

Spatial Humanities: Finding spatial and time narratives in corpus data


@ Corpus Linguistics 2023 conference 2nd July 2023

Slides at <https://github.com/SpaceTimeNarratives/demo>

Prof Paul Rayson

Director of UCREL research centre and CLARIN ambassador
School of Computing and Communications

Lancaster University, UK

 @perayson

 @perayson@fediscience.org



What is this workshop about?

- Spatial Humanities = Digital Humanities + Natural Language Processing + Corpus Linguistics
- Methods developed in our spatial narratives project to analyse space & time references in corpus data (Lake District corpus, Holocaust survivor testimonies)
- Hands-on activities with our Python Notebooks which will guide you step by step through a variety of methods to analyse space & time in narratives
- Aims
 - You will learn what's possible with SOTA NLP/CL/DH methods on our corpora
 - You will be able to apply the techniques to your own data to see how accurate they are and how useful they are for you

Spatial Narratives Project team and funders

- Lancaster University: Ian Gregory, Paul Rayson, Ignatius Ezeani



- University of Leeds: Tony Cohn, John Stell, Erum Haris
- University of Bristol: Tim Cole
- University of Manchester: Jo Taylor
- Stanford University: Zephyr Frank, Erik Steiner
- Indiana University Purdue University, Indianapolis: David Bodenhamer, Neil Devadasan
- The project is funded from 2022 to 2025 by ESRC (ES/W003473/1), and US National Science Foundation (NSF#2225179)

Schedule for the workshop ...

Time	Activity
13:30	Welcome, introductions, and aims for the workshop (Paul Rayson)
13:45	What is spatial humanities and introduction to the ESRC-NSF spatial narratives project (Ian Gregory)
14:15	Overview of tools and data available via Streamlit and a Python Notebooks Basic How To Guide (Ignatius Ezeani)
14:45	Break
	30 minutes per notebook including a five-minute introduction by Ignatius then hands on support from all of us
15:00	Python Notebook 1: Rule based methods for named entity recognition (NER)
15:30	Python Notebook 2: NLP based methods (spaCy and PyMUSAS) for extending NER
16:00	Python Notebook 3: Advanced NLP based methods (spaCy entity ruler) for extending NER
16:30	Finish and Feedback

Things you will need (and not need) ...

- A login to the lab computers
 - Either from the registration desk or from us now
- A web browser for the practical elements
- All materials are open source (code) and open access (data)
- No prior programming experience is assumed and we will guide you through how to use the Python Notebooks
- We'll provide the corpus data, but you can test the methods and tools on your own data if you wish

Any questions before we start?

- Link to slides and materials:
 - <https://github.com/SpaceTimeNarratives/demo>
- Project website:
 - <https://spacetime narratives.github.io/>
- Contact:
 - Email: p.rayson@lancaster.ac.uk
 - Twitter: [@perayson](https://twitter.com/perayson)
 - Mastodon: [@perayson@fediscience.org](https://mastodon.social/@perayson)