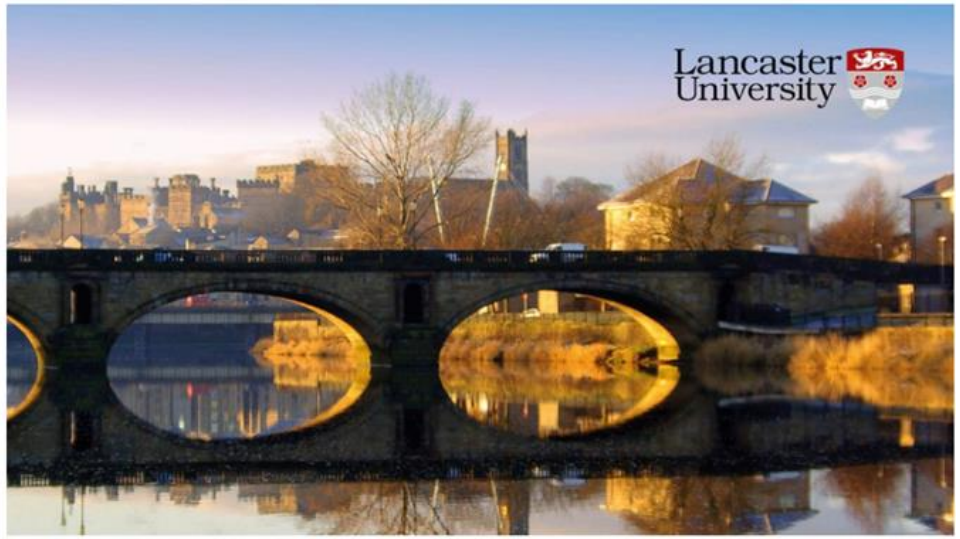
Economic and Social Research Council

The National Science Foundation

# CL2023: Spatial Humanities

# FINDING **SPATIAL** AND **TIME** NARRATIVES IN CORPUS DATA

Pre-Conference Workshop 2nd July 2023




**Dr Ignatius Ezeani**

**Senior Research Associate**
School of Computing and Communication
Lancaster University

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/

# Spatial and Temporal Relationships in Text

Lancaster University

- . . . this point , on the left , are the druidical remains called King Arthur’s Round Table, and Mayborough. Immediately after crossing Pooley Bridge the road runs along the western shore of Ulleswater to Patterdale, a distance of ten miles; but, before . . .

# Spatial and Temporal Relationships in Text

Lancaster University

- . . . this point, on the left, are the druidical remains called King Arthur's Round Table, and Mayborough. Immediately after crossing Pooley Bridge the road runs along the western shore of Ulleswater to Patterdale, a distance of ten miles; but, before. . .

# Objective

Lancaster University

Given a piece of writing, can we build a narrative around:

- Locations:
  - Texts often use toponyms (place names) for locations often reduced to a coordinate for mapping.
- Locale:
  - Usually nouns for geographical features such as ‘house’, ‘lake’, ‘mountain’ etc. often vague or ambiguous and cannot be directly mapped in Euclidean space.
- Sense of place:
  - subjective and emotional expressions by a person that make the place a unique community or landscape.
  - Events that occur at a place will contribute to this.
  - Various forms of date, time, and other temporal references

# Objective

Lancaster University

Given a piece of writing, can we build a narrative around:

- Locations:
  - Texts often use **toponyms** (place names) for locations often reduced to a coordinate for mapping.
- Locale:
  - Usually nouns for **geographical features** such as ‘house’, ‘lake’, ‘mountain’ etc. often vague or ambiguous and cannot be directly mapped in Euclidean space.
- Sense of place:
  - subjective and **emotional** expressions by a **person** that make the place a unique community or landscape
  - **events** that occur at a place will contribute to this.
  - various forms of **date**, **time,** and other temporal references.

# Entity extraction…

Lancaster University

- **Task:** Extract from this text, *place names*, *geographical feature nouns*, *events*, *emotions*, *dates and times*, *people*

From Penrith two roads lead to Pooley Bridge, about six miles distant, which spans the Eamont just at its issue from Ulleswater. Either road may be taken, be we recommend that which follows the Shap road to Eamont Bridge. Carleton Hall is near to it on the left. Cross the bridge, and take the first road to the right. At this point, on the left, are the druidical remains called King Arthur's Round Table, and Mayborough. Immediately after crossing Pooley Bridge, the road runs along the western shore of ULLESWATER to Patterdale, a distance of ten miles; but, before proceeding along it, the tourist would do well to take a walk of a few miles along the eastern shore, in the direction of Martindale, from several points on which he will obtain a good view of the lake.

# … and visualisation

Lancaster University

From Penrith PL-NAME two CARDINAL roads GEO-FEATURE lead to Pooley Bridge, PL-NAME about six miles QUANTITY distant, which spans the Eamont PL-NAME just at its issue from Ulleswater. PL-NAME Either road GEO-FEATURE may be taken, be we recommend that which follows the Shap PL-NAME road GEO-FEATURE to Eamont Bridge. PL-NAME Carleton Hall PL-NAME is near to it on the left. Cross the bridge, GEO-FEATURE and take the first ORDINAL road GEO-FEATURE to the right. At this point, GEO-FEATURE on the left, are the druidical remains called King Arthur's Round Table PL-NAME , and Mayborough PL-NAME . Immediately after crossing Pooley Bridge, PL-NAME the road GEO-FEATURE runs along the western shore GEO-FEATURE of ULLESWATER PL-NAME to Patterdale, PL-NAME a distance of ten miles QUANTITY ; but, before proceeding along it, the tourist would do well to take a walk of a few miles QUANTITY along the eastern shore, GEO-FEATURE in the direction of Martindale, PL-NAME from several points GEO-FEATURE on which he will obtain a good view of the lake. GEO-FEATURE

# Spatial Entities

Lancaster University

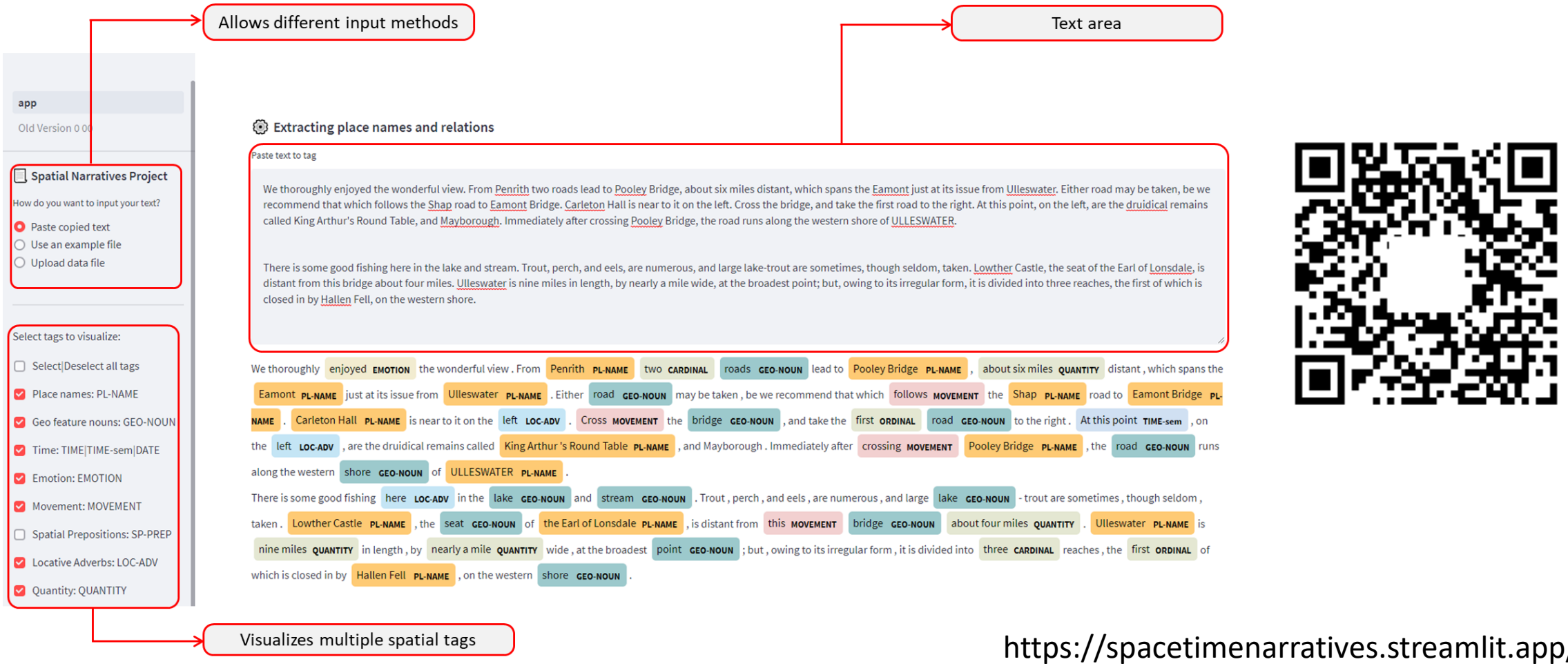
- **PLNAME:** Names of places toponyms
  - From gazetteer, annotated corpus, spacy NER
- **GEONOUN:** Geographical feature nouns
  - From 128 nouns manually created (river, road, waterfall, etc) + their inflections (*road* + *roads*)
- **EVENT:** Any description of an event or activity
  - e.g. delightful *excursions*. From manual annotations, PyMUSAS (movement)
- **EMOTION:** Sentiments and emotional expressions
  - *delightful* excursions, manual annotations, PyMUSAS (emotion)
- **DATE + TIME:** All date and time references
  - Spacy NER, manual annotation, PyMUSAS (Time)
- **PERSON:** Names and references to people
  - Spacy NER and manual annotation

*We also considering:

- **spatial prepositions** (*above*, *behind*, *below*, *through*, *within*)
- **locative adverbs** (*alongside*, *downtown*, *homewards*, *southward*, *up-country*)
- **sentiment classification** of sentences beyond identifying the emotion-bearing words

# Demo

Lancaster University




https://spacetimenarratives.streamlit.app/

# CL Pre-conference Workshop

Lancaster University

Preview | Code | Blame | 1 lines (1 loc) · 29.6 KB | Raw

Open in Colab

## Extracting Spatial Entities from text (1)

### Task Description:

From Penrith two roads lead to Pooley Bridge, about six miles distant, which spans the Eamont just at its issue from Ulleswater. Either road may be taken, be we recommend that which follows the Shap road to Eamont Bridge. Carleton Hall is near to it on the left. Cross the bridge, and take the first road to the right. At this point, on the left, are the druidical remains called King Arthur's Round Table, and Mayborough. Immediately after crossing Pooley Bridge, the road runs along the western shore of ULLESWATER to Patterdale, a distance of ten miles; but, before proceeding along it, the tourist would do well to take a walk of a few mile

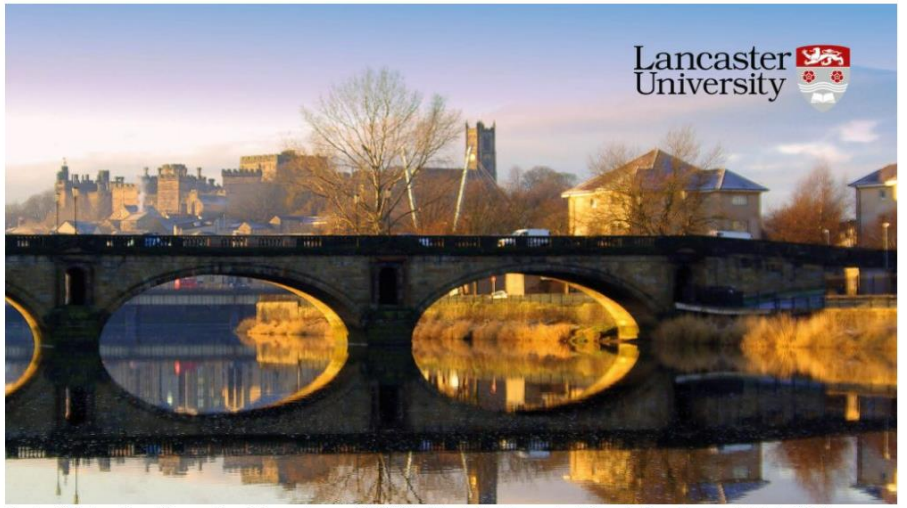
points on which he will obtain a gc

From Penrith PL-NAME two CARDINAL roads GEO-FEATURE lead to Pooley Bridge, PL-NAME about six miles QUANTITY distant, which spans the Eamont PL-NAME just at its issue from Ulleswater. PL-NAME Either road GEO-FEATURE may be taken, be we recommend that which follows the Shap PL-NAME road GEO-FEATURE to Eamont Bridge. PL-NAME Carleton Hall PL-NAME is near to it on the left. Cross the bridge, GEO-FEATURE and take the first ORDINAL road GEO-FEATURE to the right. At this point, GEO-FEATURE on the left, are the druidical remains called King Arthur's Round Table PL-NAME , and Mayborough PL-NAME . Immediately after crossing Pooley Bridge, PL-NAME the road GEO-FEATURE runs along the western shore GEO-FEATURE of ULLESWATER PL-NAME to Patterdale, PL-NAME a distance of ten miles QUANTITY ; but, before proceeding along it, the tourist would do well to take a walk of a few miles QUANTITY along the eastern shore, GEO-FEATURE in the direction of Martindale, PL-NAME from several points GEO-FEATURE on which he will obtain a good view of the lake. GEO-FEATURE

Corpus Linguistics 2023
Lancaster University

Call for papers | Registration | Venue and travel | Plenary speakers | Pre-conference workshops | Programme | Committees

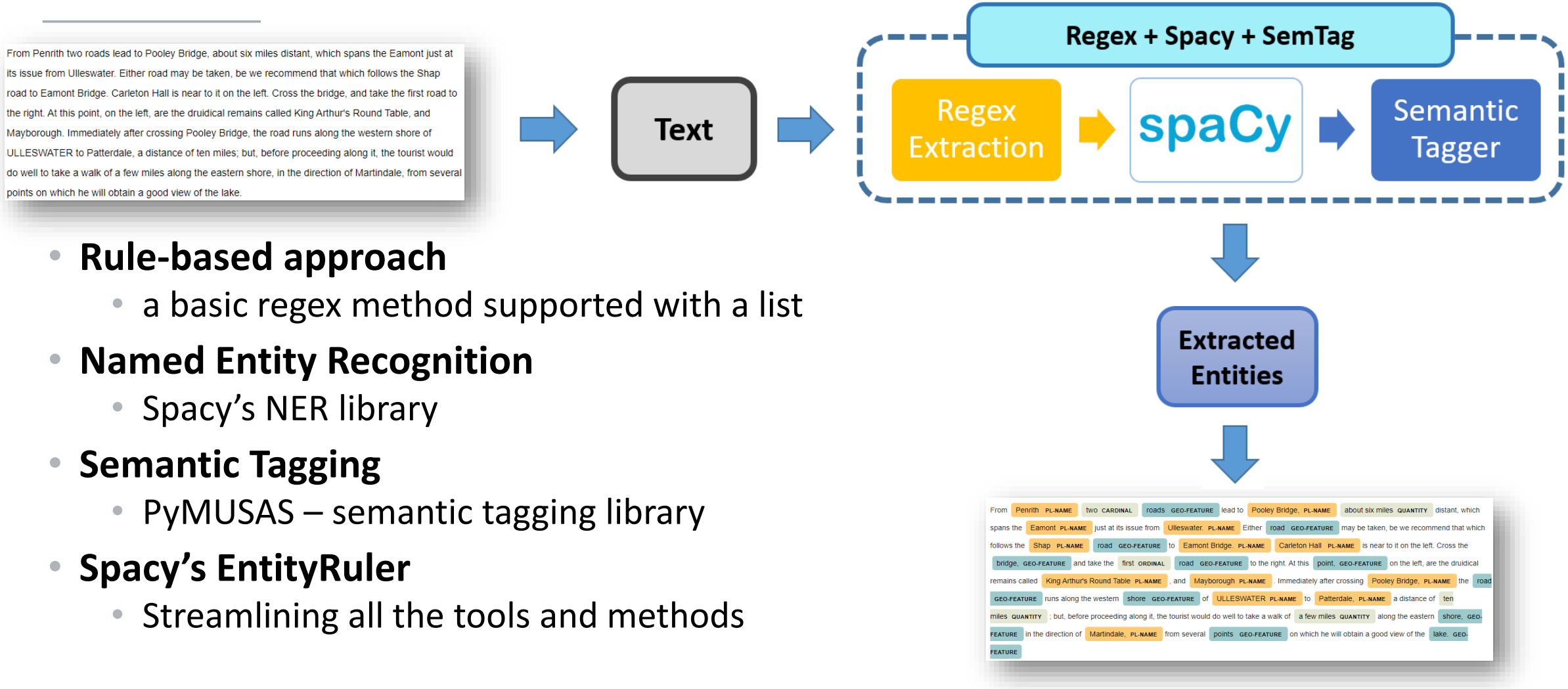
## About CL2023




The twelfth international Corpus Linguistics conference (CL2023) will be held at Lancaster University from **Monday 3rd July 2023 to Thursday 6th July 2023**. The main conference will be preceded by a workshop day on **Sunday 2nd July**.

**GitHub**: https://github.com/SpaceTimeNarratives/demo
**Workshop website**: https://wp.lancs.ac.uk/cl2023/pre-conference-workshops/

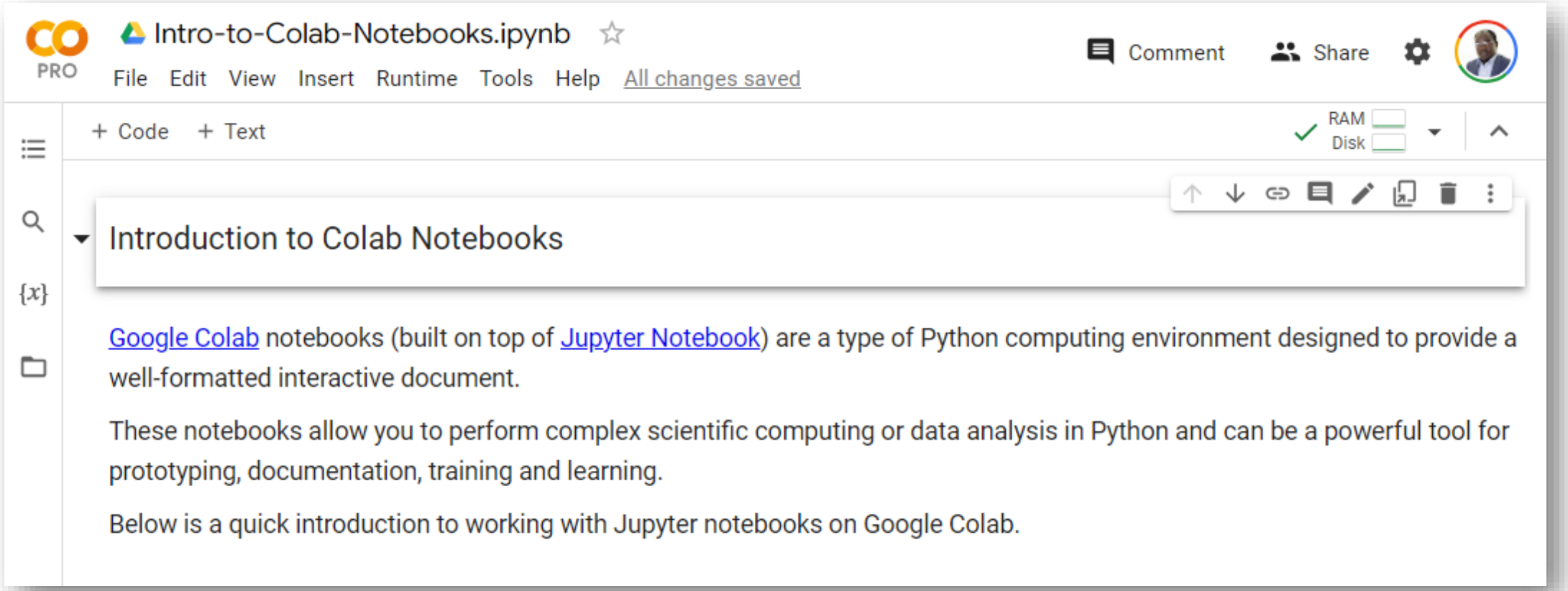
# For the workshop…




- **Rule-based approach**
  - a basic regex method supported with a list
- **Named Entity Recognition**
  - Spacy's NER library
- **Semantic Tagging**
  - PyMUSAS – semantic tagging library
- **Spacy's EntityRuler**
  - Streamlining all the tools and methods

# Intro to Colab Notebooks

Lancaster University




Link to Colab notebook:
https://colab.research.google.com/drive/1YqAfJhI2_AhOxaZCFGiLPW8xSaybv2bl?usp=sharing

Economic and Social Research Council

The National Science Foundation

# 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/