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

Julia Neugarten[1,*], Tess Dejaeghere[2,‡], Pranaydeep Singh[3,‡], Amanda Robin Hemmons[2] and Julie M. Birkholz[2,4]

[1]Arts and Culture Studies, Radboud University, The Netherlands
[2]GhentCDH-Ghent Center for Digital Humanities and Department of History, Ghent University, Belgium
[3]LT3 -Department of Translation, Interpreting and Communication, Ghent University, Belgium
[4]KBR- Royal Library of Belgium, Belgium

Abstract

The application of sentiment analysis in literary studies has been limited and often criticized, yet aspect-based sentiment analysis (ABSA) offers interesting applications in this domain because it addresses some limitations of traditional SA tools and provides a more detailed and context-sensitive analysis of sentiment. To investigate its usage in literary reception studies, we apply ABSA to a corpus of ±25,000 comments written by readers in response to fanfiction about Greek mythology on fanfiction website *Archive of Our Own* (AO3), one of the largest platforms for fanfiction in English. Our ABSA pipeline detects sentiment (positive/negative) associated with eight aspects of fanfiction stories (general evaluation, Greek mythology, character, character emotion, reading experience, writing style, events and storyworld, and non-specific sentiment). We explain the process of data collection and annotation and present a small inter-annotator agreement study (Pairwise Cohen's κ 0.86 for aspects and 0.88 for sentiments). We develop, evaluate, and fine-tune a machine-learning pipeline for ABSA, tackling the aspect extraction and sentiment analysis tasks respectively. We obtain the best results using NuNER for aspect extraction (0.5 macro F1) and Twitter-roBERTa-sentiment for sentiment analysis (0.75 macro F1). Finally, we outline some avenues for future research and reflect on the generalizability of our method to other domains, especially to fanfiction from other fandoms and platforms but also other social media.

Keywords

aspect-based sentiment analysis (absa), online discourse, fanfiction, reader response, classical reception, web data

1. Introduction

Fanfiction – stories inspired by existing works, written by and for an audience of fans, and published for free online – is a burgeoning domain of online cultural activity. *Archive of Our Own* (AO3), one of the largest English-language websites for publishing and reading fanfiction, recently reported hosting over 13 million works [31]. Fan studies scholars emphasize the need to understand fanfiction in relation to its community contexts, because fanfiction is usually

CHR 2024: Computational Humanities Research Conference, December 4–6, 2024, Aarhus, Denmark
*Corresponding author.
†These authors contributed equally.
julia.neugarten@ru.nl (J. Neugarten); tess.dejaeghere@ugent.be (T. Dejaeghere); pranaydeep.singh@ugent.be (P. Singh); amanda.hemmons@ugent.be (A. R. Hemmons); julie.birkholz@ugent.be (J. M. Birkholz)
0000-0003-3314-9445 (J. Neugarten)

written ”within and to the standards of a particular fannish community [9]” and ”is often so deeply embedded within a specific community that it is practically incomprehensible to those who don't share exactly the same set of references [35].” As such, ”a fan fiction story cannot be viewed as a wholly self-contained object, a text delimited by the first and last words that appear on the screen, the way that readers of novels and other genres of print culture conventionally read books as bounded by their covers [11].” Due to the deeply intertwined nature of fanfiction texts and the shared practices and contexts of their online community – which fans call *fandom* – fanfiction texts must be approached ”as both literary and digital objects [4].” Scholars need methods to study the ways fanfiction is received, embedded and evaluated in online spaces like AO3 that account for its hybrid literary/digital nature.

Additionally, fanfiction often revolves around emotional trajectories. Love, heartbreak, mental illness and other emotional highs and lows are frequently central to fanfiction, which has been described as ”emotional landscapes of reading [33]” both because of its intradiegetic emphasis on emotionality and because of its capacity to evoke strong emotions in readers.

The comment-section on AO3 lets readers share their affective responses to fanfiction reading, which we conceptualize as sentiments in the current paper. Sentiments can develop over the course of the reading and may extend or endure into the time after. Through comments, readers express these sentiments in the social context of their fan community. Sentiments are important for understanding which aspects of fanfiction are loved, enjoyed, and appreciated. The hybrid digital/textual nature of fanfiction comment, the enormous scale of fanfiction and comment production and exchange as born-digital objects, and the fanfiction community's emphasis on both intradiegetic and extradiegetic emotion make fanfiction comments suitable for computational approaches such as sentiment analysis (SA).

The application of Natural Language Processing (NLP) methodologies to date in literary studies has been limited [5, 23]. Specifically, SA, which categorizes texts as positive, neutral, or negative, is often criticized and considered inadequate for the detail-oriented research needs of literary scholars [20, 34]. The sentiments present in literary or narrative text often cannot be reduced to a negative/positive binary. To address some of the limitations of existing SA tools, aspect-based sentiment analysis (ABSA) has gained attention. Unlike conventional SA that labels the sentiment of entire documents, paragraphs, or sentences, ABSA operates at the aspect level, extracting specific aspects and their corresponding sentiment polarities [3, 36]. Although ABSA offers a more granular approach to sentiment mining, its application has largely been confined to commercial domains such as customer reviews [36] and applications in (computational) literary studies still need to be explored.

To fill these research gaps – the application of ABSA to literary studies data and the use of NLP to analyze textual/digital fanfiction comments – we employ ABSA to detect the aspects of fanfiction that readers refer to when writing comments and the sentiments commenters attach to these aspects, conceptualized in a binary fashion (positive/negative). We address a methodologically-oriented research question about fanfiction reception on AO3:

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

In other words, we investigate whether ABSA is an effective approach to analyzing the aspect-sentiment combinations in a dataset of fanfiction comments. We limit this research

Table 1
Descriptive statistics of character- and wordcount in the comment-corpus

stat	characters	words
mean	210.11	37.91
std	388.78	68.25
min	1	1
25%	44	8
50%	101	18
75%	223	40
max	9,088	1,659
sum	5,456,580	984,400

to one platform (AO3) for reasons of scope. We focus on the Greek mythology fandom because *Catching Feelings* is part of the larger project *Anchoring and Innovating Greek Myth in Contemporary Online Fanfiction* (2022-2026), which focuses on the way contemporary online fanfiction transforms cultural material about Greek mythology and the way this cultural material resonates with online fan communities.[1]

Section 2 outlines the processes of data collection and annotation. Section 3 describes our ABSA pipeline. Section 4 reports on the inter-annotator-agreement of the annotated dataset, and presents the results of various ABSA pipelines. Section 5 examines these results. Finally, Sections 6 reflects on the reproducibility, generalizability and limitations of our approach to ABSA for fanfiction comments and outlines some avenues for future research, and in Section 7 we answer our research question.

2. Data

2.1. Data Collection

We based our comment dataset on *MythFic Metadata* [29], an existing collection of metadata for 5.155 works of fanfiction about Greek mythology that were published on AO3 between 2008 and 2022.[2] We used the list of unique identifiers for the stories in *MythFic Metadata* as input for the AO3-api [14] to collect comments on those stories. The resulting dataset contains 25.970 comments. Table 1 provides descriptive statistics of the comments in terms of character- and word count.

We must address three limitations of this dataset. First, fanfiction comments, like any type of textual review or comment, do not fully capture the complex mental and emotional processes of reading experience and evaluation. Instead, these comments reflect only the elements of the reading experience that commenters are consciously aware of and chose to comment on, in the way they chose or were led by discursive norms to represent these elements textually. Second, commenters are only part of the users of AO3, and may not be representative of the reader reception happening all over the platform, let alone the fanfiction reception happening

[1] To find out more about this project, visit the *Anchoring Innovation* website.
[2] 2008 is the year that AO3 went into open beta.

on other platforms. In a 2013 census of AO3, 43.6% of respondents reported commenting on stories they had read (n = 4,358) [7]. When *MythFic Metadata* was collected, 1,584 stories in the fandom *Ancient Greek Religion and Lore*, or 30.7%, of the fandom's works, had not been commented on at all. Additionally, commenters can be characterized as above-averagely engaged or committed readers, since they invest the time and effort to comment. Finally, the content of fanfiction comments is co-shaped by the culture and affordances of AO3. Like with online book reviews, social factors like "the interactive nature of the reviewing platform or reviewers' desire to cultivate a persona or gain followers [22]" influence these texts. Additionally, fan communities differ between platforms because "the relationship that fans have to technology and online platforms is integral to the culture of these communities [13]." The exact dynamics of such relationships remain largely unknown and so the generalizability of our findings to other platforms may be limited.

2.2. Annotation

Of the 25,970 comments, 25,240 contain text and 730 contain only emojis. The emoji-only comments were discarded in our subsequent analysis. From the textual comments, we created an annotated subset using the following criteria:

- Length: 100 - 4,000 characters.[3]
- Language: English as predicted by the Python package langid [25] with high confidence (>0.9).

From the remaining 12,866 comments, we randomly sampled 1,000 to create the annotations.

Following existing research, we define sentiments as "social constructs of emotions that develop over time and are enduring [10]." Since we conceptualize aspects here as particular features of a fanfiction story that elicit sentiments, it becomes possible to identify and computationally extract positive and negative sentiments as they relate to particular aspects of fanfiction reception.

To develop an annotation guide we conducted exploratory analysis of the comment data, such as an analysis of the most-frequent noun chunks as detected by SpaCy [17] and a topic model of longer comments (>300 characters) created with Top2Vec [1]. Based on these explorations, we formulated a scheme of 8 aspects of fanfiction that were frequently commented on. Table 2 provides examples of each aspect-category. Following the methodological principle of ethical fabrication [19] these examples were fabricated to protect the privacy of commenters while still mirroring as closely as possible the content of the data. Some aspects were chosen bottom-up, i.e. based on the exploratory analysis, and others were chosen top-down, i.e. based on aspects of comment-data that interest us.

The eight resulting aspect-categories are:

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

[3]Shorter comments convey little to no evaluation, longer comments tend to go off-topic or quote extensively from the fanfiction. The length-limitation for writing comments on AO3 is 10,000 characters.

Table 2
Aspect Example Sentences

aspect	example sentence
NULL	'OMG amazing!!'
canon	'a great retelling of *The Iliad*.'
character	'Achilles is my favorite.'
emotion	'aawh it's great to see Odysseus happy.'
events and storyworld	'the Underworld is soooo spooky.'
general	'Great story!'
reading experience	'I couldn't stop reading this.'
style	'Your writing style is wonderful.'

- **Character**: reference to and assessment of character, characterization, character appearance and relationships between characters from the story.
- **Emotion**: references to the emotions experienced by characters intradiegetically.
- **Events and Storyworld**: reference to plot events and settings, specific scenes, world building, story content like tropes, and general plot elements like twists or endings.
- **General**: reference to the story as a whole or in general terms.
- **Reading Experience**: reference to reading experience, such as emotional engagement (of the reader, not the characters), absorption or narrative tension.
- **Style**: reference to how the story was written down or rendered, including writing style, word choices, metaphors, turns of phrase, voice, perspective.
- **NULL**: expressions of sentiment that do not refer to specific words or aspects.

One central aspect of fanfiction comments we exclude is expressions that foster a bond with the author. Because fanfiction exchange relies on a gift economy [16], expressions of gratitude, appreciation and encouragement are an important part of commenting, and phrases like "Thank you for writing this" or "Please write more!" occur very frequently. An existing LDA topic model of fanfiction comments showed that "author encouragements" and "requests for story" were among the most prominent topics in comments on another fanfiction website, Fanfiction.net [26]. However, because these types of expressions do not explicitly evaluate aspects of stories themselves, but rather relate to the status and position of authors and readers in the community, we disregard them here.

We used INCepTION [21] to annotate (see Figure 1 for an example). The first author annotated all 1,000 comments in the annotation-set. One co-author annotated an overlapping 100 comments to determine inter-annotator agreement (Section 4.1).

3. Aspect-Based Sentiment Analysis

Aspect-based sentiment analysis (ABSA) is a type of information extraction (IE) task consisting of two subtasks: aspect extraction (AE) and sentiment analysis (SA) of the sentiments associated with these aspects. Approaches to these tasks in and beyond computational humanities include rule-based models [22], machine learning models [3, 10, 36, 30] and, more recently,

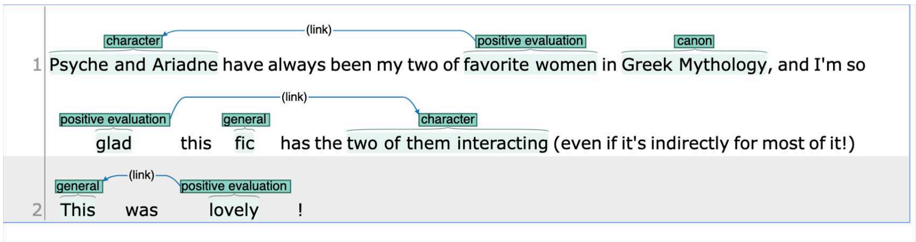

Figure 1: Example of Annotation in INCepTION

generative large language models [12, 18, 8]. We selected a machine learning approach because (a) formulating domain-specific rules is time-consuming and would not generalize well over other datasets, and (b) while approaches using generative Large Language Models (LLMs) are currently in full development, the effects of their opaque nature, tendency to propagate bias and hallucinate are not yet fully understood in IE-applications for the literary domain [12]. However, exploring the use of LLM's for ABSA presents an interesting avenue for future research.

Given the evaluative nature of the comment section on AO3, ABSA could be a good fit for assessing the sentiments of fanfiction readers on specific aspects of these texts. At the same time, the strong influence of subcultural community norms on the description of aspects and the expression of sentiments in these comments may make it challenging to apply an out-of-the-box method.

4. Results

4.1. Evaluating the Annotations

Before training a model on our annotations, we conducted a small-scale inter-annotator agreement study of the 100 comments annotated by both annotators. Our inter-annotator agreement was exceptionally high: pairwise Cohen's κ of 0.86 (aspects) and 0.88 (sentiments). This is likely due to the intensive collaboration and discussion between the annotators on the annotation guidelines and because of our shared familiarity with the discursive and social norms of fanfiction communities. The only categories showing notable confusion were **Events and Storyworld** and **Character**, with 9 instances of confusion between annotators. This is somewhat to be expected since characters often participate in story-events and always exist in the storyworld, so the distinction between these aspects can be blurry. Both annotators found a clear prevalence of positive evaluations (n=206 agreed occurrences in the annotation-set) of aspects over negative evaluations (n=4). This aligned with our expectations based on the exploratory data analysis and on a review of the existing literature, where research has shown that fanfiction comments are often written in a "positive style [28]."

Table 3
Annotation Frequencies for Aspects

aspects	frequency inter-annotator set	frequency full set
NULL	173	736
canon	87	328
character	419	1.061
emotion	60	143
events and storyworld	136	550
general	229	849
reading experience	97	521
style	87	406

Table 4
Annotation Frequencies for Sentiments

sentiment	frequency inter-annotator set	frequency full set
NULL	12	39
negative	39	146
positive	768	3,532

Table 5
Evaluation of Aspect-Extraction Models

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

4.2. Results of the Aspect Extraction

We tested three models for joint aspect extraction and categorization: roBERTa-base [24] (Macro F1 0.25), Twitter-roBERT-sentiment [2] (Macro F1 0.31), and NuNER [6] (Macro F1 0.45). Evaluation metrics for each of these models are listed in Table 5 and per aspect in Table 7. Since NuNER performed best, we tried to improve its results further by splitting the tasks of aspect extraction and categorization, to minimal effect (from 0.45 to 0.50 Macro F1).

Note that the numbers for support for split- and joint task approaches are different because in the split task, aspect extraction is a separate first step where some aspects may not have been detected. Additionally, the support-numbers in Table 7 do not perfectly match the number of annotations reported in Table 3, because in a handful of instances the aspect-term in an annotated sentence proved difficult to locate due to variations in spelling, punctuation, capitalization or even spaces.

Table 6
Evaluation of Sentiment Analysis Model

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

4.3. Results of the Sentiment Analysis

We used the Twitter-roBERTa-sentiment model since it has already been fine-tuned for sentiment detection, so we were expecting it to perform well on a polarity detection task. These results were decent (0.75 macro F1) (Table 6). This may be because sentiment polarity does not deal with as many categories as the aspect-task (two instead of eight), and because the model has been trained on Twitter-data, which is in some ways similar to our dataset of fanfiction-comments. Both types of texts are relatively brief, use internet jargon and often convey strong sentiments or opinions.

5. Analysis

Based on these results, the affordances and limitations of ABSA for analyzing fanfiction comments seem strongly differentiated by aspect category. As expected, the models perform best for those aspect categories that have high support in the annotations, especially **General** and **Style**. **Emotions** are particularly problematic, perhaps due to their low support but perhaps also because emotions are not commonly considered entities or aspects in NER-tasks.

Figure 2 is a confusion matrix between the best-performing aspect-extraction model (NuNER) and our gold standard annotations. As you can see, the only significant source of confusion was between **Character** and **Event & Storyworld** aspects, which makes sense as characters usually participate in events and interact with story worlds. Entities that the model failed to detect were left out of Figure 2 for the sake of interpretability. Manual inspection of these left-out results reveal that many **Events** were not detected. **Emotion** was also difficult for all models to detect, perhaps because words that refer to the **Emotion**-aspect also often occur in descriptions of **Reading Experience**. Detection of **Canon** was often partial, but detecting these partial matches (often revolving around words like 'retelling,' 'original' 'allusion' or 'portrayal') is still useful in mapping on a large scale how **Canon** is discussed in the comments. **Characters** that were very frequent in the data (such as Achilles) were more often accurately detected than less prominent characters (such as Demeter). **General** aspects were captured well in most models, with the exception of some relatively long and unique figures of speech, such as 'I take my hat off to you.' Non-specific expressions of affect were also detected relatively well, including the kinds of borderline-incoherent expressions that are specific to fanfiction communities. Examples of these include 'WAHHH' and 'HNNG'. NuNER even identified 'skdjahskdjhsd' as a NULL-aspect; correctly, in our opinion, as keyboardsmashes often convey strong but non-specific sentiments.

Figure 3 is a confusion matrix between the two annotators, for whom the **Character** and **General** aspects were easiest to agree on. With the **General** aspect, then, the skills of hu-

Table 7
Evaluation of Aspect-Extraction Models per Aspect

roBERTa-base	precision	recall	f1-score	support
character	0.18	0.33	0.23	110
general	0.64	0.72	0.68	112
event	0.04	0.08	0.05	66
NULL	0.24	0.31	0.27	71
reading	0.16	0.16	0.16	57
style	0.41	0.48	0.44	50
canon	0.16	0.21	0.18	39
emotion	0.00	0.00	0.00	11
micro avg	0.26	0.36	0.30	516
macro avg	0.23	0.28	0.25	516
weighted avg	0.29	0.36	0.32	516
Twitter-roBERTa-sentiment	**precision**	**recall**	**f1-score**	**support**
character	0.21	0.35	0.27	110
general	0.69	0.72	0.71	112
event	0.05	0.08	0.06	66
NULL	0.48	0.41	0.44	71
reading	0.31	0.19	0.24	57
style	0.42	0.46	0.44	50
canon	0.14	0.26	0.18	39
emotion	0.50	0.09	0.15	11
micro avg	0.32	0.39	0.35	516
macro avg	0.35	0.32	0.31	516
weighted avg	0.36	0.39	0.36	516
NuNER	**precision**	**recall**	**f1-score**	**support**
character	0.38	0.45	0.41	110
general	0.74	0.78	0.76	112
event	0.28	0.42	0.34	66
NULL	0.62	0.58	0.60	71
reading	0.29	0.32	0.30	57
style	0.63	0.74	0.68	50
canon	0.29	0.33	0.31	39
emotion	0.18	0.27	0.21	11
micro avg	0.46	0.53	0.50	516
macro avg	0.42	0.49	0.45	516
weighted avg	0.48	0.53	0.50	516
NuNER (split-task)	**precision**	**recall**	**f1-score**	**support**
micro avg	0.47	0.55	0.50	511
macro avg	0.47	0.55	0.50	511
weighted avg	0.47	0.55	0.50	511

man annotators overlap with those of NuNER. For **Style**, it seems that NuNER performed well because **Style** is often described in the comments using a relatively specific and limited vocabulary. In our test set, the **Style** tokens contained only 30 unique words, compared to 53 for **Emotion**, 93 for **Canon** and 287 for **Events & Storyworld**. The most common words for **Style** are also quite consistent and content-related, such as ’written’, ’writing’, ’describe’, ’language’, and ’words. In comparison, the most common words detected in other aspect-categories are often stop-words. This may make **Style** relatively easy to detect computationally.

With NuNER’s help, we also found some small errors in our annotations, such as character-names that were overlooked in the part of the annotation-set that was only annotated by one author. In future work, an annotation effort where each text is annotated by multiple people may provide better results.

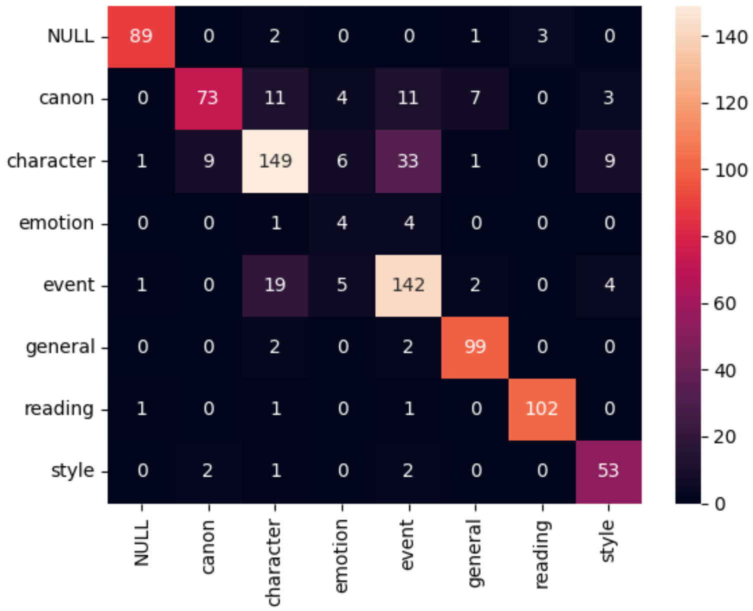

Figure 2: Confusion Matrix Gold Standard Annotation vs. best model (NuNER)

6. Discussion

The machine-learning approach to ABSA is more generalizable to other domains, especially other fandoms, than a rule-based model would be. Particularly in the relatively successful categories of **General** and **Style**, we expect no significant differences between fandoms in how commenters refer to these aspects of fanfiction. We do expect, however, that our model will be difficult to generalize across different platforms for fanfiction exchange, as research has shown that ”our lack of knowledge about platform influence on norms of commenting makes it difficult to generalize about the dynamics of fanfiction commenting beyond AO3 [28].” The importance of subcultural norms in expressing the sentiments of members of fanfiction communities like

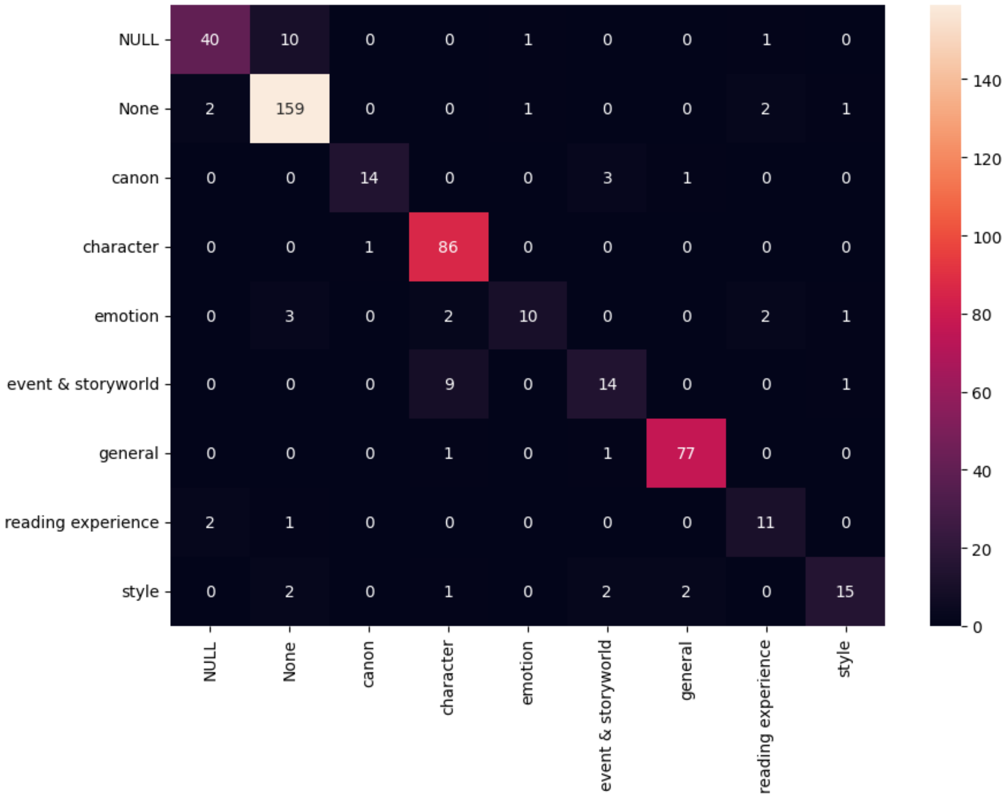

Figure 3: Confusion Matrix Between Two Annotators

AO3 may also make it difficult to apply our ABSA pipeline to other types of comments, such as book reviews on Goodreads or comments on Youtube.

We want to highlight five potential directions for future research. Firstly, it may improve results to annotate more data in a targeted way, supplementing those aspects for which support was low. Secondly, it may be of interest to analyze the 730 comments containing only emojis, using, for example, the Multidimensional Lexicon of Emojis [15]. Third, it may be interesting to tackle the task of ABSA with a generative LLM, since these models are gaining ground as information extractors. GoLLie [32] looks especially promising, as this prompting framework would allow us to feed our existing detailed guidelines for annotation directly to a generative model. Fourth, a hybrid approach to ABSA, particularly adding some rule-based elements to our existing ML pipeline, may be a productive 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 [27] to detect the **Emotion** aspect and create lists of characters and story settings based on existing metadata [29] to detect the **Character** aspect and the storyworld-dimension of the **Events and Storyworld** aspect. Finally, a fruitful next step in the field of fandom studies would be to examine the outcomes of our current best model in more detail to generate domain-specific insight into the ways sentiment is expressed and attached to particular aspects of fanfiction in the comment data.

7. Conclusion

NuNER delivered the best performance at our aspect-extraction task. Although results were poor for some aspects (especially **Emotion** but also **Events & Storyworld**), other aspects (especially **General** and **Style**) were more accurately detected. Furthermore, since the model detected many partial matches in the aspect extraction for aspects such as **Canon**, results are still very useful to researchers wanting to know which aspects are mentioned in connection to positive and negative sentiments in the fanfiction comment data. Our approach of using Twitter-roBERTa-sentiment proved sufficiently suitable to detecting positive and negative sentiment. Overall, fanfiction comments seem suited to ABSA because many of the texts explicitly express sentiments regarding specifically named aspects of fanfiction stories.

8. Data Access Statement

To protect the privacy of the fanfiction community, the dataset of fanfiction comments will not be made available for reuse. Derived data and code are available on Github.

9. Acknowledgments

This research was conducted as part of a Transnational Access Fellowship at GhentCDH & Lt3 at Ghent University in the project Computational Literary Studies Infrastructure (CLSINFRA), which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101004984.

The project *Anchoring and Innovating Greek Myth in Contemporary Online Fanfiction* is supported by the Dutch ministry of Education, Culture and Science (OCW) through the Dutch Research Council (NWO), as part of the Anchoring Innovation Gravitation Grant research agenda of OIKOS, the National Research School in Classical Studies, the Netherlands (project number 024.003.012).

References

[1] D. Angelov. *Top2Vec: Distributed Representations of Topics.* arXiv preprint. 2020. DOI: 10.48550/arXiv.2008.09470.

[2] F. Barbieri, J. Camacho-Collados, L. Neves, and L. E. Anke. "TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification". In: *CoRR* abs/2010.12421 (2020). URL: https://arxiv.org/abs/2010.12421.

[3] M. Birjali, M. Kasri, and A. Beni-Hssane. "A comprehensive survey on sentiment analysis: Approaches, challenges and trends". In: *Knowledge-Based Systems* 226 (2021), p. 107134. DOI: 10.1016/j.knosys.2021.107134.

[4] S. R. Black. "Adding a digital dimension to fan studies methodologies". In: *Transformative Works and Cultures* 33 (2020). DOI: 10.3983/twc.2020.1725.

[5] C. Blevins and A. Robichaud. *2: A Brief History » Tooling Up for Digital Humanities.* 2011.

[6] S. Bogdanov, A. Constantin, T. Bernard, B. Crabbé, and E. Bernard. *NuNER: Entity Recognition Encoder Pre-training via LLM-Annotated Data*. arXiv preprint, arXiv:2402.15343. 2024.

[7] centreoftheselights. *AO3 Census: Site Use*. https://archiveofourown.org/works/16988355/chapters/39932727. 2013.

[8] J. Chun and K. Elkins. "eXplainable AI with GPT4 for story analysis and generation: A novel framework for diachronic sentiment analysis". In: *International Journal of Digital Humanities* 5.2 (2023), pp. 507–532. DOI: 10.1007/s42803-023-00069-8.

[9] F. Coppa, ed. *The Fanfiction Reader: Folk Tales for the Digital Age*. Ann Arbor: University of Michigan Press, 2017.

[10] G. D'Aniello, M. Gaeta, and I. La Rocca. "KnowMIS-ABSA: an overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis". In: *Artificial Intelligence Review* 55.7 (2022), pp. 5543–5574. DOI: 10.1007/s10462-021-10134-9.

[11] A. De Kosnik. *Rogue Archives: digital cultural memory and media fandom*. Cambridge: MIT Press, 2016.

[12] T. Dejaeghere, E. Lefever, P. Singh, and J. Birkholz. "Evaluating generative LLMs for named entity recognition in literary-historical texts". In: *Proceedings of the 34th Meeting of Computational Linguistics in The Netherlands (CLIN 34)*. Clin. Leiden, the Netherlands, forthcoming, forthcoming.

[13] C. Fiesler and B. Dym. "Moving Across Lands: Online Platform Migration in Fandom Communities". In: *Proceedings of the ACM on Human-Computer Interaction* 4.Cscw1 (2020), pp. 1–25. DOI: 10.1145/3392847.

[14] A. Flores. *ao3_api*. https://github.com/ArmindoFlores/ao3_api. 2022.

[15] R. Godard and S. Holtzman. "The Multidimensional Lexicon of Emojis: A New Tool to Assess the Emotional Content of Emojis". In: *Frontiers in Psychology* 13 (2022). DOI: 10.3389/fpsyg.2022.921388.

[16] K. Hellekson. "A Fannish Field of Value: Online Fan Gift Culture". In: *Cinema Journal* 48.4 (2009), pp. 113–118. URL: https://www.jstor.org/stable/25619733.

[17] M. Honnibal and I. Montani. *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*. url/https://zenodo.org/doi/10.5281/zenodo.1212303. 2023.

[18] E. Hosseini-Asl, W. Liu, and C. Xiong. *A Generative Language Model for Few-shot Aspect-Based Sentiment Analysis*. arXiv preprint. 2022. DOI: 10.48550/arXiv.2204.05356.

[19] B. Kelley. *Loving fanfiction: exploring the role of emotion in online fandoms*. London ; New York: Routledge, Taylor & Francis Group, 2021.

[20] E. Kim and R. Klinger. "Who Feels What and Why? Annotation of a Literature Corpus with Semantic Roles of Emotions". In: *Proceedings of the 27th International Conference on Computational Linguistics*. Santa Fe, New Mexico, USA: Association for Computational Linguistics, 2018, pp. 1345–1359. URL: https://aclanthology.org/C18-1114.

[21] J.-C. Klie, M. Bugert, B. Boullosa, R. E. de Castilho, and I. Gurevych. “The INCEpTION platform: Machine-assisted and knowledge-oriented interactive annotation”. In: *Proceedings of the 27th international conference on computational linguistics: System demonstrations.* Santa Fe, USA: Association for Computational Linguistics, 2018, pp. 5–9. URL: http://tubiblio.ulb.tu-darmstadt.de/106270/.

[22] M. Koolen, J. Neugarten, and P. Boot. “‘This book makes me happy and sad and I love it’ A Rule-based Model for Extracting Reading Impact from English Book Reviews”. In: *Journal of Computational Literary Studies* 1.1 (2022), pp. 1–22. DOI: 10.48694/jcls.104.

[23] J. Kuhn and N. Reiter. *A Plea for a Method-Driven Agenda in the Digital Humanities.* Sydney, Australia: Adho, 2015.

[24] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. *Roberta: A robustly optimized bert pretraining approach.* arXiv preprint, arXiv:1907.11692. 2019.

[25] M. Lui and M. Baldwin. “langid.py: An Off-the-shelf Language Identification Tool”. In: *Proceedings of the ACL 2012 System Demonstrations.* Jeju Island, Korea, 2012, pp. 25–30. URL: https://aclanthology.org/P12-3005.

[26] S. Milli and D. Bamman. “Beyond Canonical Texts: A Computational Analysis of Fanfiction”. In: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing.* Austin, Texas: Association for Computational Linguistics, 2016, pp. 2048–2053. DOI: 10.18653/v1/D16-1218.

[27] S. M. Mohammad and P. D. Turney. “Crowdsourcing a Word-Emotion Association Lexicon”. In: *Computational Intelligence* 29.3 (2013), pp. 436–465.

[28] J. Neugarten. “Delicious Angst and tooth-rotting fluff: Distant reading community discourses of emotion in Harry Potter fanfiction comments”. In: *Journal of Fandom Studies* 11.2&3 (2023), pp. 205–28. DOI: https://doi.org/10.1386/jfs_00082_1.

[29] J. Neugarten and R. Smeets. *MythFic Metadata: Exploring Gendered Violence in Fanfiction about Greek Mythology.* Radboud Data Repository. 2023. DOI: 10.34973/2mye-8468.

[30] D. Nguyen, M. Liakata, S. DeDeo, J. Eisenstein, D. Mimno, R. Tromble, and J. Winters. “How We Do Things With Words: Analyzing Text as Social and Cultural Data”. In: *Frontiers in Artificial Intelligence* 3 (2020), p. 62. DOI: 10.3389/frai.2020.00062.

[31] C. Rebaza. *AO3 Celebrates 13 Million Fanworks.* webpage. 2024. URL: https://www.transformativeworks.org/ao3-celebrates-13-million-fanworks/.

[32] O. Sainz, I. García-Ferrero, R. Agerri, O. L. de Lacalle, G. Rigau, and E. Agirre. “GoLLIE: Annotation Guidelines improve Zero-Shot Information-Extraction”. In: *The Twelfth International Conference on Learning Representations.* Vienna, Austria, 2024, pp. 1–25. URL: https://openreview.net/forum?id=Y3wpuxd7u9.

[33] N. Samutina. “Emotional landscapes of reading: fan fiction in the context of contemporary reading practices”. In: *International Journal of Cultural Studies* 20.3 (2017), pp. 253–269. DOI: 10.1177/1367877916628238.

[34] T. Schmidt and M. Burghardt. “An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing”. In: *Proceedings of the Second Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. Ed. by B. Alex, S. Degaetano-Ortlieb, A. Feldman, A. Kazantseva, N. Reiter, and S. Szpakowicz. Santa Fe, New Mexico: Association for Computational Linguistics, 2018, pp. 139–149. URL: https://aclanthology.org/W18-4516.

[35] C. Tosenberger. “Mature Poets Steal: Children’s Literature and the Unpublishability of Fanfiction”. In: *Children’s Literature Association Quarterly* 39.1 (2014), pp. 4–27. DOI: 10.1353/chq.2014.0010. URL: http://muse.jhu.edu/content/crossref/journals/childrens%5C%5Fliterature%5C%5Fassociation%5C%5Fquarterly/v039/39.1.tosenberger.html.

[36] W. Zhang, X. Li, Y. Deng, L. Bing, and W. Lam. *A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges*. arXiv preprint. 2022. URL: http://arxiv.org/abs/2203.01054.