Annotation Guidelines *Catching Feelings*

The format and some of the content of these guidelines is copied from the Annotation Guidelines document created by Tess Dejaeghere, which is available here. Thanks so much, Tess ;-)

This document details everything you need to know to start annotating for the *Catching Feelings* project. First off, it's important to know **why we ask you to label this data**. This data will be used to conduct Aspect-Based Sentiment Analysis (ABSA) on a dataset of fanfiction comments from the *Archive of Our Own,* one of the largest online repositories of English-language fanfiction. ABSA is an NLP technique for determining two interrelated characteristics of texts: the *evaluations* they convey (sentiments) and the *aspects* of the thing – in this case, fanfiction texts – that this evaluation refers to.

The texts you will be annotating are all comments left on fanfiction about Greek mythology by readers of that fanfiction. You will be annotating these comments on three levels:

1. the *aspects* of stories that fanfiction-readers comment on
2. their evaluation of these aspects (sentiment)
3. the *link* between aspect and sentiment

Aspect-based sentiment analysis is currently mainly used in the context of customer reviews. Websites such as Amazon analyze their customer reviews to see how people feel about certain aspects of the company, and where they can improve their services (CUSTOMER SERVICE, PRICE, DELIVERY, ...). In the context of literature, it is currently very rarely applied! If you're interested in learning more on ABSA, certainly check out the work by our UGhent colleagues at LT3 on literary prizes in social media; and this interesting paper on the creation of an ABSA-annotated corpus for Swedish.

The software behind the NLP-tools often thrives on **Machine learning** (or **deep learning**, which is a more advanced version of ML). This means that we feed the software some examples of data annotated by humans, and the software automatically learns the patterns to extract these entities automatically. Training data is also important to evaluate these tools. Manually annotating these examples takes a lot of time and effort, but it's one of the most important steps to make the tools better – which may uncover a wealth of untapped information.

So please know: the quality of your work is very important to us!

What am I annotating and why?

You are contributing to my dissertation research. My project has the preliminary title *Gods, Heroes, Myths: Anchoring and Innovating Classical Motifs in Fanfiction,* and it is part of the larger research project *Anchoring Innovation.* My dissertation revolves around the question:

How – in what ways and to which ideological ends – is source material from Greek myth being rewritten and transformed in contemporary online fanfiction?

To answer this research question, I have collected a corpus of slightly more than 5.000 works of English-language fanfiction about Greek mythology from *Archive of Our Own,* the fanfiction-website mentioned above.

I have narrowed down my research to three themes, each of which will eventually be a chapter in the dissertation: power, emotion, and repetition. The *Catching Feelings* project focuses on analyzing reader response to fanfiction. To find out how fanfiction's depictions of Greek myth resonate *emotionally* with their fans and readership, I am applying aspect-based sentiment analysis to the comments these readers have left on the fanfiction.

By collecting only fanfiction-comments that were written in response to fanfiction about Greek myth in English, I have gathered a corpus of 25.970 comments. To train the ABSA-pipeline, we need to manually annotate 1.000 of those comments. Note that the comments were written in response to a variety of stories.

Content warning: If you're not familiar with fanfiction, you should know that it's a pretty varied and sometimes surprising or weird corner of the internet. Some of these comments refer to story-events of a sexual nature and/or contain curse words. If you're uncomfortable with that type of content, please let me know, and we can find a solution together.

Annotating Sentiment

Annotating for sentiment happens according to a simple two-category system: positive evaluation versus negative evaluation. If you're having difficulty determining the valence of a word but nonetheless feel that it expresses sentiment, you can also select the **NULL** category.

You annotate for sentiment if there is an explicit or implicit clue in the text suggesting that the speaker is **positive** (i.e., happy, admiring, relaxed, forgiving, etc.) or **negative** (i.e., sad, angry, anxious, violent, etc.) regarding the entity or regarding the work of fanfiction overall.

Watch for textual clues and intensifiers e.g.: "very", "extremely", "exceptionally", "unusually", ... Please ==include== those in your annotation.

You might run into an expression of a sentiment where the entity is not explicitly mentioned, but it is clear that the sentence/group of words is *about* an aspect of the fanfiction-text. For example, if a commenter writes only 'Fantastic', they are evaluating the story positively even if there is no explicit reference to the story itself in the text of the comment. In such cases, label the evaluation, and also add a **NULL** aspect.

Fanfiction-communities are characterized by their own subcultural norms of expression and language use. Here are some general pointers.

- You will encounter the use of emojis and other non-specific expressions of sentiment, such as strings of random capitalized letters. Annotate them with the appropriate valence and, if it is not clear which aspect of the fanfiction they are referring to, a **NULL** aspect.

- Be aware that not all negative words convey negative value judgments. For example, a 'heartbreaking' reading experience can actually be positive. Use your own judgment to evaluate this.

- The comments sometimes refer to fans' creations that are not texts, such as a recorded audio version of fanfiction or an illustration that accompanies the story. If these comments contain value judgements and/or mention particular aspects, annotate them as you would elsewhere.

If you are confused about any other community norms or jargon, it can be useful to consult Fanlore.org, or you can email me and ask at julia.neugarten@ru.nl.

Annotating Aspects

We are annotating 8 types of aspects. Some of these were devised bottom-up, i.e. based on what was found in the data, and others were devised top-down, i.e. based on what I was looking for, hoping to find, or particularly interested in in the data. The 8 aspect-categories are:

Canon – reference to the way the fanfiction transforms, critiques, or engages with canonical material, including references to Greek mythology but also any other source material the fanfiction is referencing or making use of, such as for example elements of popular culture.

If you are not sure whether something originates from other material, Googling it can be a good idea. A fanfiction-writer's Original Character (OC) is not likely to have a lot of search engine hits, but a Greek myth or a pop culture location like Hogwarts will.

Character - reference to character, characterization, character appearance, relationships between characters and assessment of characters that are *in* the story; so a comparison to a character from another story would go under **Canon**.

Note that fans often call romantic relationships between characters *ships*.

Emotion – references to the emotions experienced by characters within the story.

Events and Storyworld – reference to plot events and their setting, specific scenes, world building, story content like concepts, ideas or tropes, and also more general plot elements like plot twists or endings.

I am *excluding* requests for specific plot developments from this category, so "I hope they end up traveling to Rome" is not considered a reference to **Events and Storyworld** because the event (a trip to Rome) does not occur in the story being commented on.

General – reference to the story as a whole or in general terms, including references to a chapter, update or installment, but also headcanons.

Reading Experience – reference to the reading experience, such as emotional engagement (of the reader, not the characters), absorption or narrative tension.

I am *excluding* expressions of desire for future works, sequels and specific plot developments from this category, so "I wonder what happens next" is not considered a reference to **Reading Experience**.

Style – reference to any aspect of how the story was written down or rendered, including writing style, word choices, metaphors, turns of phrase, voice, perspective.

NULL - This category is for expressions of sentiment that don't refer to a specific word or aspect in the text.

Some Overall Tips

Please annotate the smallest span possible, so annotate "characterization" and not "the characterization."

Do not include punctuation (question marks, periods, exclamation marks) in your annotations, not even when there are very many exclamation marks, for example. You may sometimes include commas, as in "very, very, good" for example. Occasionally, commenters use an asterisk to communicate an action, like for example "*applause*". In those cases, include the asterisks in the annotation.

Sometimes spaces are missing, so a comment may say "greatcharacter" all in one. You can annotate only the part of the word that is relevant, so in this case you could annotate "great" as "positive evaluation" and "character" as a **Character** aspect, then link the two together.

In cases of summation of *aspects*, so for example "Ariadne and Theseus" annotate each aspect separately. Sometimes I've struggled because a character's relationship, which is also included in the **Character** category, can be referred to in this manner. When it's clear the positive evaluation is linked to the relationship *between* the characters rather than to the presence of the individual characters, I have annotated multiple words together as *one* aspect, so for example "dynamic between Hades and Persephone" as a single entity.

In cases of summation of *sentiments,* for example "great and fantastic", annotate them together so it counts as *one* occurrence of **Positive Evaluation.** It is also permissible to annotate punctuation marks this way, like in the above example of "very, very, good".

These comments contain a lot of positive expressions that do not relate to the story itself but to its context of production and exchange, such as "Thank you for writing/sharing/posting this" or "Please write more of this story." You do not need to annotate those. If they contain explicit evaluations, such as "Thank you for sharing this fantastic story," you can annotate the evaluation. In this example, you could link it to the **General** aspect 'story'.

The comments also sometimes contain quotes from stories. Do **not** annotate those, because it is usually unclear which aspects are being evaluated through quoting. You can often tell if something is a quote because its tone/style is very different from the comments, or because it is in quotation marks,

You *can* annotate a word or phrase as both an aspect and an evaluation. So far I have mostly done this for the **Reading Experience** category. For example, if the comment is "this fic was really moving" then I have annotated "really moving" as a positive evaluation of the **General** aspect "fic" but also as the **Reading Experience** aspect.

Annotating Aspect-Sentiment Relations

If you've annotated an aspect and a positive or negative evaluation that you believe refers to that aspect, please create a link in Inception between the two, with the arrow pointing from the evaluation to the aspect. More than one link can exist between a sentiment and the aspects it qualifies. For example, if the comment says "I love Ariadne and the writing style in this," the word "love" qualifies both "Ariadne" (a **Character** aspect) and "writing style" (a **Style** aspect), so you can create two arrows from "love" to those two separate aspects.

Lexicon of Fanfiction Community Terms

Angst	Sad events in fanfiction-stories, or a sad genre of fanfiction

Note: angst is often valued positively despite – or maybe because of – its sad story-content

AU	Alternate Universe, a subgenre of fanfiction set in a non-canon world
Canon	The source material fanfiction is based on
Crossover	Fanfiction that blends elements from various canons.

For example, a story about the Greek gods at Hogwarts is a Greek myth/Harry Potter crossover.

Fanon	Non-canonical ideas about canon shared or appreciated by the fan community
Feels	A generalized sense of fannish attachment or feelings, usually positive
Fic	Short for fanfiction
Fluff	Happy events in fanfiction-stories, or a happy genre of fanfiction
Headcanon	Non-canonical ideas about canon appreciated by one fan (in their head)
Kudos	Positive evaluation, like a 'Like' on Facebook
OC	Original character, so a character not from canon
OOC	Out of character, meaning characterization does not meet canon expectations

Note: writing an out of character characterization is usually evaluated negatively

One-Shot	A work of fanfiction that consists of only one chapter
OTP	One True Pairing: a *ship* that fans have a strong investment in
Prompt	A request to write a story that meets certain readers' specifications
Prompt fill	A story that meets the specifications in a prompt
Ship	Verb – to be deeply invested in the romance between two fictional characters
	Noun – the (fictional) romance between two fictional characters

Squee   A generalized sense of fannish attachment or feelings, very positive

AU   Alternate Universe, or fic set in a different storyworld from canon

WIP   Work in progress

Rolling notes and clarifications based on the annotation process.

- Not everything that is annotated needs to be annotated with both aspect and evaluation.

- If the evaluation doesn't have a link to any aspect, then you should give it its own aspect: **NULL** (or something else if applicable). If the aspect doesn't have an evaluation, you can leave it unlinked and don't need to add a **NULL**. This is to save work.

- You can have one aspect inside another, so a **Character** can be tagged within the span of an **Event and Storyworld** annotation, for example.

- Character **Emotion** is not character motivation.

- Sorrow-related descriptions of **Reading Experience** is usually positive.

- I've not tagged any explicit references to the author as a person.

- Should we exclude the comments that we can clearly tell are about artworks rather than stories?

- **Reading Experience** can be divided into emotions experienced during reading, and cognitive processes for the reader.

- I've annotated intensifiers, so "totally great" rather than "great" *except* in cases where the intensifier occurs separately in the text. So in "you've captured the vibes so well" I would annotate "captured" as a positive evaluation of the **Reading Experience**-aspect "vibes" and leave out "so well" because attaching it means a whole bunch of extra actions/clicks and if you don't attach it you still include the fact that there's positive evaluation going on.