U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Historical Conflicts Search Engine

Information Retrieval

Dinis Sousa (up202006303)
Gabriel Ferreira (up201906072)
João Matos (up202006280)
João Pinheiro (up202008133)

Table of contents

01. Introduction

02. Past Results

03. Semantic Search

04. Learning To Rank

05. Old vs New

06. Interface

07. Conclusions

01

Introduction

Introduction

In this milestone, we will explore improvements in our search engine as building an interface.

Historical Conflicts

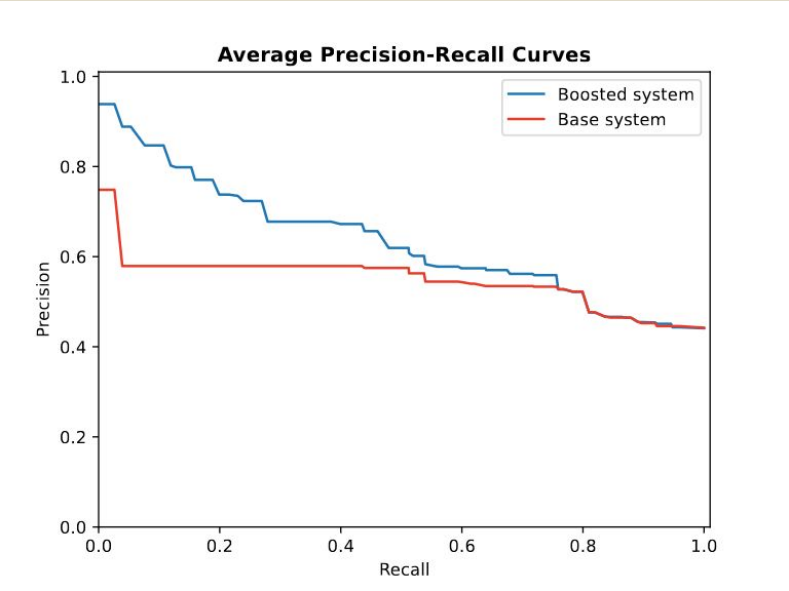
Remembering...

We collected our data from Wikidata to build our database, and with the help of Solr we built our search engine

Past Results

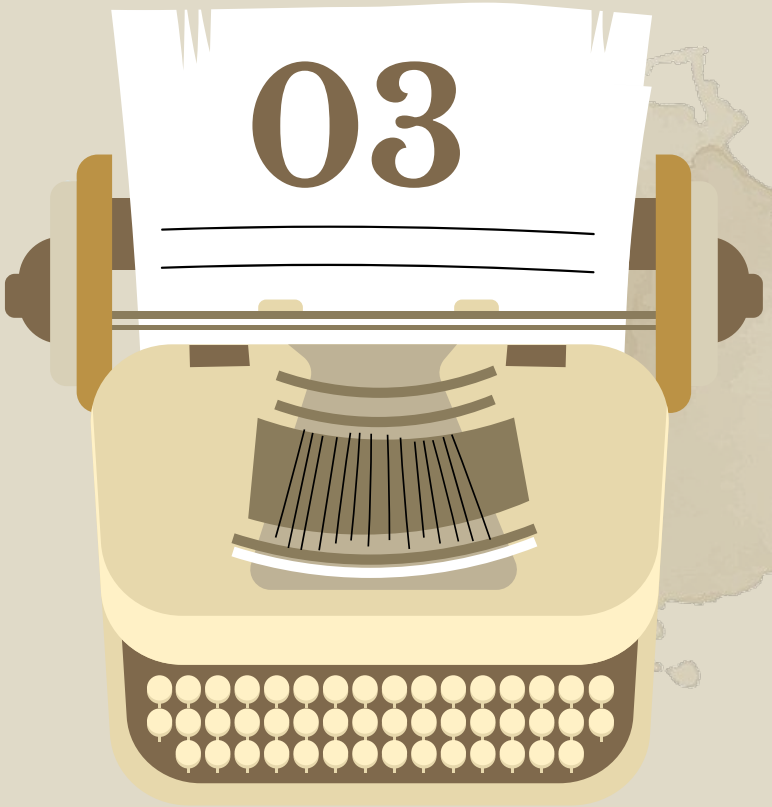
Past Results

Boosted System outperformed Base System



Boosted System performed better for all queries.
We used query geral filters and query specific.

Semantic Search



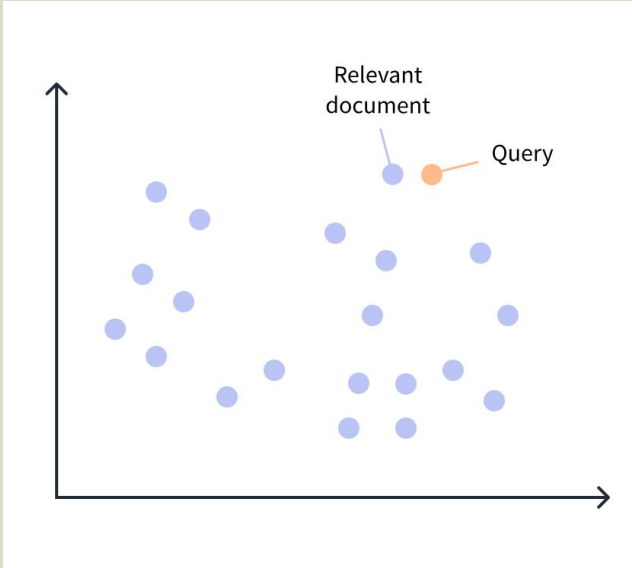
Semantic Search

Using the Solr tool, we implemented semantic search for our documents. The vector is created using the label and the summary.

We created a schema for semantic, as well as creating the vectors for all the documents using a python script.

We aim achieving:

- Contextual Understanding
- Context Recognition
- Nuanced Query Interpretation
- Enhanced Relevance



Learn To Rank

Learning To Rank

Learn To Rank (LTR) leverages machine learning algorithms to improve the ranking of search results.

We manually attributed a relevance score to each of the documents gathered for each information need.

Using a linear classifier, we got the weights for each one of the features chosen.

Feature	Weight
Original Solr Ranking Score	0.0
Is Battle (binary)	-0.94669
Is War (binary)	-0.32686
Is Revolution (binary)	0.85364
Summary Length	-0.00055
"Part Of" Length	0.03309
Participants Count	0.02995
Date	-6.62141e-14

Table 6: Features and their weights of the Learning to Rank model

Features were limited by Solr, which were mainly, features that use the value of a field, the length of a field, or some basic functions like if conditions.

Old vs New

Evaluation Process

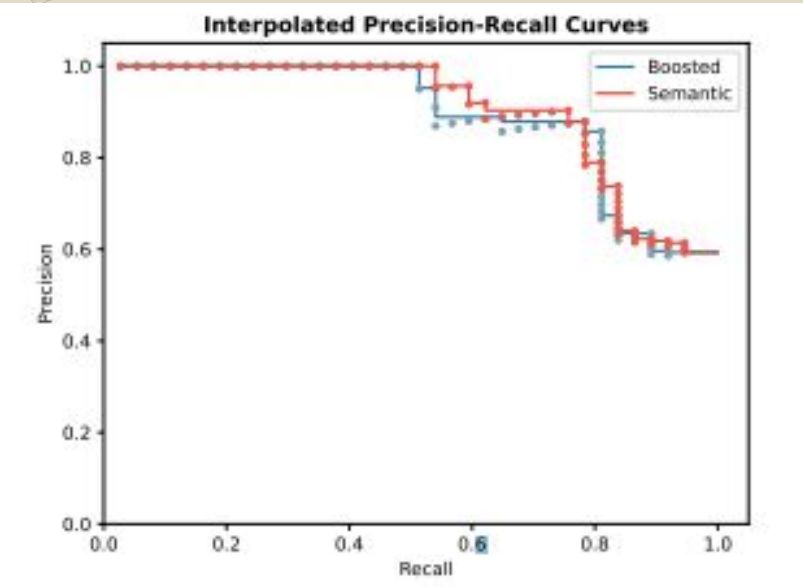
Recalling what we have done in the previous milestone, the metrics used for evaluation are below. We evaluated each one of the new systems, the system using semantic search and the system using LTR with our current best system, the boosted system.

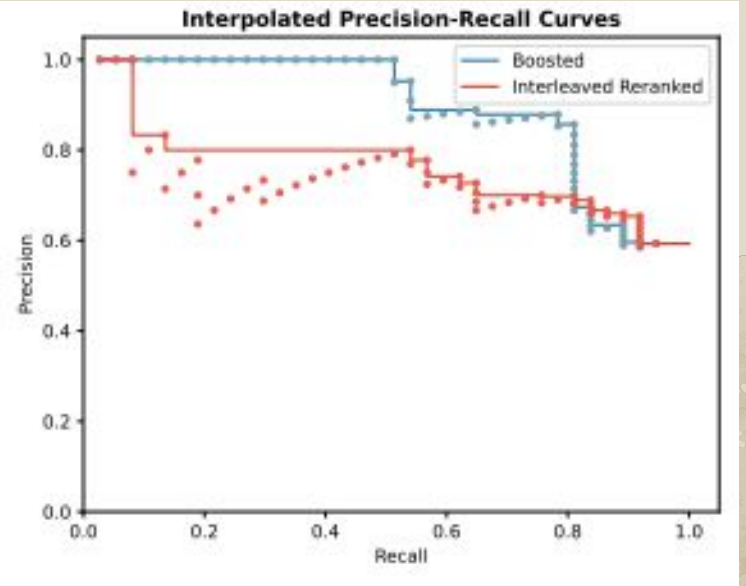
Metrics:

- Precision at 10 (P@10): The proportion of relevant results among the top 10 retrieved documents.
- Average Precision (AvP): The mean precision across different recall levels.
- Precision-Recall Curves (P-R curves): Visual representation of precision at varying recall levels.

Evaluation

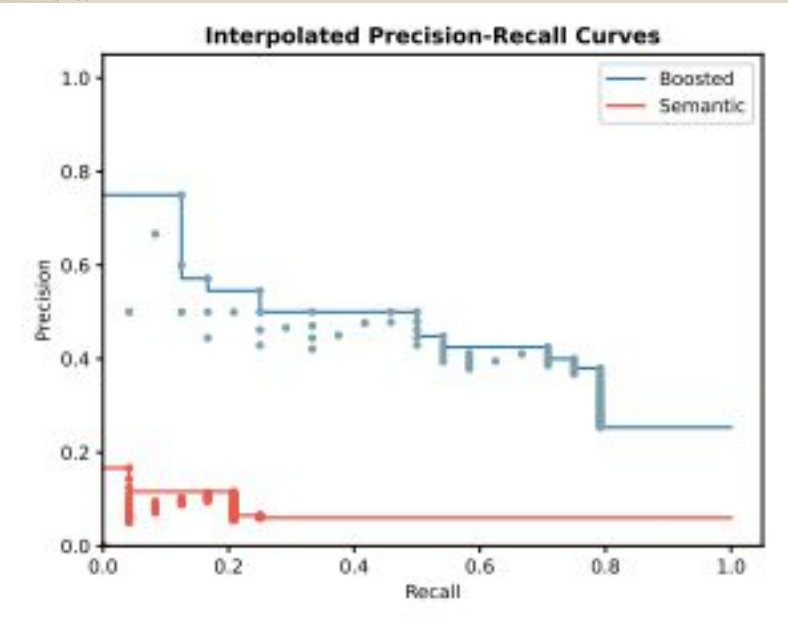
1 What battles took place near a **river** between **1700 and 1775**?

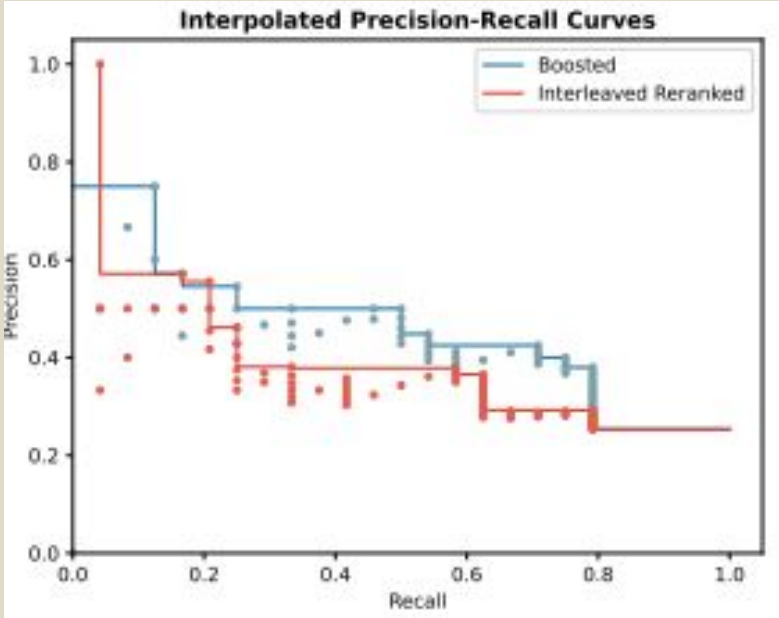




Evaluation

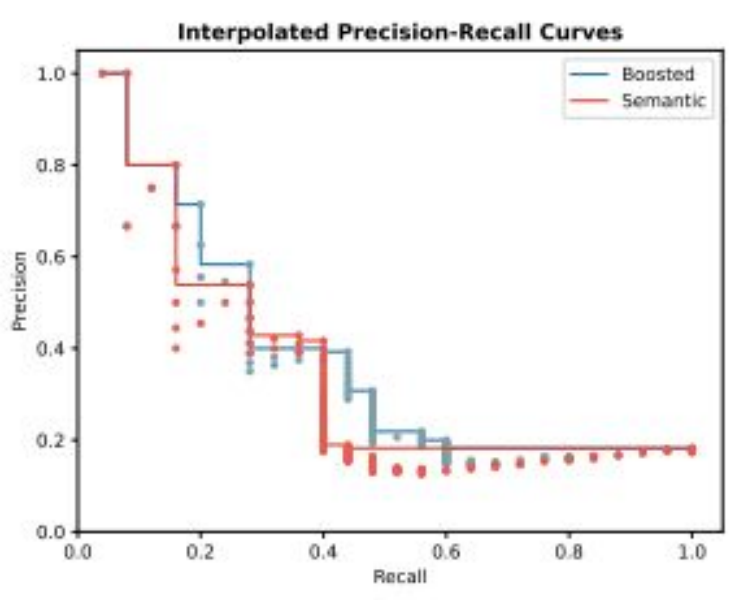
2 What were the most **destructive** battles in **Europe** during **World War I**?

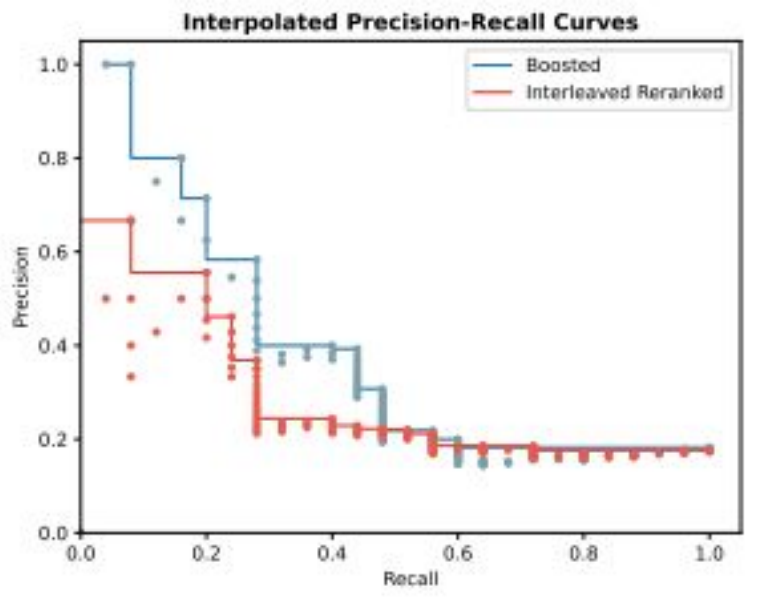




Evaluation

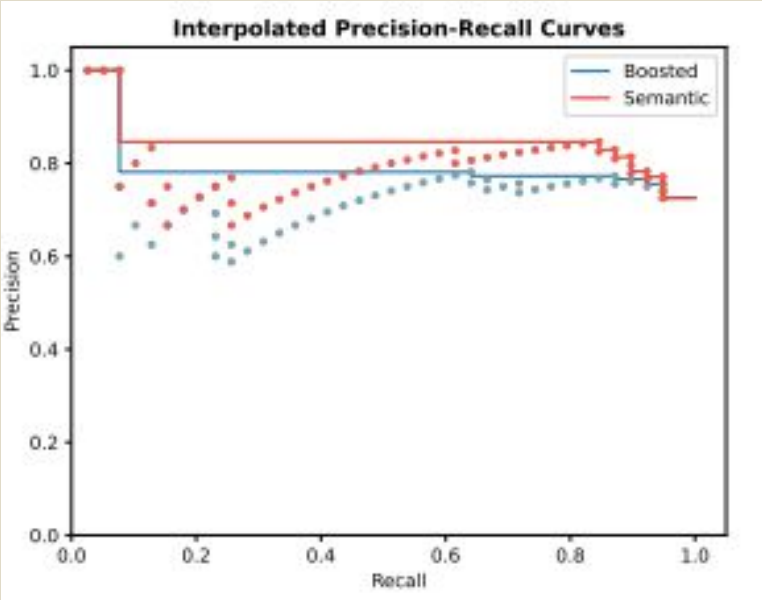
3 In which battles did the **Portuguese** participate between **1300 and 1800**, where they were not the main participant?

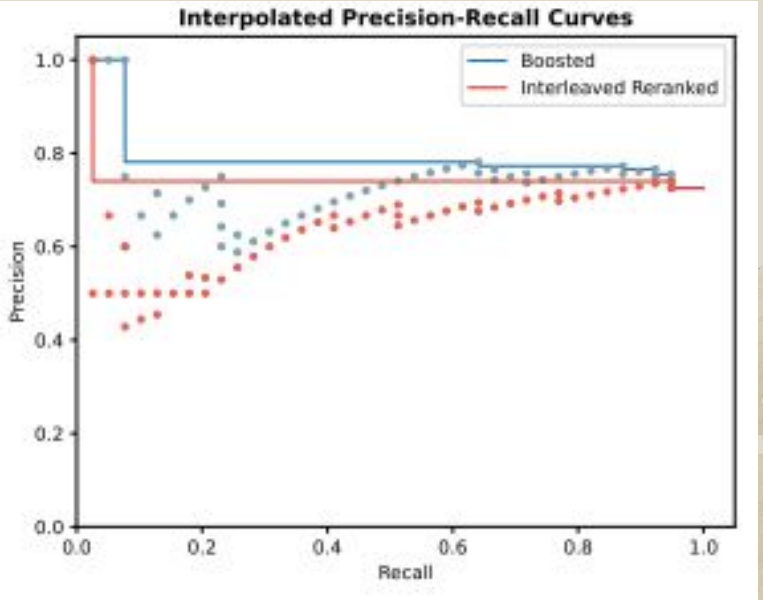




Evaluation

4 What **battles**, that are **part of revolutions**, had **economic consequences**?





Evaluation

We ended the evaluation by comparing the overall average of the precision-recall curves for the 4 information needs, for the 4 systems

The boosted system is consistently better than the other systems

Which was expected, as the boosted system is not an definitive answer, as it is not a solution that fits all needs, being too specific

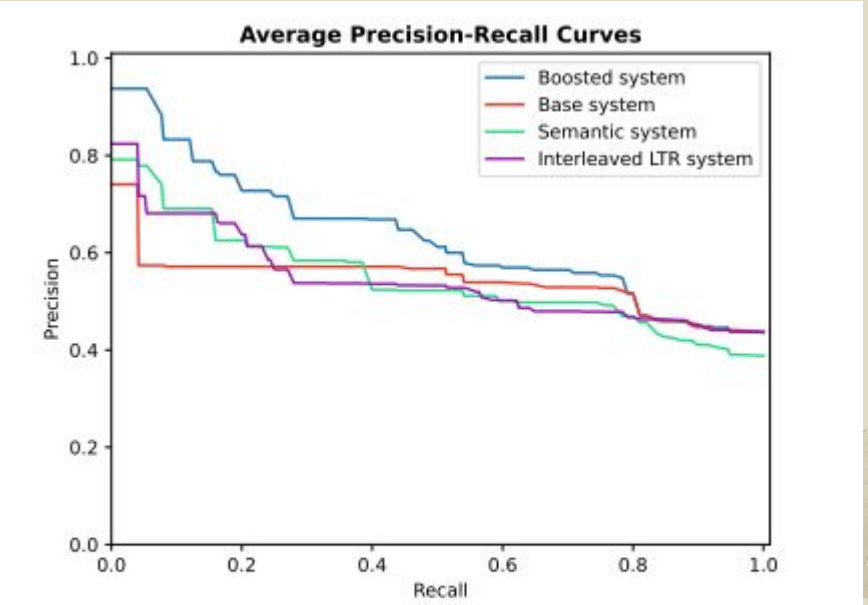


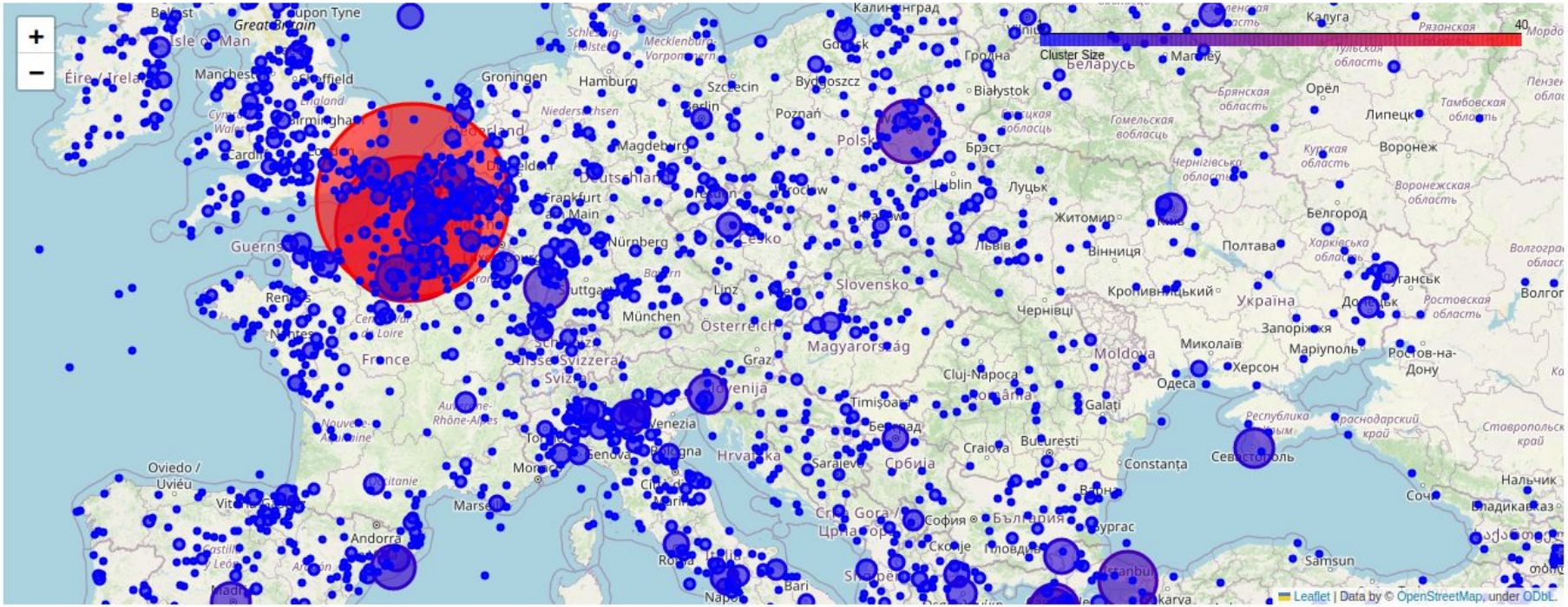
Figure 18: Average Interpolated Precision-Recall of the 4 systems

Interface

All Fields | Search... | More options

Historical Conflicts

Search for more than 9000 battles, wars and other of the most relevent conflicts in the world history



All Fields | Which battles did Spanish participated? | More options

Start Over | Which battles did Spanish participated? ✖

Limit your search

Date ❯

Label ❯

Country ❯

« Previous | **1** - **10** of **8,321** | Next »

Sort by none | 10 per page

1. Spanish American wars of independence

☐ Bookmark

Date: 1808-09-25

Summary: The Spanish American wars of independence (25 September 1808 – 29 September 1833; Spanish: Guerras de independencia hispanoamericanas) were numerous civil wars in Spanish America with the aim of political independence from Spanish rule during the early 19th century. These began shortly after the start of the French invasion of Spain, during the Napoleonic Wars, as a struggle for sovereignty in both hemispheres, between those who wanted a unitary monarchy (royalist) rather than plural monarchies or republics (patriots). Thus, the strict period of military campaigns would go from the battle of Chacaltaya (1809), in present-day Bolivia, to the Battle of Tampico (1829), in Mexico.In 1808, the abduction of the Spanish royal family by Napoleon Bonaparte, the Abdications of Bayonne, gave rise tw... Read More

Participants: Bolivia, Gran Colombia, Chile, Peru, First Mexican Empire, United Provinces of the Río de la Plata, and Spanish Empire

Participants Count: 7

Instance Of: war of national liberation and historical event

Location: Hispanic America

Part Of: Latin American wars of independence

Tools

☐ Bookmark

Spanish American wars of independence

Date: 1808-09-25

Wikipedia Article: Spanish American wars of independence on Wikipedia

Wikidata Event: Spanish American wars of independence on Wikidata

Summary: The Spanish American wars of independence (25 September 1808 – 29 September 1833; Spanish: Guerras de independencia hispanoamericanas) were numerous civil wars in Spanish America with the aim of political independence from Spanish rule during the early 19th century. These began shortly after the start of the French invasion of Spain, during the Napoleonic Wars, as a struggle for sovereignty in both hemispheres, between those who wanted a unitary monarchy (royalist) rather than plural monarchies or republics (patriots). Thus, the strict period of military campaigns would go from the battle of Chacaltaya (1809), in present-day Bolivia, to the Battle of Tampico (1829), in Mexico.In 1808, the abduction of the Spanish royal family by Napoleon Bonaparte, the Abdications of Bayonne, gave rise two years later to an emergence of liberalism and desire of Popular sovereignty throughout the Spanish Empire. At first, some major cities or capitals formed local Juntas on the basis of laws from the Hispanic tradition. The violent conflicts started in 1809, with short-lived governing juntas established in Chuquisaca, La Paz and Quito opposing the government of the Supreme Central Junta of Seville. At the beginning of 1810, numerous new juntas appeared across the Spanish domains in the Americas when the Central Junta fell to the French invasion. Although various regions of Spanish America objected to many crown policies, "there was little

Search History

Your recent searches

Clear Search History

Worl War I

Spanish Colonization of Hispaniola

"Spanish Colonization of Hispaniola"

Part Of:"Spanish Colonization of Hispaniola"

Part Of:"Latin American wars of independence"

Which battles did Spanish participated?

Which battles did Spanish participated?

Summary:revolutions that had several economic consequences

What were the revolutions that had several economic consequences?

*

Summary:battles by river in 18th century Date:Last 500 Years

All Fields | Search... | More options

Advanced search

Match all of the fields below

All Fields

Label

Part Of

Summary

AND have these attributes

Label

Country

Sort results by none

Search | Start over

Search tips

- Select "match all" to require all fields.
- Select "match any" to find at least one field.
- Combine keywords and attributes to find specific items.
- Use quotation marks to search as a phrase.
- Use "+" before a term to make it required. (Otherwise results matching only some of your terms may be included).
- Use "-" before a word or phrase to exclude.
- Use "OR", "AND", and "NOT" to create complex boolean logic. You can use parentheses in your complex expressions.
- Truncation and wildcards are not supported - word-stemming is done automatically.

Conclusions

Conclusions

- We couldn’t achieve a better performance with Semantic Search or Learning to Rank.
- Continuing to Refine and Expand the search engine’s capabilities to accommodate a broader spectrum of user queries.
- Exploring Solr Geospatial search capabilities to retrieve conflicts that occurred in proximity to the user’s current location, providing a personalized historical context would be a good improvement.
- Ongoing work on improving the interface and interactivity of the search engine to ensure a seamless and informative user experience, enhancing accessibility and engagement.

Thanks

Do you have any questions?