Economic and Social Research Council

The National Science Foundation

The Spatial Narratives Project

Accessing and Interpreting Space and Time in Large Corpora

THE GEOGRAPHY OF FEAR, SADNESS, ANGER & JOY

Exploring the Emotional Landscapes in the Holocaust Survivors' Testimonies

7th International Workshop on the Narrative Extraction from Text (Text2Story 2024)

Co-located with the 46th European Conference on Information Retrieval

March 24th 2024

Ignatius Ezeani, Paul Rayson, Ian Gregory, Tim Cole, Erik Steiner and Zephyr Frank

Lancaster University

UNIVERSITY OF LEEDS

University of BRISTOL

MANCHESTER 1824 The University of Manchester

Stanford University

IUPUI

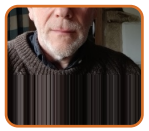
Project Team

Ian Gregory (PI)
Lancaster University, UK

Paul Rayson
Lancaster University, UK

Ignatius Ezeani
Lancaster University, UK

Tony Cohn
Leeds University, UK



John Stell
Leeds University, UK

Tim Cole
University of Bristol, UK

Erik Steiner
Stanford University, UK

Jackie Olson
Stanford University, US

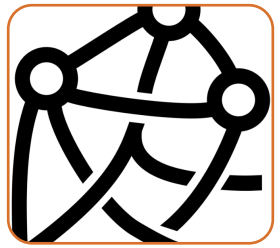
David Bodenhamer
Indiana University, US

Neil Devadasan
Indiana University, UK

Zephyr Frank
Stanford University, US

Erum Haris
Leeds University, UK

Joanna Taylor
Manchester University, UK



https://spacetimenarratives.github.io/

The Project

- Approaches to 'textual geographies' focus on geoparsing - identifying place names and allocating coordinates to them.
- Limited to named places with coordinates; little to no concept of time.
- Omits vague, imprecise, and ambiguous, e.g. "*the camp*", or "*the hills in the distance*", and relative locations such as "*near to*", "*on the left*", "north of" or "*a few hours' journey from*".
- Leaves out events, experiences, emotions, and memory accompanying the geography.
- In this project, we explore ways to identify, extract, visualise and analyse qualitative and quantitative reference to place and time.

Corpora: CLDW and HST

Corpus of Lake District Writing

- Comprises 80 texts and around 1.5 million
- Texts created between 1622 – 1900
- Containing travel literature, fiction, histories, letters, and diaries.
- Includes works by well-known Lake Poets
  - e.g. William Wordsworth and Samuel Taylor Coleridge.
- Also accounts by prominent writers such as Daniel Defoe and Celia Fiennes

Holocaust Survivor's Testimonies

- ≈1k transcripts of survivor interviews, tokens
- Each comprises a series of interviewer's questions and the survivor's answers
- Chronicles the survivor's Holocaust experience
- Duration ≈ 2hrs: 20% pre-war life; 60% wartime experience; 20% post-war
- Highlights sites of incarceration or hiding serving as anchors that describe survivors' wartime trajectories.

Data sources:

- CLDW: Lake District Corpus Repository (https://github.com/UCREL/LakeDistrictCorpus)
- HST: United Holocaust Memorial Museum (https://www.ushmm.org/)

Project Objective

Given a transcript, we want to extract and possibly build a narrative around

- Locations:
  - Names of places – cities, countries, camps – mentioned in the testimony. May be converted to coordinates for mapping e.g. **Poland**, **Utrecht**, **Auschwitz-Birkenau**
- Locale:
  - events that occur at a place will contribute to this
  - various forms of date, time, and other temporal references
  - as well as spatial and geographical features, ‘camp’, ‘ghetto’, ‘gas chamber’, ‘school’, `train` etc. often vague or ambiguous and cannot be directly mapped in Euclidean space.
- Sense of place:
  - subjective sentiments and emotional expressions by the survivor about a place

Project Objective

Given a transcript, we want to extract and possibly build a narrative around

- Locations:
  - **Names of places** – cities, countries, camps – mentioned in the testimony. May be converted to coordinates for mapping e.g. Poland, Utrecht, Auschwitz-Birkenau
- Locale:
  - **events** that occur at a place will contribute to this
  - various forms of **date**, **time**, and other **temporal references**
  - as well as **spatial and geographical features**, ‘camp’, ‘ghetto’, ‘gas chamber’, ‘school’, `train` etc. often vague or ambiguous and cannot be directly mapped in Euclidean space.
- Sense of place:
  - subjective **sentiments** and **emotional expressions** by the survivor about a place

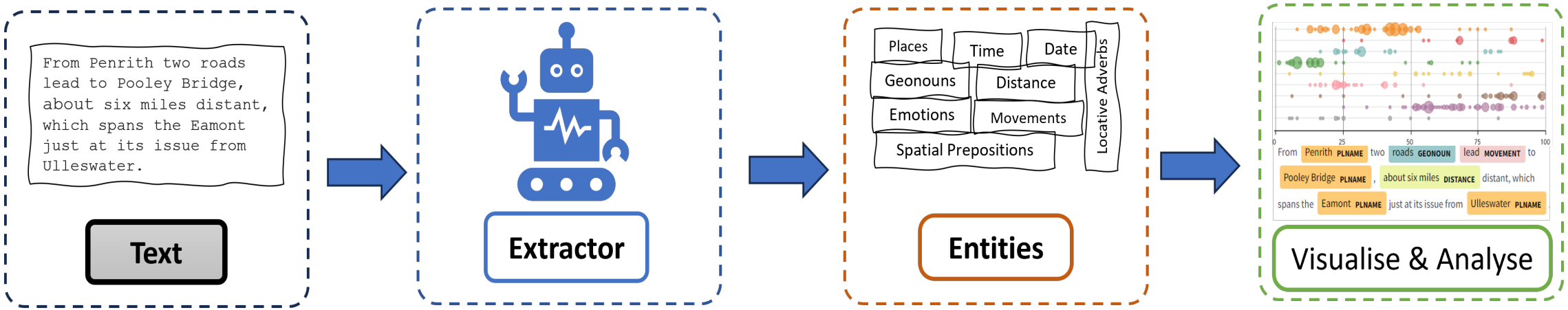
Key Tasks

Corpus annotation	Annotate the corpus to reflect the occurrence of spatial elements of interest: *toponyms, geo nouns, emotions, events, time*
Spatial entity extraction	Develop NLP techniques to appropriately identify, extract, and categorize these elements.
Spatial visualization	Add a user-friendly visualization interface to the extraction tool to look at the special entities
Analysis	Perform analysis on the extracted entities to build a narrative

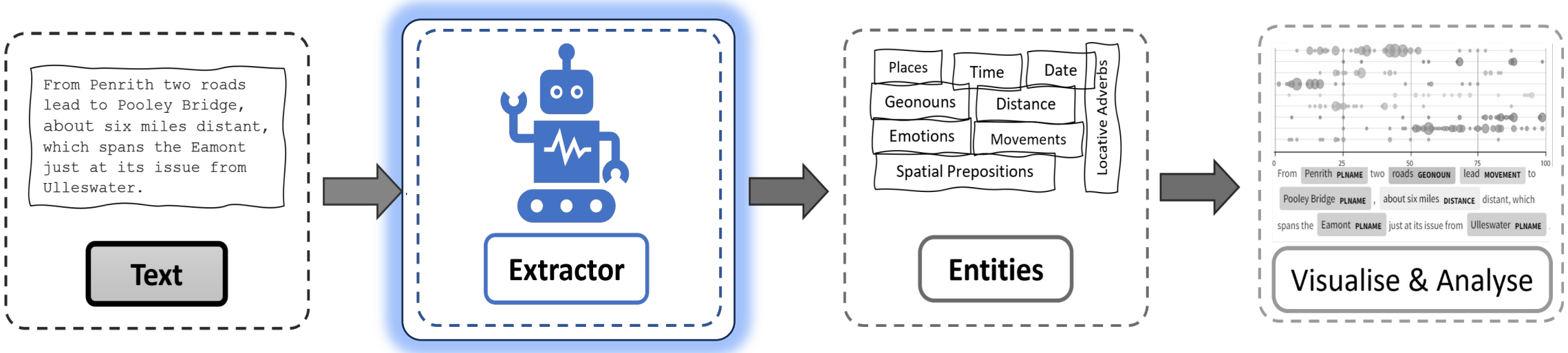
Labelling Spatio-Temporal Entities

- Names of places:
  - e.g. *Poland*:**COUNTRY**; *Utrecht*:**CITY**, *Auschwitz-Birkenau*:**CAMP**, *Europe*:**CONTINENT**
- Spatial or geographical features:
  - **GEONOUN** – concentration camp, ghetto, factory, coal mine.
- Temporal references:
  - *October 7th, 1925*: **DATE**; *the summertime*: **DATE**
  - *about 6 o'clock in the morning*-**TIME**; *less than half an hour*-**TIME**
- Events:
  - **EVENT** – *Hanukkah, World War II, Holocaust, Passover*
- Sentiments:
  - Uses a sentiment lexicon e.g. *delightful*:**+EMOTION**; *pleasant*:**+EMOTION**; *horrible*:-**EMOTION**

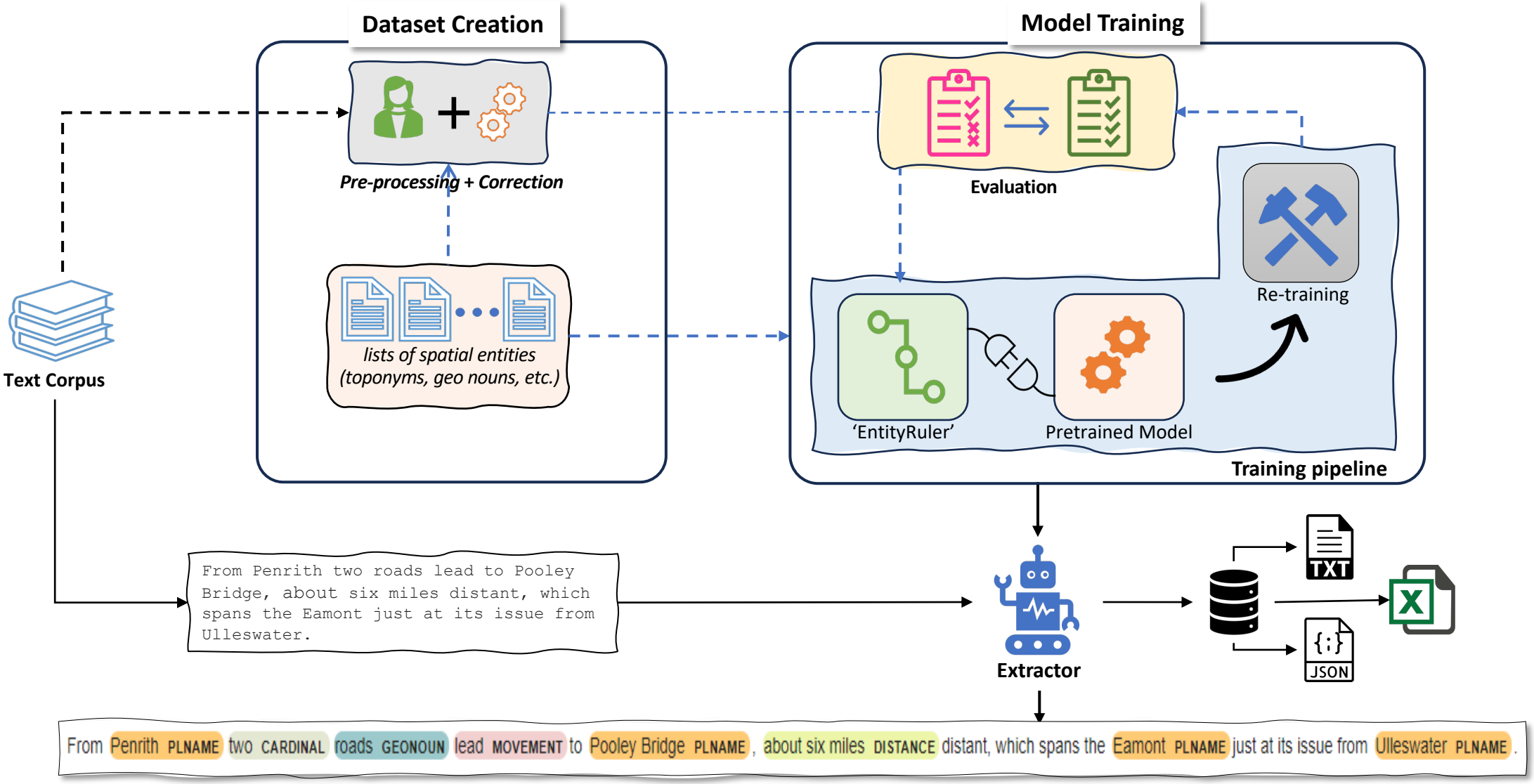
Annotation Pipeline



Annotation Pipeline



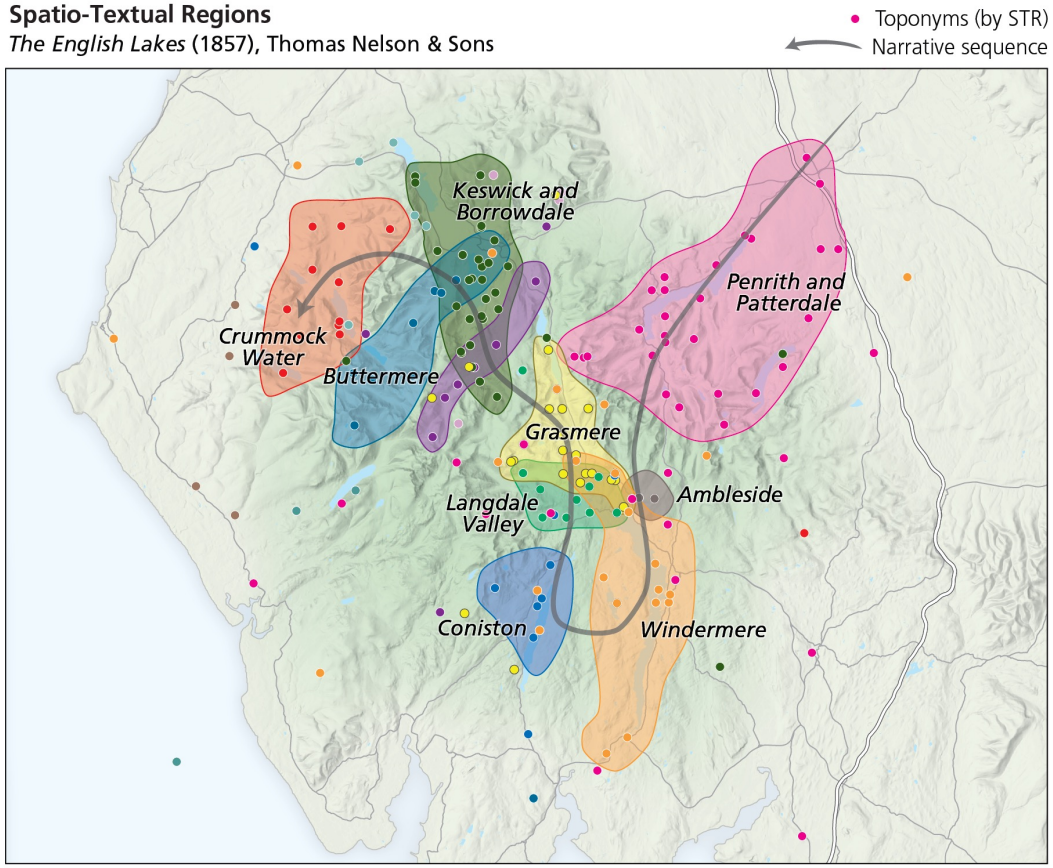
Annotation Model



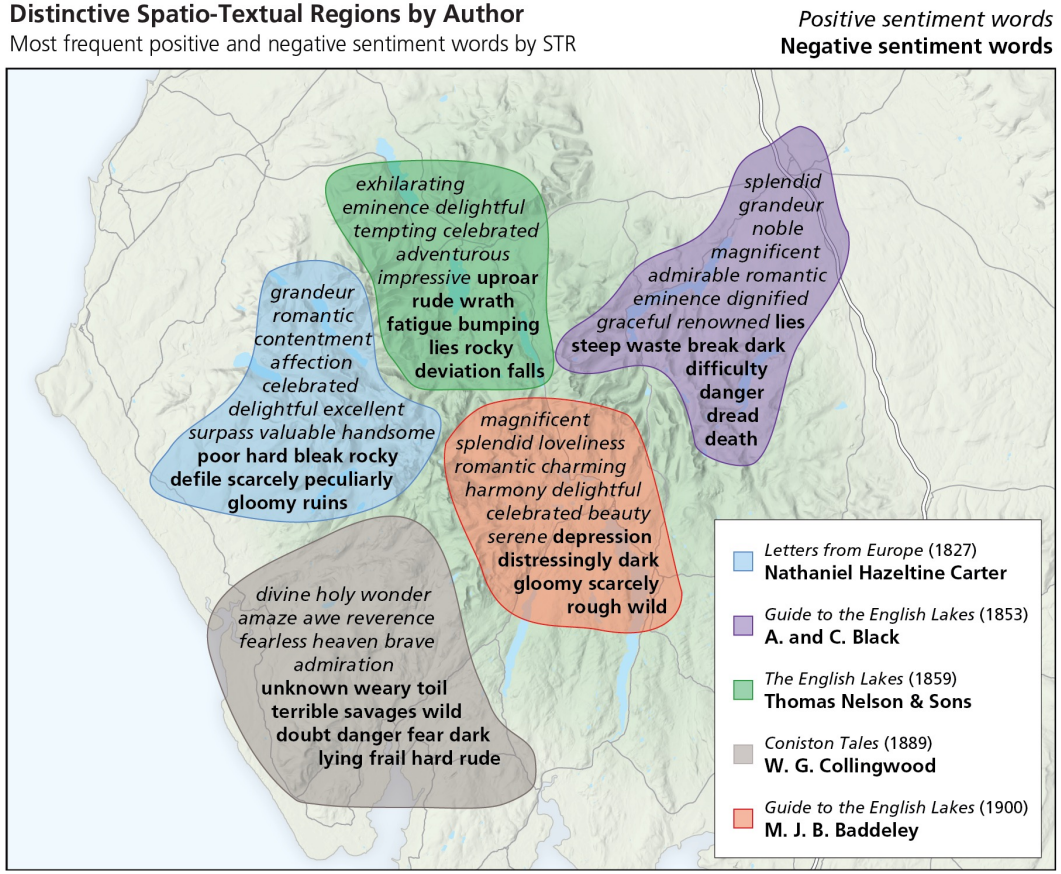
The Spatial Narratives Project
Accessing and Interpreting Space and Time in Large Corpora

CLDW: Narrative sequence and sentiments

Spatio-Textual Regions
The English Lakes (1857), Thomas Nelson & Sons

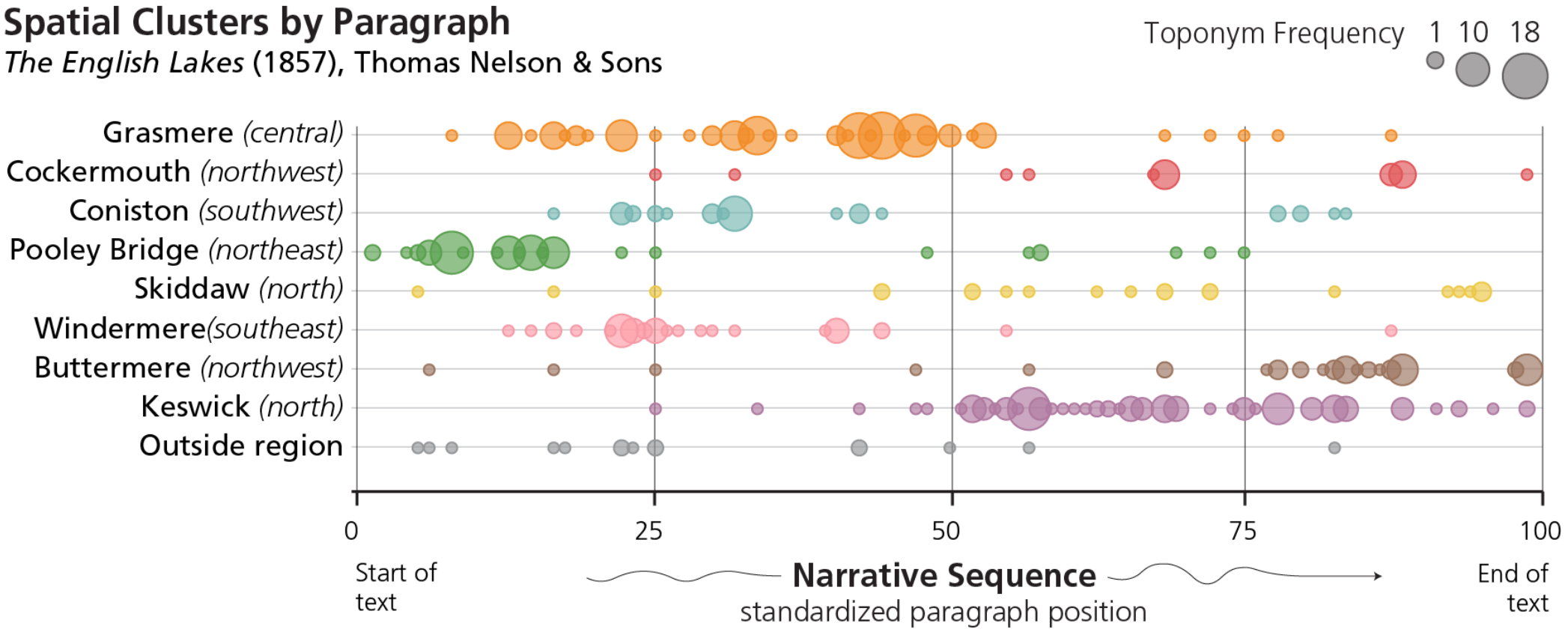


Distinctive Spatio-Textual Regions by Author
Most frequent positive and negative sentiment words by STR



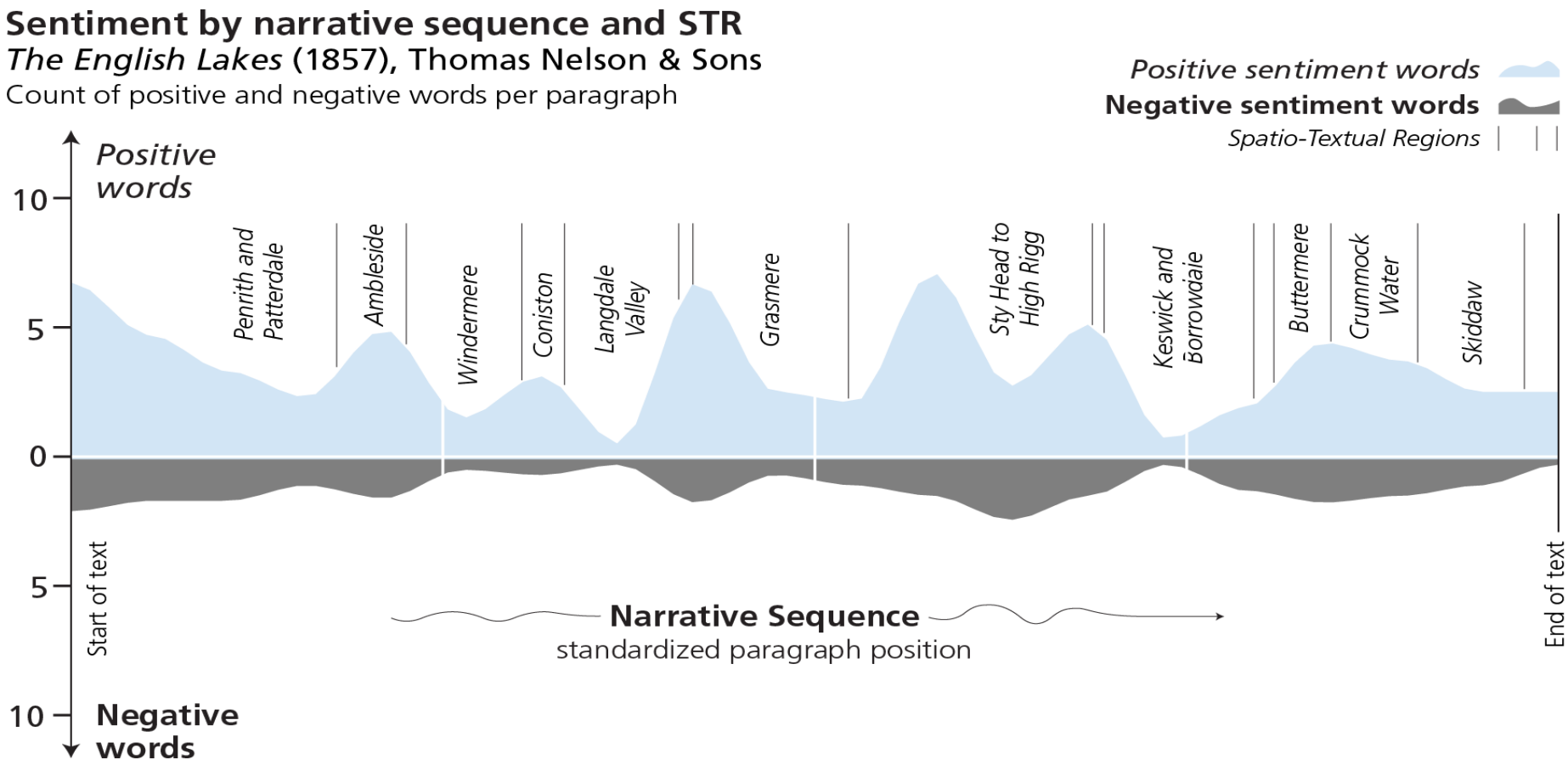
Source: Steiner, et al, Spatio-Textual Regions: Extracting Sense of Place from Spatial Narratives PLATIAL'23

CLDW: Narrative sequence and regions



Source: Steiner, et al, Spatio-Textual Regions: Extracting Sense of Place from Spatial Narratives PLATIAL'23

CLDW: Narrative sequence and sentiments



Source: Steiner, et al, Spatio-Textual Regions: Extracting Sense of Place from Spatial Narratives PLATIAL'23

How ‘fear’ was expressed…

'If they find me now, they sure will kill me.'

"Don't shoot! Don't shoot!",

'And of course, I was terrified.',

'And I said, please, can you help me?'

"I knew that I couldn't last very much longer there." ,

'Two or three shots rang out.'

"I couldn't make it to the attic.",

'And they would probably kill me.',

'As we were marching forward the sounds came from behind us.',

"I said no, no, no, it's not me."

"If I go snitch on him, he says, they'll kill me."

"You're going to get us all killed, he said."

"I couldn't jump out the window."

GPT4 on *testimony_id*: 268

In this work…

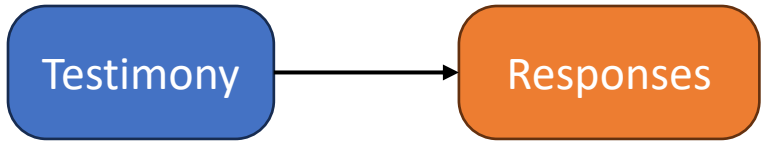
- We initially focussed on the emotion ‘fear’ in the HST.
- But later considered ‘sadness’, ‘anger’, and ‘joy’.
- How much ‘fear’ (or other emotions) is expressed in each testimony?
- How does the intensity of the emotion vary across the narrative sequence?
- Are specific places or geographies linked in some way with individual emotions?
- Only 10 testimonies were used for this exploratory work.

Classification Pipeline

Testimony

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.

Classification Pipeline



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.

Classification Pipeline

index	FileID	AnswerID	SentID	Text
0	268	0	0	My name is Henry Rosmarin.
1	268	1	0	Yes.
2	268	1	1	R-O-S-M-A-R-I-N.
3	268	2	0	At birth it was Henryk, H-E-N-R-Y-K Rozmaryn, R-O-S-M-A-R-Y-- I'm sorry, English spelling.
4	268	2	1	Polish spelling, R-O-Z-M-A-R-Y-N, Rozmaryn.
5	268	3	0	October 7th, 1925.
6	268	4	0	My present age is 73.
7	268	5	0	I was born in a little town called Czeladz, in Poland.
8	268	5	1	It's southwestern corner of Poland.
9	268	5	2	I spell it for you, C-- capital C-Z-E-L-A-D-Z.
10	268	6	0	The nearest town was Bedzin, Sosnowiec.
11	268	6	1	And towards the German border was Katowice and Siemianowice.
12	268	6	2	Incidentally, it was in Siemianowice that I lived.

Classification Pipeline

Testimony → Responses → Sentences → Emotions

1	268	1	0	Yes.
2	268	1	1	R-O-S-M-A-R-I-N
5	268	3	0	October 7th, 1925.
9	268	5	2	I spell it for you, C-- capital C-Z-E-L-A-D-Z.

And I used to be very afraid because I was the only Jew on the… on the street → ***Fear*** 😨

And at that time, I lost my mother and my sister and my little nephew → ***Sadness*** 😭

But still, they were incensed that this kind of jazz would be played there → ***Anger*** 🤬

So they were relieved to see us in the morning, coming back home → ***Joy*** 😀

Emotion Classifier

- Base model is a fine-tuned checkpoint of the DistilRoBERTa-base model
- Trained on 6 diverse datasets to predict Ekman's 6 basic emotions:
  - ***fear***, ***anger***, ***disgust***, ***joy***, ***sadness***, ***surprise*** plus ***neutral***
- Our work focused on only four of the emotions;
  - ***fear***, ***anger***, ***sadness*** and ***joy***
- The outputs from the models are validated by LLMs: GPT 3.5, GPT4
- The 'voted' emotion must be predicated by at least 2 out of the 3 models

Prompting LLMs

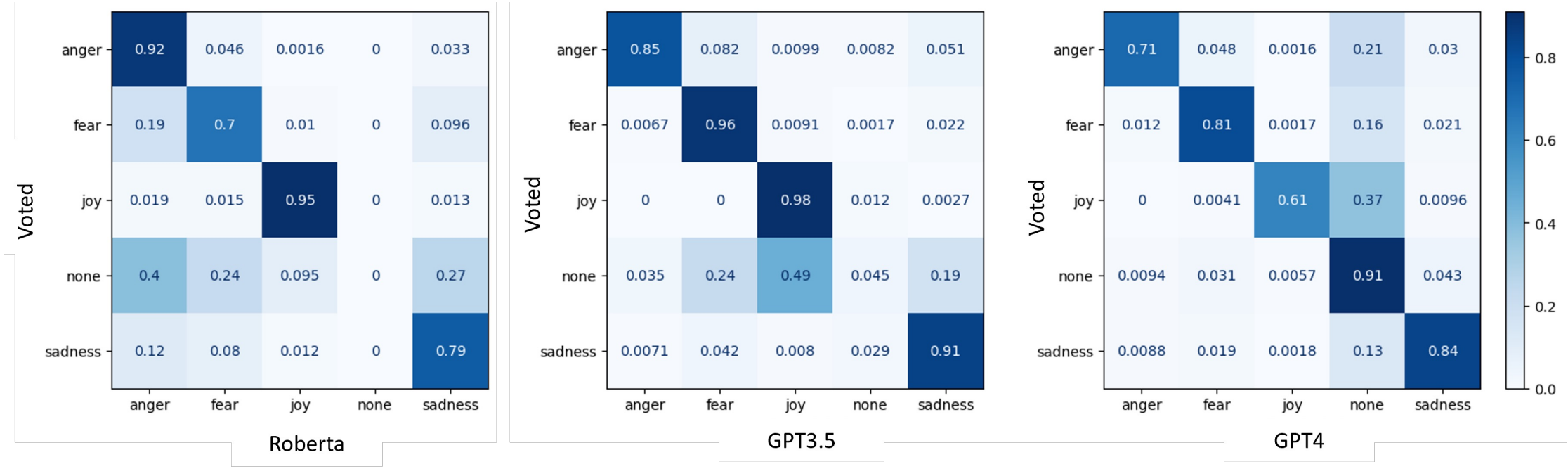
```
SYSTEM="""You're an emotional analyst that identifies emotions from text.
          Identify the emotion in this as 'fear', 'sadness', 'anger', or 'joy'
          and return ONLY one of the labels, otherwise, return 'none'"""
text='And I used to be very afraid because I was the only Jew on the street.'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

Result: ***Fear*** 😨

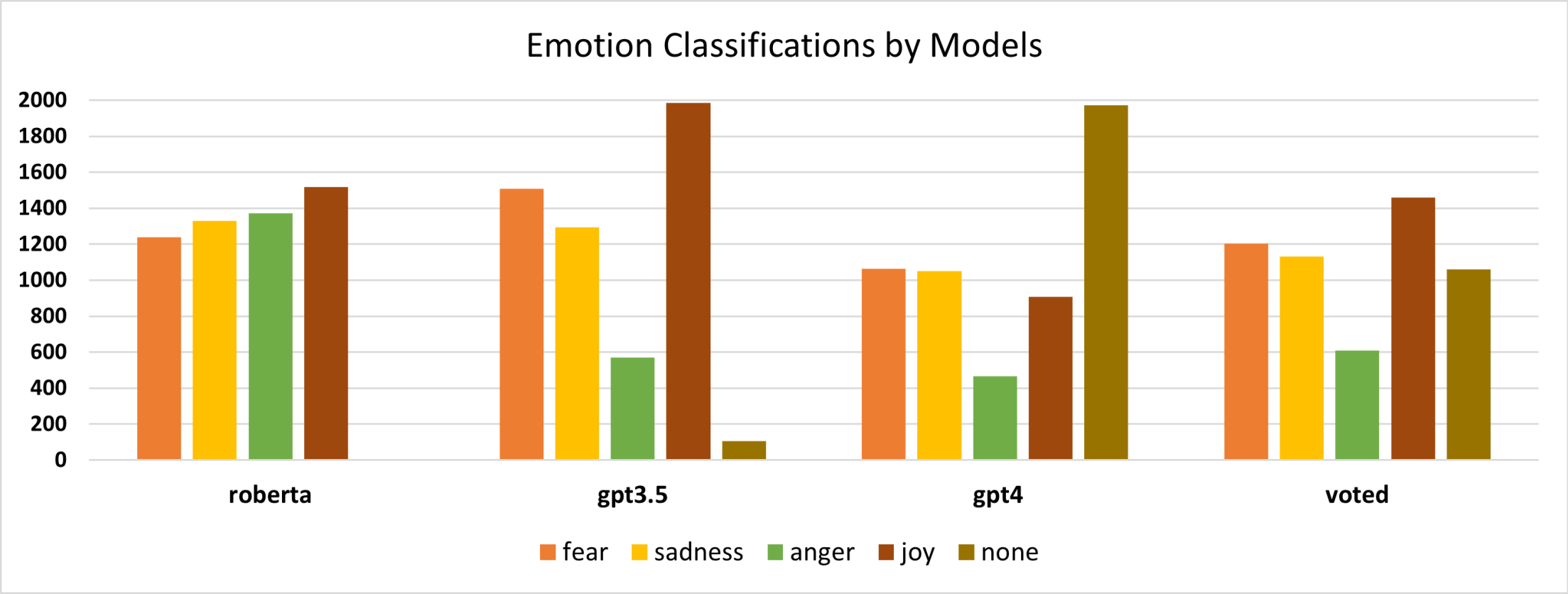
Classification examples

Sentence	fear	sad	anger	joy	roberta	gpt3.5	gpt4	voted
And I used to be very afraid because I was the only Jew on the on the street	**0.992**	0.002	0.001	0.001	fear	fear	fear	***fear***
And at that time, I lost my mother and my sister and my little nephew	0.003	**0.981**	0.002	0.001	sad	sad	sad	***sad***
But still, they were incensed that this kind of jazz would be played there	0.002	0.003	**0.971**	0.001	anger	anger	anger	***anger***
So they were relieved to see us in the morning, coming back home	0.000	0.004	0.002	**0.970**	joy	joy	joy	***joy***
They drove him out, out of the camp, and they said, go back	0.067	0.023	0.353	0.007	anger	fear	sad	***none***

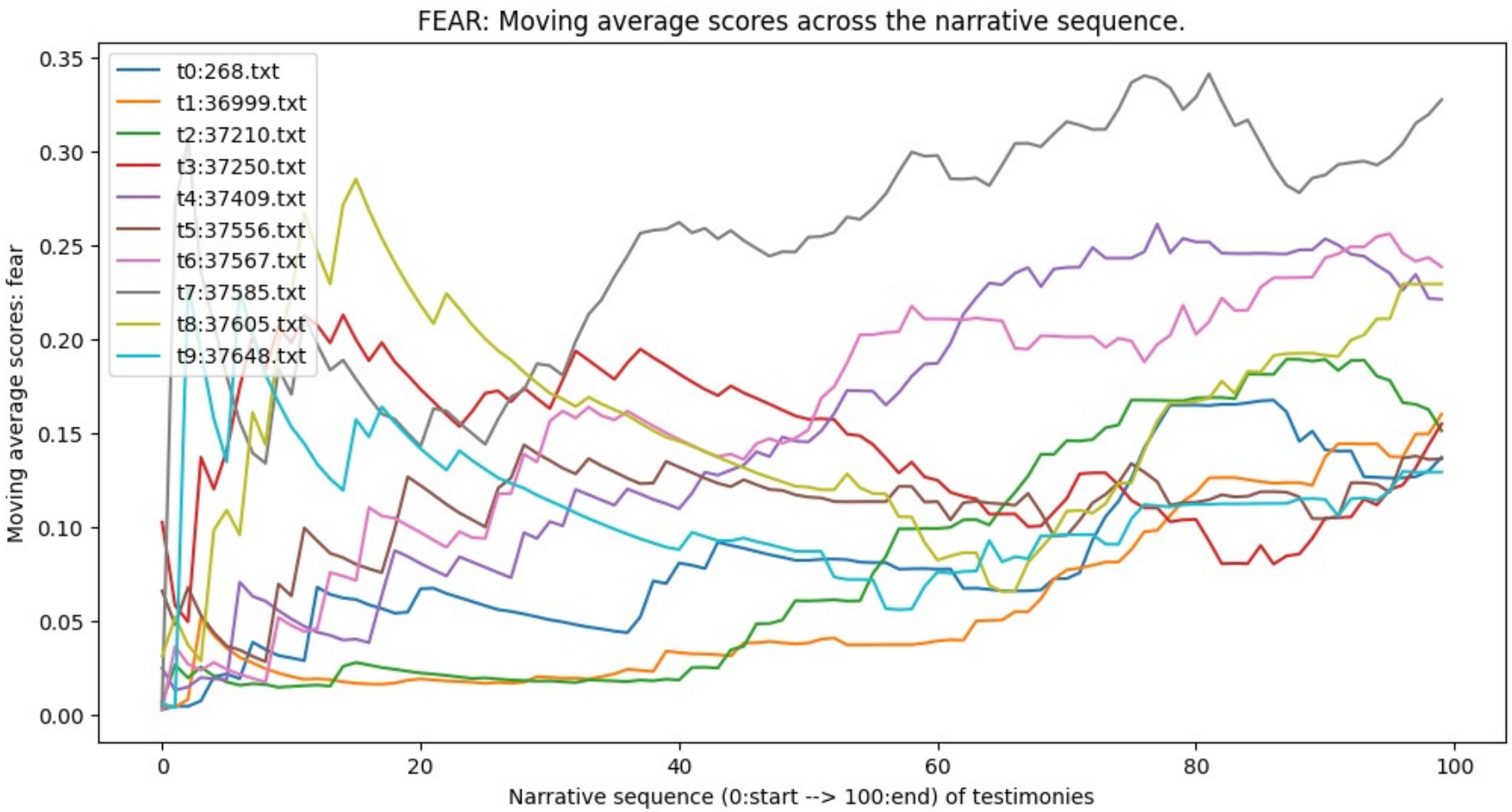
‘Voted’ vs other models



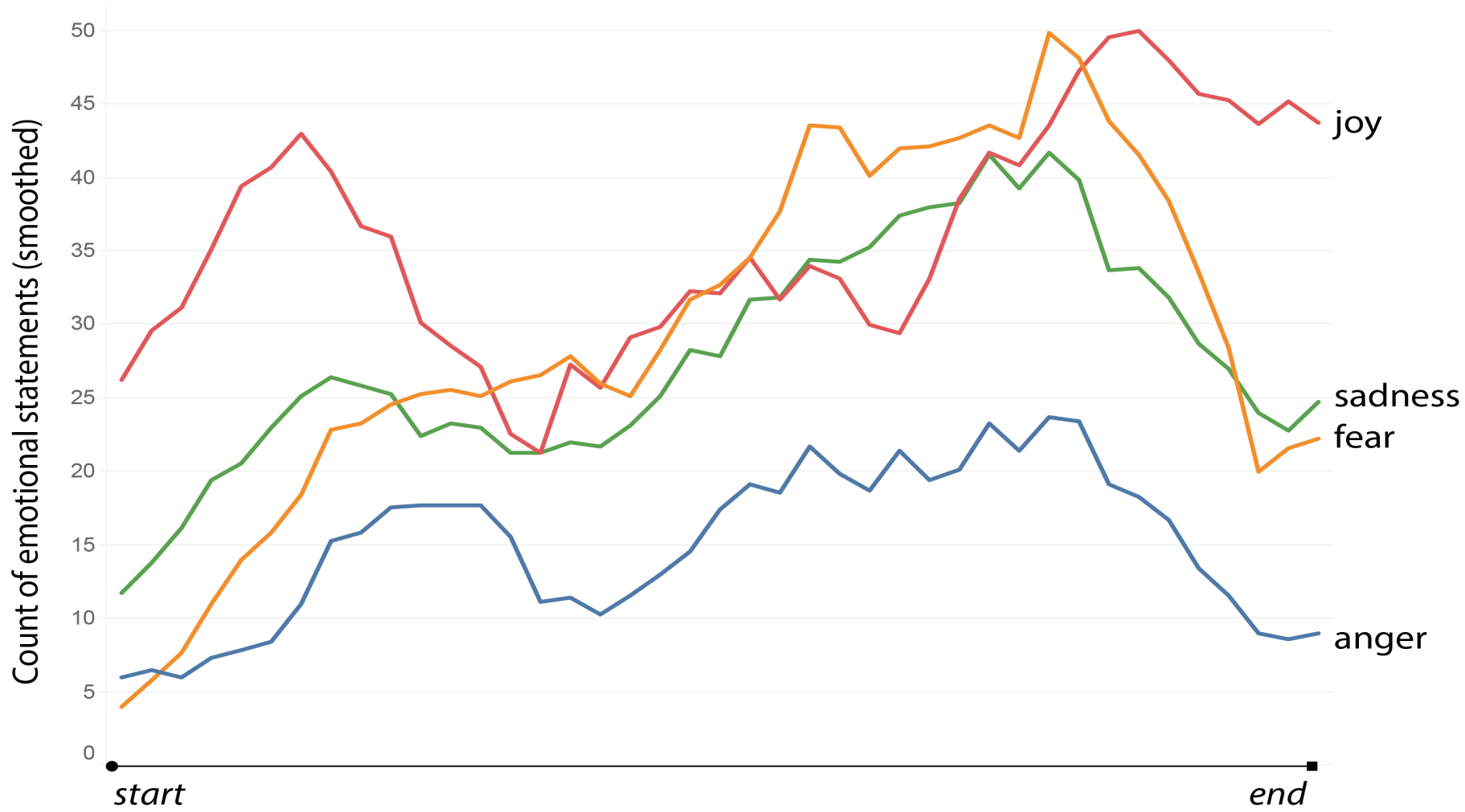
Emotion counts by models



The journey of ‘fear’



The journey of all emotion



Observations

- There seem to be variations of emotional intensities across the narrative sequence and well as between survivors.
- Survivors tend to express strong emotions at the beginning and end of the interviews
- 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 children.
- Across all interviews, positive and negative emotions are inversely correlated, with negative emotions of fear and sadness appearing together most often.

Also…emotions and places???

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%)

Conclusion

- This work delves into the concept of ‘emotional geography’ of the holocaust survivor's testimonies
- It explores the trajectories of 4 emotions (fear, anger, sadness, joy)
- Part of a framework for understanding imprecise references to space and time in the testimony data
- Key limitations to consider for future work:
  - Lack of a gold-standard dataset to evaluate/ fine-tune our models
  - Not enough testimonies to achieve reliable estimates from the emotion scores
  - Experiment with other contexts e.g. including the interview questions

Economic and Social Research Council

The National Science Foundation

The Spatial Narratives Project

Accessing and Interpreting Space and Time in Large Corpora

Thanks for your attention

Spatial Narratives Project

Lancaster University

UNIVERSITY OF LEEDS

University of BRISTOL

MANCHESTER 1824 The University of Manchester

Stanford University

IUPUI

Funded by: ESRC and NSF; **Duration:** 2022 to 2025; **Reference:** ES/W003473/1 **Website:** https://spacetimenarratives.github.io/