TNA ACTIVITY REPORT

Catching Feelings: Aspect-Based Sentiment Analysis for Comments on Fanfiction about Greek Myth

Author: Julia Neugarten

Current position: PhD Candidate

Affiliation: Radboud University

Host institution: Ghent Center for Digital Humanities

Mentor: Julie Birkholz

Period of stay: May 13th – June 30th 2024

Introducing the *Catching Feelings* Project

Fanfiction – stories inspired by existing source material, written by and for fans, and published online for free – often revolves around emotional trajectories. Love, heartbreak, mental illness and other emotional highs and lows are frequently central to fanfiction plots, which have therefore been described as "emotional landscapes of reading" (Samutina 2017). Additionally, because of their online context of production, fanfiction texts must be approached "as both literary and digital objects" (Black 2020). Combined, fanfiction's digital and textual nature and its emphasis on emotion make it suitable for computational approaches such as sentiment analysis (SA).

However, within literary studies, sentiment analysis has often been critiqued, as "most tools and methods still lack a validation" and "connections with literary theory

are frail or disputable" (Rebora 2023). Additionally, many approaches to SA fail to account for the complex ways sentiments take shape textually. For example, is a sentiment analysis of literary or narrative texts concerned with the sentiments of fictional characters intradiegetically, with readers' response to text, or with some sort of sentiment imbued in the text through the craft or intent of the author? Furthermore there is the issue of scale: many methods operationalize sentiments at the level of entire documents, paragraphs, or sentences, failing to account for the ways sentiment may shift subtly from one word or moment of reading to the next.

The method of aspect-based sentiment analysis (ABSA), where sentiments in text are attached to particular aspects of the object of evaluation, addresses this problem of granularity or scale that many existing approaches to SA have.

In the *Catching Feelings* project, we used ABSA to measure textual reception, i.e. the ways stories are evaluated by their readership. This focus circumvents the problem of theorizing the role of sentiment in narrative text itself. In comments on fanfiction, it is clear that fanfiction is being evaluated or responded to and so the sentiments expressed in the comments can be assumed to reflect evaluations from the perspective of the commenter of the fanfiction texts the comment is posted under.

The aim of this TNA fellowship, then, was to apply aspect-based sentiment analysis to a corpus of over 25.000 comments on fanfiction-texts. For reasons of scope, we focused on fanfiction about Greek mythology on the popular English-language platform *Archive of Our Own* (AO3). This analysis fits within the PhD-project *Anchoring and Innovating Classical Motifs in Fanfiction,* which analyzes how source material from Greco-Roman Antiquity is transformed in contemporary online fanfiction on AO3.[1] The fellowship addressed the following research questions:

1. In reader response to fanfiction about Greek mythology, which aspects of the source material are loved, enjoyed, and appreciated?

---

[1] Find out more about the project here:
https://anchoringinnovation.nl/projects/gods-heroes-myths-anchoring-and-innovating-classical-motifs-in-fanfiction.

2. What are the affordances and limitations of aspect-based sentiment analysis (ABSA) for analyzing this corpus of fanfiction comments?

While the first research question is oriented towards the content of the data, the second question is methodological.

Methodology

Our approach to ABSA consisted of five steps.

1. Data Collection

Based on an existing metadata-set of fanfiction about Greek mythology (Neugarten & Smeets 2023, Neugarten 2024), we scraped fanfiction comments from the web using the AO3-api (Rodrigues 2021). We collected 25,970 comments (984,400 words in total).

2. Annotation

To develop an annotation guide, we conducted exploratory analysis of the data using SpaCy (Montani et. al. 2023) and Top2Vec (Angelov 2021). Based on this exploration, we formulated a scheme of eight aspects of fanfiction that we found were frequently commented on in the dataset:

1 **Canon**: reference to how fanfiction transforms, critiques, or engages with canonical material, including references to Greek mythology but also source materials from popular culture.

2 **Character**: reference to and assessment of character, characterization, character appearance and relationships between characters from the story.

3 **Emotion**: references to the emotions experienced by characters intradiegetically.

4 **Events & Storyworld**: reference to plot events and settings, specific scenes, worldbuilding, story content like tropes, and general plot elements like twists or endings.

5 **General**: reference to the story as a whole or in general terms.

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

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

8 **NULL**: expressions of sentiment that do not refer to specific words or aspects.

I annotated 1.000 comments for occurrences of these aspects and expressions of sentiment (positive/negative) using the INCePTION platform (Klie et. al. 2018). Amanda Hemmons, an intern at the Ghent Center for Digital Humanities during the period of the fellowship, annotated an overlapping 100 comments so we could calculate inter-annotator agreement.

3 Inter-Annotator Agreement

Before training any machine learning models on these annotations, we calculated inter-annotator agreement to evaluate how consistently aspects and sentiments had been annotated, using INCePTALYTICS (Hamacher & Zesch 2022). With a Pairwise Cohen's κ of 0.86 of for aspects and 0.88 for sentiments, inter-annotator agreement was high. This is probably due to the intensive collaboration on the annotation guidelines and both annotators' familiarity with the discursive norms of fanfiction communities. Figures 1 and 2 provide confusion matrices between annotators.

Figures 1 and 2: Confusion Matrices between annotators for aspects (left) and sentiments (right)
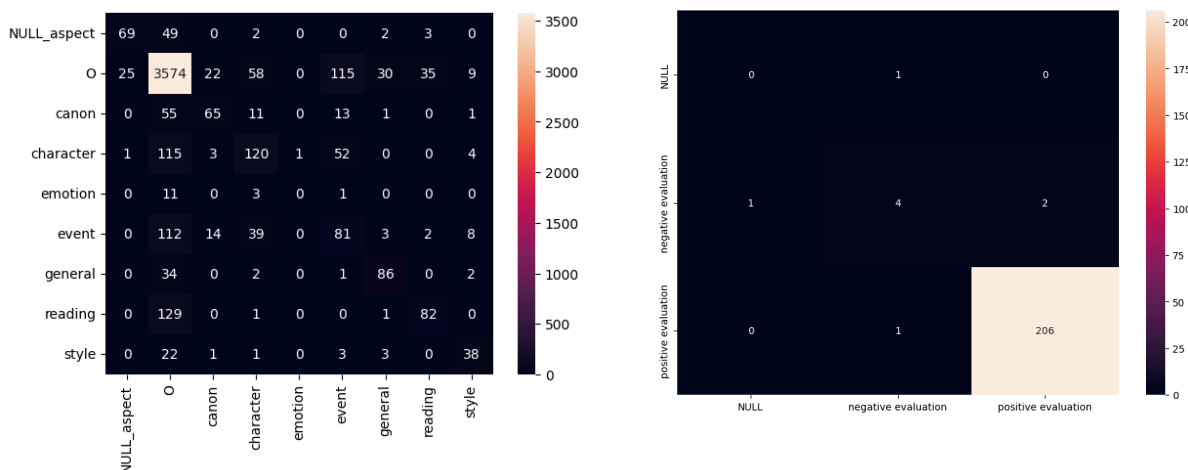
4 Aspect Extraction

For the aspect extraction task of identifying our eight aspects, we tried out three machine learning models: roBERTa-base (Liu et. al. 2019), Twitter-roBERTa-sentiment (Barbieri et. al. 2020) and NuNER (Bogdanov et. al. 2024). We obtained the best results with NuNER (Figure 3). However, results differed quite a lot between aspects. Aspects that were more frequent in the annotation-set, such as Style and General, turned out to be much easier for the models to detect than rarer aspects such as Emotion or Event & Storyworld.

model	F-score (micro)	F-score (macro)	accuracy
roBERTa-base	0.30	0.25	0.18
Twitter-roBERTa-sentiment	0.35	0.31	0.22
NuNER	0.50	0.45	0.34
NuNER (split task)	0.50	0.50	0.34

Figure 3: Evaluation of Aspect-Extraction Models

5 Sentiment Analysis

We used Twitter-roBERTa-sentiment for the sentiment analysis, because that model has already been fine-tuned for sentiment detection. Results were good (Figure 4).

model	F-score (micro)	F-score (macro)	accuracy
Twitter-roBERTa-sentiment	0.95	0.75	0.96

Figure 4: Evaluation of Sentiment Analysis Model

Description of the Research Visit

The table below describes the tasks and activities conducted during the research visit and their ties to the deliverables enumerated below.

	Week	Tasks	Deliverables
Pre-fellowship		Annotation	3
	May 13 - 19	Annotation + discussion of annotation guidelines exploration of existing ABSA notebooks	3, 1
	May 20 - 26	Inter-annotator agreement calculation	3
	May 27 - June 2	Attended DH Virtual Discussion Group, meetings with Marianne van Remoortel and Gunter Martens	4
	June 3 - 9	Meeting Elly McCausland Attended DH Benelux	4
	June 10 – 16	Visit to Vienna: Training school & CCLS	4, 5
	June 17 -23	Wrap-up presentation	5
	June 24 - 30	Evaluation ABSA results	3
	Post-fellowship	Write and submit paper to CHR	2

Outcomes

The research conducted during the fellowship has resulted in five types of deliverables.

1. Insight

I have gained an understanding of ABSA and some skill in applying it. This has contributed to my development as a CLS-scholar. I now understand the pros and cons of several approaches to ABSA, including rule-based methods, machine learning approaches and the use of LLMs.

2. Publication

We have had a short paper accepted to Computational Humanities Research 2024 as a poster. In this paper, we were able to answer research question 2 – regarding the methodological suitability of ABSA to fanfiction comment data – with an affirmative. To quote from the paper: "fanfiction comments seem suited to ABSA

because many of the texts explicitly express sentiments regarding specifically named aspects of fanfiction stories." The paper is available in the conference proceedings.

3. Documentation and Code

Our research process has been documented in a Github repository containing:
- A set of guidelines for annotating fanfiction comments in INCePTION.
- A collection of code and Jupyter Notebooks to guide researchers through the processes of data collection, exploratory data analysis, transformation of INCePTION annotations into the IOB-format (Ramshaw & Marcus 1995), calculation of inter-annotator agreement using Inceptalytics (Hamacher & Zesch 2022), the creation of confusion matrices between annotators, and evaluation and exploration of results.
- Results of our ABSA pipeline using three different models for the aspect extraction task (roBERTa-base, Twitter-roBERTa-sentiment and NuNER) and one model for sentiment analysis (Twitter-roBERTa-sentiment).

4. Networking

During the fellowship, I expanded my professional network by attending the Digital Humanities Benelux conference in Leuven, the *ExploreCor* training school organized by CLS Infra in Vienna, and the Conference of Computational Literary Studies (CCLS), also in Vienna. Additionally, I attended the yearly in-person meeting of the Digital Humanities Virtual Discussion Group for Early Career Researchers in Belgium, organized by Dr. Leah Budke. I also spoke to several colleagues in my field and adjacent disciplines at Ghent University, including Dr. Gunther Martens, Dr. Marianne Van Remoortel and Dr. Elly McCausland.

5. Valorization

With the help of Dr. Sarah Hoover, I created two videos to communicate to a wide online audience about the outcomes of my TNA. The first video is an interview in which I reflect on the experience and usefulness of the fellowship, the second is a recording of the presentation I gave towards the end of my fellowship to give insight into the research process and some preliminary results. You can find the video of my

interview about the fellowship here and the wrap-up presentation is here, with the PowerPoint available on Zenodo as well.

Future Work

The research conducted over the course of the TNA offers three broad avenues for future research: finalizing the ABSA workflow, improving the ABSA workflow and linking the outcomes of this analysis of comments to an analysis of fanfiction texts themselves.

When it comes to finalizing the ABSA workflow, we still need to connect the outcomes of our aspect extraction step to the outcomes of the sentiment analysis to see which aspects are evaluated in what ways by commenters, i.e. to answer research question 1. These findings must then also be contextualized in the wider field of fanfiction production; the popularity of a character in comments may be impacted by their popularity in fanfiction itself, and the way these two types of popularity interact remains largely opaque.

We have five ideas to improve and elaborate on the ABSA-workflow:

1. To improve results, it could be useful to annotate more data in a targeted way, supplementing those aspects for which support was low and ensuring a larger part of the training data has been assessed by more than one annotator.

2. It may be interesting to analyze the comments containing only emojis, which were not part of the text-focused analysis conducted during the fellowship.

3. The usefulness of LLMs for ABSA requires further attention, since these models are gaining ground as information extractors. For example, GoLLie (Sainz et. al. 2024) looks interesting. This prompting framework would let us feed our annotation guidelines directly into a generative model.

4. A hybrid approach to ABSA, particularly adding some rule-based elements to our existing pipeline, may be a way to improve outcomes for some of the aspects that were difficult to detect using our current setup. For example, we could use an emotion lexicon like EmoLex (Mohammad & Turney 2013) to detect the Emotion aspect and create lists of characters and story settings

based on existing metadata (Neugarten & Smeets 2023) to detect the Character-aspect and the storyworld-dimension of Events & Storyworld.

5. Examining the outcomes of our current best model in more detail could lead to domain-specific insight into the ways sentiment is expressed and attached to particular aspects of fanfiction in comments.

Finally, an exciting next step in reception studies based on online lay reviews of or comments on reading experiences, of which the *Catching Feelings* project is but one example, would be to connect the sentiments and evaluations expressed in comments to text-specific features of the stories being commented on. In the broader context of my dissertation, I am hoping to connect the aspect-sentiment pairs detected through ABSA to some textual features or trends in the fanfiction data.

Evaluation of NLP Tools

Overall, the NLP tools we applied to conduct ABSA of fanfiction comments seemed well-suited to the task. Additionally, during the fellowship I played around with some of the Jupyter Notebooks for aspect-based sentiment analysis created by Tess Dejaeghere and Pranaydeep Singh (2024) and found them to be clear and accessible introductions to the practical and methodological considerations that come into play when applying ABSA to data from the literary domain.

Acknowledgements

References

Angelov, Dimo. *Top2Vec: Distributed Representations of Topics*. arXiv:2008.09470, arXiv,
19 Aug. 2020. *arXiv.org*, https://doi.org/10.48550/arXiv.2008.09470.

Barbieri, Francesco, et al. TweetEval: Unified Benchmark and Comparative Evaluation for
Tweet Classification. 26 Oct. 2020. arXiv.org, http://arxiv.org/abs/2010.12421.

Black, Suzanne R. "Adding a Digital Dimension to Fan Studies Methodologies."
Transformative Works and Cultures, vol. 33, June 2020.
https://doi.org/10.3983/twc.2020.1725.

Bogdanov, Sergei, et al. *NuNER: Entity Recognition Encoder Pre-Training via
LLM-Annotated Data*. 1, 2024. https://doi.org/10.48550/ARXIV.2402.15343.

Dejaeghere, Tess, and Pranaydeep Singh. *NER and ABSA for literary-historical corpora.*
2024. Github repository, https://github.com/GhentCDH/CLSinfra.

Hamacher, Marius, and Torsten Zesch. *INCEpTALYTICS - An Easy-to-Use API for
Analyzing INCEpTION Annotation Projects*. 0.1.1, Zenodo, 20 Sept. 2022.
https://doi.org/10.5281/ZENODO.7095346.

Klie, Jan-Christoph, et al. "The INCEpTION Platform: Machine-Assisted and
Knowledge-Oriented Interactive Annotation." *Proceedings of the 27th International
Conference on Computational Linguistics: System Demonstrations*, Association for
Computational Linguistics, 2018, pp. 5–9, http://tubiblio.ulb.tu-darmstadt.de/106270/.

Liu, Yinhan, et al. *RoBERTa: A Robustly Optimized BERT Pretraining Approach*.
arXiv:1907.11692, 26 July 2019. *arXiv.org*, http://arxiv.org/abs/1907.11692.

Mohammad, Saif M., and Peter D. Turney. *Crowdsourcing a Word-Emotion Association
Lexicon*. 28 Aug. 2013. arXiv.org, http://arxiv.org/abs/1308.6297.

Montani, Ines, et al. *SpaCy: Industrial-Strength Natural Language Processing in Python*.
v3.7.2, Zenodo, 16 Oct. 2023. https://doi.org/10.5281/ZENODO.1212303.

Neugarten, Julia. "MythFic Metadata: Gendered Power Dynamics in Fanfiction
about Greek Myth." *Digital Humanities Benelux Journal,* vol. 6, pp. 133-153, 2024.

Neugarten, J., Dejaeghere, T., Singh, P., Hemmons, A., Birkholz, J. "Catching Feelings:
Aspect-Based Sentiment Analysis for Fanfiction Comments about Greek Myth" *CHR
2024: Computational Humanities Research Conference,* pp. 217-231, 2024.

Neugarten, Julia, and Roel Smeets. *MythFic Metadata: Exploring Gendered Violence in
Fanfiction about Greek Mythology*. Radboud University, 28 Apr. 2023.
https://doi.org/10.34973/2MYE-8468.

Ramshaw, Lance A., and Mitchell P. Marcus. Text Chunking Using Transformation-Based Learning. 23 May 1995. arXiv.org, http://arxiv.org/abs/cmp-lg/9505040.

Rebora, Simone. "Sentiment Analysis in Literary Studies. A Critical Survey." *Digital Humanities Quarterly*, vol. 017, no. 2, July 2023.

Rodrigues, Francisco. *AO3 API 2.07*. 2021, https://github.com/ArmindoFlores/ao3_api.

Sainz, Oscar et. al. "GoLLIE: Annotation Guidelines improve Zero-Shot Information-Extraction." *The Twelfth International Conference on Learning Representations*. 2024. https://openreview.net/forum?id=Y3wpuxd7u9.

Samutina, Natalia. "Emotional Landscapes of Reading: Fan Fiction in the Context of Contemporary Reading Practices." *International Journal of Cultural Studies*, vol. 20, no. 3, May 2017, pp. 253–69. https://doi.org/10.1177/1367877916628238.