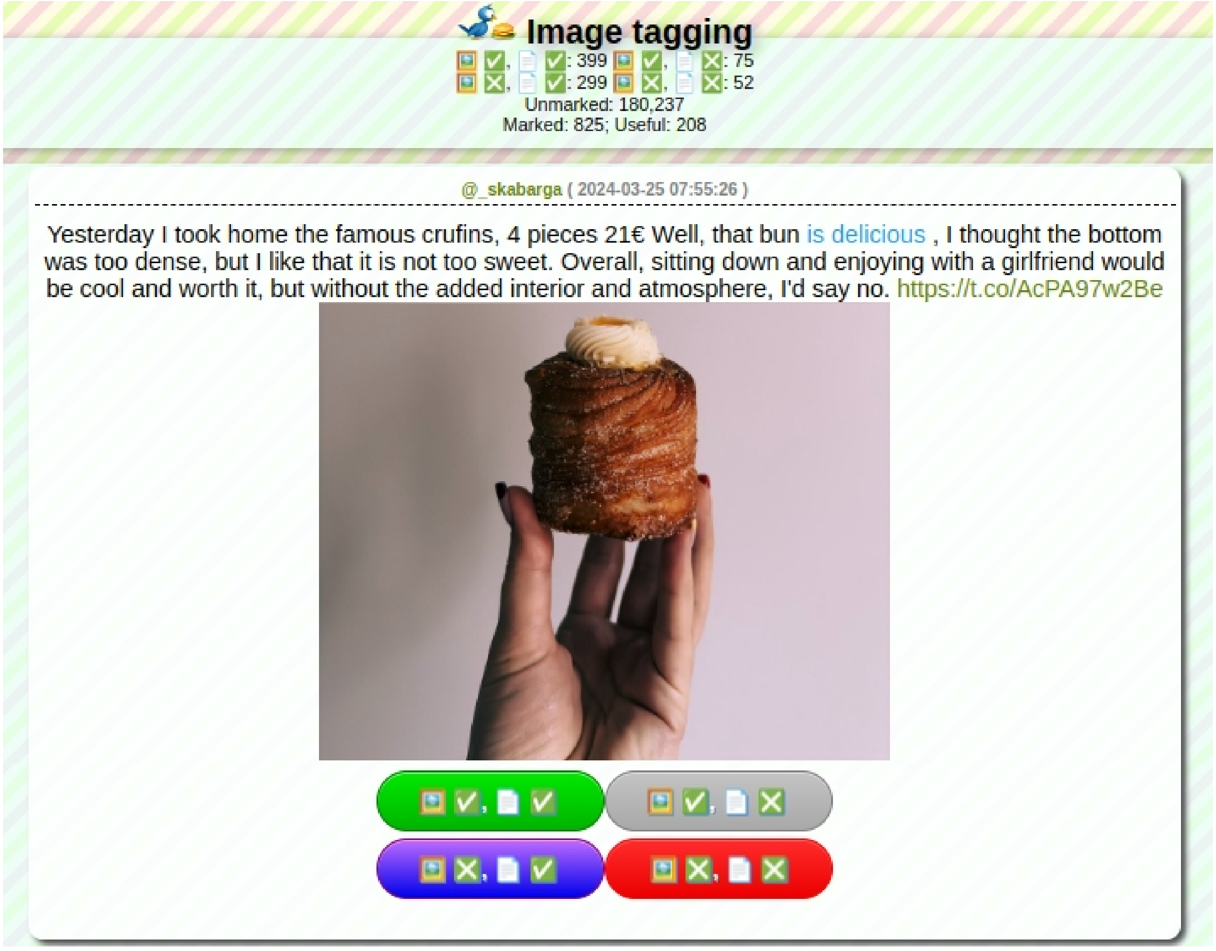
Appendix

A Annotation Environment and Instruction Examples

In this section, we present examples of annotation environments for the text-image relation task, sentiment analysis and NER, as well as introduce the general annotation instructions given to annotators.

A.1 Text-image Relation Annotation

Figure 1 shows an example of how the annotators were presented with text-image pairs from tweets. Annotators were instructed to click one of the following: 1) the green button in cases when what was written in the tweet text was represented in the image AND the image added to the meaning of the text; 2) the grey button when what was written in the tweet text was NOT represented in the image, BUT the image added to the meaning of the text; 3) the purple button when what was written in the tweet text was represented in the image BUT the image did NOT add to the meaning of the text, and 4) the red button when NEITHER when what was written in the tweet text was represented in the image NOR the image added to the meaning of the text. The top part of the annotation environment summarises the annotation statistics of the current annotator.



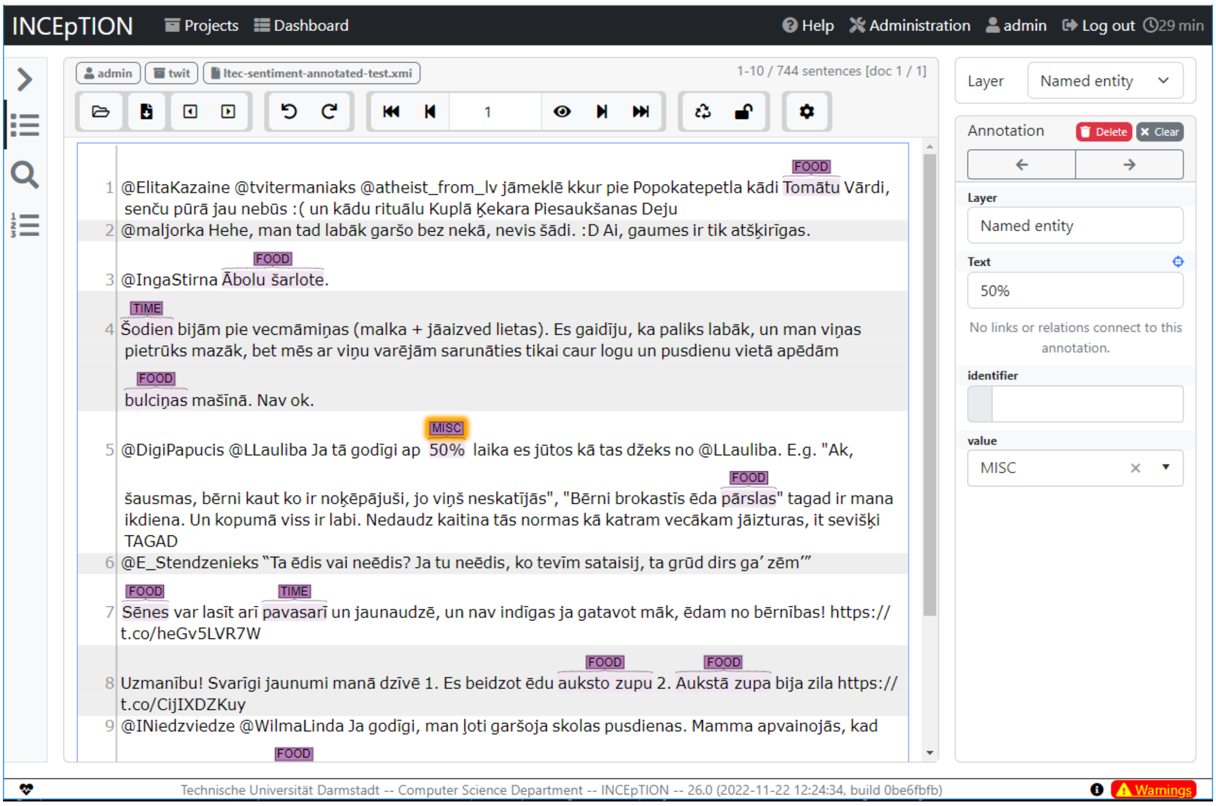
@_skabarga (2024-03-25 07:55:26)

Yesterday I took home the famous crufins, 4 pieces 21€ Well, that bun is delicious , I thought the bottom was too dense, but I like that it is not too sweet. Overall, sitting down and enjoying with a girlfriend would be cool and worth it, but without the added interior and atmosphere, I'd say no. https://t.co/AcPA97w2Be

Figure 1: An example of the text-image relation annotation environment.

A.2 Named Entity Annotation

Figure 2 shows an example of how the annotators were presented with tweets for named entity annotation. Tweets are presented each on a separate line, with 10 tweets per screen. The task of annotators was to select appropriate text spans and assign a named entity category to the selected span. The named entity categories are already set, and the annotator needs to select correct category from drop-down menu.



1 @ElitaKazaine @tvitermaniaks @atheist_from_lv jāmeklē kkur pie Popokatepetla kādi Tomātu Vārdi, senču pūrā jau nebūs :(un kādu rituālu Kuplā Ķekara Piesaukšanas Deju

2 @maljorka Hehe, man tad labāk garšo bez nekā, nevis šādi. :D Ai, gaumes ir tik atšķirīgas.

3 @IngaStirna Ābolu šarlote.

4 Šodien bijām pie vecmāmiņas (malka + jāaizved lietas). Es gaidīju, ka paliks labāk, un man viņas pietrūks mazāk, bet mēs ar viņu varējām sarunāties tikai caur logu un pusdienu vietā apēdām bulciņas mašīnā. Nav ok.

5 @DigiPapucis @LLauliba Ja tā godīgi ap 50% laika es jūtos kā tas džeks no @LLauliba. E.g. "Ak, šausmas, bērni kaut ko ir noķēpājuši, jo viņš neskatījās", "Bērni brokastīs ēda pārslas" tagad ir mana ikdiena. Un kopumā viss ir labi. Nedaudz kaitina tās normas kā katram vecākam jāizturas, it sevišķi TAGAD

6 @E_Stendzenieks "Ta ēdis vai neēdis? Ja tu neēdis, ko tevīm sataisij, ta grūd dirs ga' zēm'"

7 Sēnes var lasīt arī pavasarī un jaunaudzē, un nav indīgas ja gatavot māk, ēdam no bērnības! https://t.co/heGv5LVR7W

8 Uzmanību! Svarīgi jaunumi manā dzīvē 1. Es beidzot ēdu auksto zupu 2. Aukstā zupa bija zila https://t.co/CijIXDZKuy

9 @INiedzviedze @WilmaLinda Ja godīgi, man ļoti garšoja skolas pusdienas. Mamma apvainojās, kad

Layer: Named entity

Layer: Named entity

Text: 50%

identifier:

value: MISC

Technische Universität Darmstadt -- Computer Science Department -- INCEpTION -- 26.0 (2022-11-22 12:24:34, build 0be6fbfb)

Figure 2: An example of the named entity annotation environment.

A.3 Sentiment Annotation

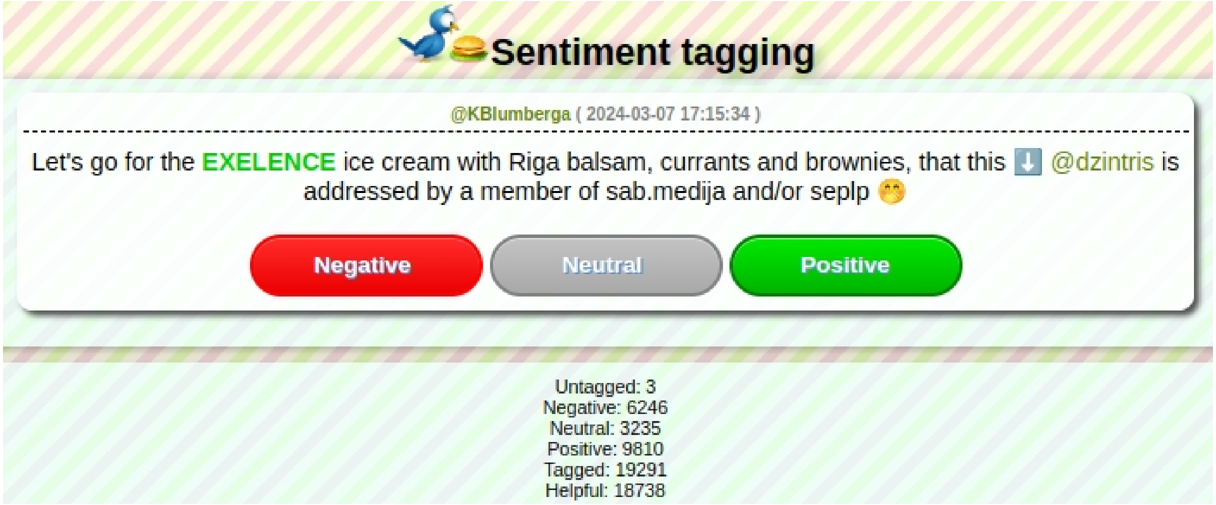
Figure 3 shows an example of how the annotators were presented with tweets for sentiment annotation. Tweets were displayed one by one and annotators were given instructions to click either the red 'Negative', the grey 'Neutral', or the green 'Positive' button as an indication of which sentiment they believe the tweet text represents. The bottom part of the annotation environment summarises the annotation statistics of the current annotator.

B Annotator Profile

Annotators for the named entity recognition (NER) and the text-image relation task were native Latvian speakers in their 20s-30s with at least a master's degree and work experience in the field of natural language processing (NLP) or linguistics. For sentiment analysis, annotators were in the age ranges between 20 and 50 with at least a bachelor's degree in various fields, and the translator and post-editor were professional linguists with a master's degree in their respective fields in their 30s-40s.

C Text-image Relation Ablation Experiments

Since the LLaVA models can be sensitive to the provided random seed, we performed additional ablation experiments with 10 different random seeds as well as the latest available 1.5 version, and both 7B and 13B model sizes. Results summarised in Table 1 show that the scores are slightly lower than in the original experiments with the 1.3 version.



@KBlumberga (2024-03-07 17:15:34)

Let's go for the EXELENCE ice cream with Riga balsam, currants and brownies, that this ⬇️ @dzintris is addressed by a member of sab.medija and/or seplp 🤭

Untagged: 3
Negative: 6246
Neutral: 3235
Positive: 9810
Tagged: 19291
Helpful: 18738

Figure 3: An example of the sentiment annotation environment.

	Average	+/-
EN-13b	19.78%	0.86%
EN-7b	22.13%	2.46%
LV-13b	12.54%	1.29%
LV-7b	17.82%	1.60%

Table 1: Text-image relation classification ablation experiment results with Llava 1.5 models.