Cross-Modal Modeling of Emotional and Thematic Trajectories in Holocaust Survivor Oral Histories

Henry Gagnier
Pittsford Sutherland High School
Pittsford, NY, USA
henrygagnier9@gmail.com

Abstract

Large-scale corpora of Holocaust testimonies preserve vast amounts of historical, emotional, and narrative information, but their size and complexity can make accurate, systematic analysis challenging. This paper presents a cross-modal computational analysis of emotional and thematic trajectories in the CORHOH corpus, containing 500 Holocaust survivor testimonies as a language resource for computational analysis. We segment each testimony into ten segments and apply sentiment analysis, emotion recognition, and topic modeling to each of these segments to reveal how theme and emotion evolve over time in Holocaust testimonies. Results reveal a sharp decline from pre-war life in wartime and camp experiences, with sentiment and emotion remaining negative in post-war segments. Emotion analysis reveals decreasing joy and increasing sadness and fear during segments related to deportation and concentration camps, with limited emotional recovery. Topic modeling identifies coherent themes that align closely with sentiment and emotional patterns. We systematically examine correlations between sentiment, emotion, and topic trajectories, which demonstrate many strong associations between topic and emotion. This work demonstrates that combining sentiment analysis, emotion recognition, and topic modeling can reveal systematic patterns in large oral history corpora, and shows the value of computational approaches for studying historical narratives like the Holocaust.

Keywords: Emotion Analysis, Sentiment Analysis, Topic Modeling, Holocaust, Testimonies

1. Introduction

Major efforts have been made to collect and digitize testimonies of Holocaust survivors, creating vast corpora of important information. This presents challenges, making it difficult to attend to all testimonies while preserving their narrative integrity and emotional dimensions (Wagner et al., 2024b; Ifergan et al., 2024). Advances in computational linguistics and natural language processing (NLP) offer new possibilities to understand these narratives at scale while respecting the integrity and uniqueness of each story.

Holocaust testimonies are a unique type of language resource as first-person oral narratives that document historical trauma at a large scale. By treating Holocaust testimonies as language resources rather than simply historical documents, we can apply NLP to reveal patterns across hundreds of documents while preserving the integrity of these testimonies. This work demonstrates how computational linguistic methods can systematically extract insights from Holocaust testimony corpora, validating their status as valuable language resources for historical research and NLP methodological development.

NLP has increasingly been applied to Holocaust testimonies as language resources to extract insights more efficiently and effectively than traditional reading. Work has focused on topic-based segmentation to identify narrative structures (Wagner et al., 2024b), topic modeling with BERTopic (Ifergan et al., 2024), and analyzing character development trajectories (Shizgal et al., 2025). Research has also focused on spatial trajectory mapping (Wagner et al., 2024a). Blanke et al. (2019) applies sentiment analysis to Holocaust testimonies and finds that testimonies unsurprisingly have negative sentiment. However, less work has been done on the emotional and effective dimensions of Holocaust testimonies as they occur over time.

Sentiment analysis identifies positive or negative attitudes in text, while emotion analysis identifies specific emotions such as sadness, anger, or joy (Plaza-del Arco et al., 2024). In narrative contexts such as Holocaust testimonies, sentiment and emotion analysis can reveal narrative trends and structure correlating to narrative progression, and assist in narrative understanding (Min and Park, 2019; Zad and Finlayson, 2020). Understanding emotional trajectories in Holocaust testimonies can be a new side to thematic and narrative understanding, revealing how Holocaust victims and survivors retell their histories and construct meaning.

We apply sentiment analysis, emotion recognition, and topic modeling to the CORHOH (Text CORpus of HOlocaust Oral Histories), consisting of over 500 oral histories from Holocaust survivors. The purpose of this study is to (1) explore how sentiment evolves over the temporal arc of Holocaust testimonies, (2) find how emotion characterizes different phases of the survivors' testimonies, and (3)

Proceedings of The Second Workshop on Holocaust Testimonies as Language Resources (HTRes) @ LREC 2026, pages 93–100
11 May 2026.

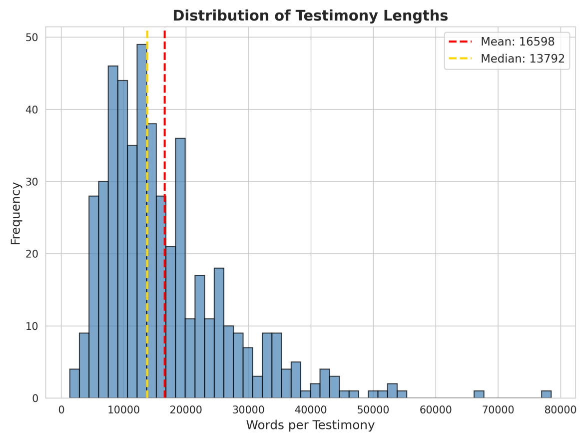

Figure 1: Distribution of CORHOH testimony lengths in words

discover what thematic topics occur with different sentiment and emotional patterns through cross-modal trajectory alignment. We also aim to encourage and continue the computational research of Holocaust testimonies and provide methodologies applicable to other oral history collections.

2. Data

2.1. CORHOH Corpus

We use the CORHOH (Text CORpus of Holocaust Oral Histories) (Jaff, 2025), which consists of 500 oral histories in English, with each narrative from one survivor. The transcripts have been preprocessed and annotated, making them suitable for topic modeling and sentiment analysis. We further isolate the survivor's speech for analysis, which is labeled in CORHOH. Testimonies ranged from 1,365 to 78,514 words, with an average of 16,598 words in a testimony. Figure 1 visualizes the distribution of the length of testimonies.

2.2. Resource Evaluation and Challenges

As a language resource, the CORHOH was very well suited to this research. Its scale of 500 testimonies provided significant volume for stable and meaningful results. The pre-processed, speaker-isolated transcripts allowed us to analyze survivor speech without extensive preprocessing, and the English-language format of all testimonies ensures compatibility with all models used in the analysis. The diversity of survivor experiences and emotions also strengthened the generalizability of our findings across different Holocaust experiences.

The CORHOH also had several challenges as a language resource. The variety of testimonial length, ranging from 1,365 to 78,514 words, meant that equal-length segmentation produced segments of substantially different sizes, which may introduce noise into segment-level comparisons. The testimonies also contain features characteristic of spoken languages, such as false starts, self-corrections, repetition, hesitation, and informal connective structures. A representative example of this is "he regist-... because he was the big shot. He was at that time a Polish officer." Standard NLP models are not designed to handle such text, and their presence may degrade classification accuracy.

2.3. Data Preprocessing

To track the emotional and thematic trajectories of testimonies, we divided each testimony into 10 equal segments based on word count. We do this for multiple reasons. The substantial variation in testimony length makes narrative-phase-based segmentation impractical without ground-truth phase boundary annotations, which are not available at this scale. Equal-length segmentation enables direct comparison between testimonies of segment-level scores at equivalent relative positions in the narrative arc. We acknowledge that narrative phases do not fall neatly at fixed word-count percentages, and that this introduces some noise into segment-level interpretations.

3. Methodology

Using a multifaceted computational approach to analyze sentiment, emotion, and topic in Holocaust survivor testimonies, we applied multiple NLP techniques to each segment to capture distinct elements of the narratives.

3.1. Sentiment Analysis

We employed two methods for sentiment analysis: lexicon-based and deep-learning approaches, applied to each segment.

We used Valence Aware Dictionary and sEntiment Reasoner (VADER) from nltk (Hutto and Gilbert, 2014), a lexicon-based sentiment analysis tool that is specifically designed for social media texts but performed well on many text types. We included VADER as a lexicon-based baseline to contrast with transformer-based sentiment models and to examine how genre mismatch affects sentiment estimation in Holocaust testimonies. We also employed the DistilBERT model fine-tuned on the Stanford Sentiment Treebank (SST-2) (`distilbert-base-uncased-finetuned-sst-2-english`) (Sanh et al., 2019) using the HuggingFace Transformers library. Both models output a score between -1 (negative sentiment) and 1 (positive sentiment) for each segment. In both models, we applied the sentiment scoring function to each testimony segment and averaged the scores across all segments at a given position in the testimony to identify the overall sentiment and its overall trajectory. Both models contain a significant risk of domain mismatch as they were trained on movie reviews and social media texts. These domains do not represent the complex, spoken, trauma-centered language of Holocaust testimonies and may affect the reliability of sentiment scores.

3.2. Emotion Analysis

In order to reveal more emotional dimensions than sentiment analysis, we use a multi-class emotion recognition model.

We used the `emotion-english-distilroberta-base` (Hartmann, 2022) model from HuggingFace Transformers, a DistilRoBERTa-based classifier that has been fine-tuned to recognize seven emotions: anger, disgust, fear, joy, neutral, sadness, and surprise. This model was fine-tuned on data from social media sources, representing a domain mismatch when applied to Holocaust testimonies. This mismatch should be considered when interpreting emotion, given the trauma-specific language of the CORHOH corpus. For each segment, we obtained probability scores for each of the seven emotions and assigned the emotion with the highest probability to the text. Next, we calculated the percentage of text at each narrative segment that is of a certain emotion and used this data to analyze how different emotions evolve throughout the narrative arcs of Holocaust testimonies.

3.3. Topic Analysis

To analyze the thematic content of the testimonies and to see how emotions and sentiment correlate with direct topics, we use BERTopic (Grootendorst, 2022), a neural topic modeling model that uses transformer-based embeddings and clustering algorithms.

We used the Sentence-BERT model (`all-MiniLM-L6-v2`) (Wang et al., 2020) to produce meaningful embeddings that capture contextual information from the text. The embeddings were reduced using UMAP (McInnes et al., 2018) with `n_neighbors` as 25, `n_components` as 5, `min_dist` as 0.0, `metric` as cosine, and `random_state` as 42. Documents were clustered using HDBSCAN (McInnes et al., 2017) with `min_cluster_size` as 25, `min_samples` as 10, `metric` as euclidean, `cluster_selection_method` as eom (excess of mass), and `prediction_data` as True. Then, we used CountVectorizer to generate interpretable topic representations with `stop_words` as English to remove common English stop words from topic analysis, `min_df` as 2, `max_df` as 0.95, `ngram_range` as (1,2), and `top_n_words` as 10. Finally, we reduced the topics to 5 using BERTopic's topic reduction functionality to ensure themes were manageable and coherent.

3.4. Cross-Modal Trajectory Alignment

To examine the relationship between sentiment, mood, and topic over time, we calculate Pearson correlation coefficients, integrating sentiment, mood, and topics at the segment level. Instead of analyzing these features independently, we quantify how emotional and thematic signals vary across the progression of testimonies.

Using the mean sentiment scores from DistilBERT, mean emotion scores, and topic prevalence, we measure the alignment between these modalities using a correlation-based analysis. DistilBERT sentiment scores are used because transformer-based models capture contextual and implicit details in narrative text and are conceptually aligned with the embedding-based topic representations generated by BERTopic. We compute the Pearson correlation coefficients between sentiment and topic, and emotion and topic, to assess sentiment-topic alignment and emotion-topic alignment, respectively. These correlations quantify the strength and direction between emotional signals and thematic content across the testimonies. As we analyze alignment over 10 segments, correlation coefficients are interpreted as indicators of effect size rather than statistical tests.

This cross-modal trajectory alignment analysis enables us to analyze Holocaust testimonies with greater depth, integrating both emotion and theme, and complementing our trajectory analyses of sentiment, emotion, and topics independently.

4. Results

4.1. Sentiment Analysis

We first look at the trajectory of sentiment in the Holocaust testimonies using VADER and DistilBERT independently (Figure 2). Despite extremely different model architectures, both models exhibit a very similar temporal pattern of sentiment, but the sentiment output is very different. As expected, given the domain mismatch, VADER classifies the sentiment as positive throughout the entire testimony. All subsequent sentiment analysis, therefore, focuses on DistilBERT results. DistilBERT classifies the sentiment as negative throughout the entire testimony, with the exception of the first segment, which mainly corresponds to pre-war life. This divergence reflects the limitations of VADER when

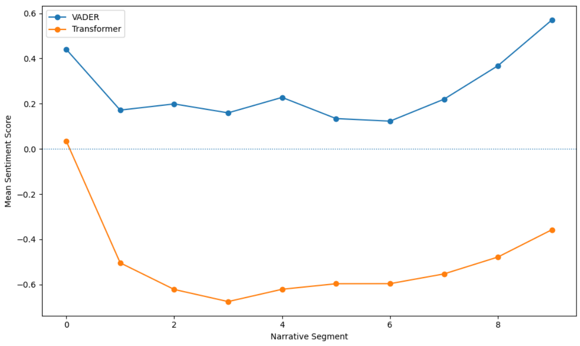

Figure 2: Trajectory of sentiment in Holocaust survivor testimonies using VADER and Transformer-based models

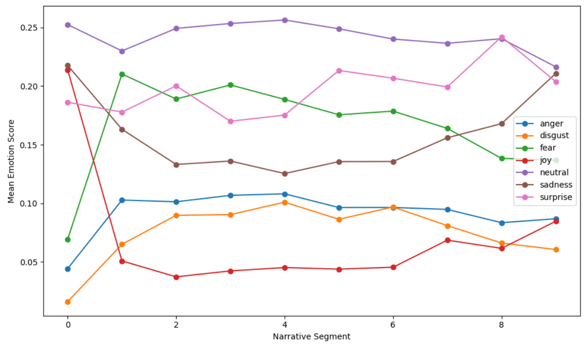

Figure 3: Trajectory of emotion in Holocaust survivor testimonies using DistilRoBERTa-base

applied to domain-specific language and context-dependent sentiment (Villanueva-Miranda et al., 2025). Both models show a rapid decrease in sentiment from the first and second segments, as the testimony transitions in themes of deportation and forced movement. In DistilBERT, sentiment remains extremely negative, ranging from -0.67 to -0.55 from the third to the eighth segment. In the final two segments, sentiment increases to -0.47 to -0.35, while not recovering to pre-war life, reflecting long-term impacts rather than emotional resolution. The similar sentiment trajectories but varying overall sentiment shows that while trajectories may be robust to model choice, overall sentiment may not be robust.

4.2. Emotion Analysis

We now analyze the emotional trajectories of the Holocaust sentiments using DistilRoBERTa-base (Figure 3). Clear phases of emotion are observed, with joy and sadness being prevalent in the first segment, corresponding to pre-war life. Surprisingly, sadness and joy are of relatively equal prevalence in the first segment, and both decrease in prevalence in the second and third segments. In the second segment, fear is the most common sentiment other than neutrality, and fear declines throughout the testimonies but remains highly prevalent even in the last segments. Conversely, sadness increases in the last segments. Neutrality and surprise both remain highly prevalent throughout the entire testimony, with scores from 0.17 to 0.25 across segments. The prevalence of sadness, fear, and lack of joy in the later segments reveals the absence of emotional recovery in Holocaust trajectories, and the high prevalence of fear and surprise throughout the central testimonies highlights the traumatic and emotional experiences throughout these testimonies.

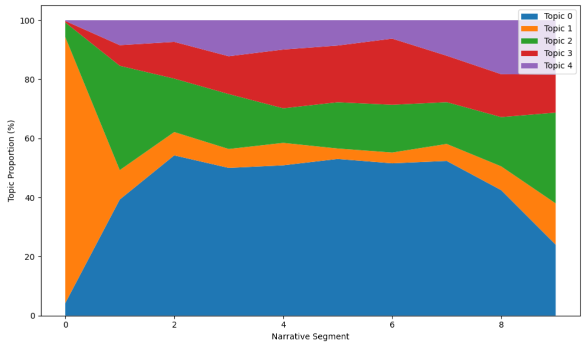

Figure 4: Topic distribution over time in Holocaust survivor testimonies excluding outliers using BERTopic

4.3. Topic Analysis

We look at the topics identified by BERTopic across the testimonies (1). The five topics identified were (0) deportation, forced movement, and camp transitions; (1) pre-war life, family, and early education; (2) community, cultural life, and social relationships; (3) camp conditions, illness, and survival practices; and (4) flight, rescue networks, and post-war displacement. In the first segment, pre-war life was the most prevalent by far, with prevalence decreasing in the following segments. Next, deportation and forced movement, and community and cultural life rise in the second segment and remain prevalent throughout the entire narrative. Camp conditions increase in prevalence throughout the first four segments, then remain constant throughout the remainder of the narratives. Rescue networks increase in prevalence throughout the narratives, with prevalence greatest in the final segments. This topic analysis closely mirrors the historical chronology of Holocaust survivor experiences.

Topic	Interpretive Label	Representative High-Weight Terms
0	Deportation, Forced Movement, and Camp Transitions	wagon; Birkenau; Vilnius; Russian army; kapos; running away; Krakow
1	Pre-War Life, Family, and Early Education	cheder; father; 1918; 1935; grades; congregation; governess
2	Community, Cultural Life, and Social Relationships	congregation; theatre; lesson; relationships; compete; culture; Radom
3	Camp Conditions, Illness, and Survival Practices	Birkenau; latrine; epidemic; inoculated; doctors; coffee beans; barracks
4	Flight, Rescue Networks, and Post-War Displacement	Vichy; Portugal; Lisbon; OSE; passengers; State Department; Le Chambon

Table 1: Interpretive labels and representative high-weight terms for topics identified using BERTopic

4.4. Cross-Modal Trajectory Alignment

4.4.1. Sentiment-Topic Alignment

To examine how sentiment relates to topic analysis in the trajectory of Holocaust testimonies, we look at the correlation coefficients between sentiment and topic prevalence (Table 2). With only ten data points, these coefficients should be interpreted as indicators of effect size, and results with $p > 0.05$ should be treated with caution. Strong negative correlations existed between sentiment and most topics. Pre-war life showed a strong negative correlation of -0.964, indicating that as sentiment becomes more negative throughout testimonies, discussion of pre-war life decreases dramatically. This reflects the progression from positive pre-war memories to negative memories of deportation and concentration camp conditions.

Deportation and forced movement also had a strong negative correlation with sentiment, indicating that in narratives of deportation, sentiment reaches extremely low values. Community and cultural life showed a moderate negative correlation of -0.753, showing that as discussions of community became less prevalent, sentiment decreased. Camp conditions and post-war displacement both produced non-significant results, and given the small number of segments, no meaningful association between these topics and sentiment trajectory should be inferred.

Topic Label	Pearson r	p-value
Pre-war Life	-0.964	0.000
Deportation	-0.931	0.000
Camp Conditions	-0.244	0.497
Community and Cultural Life	-0.753	0.012
Post-war Displacement	-0.343	0.333

Table 2: Correlation of average sentiment scores with topic prevalence across narrative segments.

4.4.2. Emotion-Topic Alignment

We now look at the emotion-topic alignment across the narrative segments to see how topic and emotion relate in Holocaust testimonies (Figure 5). As with the sentiment-topic alignment analysis, these correlations are computed across only ten segments and should be interpreted as descriptive effect-size indicators rather than statistically robust findings. This analysis reveals emotional associations with different phases of the testimonies. Pre-war life showed a strong positive correlation with joy ($r = 0.97$) and strong negative correlations with anger ($r = -0.91$), disgust ($r = -0.87$), and fear ($r = -0.83$). This reflects the emotional positivity of pre-war life, which is characterized by family life and childhood experiences. Deportation and forced movement topics had strong positive correlations with fear ($r = 0.83$), disgust ($r = 0.94$), and anger ($r = 0.87$), while having strong negative correlations with sadness ($r = -0.93$) and joy ($r = -0.91$), showing the traumatic experiences of forced displacement and separation. Camp conditions demonstrated strong positive correlations with disgust ($r = 0.86$) and negative correlations with both joy (-0.73) and sadness (-0.70). Community and cultural life showed moderate positive correlations with fear ($r = 0.47$) and anger ($r = 0.43$) and a moderate negative correlation with joy ($r = -0.42$), potentially showing that community was still a primary area of fear during the Holocaust. Flight and post-war displacement had a moderate negative correlation with joy ($r = -0.45$) and a moderate positive correlation with surprise ($r = 0.42$), revealing that joy does not recover as survivors speak about their post-war experiences.

5. Discussion

We conduct a computational analysis of Holocaust testimonies from the CORHOH, revealing systematic emotional and thematic trajectories in Holocaust survivor testimonies. We applied sentiment analysis, emotion analysis, and topic modeling to 500 testimonies and examined how these dimensions evolve over narrative progression. Testimonies exhibit a sharp decline in sentiment from pre-war life to wartime segments and remain neg-

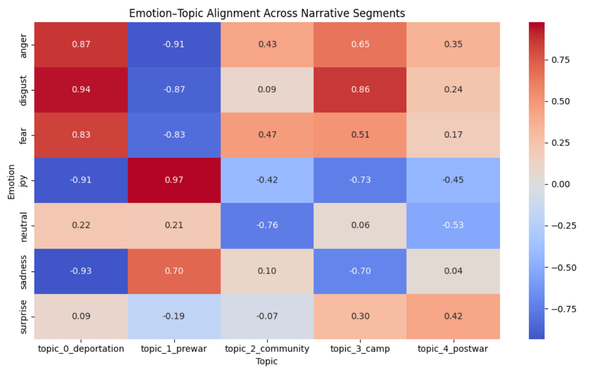

Figure 5: Heatmap showing correlations between emotion and topic prevalence across Holocaust testimony narrative segments

ative in post-war segments. Emotionally, joy decreases and fear increases during wartime segments. Cross-modal trajectory alignment reveals strong correlations between topics, such as a correlation of 0.97 between pre-war topics and joy.

Unexpectedly, joy and sadness were balanced in pre-war life, likely reflecting bittersweet pre-war memories. Sadness negatively correlated with deportation topics while fear, disgust, and anger had strong positive correlations, suggesting that survivors recall deportation through trauma and anger rather than sadness. Camp conditions negatively correlated with joy and sadness, potentially indicating emotionally detached and traumatic language when speaking about camp conditions.

Future work can replace fixed-length narrative segmentation with topic-based segmentation to allow emotion and sentiment analysis to align more with narrative structure and topics. Using different, finer emotion models or models fine-tuned for historical or trauma-related corpora could enrich emotion modeling. Survivor-level trajectory clustering could also reveal distinct emotional types rather than corpus-level averages.

We use fine-tuned transformer models in this work instead of more recent large language models (LLMs) with zero-shot or few-shot prompting. At a scale of 500 testimonies divided into ten segments each, LLM-based inference would introduce substantial computational cost. Fine-tuned models also offer greater reproducibility, as their outputs are deterministic, whereas LLM results may vary across runs and API versions. Future work should explore whether zero-shot or few-shot prompting with LLMs allows for better-suited emotion or sentiment classifications for trauma-domain oral history text, given the domain mismatch limitations of the models used in this study.

Methodologically, this work demonstrates that analyzing sentiment, emotion, and topic in combination provides richer insights than analyzing these dimensions independently. Cross-modal alignment provides a richer lens to understanding how survivors tell their narratives over time. Beyond Holocaust trajectories, this methodology is applicable to other testimonies and oral histories where emotional and thematic structures are central.

6. Conclusion

This paper presents a computational analysis of Holocaust testimonies using sentiment analysis, emotion analysis, topic analysis, and cross-modal trajectory alignment. We analyze 500 testimonies from the CORHOH and segment each testimony in ten segments for trajectory analysis.

We find that Holocaust testimonies often follow a path from pre-war life to deportation, camp experiences, to post-war displacement, with an absence of emotional recovery. By using a cross-modal trajectory alignment, this work shows that emotional expression is often closely related to topics across narratives.

These findings demonstrate the potential of computational methods to create large-scale analyses of oral histories while preserving narrative integrity and emotional dimensions, and offer a transferable method for studying other trauma-centered historical corpora. This study also furthers the inclusion and use of Holocaust testimonies as language resources in NLP.

7. Limitations

Several limitations should be considered in this study. First, equal-length segmentation may not align completely with narrative boundaries. Second, this study does not manually validate the emotion or sentiment labels assigned by the automated models, and given the domain mismatch between model training data and the CORHOH corpus, systematic misclassification is possible. Third, seven emotions may not be enough to capture fine trauma-specific emotions in the context of the Holocaust testimonies. Fourth, the correlation-based cross-modal analysis is descriptive rather than causal. With only ten segments, the correlations must be interpreted as indicators of the association and effect size, not as statistically robust claims.

8. Ethics

Ethical considerations are vital in the computational analysis of Holocaust testimonies. We treat the testimonies as records of individual human experience rather than as data points. Aggregating testimonies into corpus-level trajectories risks decreasing the diversity of individual survivor experiences, so we present our findings as patterns rather than definitive feelings or memories of survivors. We acknowledge that sentiment and emotion classification tools may misclassify or misrepresent the experiences described, given the trauma-centered language of much of these testimonies. De Bruyne (2023) highlights that there is low diversity in emotion conceptualization in emotion recognition, so the seven emotions used in this study likely do not capture the fine-grained emotions in the testimonies. Mohammad (2022) cautions that transferring emotion and sentiment analysis models to unseen domains may result in poor accuracy. These limitations risk misrepresenting survivor experience and should continue to be improved upon with future work. The goal of this work is to support historical understanding and encourage computational engagement with Holocaust testimonies as language resources, not to reduce survivor narratives to emotional or thematic labels.

9. Bibliographical References

Tobias Blanke, Michael Bryant, and Mark Hedges. 2019. Understanding memories of the holocaust—a new approach to neural networks in the digital humanities. *Digital Scholarship in the Humanities*, 35(1):17–33.

Luna De Bruyne. 2023. The paradox of multilingual emotion detection. In *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pages 458–466, Toronto, Canada. Association for Computational Linguistics.

Maarten Grootendorst. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.

Jochen Hartmann. 2022. Emotion english distilroberta-base. `https://huggingface.co/j-hartmann/emotion-english-distilroberta-base/`.

C. Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1):216–225.

Maxim Ifergan, Omri Abend, Renana Keydar, and Amit Pinchevski. 2024. Identifying narrative patterns and outliers in holocaust testimonies using topic modeling. In *Proceedings of the First Workshop on Holocaust Testimonies as Language Resources (HTRes) @ LREC-COLING 2024*, pages 44–52, Torino, Italia. ELRA and ICCL.

Leland McInnes, John Healy, and Steve Astels. 2017. hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11):205.

Leland McInnes, John Healy, Nathaniel Saul, and Lukas Grossberger. 2018. Umap: Uniform manifold approximation and projection. *The Journal of Open Source Software*, 3(29):861.

Semi Min and Juyong Park. 2019. Modeling narrative structure and dynamics with networks, sentiment analysis, and topic modeling. *PLOS ONE*, 14(12):e0226025.

Saif M. Mohammad. 2022. Ethics sheet for automatic emotion recognition and sentiment analysis. *Computational Linguistics*, 48(2):239–278.

Flor Miriam Plaza-del Arco, Alba A. Cercas Curry, Amanda Cercas Curry, and Dirk Hovy. 2024. Emotion analysis in NLP: Trends, gaps and roadmap for future directions. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 5696–5710, Torino, Italia. ELRA and ICCL.

Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *ArXiv*, abs/1910.01108.

Esther Shizgal, Eitan Wagner, Renana Keydar, and Omri Abend. 2025. Computational analysis of character development in holocaust testimonies. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 22710–22734, Suzhou, China. Association for Computational Linguistics.

Ismael Villanueva-Miranda, Yang Xie, and Guanghua Xiao. 2025. Sentiment analysis in public health: a systematic review of the current state, challenges, and future directions. *Frontiers in Public Health*, 13.

Eitan Wagner, Renana Keydar, and Omri Abend. 2024a. Zero-shot trajectory mapping in holocaust testimonies. In *Proceedings of the First Workshop on Holocaust Testimonies as Language Resources (HTRes) @ LREC-COLING 2024*, pages 63–70, Torino, Italia. ELRA and ICCL.

Eitan Wagner, Renana Keydar, Amit Pinchevski, and Omri Abend. 2024b. Automatic topic-guided segmentation of holocaust survivor testimonies. *Journal of Computational Literary Studies Volume 2 Issue 1 2023*.

Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. 2020. Minilm: Deep self-attention distillation for task-agnostic compression of pre-trained transformers.

Samira Zad and Mark Finlayson. 2020. Systematic evaluation of a framework for unsupervised emotion recognition for narrative text. In *Proceedings of the First Joint Workshop on Narrative Understanding, Storylines, and Events*, pages 26–37, Online. Association for Computational Linguistics.

10. Language Resource References

Jaff, Daban Q. 2025. *CORHOH: Text corpus of holocaust oral histories*. Elsevier BV.