The Geography of ‘Fear’, ‘Sadness’, ‘Anger’ and ‘Joy’: Exploring the Emotional Landscapes in the Holocaust Survivors’ Testimonies

Ignatius Ezeani[1,*], Paul Rayson[1], Ian Gregory[2], Tim Cole[3], Erik Steiner[4] and Zephyr Frank[4]

[1]*UCREL, School of Computing and Communications, InfoLab21, Lancaster University, Lancaster, LA1 4WA, UK*
[2]*Department of History, Lancaster University, Lancaster, LA1 4YT, UK*
[3]*Department of History, University of Bristol, BS8 1TB, UK*
[4]*Department of History, School of Humanities and Sciences, Stanford University, USA*

Abstract

Holocaust survivors’ testimonies provide a rich source of evidence about the personal experiences of survivors who witnessed and endured horrors during the Nazi genocide of Jews and other persecuted groups. The narratives contain references to the emotions experienced when describing memories of people, places, and events. Analysing the spatiality of these human emotions enables us to understand how they are connected to the places around them. We focus on *fear*, *sadness*, *anger*, as well as *joy* to examine the interplay of these emotional experiences by multiple individuals at different places and times and in different circumstances. Understanding these complex emotional landscapes, especially from very large collections of textual data requires a carefully designed technique that can effectively and efficiently apply existing and new technologies. In this work, therefore, we explore the possibility of extracting and analysing these emotions as well as their related geographies by applying a combination of natural language processing methods including large language models.

Keywords

spatial narratives, holocaust testimonies, large language models, emotion geography, spatial emotion classification

In: R. Campos, A. Jorge, A. Jatowt, S. Bhatia, M. Litvak (eds.): Proceedings of the Text2Story’24 Workshop, Glasgow, Scotland, 24-March-2024

*Corresponding author.
i.ezeani@lancaster.ac.uk (I. Ezeani); p.rayson@lancaster.ac.uk (P. Rayson); i.gregory@lancaster.ac.uk (I. Gregory); tim.cole@bristol.ac.uk (T. Cole); ebs110@stanford.edu (E. Steiner); ebs110@stanford.edu (Z. Frank)
https://www.lancaster.ac.uk/scc/about-us/people/ignatius-ezeani (I. Ezeani);
https://www.lancaster.ac.uk/scc/about-us/people/paul-rayson (P. Rayson);
https://www.lancaster.ac.uk/staff/gregoryi/ (I. Gregory);
https://www.bristol.ac.uk/people/person/Tim-Cole-7e176567-189f-4f5e-b62e-c6a3e7d5e84a/ (T. Cole);
https://web.stanford.edu/group/spatialhistory/static/ (E. Steiner); https://history.stanford.edu/people/zephyr-frank (Z. Frank)
0000-0001-8286-9997 (I. Ezeani); 0000-0002-1257-2191 (P. Rayson); 0000-0001-8745-2242 (I. Gregory); 0000-0001-5711-8200 (T. Cole); 0009-0001-4407-8956 (E. Steiner); 0000-0003-0691-9299 (Z. Frank)

1. Introduction

During what became known as the Holocaust, the Nazi regime systematically persecuted and murdered millions of Jews and other targeted groups during World War II. The survivors of these events have shared their personal experiences and memories through various forms of testimony, such as interviews, memoirs, and artworks. These provide valuable sources of historical and cultural knowledge, as well as emotional and psychological insight into the human condition under extreme circumstances. One aspect that can be explored in Holocaust survivors' testimonies is the spatial and temporal dimensions of the emotions expressed about people, places, and events, otherwise known as *emotional geography*. As Guy Miron signals in the case of German Jews, individuals experienced Nazi spatial control "both as a feeling and as a physical reality" [1]. Just as spatial experiences had an emotional dimension, so too did emotions have a spatiality or geography. Emotional geography is a concept that helps us understand how people feel about and react to, their environment, and how their environment influences their identity and memory. It also allows us to examine the interplay of different emotional experiences, such as fear, anger, surprise, sadness, disgust, and even joy, by multiple individuals at different places and times and in different situations.

However, studying the emotional geography of Holocaust survivors is not an easy task, especially when dealing with very large collections of video, audio, and textual data (the collection we work with includes over 55,000 recorded interviews). Therefore, there is a need for effective methods and tools to extract and analyse the emotions and geographies from the Holocaust survivors' testimonies and to visualise and present the results in a meaningful and accessible way. This work aims to address this challenge by applying a combination of existing natural language processing (NLP) techniques, such as sentiment analysis, emotion detection, named entity recognition, geocoding, and geovisualisation, to a corpus of video-recorded testimonies from the USC Shoah Foundation. The key research questions that drive this work include:

- How can we use NLP techniques - possibly leveraging large language models - to extract and analyse expressions of emotions in Holocaust testimonies?
- How much 'fear' compared to other emotions - 'sadness', 'anger', and 'joy' - is contained in each testimony narrative?
- How does the expression of 'fear' change across the narrative sequence of each testimony?

This paper presents the outcome of exploratory work which contributes::

- An application of NLP methods to the study of the emotional geography of Holocaust survivors, and a demonstration of the potential and challenges of using these techniques for this purpose.
- An analysis of the emotions and geographies expressed in the testimonies, and a discovery of new patterns and insights that can enhance our understanding of the Holocaust and its survivors.

2. Related Work

This work draws on and contributes to two main fields of research: emotional geography and natural language processing. Emotional geography is an interdisciplinary field that studies the relationship between human emotions and space, place, and environment [2]. It covers a wide range of topics and perspectives, such as the emotional attachment to place, the emotional impact of displacement and migration, the emotional dimensions of power and resistance, the emotional aspects of memory and identity, and the emotional expressions of culture and society [3].

Researchers have long advocated exploring the physicality of the event. Beorn *et al.* signalled that the Holocaust was 'rooted in specific physical spaces, times, and landscape' and also 'characterised by a spatiality of the process - concentration, deportation, dispersal, dislocation' [4]. Some of their key questions were: *How did one (or why would one) "map" testimony?*, *How would a typological approach to the Holocaust differ from accounts of individual experience?*, *How can the "cognitive mapping" so present in survivor and postwar SS tribunal testimony be reconciled with the physical environment of landscapes and buildings?* However in the 2000s, when these ideas were being proposed, the required technology was either not available or was not fully developed. However, multidisciplinary collaboration involving historians, historical geographers, GIScientists, and cartographers were established [5]. Advances in AI and natural language processing (NLP) enabled researchers to have the platform and the technology for a deeper investigation of the transcripts of testimonies. Recent approaches apply a variety of NLP techniques [6, 7, 8, 9] that leverage large language models (e.g GPTs) and other transformer-based models. For example, Woods *et al.* [8] investigated how sentiments associated with places vary over the narrative sequence by comparing the performance of different machine learning algorithms. Beyond geocoding, some studies [10] applied bespoke named entity recognisers to extract and analyse named places and other spatial elements - geographical features ('river', 'hill', 'road'), imprecise description of landscapes ('the majestic mountains', 'the camp') and feature relative terms ("a quick detour along the lake", "turn left after the inn") in text [11]. However, previous work has not used automated methods on a large-scale to annotate emotions in testimonies.

3. Dataset and Methods

For this work, we have focused on a small portion of the Holocaust Survivors' Testimonies (HST) which comprises a random selection of transcripts of one thousand oral history interviews (about 21 million words) undertaken by the USC Shoah Foundation Visual History Archive[1] in the 1990s.

3.1. Dataset: Holocaust Survivors' Testimonies

The transcripts follow a similar format that includes a series of questions posed by the interviewer and the corresponding answers from the interviewee who is a survivor. According to the

[1]Information about the USC Shoah Foundation Visual History Archive can be found https://sfi.usc.edu/what-we-do/collections

Table 1
Analysis of the Holocaust Survivor's Testimonies files and contents

	Holocaust Survivor's Testimonies
File count	1000
Sentence count	816,800
Words (tokens) count	21,516,122
File size range (words)	4852 – 84,051
Averages file size (words)	21,516

Table 2
Statistics of a sample of 10 Testimonies used in this work. The token (`Tokens`) and sentence (`Sents`) counts focus only on pairs consisting of the responses from the survivors and the questions from the interviewer. The IDs correspond respectively to the testimony file names

ID	QA-Pairs	Tokens	Sents	ID	QA-Pairs	Tokens	Sents
268	98	43192	4084	37556	233	57862	4450
36999	175	12965	1250	37567	253	24182	2563
37210	254	39245	3655	37585	273	12367	1354
37250	186	41199	4223	37605	273	12555	1435
37409	132	18151	1759	37648	181	19731	1850

interview guideline [12], each interview focuses on the individual's experiences during the Holocaust which are explored in a broadly chronological order. Each interview – generally of around two hours duration – devotes approximately 20 percent of the time to pre-war life, 60 percent to wartime experiences focused on the events of the Holocaust, and 20 percent to post-war life [13]. In short, these are not full life histories, but more focused interviews asking about wartime experiences across a series of sites of incarceration or hiding. These sites serve as anchors in the narratives that describe survivors' wartime trajectories.

In this exploratory work, we used only a small sample of 10 out of the 1000 testimonies described above as a preliminary test to assess the salience of the emotion of fear, its spatial distribution, and its relation to other important emotional states. For this experiment, we purposely restrict our sample size and the number of emotions examined to allow for the application of domain expert knowledge to check the consistency and accuracy of our methods.

3.2. Spatial Entity Extraction

This extraction pipeline used in this work was a version of the framework, illustrated in Figure 1, defined by [11] for extracting place names and geographical feature nouns from text through named entity recognition for the Lake District corpora, and a more generalised version of the framework described in another work [10]. The key elements of the framework include processes for enhancing an off-the-shelf named-entity recogniser, Spacy [14], with lists of items (e.g. place names, geographical feature nouns, dates and time, sentiments or emotions, locative adverbs, and spatial prepositions). This was achieved by adding the Spacy's rule-based module annotation module 'EntityRuler' to the models pipeline thereby improving the ability of its

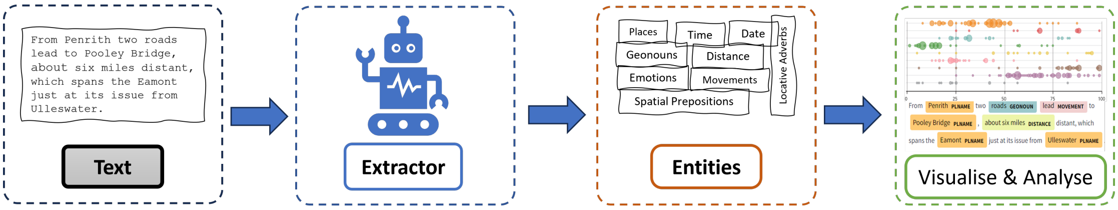

Figure 1: An overview of the spatial entity extraction framework used in this work, originally presented by Ezeani *et al.* [11, 10].

named entity recogniser to identify and extract our pre-defined spatial entities and concepts. The enhanced model is subsequently applied to similar texts to perform surface-level extraction of spatial elements and even sentiment-bearing words for basic sentiment analysis.

Using this technique, we could identify some key geospatial elements like the countries, cities, concentration camps, and even geonouns mentioned in the testimonies. Table 3 shows some of the top elements identified in the answers provided by the survivors in the 2 randomly selected testimonies - id='01' and id='02'. For this work however, we explore emotion beyond the binary classifications of positive and negative sentiments to capture more fine-grained affect analysis but restricted to mainly 'fear' but 'sadness', 'anger', and 'joy' which were also included for richer comparative analysis. These classifications are a selection from the 6 basic classes of emotion ('fear', 'anger', 'disgust', 'joy', 'sadness', 'surprise') identified by Ekman and Friesen [15] as well as the 'neutral' class, which are popularly used in NLP for studies in affect classification and analysis.

3.3. Emotion Classification

The transcribed testimonies are available in plain text formats generally structured for conversational turn-taking mainly between the interviewer and the survivor but often with interjection from crew members or family members of the survivors. We first pre-processed the files to separate the questions by the interviewer and the responses by the survivor. We then focused only on these responses (see Table 2 for details) given by the survivor for the emotional classification. Since some of these responses were quite long and span many paragraphs, we split them into sentences to keep the context to a manageable scope.

3.3.1. Classification Models

There was no available bespoke emotion model trained on the Holocaust corpora or a dataset for training one that we were aware of during this work. Besides, this was meant to be an exploratory work with no intention of building new models at the first stage. Hence, we proceeded with existing off-the-shelf emotion classification models as well as generic large language models. We therefore applied the following three models to classify the sentence contexts extracted from survivors' testimonies and select the most 'voted' class out of the three.

The first model we applied was Hartmann's transformer model [16]. The model which is a version of the DistilRoBERTa-base fine-tuned on about 20k observations extracted from six diverse emotion datasets. The selected observations are fairly distributed across the emotion

```
SYSTEM="""You're an emotional analyst that identifies emotions from text.
         Identify the emotion in this as 'fear', 'sadness', 'anger', or 'joy'
         and return ONLY one of the labels, otherwise, return 'none'"""
text='And I used to be very afraid because I was the only Jew on the street.'
prompt = f"{SYSTEM} {text}"
```

Figure 2: An example of the large language model prompt used for the classification of emotions from the Holocaust texts

Questions	Answers
INT: Could you spell that, please?	HR: Yes. R-O-S-M-A-R-I-N.
INT: And Henry, what was your name at birth?	HR: At birth it was Henryk, H-E-N-R-Y-K Rozmaryn, R-O-S-M-A-R-Y-- I'm sorry, English spelling. Polish spelling, R-O-Z-M-A-R-Y-N, Rozmaryn.
INT: When were you born, Henry?	HR: October 7th, 1925.
INT: And your present age is--	HR: My present age is 73.
INT: Where were you born?	HR: I was born in a little town called Czeladz, in Poland. It's southwestern corner of Poland. I spell it for you, C-- capital C-Z-E-L-A-D-Z.
INT: What was the nearest town or city?	HR: The nearest town was Bedzin, Sosnowiec. And towards the German border was Katowice and Siemianowice. Incidentally, it was in Siemianowice that I lived.

Figure 3: An example of the format for the pre-processed testimonies

classes. It is a lightweight transformer model that is freely available and easy to use, hence the choice. Another key contribution of this model is the provision of the emotion scores as a probability distribution across the labels.

From the output of the transformer model, we retained only the contexts that were labelled with the four classes we are studying in this paper - 'fear', 'sadness', 'anger', and 'joy'. Given the need for some form of evaluation, these classified sentences were then passed individually to a variant of the `GPT-3.5 Turbo` models [17] `gpt-3.5-turbo-instruct` as well as the `GPT-4` model [18] with the same prompt basically to get another opinion. We tried several prompts to see which works best but Figure 2 shows an example of the prompt we settled for the LLM experiments.

Finally, the accepted class for each sentence is determined by selecting the most '`voted`' class. This requires that at least two of the models will predict a particular class for it to be selected, otherwise '`none`'. While we do not expect the approach to guarantee high-quality annotations we consider it a stop-gap method that is good enough to enable us to gain some interesting insights from the Holocaust testimonies.

4. Results and Discussion

We performed the classification with the models and 'voting' process, and confusion matrices can be seen in Figure 6. Our final output is a set of 5,461 sentences with emotion scores and labels for *fear*, *sadness*, *anger*, and *joy*. Table 3 shows examples of some of the sentences and

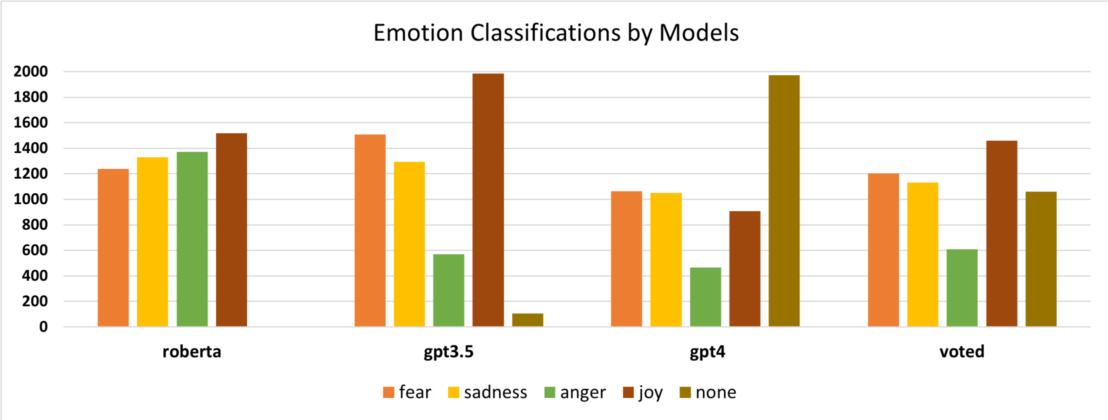

Figure 4: Graph showing the counts of sentences classified as one of the four key emotions - *sadness*, *anger*, *fear*, *joy* - by different models

the scores assigned by different models while Figure 4 is a plot of the distributions of the four emotions or 'none' as predicted by the models. The label 'none' is not present in the predictions from `roberta` because we only extracted instances of the four labels from its output. As shown in Figure 4, the emotions of 'fear' and 'sadness' tend to trend in the same pattern consistently across the model outputs.

When comparing the variations of the negative emotions *fear*, *sadness*, *anger* and positive emotion *joy* across the narrative sequence as shown in Figure 5, some interesting patterns emerge. Survivors tend to express positive emotions at the beginning and end of the interviews while negative emotions tend to follow a bimodal, increasing pattern before falling sharply towards the end of the interviews. This follows from the structure of interviews, where survivors begin by discussing pre-war childhood memories and end with reflections on more recent life events such as their own children. Across all interviews, positive and negative emotions are inversely correlated, with negative emotions of fear and sadness appearing together most often.

Finally, Table 4 shows the connections we observe in these 10 texts between named places and the emotions observed. We split the results into three groups: 1) places, 2) geographic feature nouns, and 3) camps, and for each group, we list the overall top five in each group, plus the lists associated most with each of the four emotions. This serves to illustrate the potential of the technique for linking emotions to places of the different types under analysis.

5. Conclusion

In this work, we explored the idea of developing a computational framework for analysing the emotional landscapes of a textual narrative in a more nuanced form beyond the classification into positive and negative sentiments which has been undertaken previously. We applied our method to a sample of 10 Holocaust survivors' testimonies specifically focusing on only four out of Ekmans 6 emotion classes [15] - 'fear', 'sadness', 'anger', and 'joy'. These were chosen for no particular reason other than that they seem like sensible themes for Holocaust-related study. The interaction of these emotions with the geography of the space they were connected

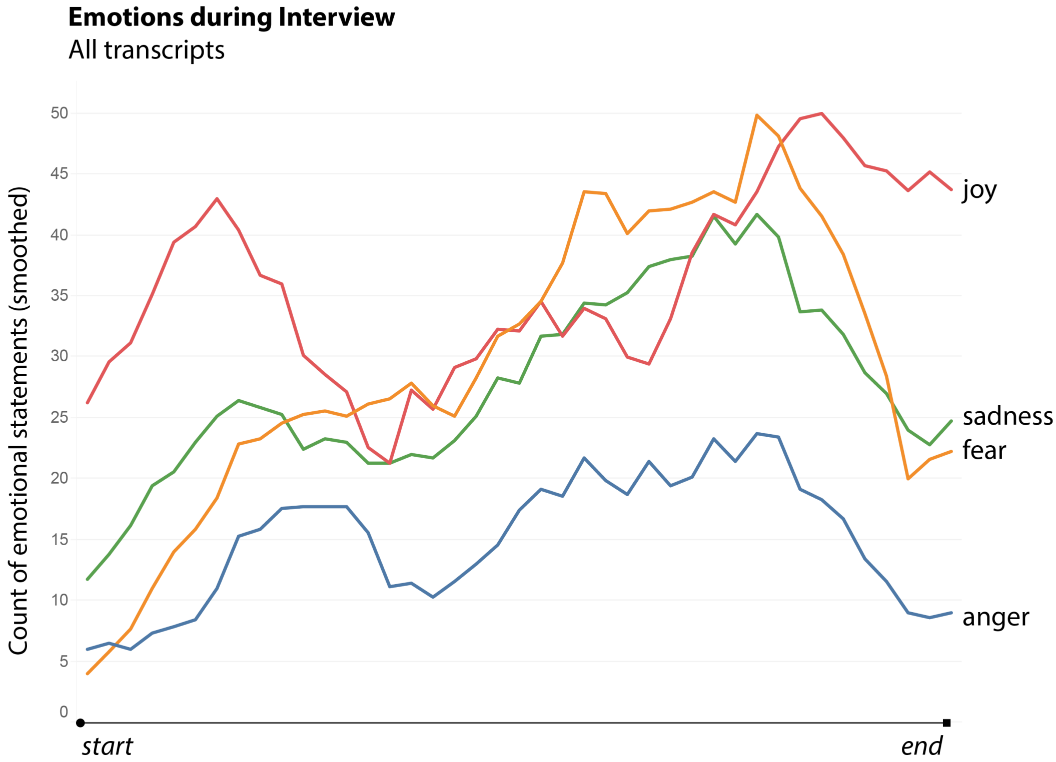

Figure 5: Graph showing smoothed emotion counts across the combined testimonies

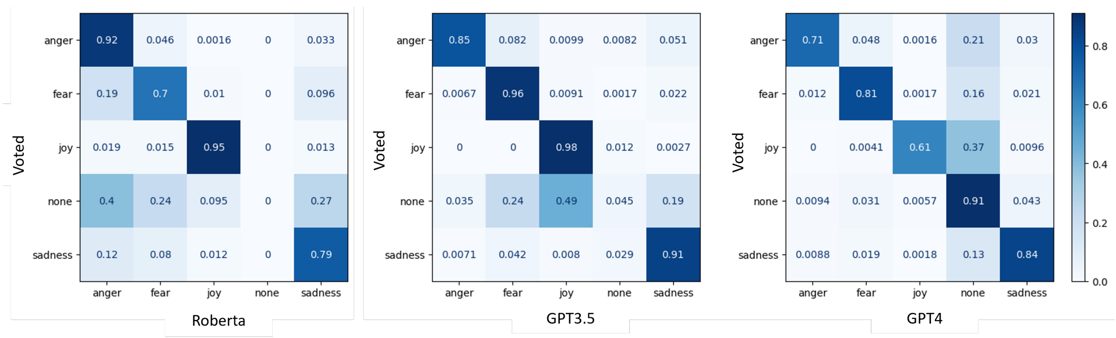

Figure 6: The confusion matrix comparing the prediction agreements between the 'voted' labels and those from the models.

to - named places, camps, and geographical features were also extracted and analysed.

Given that we had no available model or previously existing annotated dataset for our task, we applied an efficient combination of models consisting of a lightweight transformer-based model and two large language models to support our annotation process. We are aware of the limitations of this method and will focus on building a more robust, bespoke, and in-domain

Table 3
Example: Sentence-level emotion scores and model labels. 'Mod1', 'Mod2', and 'Mod3' refer to `roberta`, `gpt3.5` and *gpt4* respectively.

text	fear	sad	anger	joy	roberta	gpt3.5	gpt4	voted
1. *And I used to be very afraid because I was the only Jew on the– on the street.*	**0.9921**	0.0017	0.0012	0.0011	fear	fear	fear	fear
2. *And at that time, I lost my mother and my sister and my little nephew.*	0.0030	**0.9811**	0.0023	0.0010	sad	sad	sad	sad
3. *But still, they were incensed that this kind of jazz would be played there.*	0.0021	0.0030	**0.9709**	0.0007	anger	anger	anger	anger
4. *So they were relieved to see us in the morning, coming back home.*	0.0004	0.0037	0.0018	**0.9695**	joy	joy	joy	joy
5. *They drove him out, out of the camp, and they said, go back.*	0.0668	0.0226	*0.3528*	0.0072	anger	fear	sad	none

model for more accurate labelling and analysis. However, it produced sufficiently good results for initial exploratory research. We will fine-tune our processes and scale up our study to analyse a larger collection of Holocaust testimonies to gain more insight into the complex interplay of emotions in a spatio-temporal context, potentially fine-tune a model on Holocaust data, to link emotions to those who experienced them, and to explore zero-shot learning methodologies in the future.

Acknowledgments

We thank the anonymous reviewers for their comments on our paper submission. The project is funded in the UK from 2022 to 2025 by ESRC, project reference: ES/W003473/1. We also acknowledge the input and advice from the other members of the project team in generating requirements for our research presented here and the UCREL Hex team for providing the compute needs for this project[19]. More details of the project can be found on the website: https://spacetimenarratives.github.io/

References

[1] G. Miron, "Lately, Almost Constantly, Everything Seems Small to Me": The Lived Space of German Jews under the Nazi Regime, Jewish Social Studies: History, Culture, Society 20 (2013) 121–149.
[2] K. Anderson, S. J. Smith, Emotional geographies, Transactions of the Institute of British geographers 26 (2001) 7–10.
[3] J. Davidson, C. Milligan, Embodying Emotion Sensing Space: Introducing Emo-

Table 4
Top 5 named places (countries or cities), geographical features, and camps in the 10 testimonies compared with places associated with each emotion. Each row is the cumulative count of the entity in the collection followed by the percentage occurrence of the top ones.

Places	
Overall (443)	(Russia,11%); (Israel,7%); (Germany,6%); (York,5%); (Poland,5%)
fear (83)	(Russia,23%); (Germany,9%); (Braunschweig,5%); (Palestine,5%); (Buchenwald,3%)
sadness (68)	(Russia,9%); (Germany,9%); (Israel,9%); (York,7%); (Hungary,5%)
anger (75)	(Russia,32%); (Germany,12%); (Palestine,12%); (Poland,6%); (Boston,6%)
joy (66)	(Poland,7%); (Vienna,7%); (Israel,7%); (York,6%); (Oswego,6%)
Geographic feature nouns	
Overall (2950)	(end,6%), (school,4%); (camp,3%); (train,2%); (house,2%)
fear (118)	(train,4%); (camp,3%); (rough,3%); (side,3%); (end,3%)
sadness (119)	(camp,5%); (end,5%); (school,4%); (shed,3%); (place,3%)
anger (117)	(top,4%); (shed,4%); (head,3%); (stop,3%); (door,3%);
joy (121)	(ice,11%); (school,8%); (well,3%); (port,3%); (house,3%);
Camps	
Overall (2950)	(Russia,21%); (Auschwitz,16%); (Germany,12%); (Poland,9%); (Czechoslovakia,6%)
fear (89)	(Auschwitz,27%); (Russia,26%); (Germany,10%); (Czechoslovakia,5%); (Braunschweig,5%)
sadness (91)	(Auschwitz,24%); (Russia,14%); (Germany,14%); (Czechoslovakia,8%); (Theresienstadt,8%)
anger (82)	(Russia,46%); (Germany,17%); (Poland,8%); (Czechoslovakia,4%); (Auschwitz,4%)
joy (73)	(Poland,22%); (Russia,13%); (Bar,11%); (Germany,11%); (Auschwitz,4%)

tional Geographies, Social & Cultural Geography 5 (2004) 523–532. doi:10.1080/14649360420000317677.

[4] W. Beorn, T. Cole, S. Giggliotti, A. Giordano, A. Holian, P. B. Jaskot, A. K. Knowles, M. Masurovsky, E. B. Steiner, Geographies of the holocaust, Geographical Review 99 (2009) 563–574.

[5] A. K. Knowles, T. Cole, A. Giordano, Geographies of the Holocaust, Indiana University Press, 2014.

[6] O. Suissa, A. Elmalech, M. Zhitomirsky-Geffet, Text analysis using deep neural networks in digital humanities and information science, Journal of the Association for Information Science and Technology 73 (2022) 268–287.

[7] T. Blanke, M. Bryant, M. Hedges, Understanding Memories of the holocaust—A New Approach to Neural Networks in the Digital Humanities, Digital Scholarship in the Humanities 35 (2020) 17–33.

[8] L. Woods, L. Mardona, 2023. URL: https://www.automatedfutures.net/geographies-of-affect-place-and-sentiment-in-holocaust-testimonies/.

[9] A. K. Knowles, P. B. Jaskot, T. Cole, A. Giordano, Mind the gap: Reading across the holocaust testimonial archive, in: The Holocaust in the 21st Century: Relevance and Challenges in the Digital Age: Lessons and Legacies Volume XIV, Northwestern University Press, 2021, pp. 216–241.

[10] I. Ezeani, P. Rayson, I. Gregory, E. Haris, A. Cohn, J. Stell, T. Cole, J. Taylor, D. Bodenhamer, N. Devadasan, et al., Towards an extensible framework for understanding spatial narratives, in: Proceedings of the 7th ACM SIGSPATIAL International Workshop on Geospatial Humanities, 2023, pp. 1–10.
[11] I. Ezeani, P. Rayson, I. N. Gregory, Extracting imprecise geographical and temporal references from journey narratives., in: Text2Story@ ECIR, 2023, pp. 113–118.
[12] J. Ringelheim, A. Donahue, E. Hedlund, A. Rubin, Oral history interview guidelines, 2007. URL: https://www.ushmm.org/m/pdfs/20121003-oral-history-interview-guide.pdf, accessed: 04-01-2024.
[13] N. Shenker, Reframing Holocaust Testimony, Indiana University Press, Indiana University, 2015.
[14] M. Honnibal, I. Montani, S. Van Landeghem, A. Boyd, spaCy: Industrial-strength Natural Language Processing in Python (2020). doi:10.5281/zenodo.1212303.
[15] P. Ekman, W. V. Friesen, Constants across cultures in the face and emotion., Journal of personality and social psychology 17 (1971) 124.
[16] J. Hartmann, Emotion English DistilRoBERTa-base, https://huggingface.co/j-hartmann/emotion-english-distilroberta-base/, 2022.
[17] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language models are few-shot learners, 2020. arXiv:2005.14165.
[18] OpenAI, et al., Gpt-4 technical report, 2023. arXiv:2303.08774.
[19] J. Vidler, P. Rayson, UCREL - Hex; A Shared, Hybrid Multiprocessor System, https://github.com/UCREL/hex, Accessed: 2024.