For each tweet you will need to:

1. Read the tweet, try to understand the meaning in general.
2. For each emoji character in the tweet, decide if it has its own meaning, or if it is necessary to examine it as part of a sequence of emojis.
3. For each emoji or sequence of emojis (with a compositional meaning), decide on the relationship between the emoji and the rest of the text and add one of three main tags: multi-modal informative, content word or grammatical function.
4. Once you determined the classification, choose the tag that best represents the sub-category of the emoji.
5. Try to think of the meaning conveyed by the tweet and suggest a replacement or an interpretation for the emoji.

Steps:

Step 1- first read:

Before annotation each tweet, read it with the following questions in mind:

1. Can the tweet be understood without the emoji at all?
2. Can you think of a simple verbal replacement for the emoji?
3. How would the same message be expressed in a conversation?
   a. Would the meaning conveyed by the emoji be replaced with a physical gesture or intonation?
   b. Would the meaning conveyed by the emoji be clearer in an established context of location/previous utterance/relationship?
4. Is the use of emoji seem to add communication abilities that cannot be expressed in text alone?

Step 2- Identifying emojis as grammatical functions:

In order to decide whether an emoji is used as a grammatical function within the text, try to answer the following questions for each sub-category:

1. Function words
   a. Is the emoji used to replace a function word such as a preposition/auxiliary/connective/conjunction/determiner?
   b. What is the most obvious word represented in the emoji? Does it sound like another word? If so, does it fit the definition of a "function word"?
   c. Examples:

   I want 2 go 2 sleep.

   The word "to" is replaced by an emoji of the number 2 (since they are homophones).

   Hugs 4 everyone.

   The word "for" is replaced by an emoji of the number 4 (since they are homophones).

   I 🍩 know.

   The word "don't" is replaced by a doughnut emoji (which sounds similar).

I will 🐝 there!!

The word "be" is replaced by a bee emoji (since they are homophones).

   d. If the answer to the above question is yes, and it is similar to the examples in c, then tag the emoji is a grammatical function and the attribute "function word".
   e. In the replacement attribute, add the word or phrase you think best fit the intended meaning of the emoji ("be" for the bee emoji).

2. Punctuation marks
   a. If you see an emoji of a punctuation mark that is used as regular punctuation, add the grammatical function tag, with the punctuation attribute.
   b. In the replacement attribute, add the same punctuation marks from the regular keyboard.
   c. Example:

Stop it ‼️

Here the replacement would be "!!".

Step 3- Identifying emojis as content words:

If the emoji you are examining doesn't fall under the previous category, you will next try to decide if it's part of the main message in the text and is used as a content word. This group also mainly includes emojis that are replacing words within a sentence or textual expression. However, the emojis tagged as content words should have a semantic role similar to that of common nouns, non-auxiliary verbs and adjectives.

In order to determine if the emoji should be tagged as a content word, think of the following:

1. Is the emoji placed inside a sentence (i.e. not after a punctuation mark or at the beginning of the tweet)?
2. Try to read the tweet out loud. Did you have to include the emoji in order to read it? How did you refer to the emoji?
3. If you replaced the emoji with a simple word or phrase, what was it? Does the tweet lose any of its meaning by replacing the emoji with text?
4. Is the emoji part of a sentence or phrase structure?

If you answered "yes" to most of these questions, then you can probably tag the emoji or the sequence as a content word or phrase.

1. Tag the emoji as "content".
2. Add the type attribute and determine the part of speech that fits the perceived meaning of the emoji the most.
3. Add the replacement attribute and choose the word or phrase that can replace the emoji while maintaining the original meaning of the tweet.
4. Examples:

a. Type= "noun", replacement= "a cookie"

I want 🍪

b. Type= "verb", replacement = "ran"

I saw the bus coming so I 🏃 to the stop.

c. Type= "adjective", replacement= "engaged"

She is 💍.

d. Type= "adverb", replacement= "soon"

I'll be there ➡️

Step 4- tagging multimodal emoji uses:

If the emoji doesn't seem to belong to the other two groups, it is very likely to be tagged as a multimodal. Emojis are often used as content or functional words as abbreviations or space/time savers, but their graphics and variety allow Twitter users to express more than what they could have with only 140 alphanumeric characters. In this category you will tag emojis that add additional information and forms of communication to the text, such as implying gestures, intonation, topic and attitude.

Since these are often meta-linguistic additions to the tweet, it might be hardest category to annotate. The multimodal nature of these cases generates ambiguity that is very likely to be intended by the writer of the tweet. Try to choose a subcategory from the suggested types that you think describes the most obvious and important message intended by the writer. If you believe that there are 2 or more equally strong interpretations of the emoji, DO???????, but try to avoid this whenever possible.

<u>Types:</u>

This is not a comprehensive list or a closed definition. We expect that you will find other cases that fall under the multimodal umbrella. We have kept the 3 subtypes broad enough that you can probably find a suitable way to annotate most cases. Use the descriptions as examples and a guide, rather than restrictions on what belongs to this type.

1. <u>Topic markers</u>
    a. Emojis as graphic representations of the subject of the tweet.
    b. Emojis that serve similar functions as hashtags, as a way to provide otherwise unspecified context.
    c. Emojis that are referencing explicit words or phrases in the tweet, emphasizing the stated topic.
2. <u>Physical gestures</u>
    a. Emojis that imply a gesture made by the writer to accompany the rest of the text.
    b. Emojis that represent a facial expression or physical reaction to the text.
3. <u>Attitude</u>

a. Emojis meant to express the writer's attitude towards the text.
b. Emojis that add sentiment towards a situation or a person (possibly the one the tweet is addressed to).
c. Emojis that suggest the writer is joking or being sarcastic.
d. Emojis that represent the writer's state of mind or mood.

Once you have chosen the mm (multimodal) tag and the type attribute, suggest a short interpretation of the emoji.

<u>Examples:</u>

1. Type= "attitude", interpretation= "joking"

I hate you! 😜

2. Type= "attitude", interpretation= "I'm miserable"

I'm so tired... 😩😩😩

3. Type= "topic", interpretation= "co-reference: NYC"

Going to NYC tomorrow 🗽

4. Type= "topic", interpretation= "bowling"

I'm really bad at this game. 🎳

5. Type= "physical gesture", interpretation= "blowing a kiss"

I'll see you soon 😘💋