

# Advancing the Spatial Humanities: A Spanish Adaptation of a Temporal and Geographical Reference Extractor on *Trafalgar*, by Benito Pérez Galdós



Sofía Micaela Roseti, Yanco Amor Tortero Orta, Antonio Moreno Sandoval  
Laboratorio de Lingüística Informática - Universidad Autónoma de Madrid



## METHODOLOGY



1. To make some lists with characters, ships, locations, etc. that are mentioned in the book



2. To preannotate the text using the lists and correct the tags. To annotate non-automatable info (dates...)



3. To transform the tags into a CSV file compatible with Neo4j's importation options using Python (BeautifulSoup)



4. To import the CSV file using the Neo4j Python API to automate the process of



5. To query the database using Cypher in order to obtain visual insights and extract conclusions



6. To obtain the coordinates of every mentioned place (when possible) and project them onto a NeoDash map

## RESULTS

The result is a Neo4j database containing **1106** unique nodes and **1885** edges.

Abstract node representations can be obtained in **Neo4j Browser** by querying with Cypher. Maps can be built and interacted with using **NeoDash**.

### Unique insights

The abstract representation can show a wide array of –previously **unpredicted**– relationships and connections. This methodology is fully **customizable**, with the structure of the annotated data shaping the nodes and edges that are linked and displayed. Furthermore, this marks the starting point, as Neo4j offers countless possibilities beyond, like querying using natural language, vectorizing your database, implementing a ChatBot for data querying, serve RAG purposes, etc.

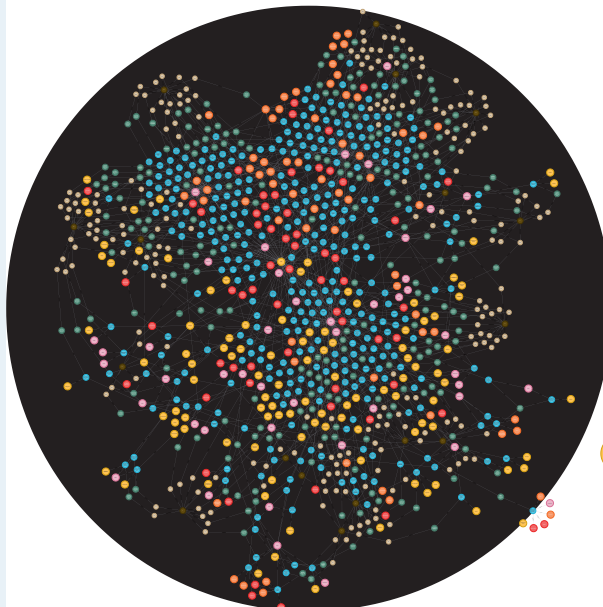
## CONCLUSION

Despite their limited use in this context, **Knowledge Graphs** have proven to be a powerful tool for **literary analysis**. The nodes and edges obtained by querying can reveal valuable insights into the book, and the maps are essential for comprehending the breadth of locations where the story unfolds. *Trafalgar* was an ideal choice for this analysis, given the richness of its depicted locations and their significant role in shaping the plot.

## INTRODUCTION

This is our first approach to literary analysis using **Knowledge Graphs**, namely Neo4j. An initial inspiration has been taken from Lancaster University's UCREL group, as they work with geographical features. Thus we aimed at extracting and representing the **geographical mentions** from *Trafalgar*, given its rich portrayal of recognizable Spanish locations. However, our scope extends beyond geography to include characters, ship names and the dates, all vital for constructing a coherent narrative.

## VISUALIZATION



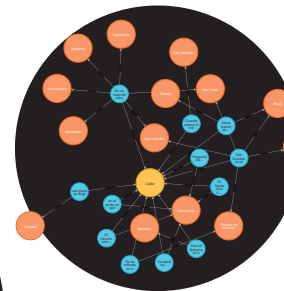
All nodes and edges

## GOALS

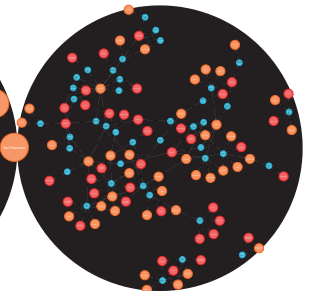
■ To depict an **abstract representation** of the relationships among characters, locations, ships and dates where the actions take place. The hierarchical structure of *Trafalgar* was preserved by also taking into account the chapters, paragraphs and units (segments containing a location/ship). This enables us to query the database where these relationships are stored in order to gain fresh insights into the story.

■ To project the places mentioned in the book onto an actual **map**, thereby bringing Benito Pérez Galdós's vivid depictions closer to the reader. Visualizing these geographical details enhances our understanding of the setting and enriches our experience of this literary masterpiece.

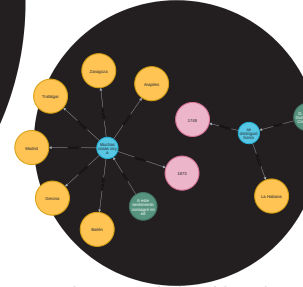
● Locations ● Characters ● Unit ● Para. no. ● Ships ● Date ● Para. cont. ● Chapter



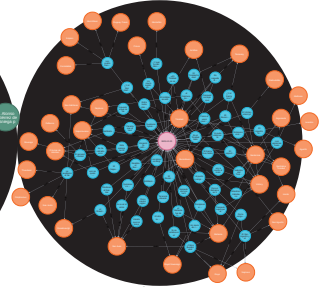
The location with the biggest number of ship mentions (Cádiz)



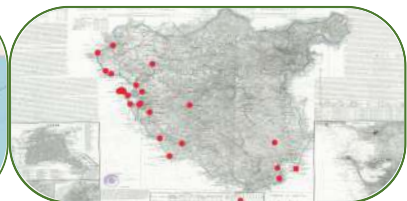
Character mentions alongside ship mentions



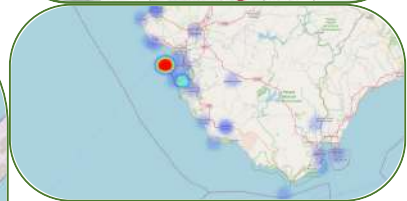
The unit with the oldest date (1748) vs. the most recent (1873)



Every ship mention on 21/10/1805



CÁDIZ



## AKNOWLEDGEMENTS

The participation of Yanco Amor Tortero Orta and Sofía Micaela Roseti has been sponsored by the Chair UAM-IIC on Computational Linguistics.



Visit our website