

# Prototyping an Atlas of Early Modern English Drama: An Experiment on DraCor Data

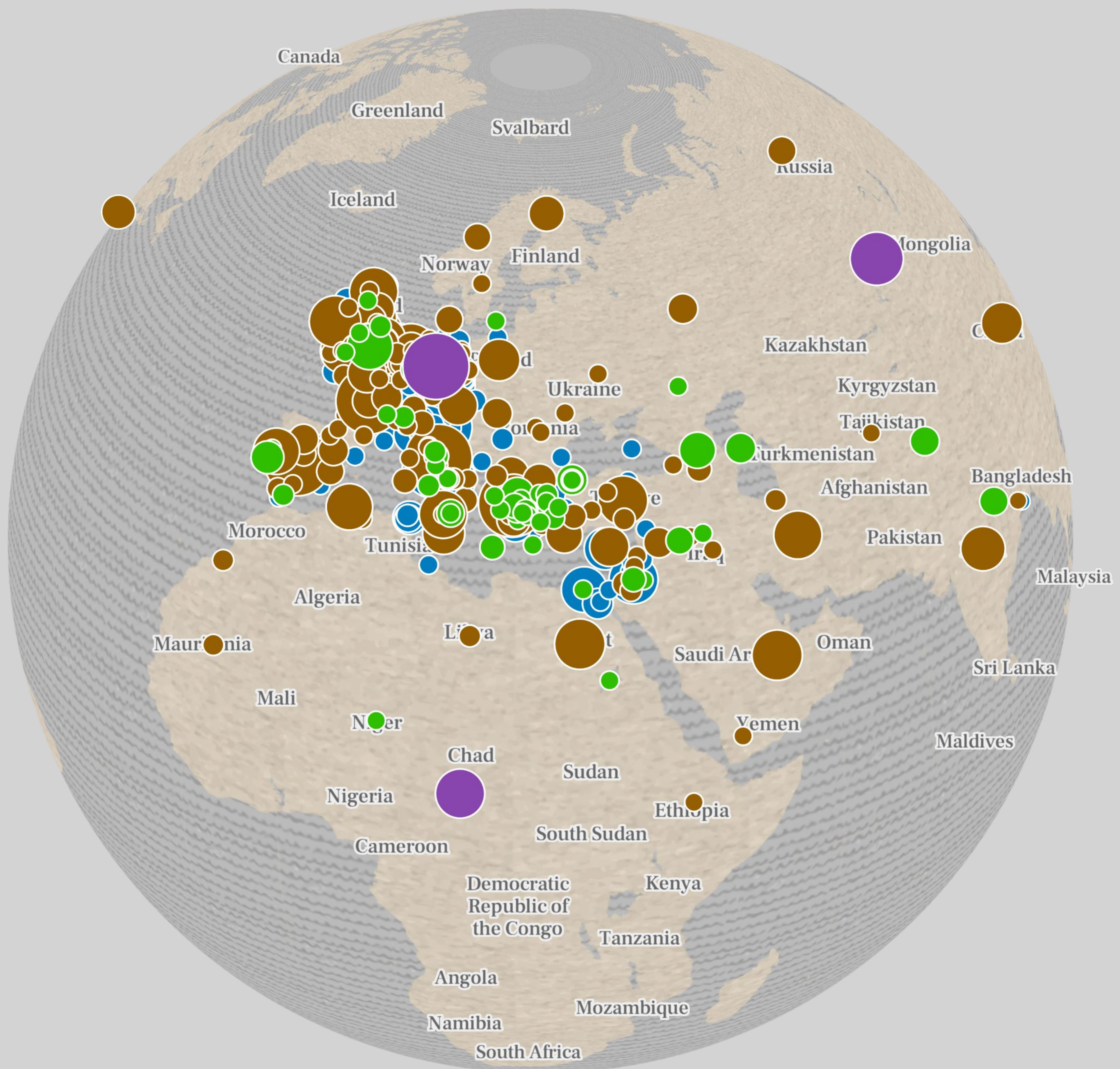
## Introduction

The "spatial turn" in literary studies and the wide availability of digital corpora now allow computational approaches to geographical data in literature.

This project attempts to map geographic mentions in an English-language corpus of early modern dramatic texts from the DraCor project (<https://dracor.org>).

## Pipeline

1. Retrieval of all spoken text from EngDraCor plays via DraCor API
2. Named Entity Recognition via *spacy* (after various experiments)
3. Post-processing and cleanup with *OpenRefine* and *GPT-4*
4. Geocoding via *Nominatim* with extensive manual correction
5. Classification (locations labelled as **cities**, **countries and regions**, **continents**, and **natural features**).



## Features

- explore more than 500 geographical references within 434 early modern English plays
- click on a location to see all plays mentioning it (in the form of "drama cards" with information on genre, date, and a direct link to the text in DraCor)
- filter locations by type (city, country/region, continent, natural feature)
  - select an individual play to see all locations linked to it

## Outlook

- The map allows to visualise the "cartographic imagination" of early modern English dramatist and can be used for teaching and research.
- It also highlights the need for tailored NLP tools for early modern texts, as current NER solutions are insufficient and pipelines are labor-intensive and not easily scalable.
- A future goal would be including other DraCor corpora in the project to enable transnational, diachronic analysis of geographic discourse.

## Authors:

Luca Giovannini (Potsdam),  
Andreas Wagner (Erlangen-Nuremberg)



Project website:  
<https://atlas-emed.github.io>