

# Team Bruce Banner at Touché 2022: Argument retrieval for controversial questions

**Team** : Bernardo Moreira, Bruno Martins,  
Fábio Goularte, and Henrique Cardoso

8<sup>th</sup> september 2022

# Overview

- Introduction
- Background
- Problem
- Solution
- Experiments
- Conclusions

## Context

- Argument retrieval focuses on retrieving the most pertinent information, given a query by a user;
- Controversial topics are often sensitive topics;
- To defend a point of view, one needs access to pertinent and reliable information;

## Context

- The Touché task was presented at CLEF and focuses on retrieving arguments for controversial topics.

The screenshot shows the 'args' search engine interface. The search bar contains 'the war in iraq'. Below the search bar, there are navigation links: 'All', 'Discussions', 'News', and 'People'. On the right, it says 'Pro vs. con view' and '2265 arguments retrieved in 721.0 ms'.

The results are divided into two columns: 'PRO' (Pro) and 'CON' (Con).

**PRO Results:**

- Also, the war in Iraq is step one in the so called "War..."**
  - Show full argument
  - Also, **the war in Iraq** is step one **in the** so called "**War on Terror**". ... **The UN** also found no WMD's **in Iraq**, and **the US** superseded their authority and invaded anyway.
  - <https://www.debate.org/debates/The-war-in-iraq-is-not-a-justified-war/1/> score ▾
- The United Nations asked Iraq to cooperate with weapons...**
  - Show full argument
  - The United Nations** asked **Iraq** to cooperate with weapons inspectors and verify their possession of cruise missiles and other weapons, but **the nation's** officials were noncompliant. ... However, others argue against **the war** because ...
  - <https://www.debate.org/debates/Iraq-war/3/> score ▾
- If we did not attack when we did, then we would have to...**
  - Show full argument
  - If we did not attack when we did, then we would have to attack later after we were nuked and **the international** prestige would be lowered then." ... We don't have magical levitation powers to lift them out. ...
  - <https://www.debate.org/debates/War-in-Iraq/1/> score ▾

**CON Results:**

- When you go to war you are aggressive. ... Saddam...**
  - Show full argument
  - When you go to **war** you are aggressive. ... **Saddam** admitted during interrogation that he had planned to renew his development of WMDs, including nuclear weapons.
  - <https://www.debate.org/debates/The-war-in-iraq-is-not-a-justified-war/1/> score ▾
- I will be arguing to end the Iraq war while the opponent...**
  - Show full argument
  - I will be arguing to end **the Iraq war** while **the opponent** must give reason to stay **in the Iraq war**. ...
  - <https://www.debate.org/debates/Iraq-War/1/> score ▾
- If you did hear this, I'd like to see proof. ... Sources:...**
  - Show full argument
  - If you did hear this, I'd like to see proof. ... Sources: [1] <http://www.msnbc.msn.com...> [2] <http://news.bbc.co.uk...> [3] <http://www.defense.gov...> [4] <http://www.iraqbodycount.org...> [5] <http://www.uslaboragainstarwar.org...> [6] ...
  - <https://www.debate.org/debates/The-war-in-iraq/4/> score ▾

## Goals

- Despite the focus for the 2022 edition of the Touché Task 1 being different from the 2021 edition, can a **simple IR approach** based on the construction of documents as **sentence pairs** obtain a competitive model for the Touché Task 1?

## Contributions

- With an IR approach based on sentence pairs, we achieved **competitive** results for the **2022** edition of the Touché Task 1;
- Submitted a paper for the **CLEF** conference thoroughly describing our approach.

## Approaches for the Touché Task 1

### BM25



Estimate relevance of documents to an input query.

### TF-IDF



Measure used to evaluate the relevance of a word in a document.

### DirichletLM



The Dirichlet function avoids unrecognized terms having estimated zero probability.

### WordNet



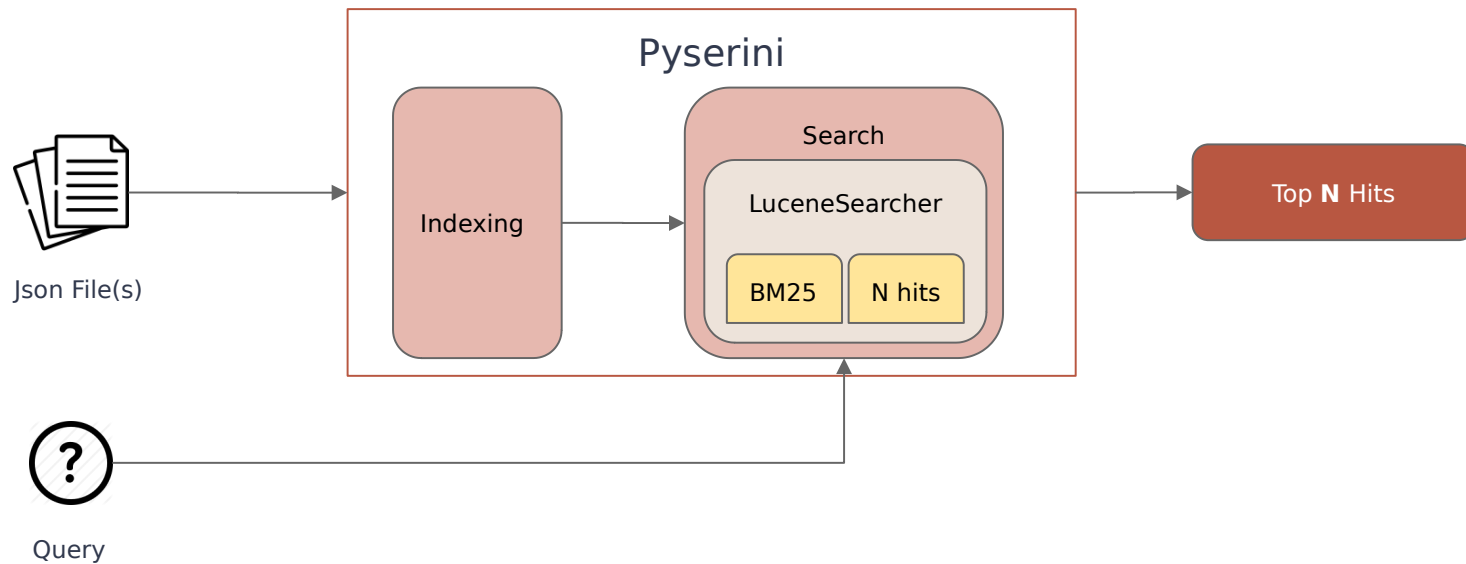
Used for query expansion.

## Pyserini

- Python toolkit for information retrieval search;
- Sparse search using **Lucene** search library;
- Dense search using **Faiss** search library;
- Hybrid search using **HybridSearcher** which combines **Lucene** and **Faiss**.

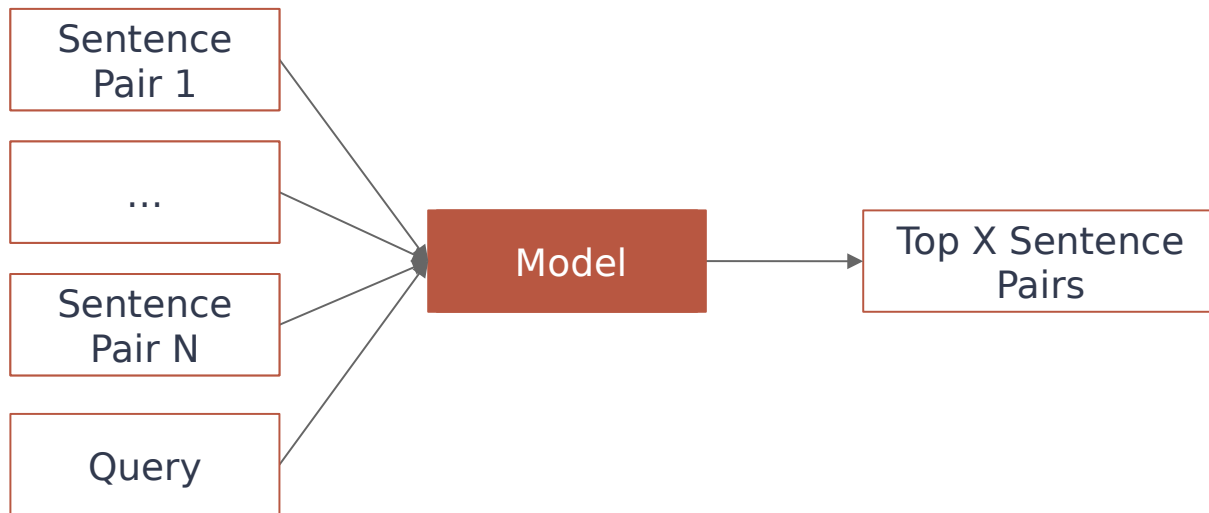


# Pyserini sparse search

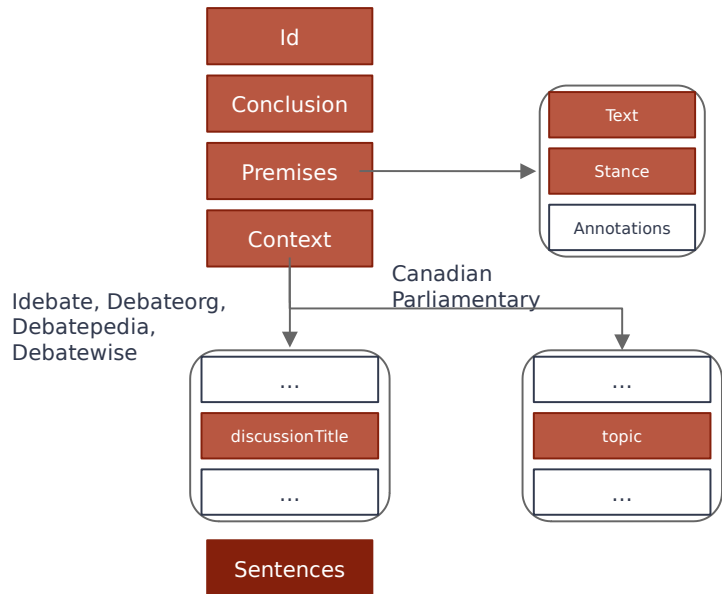


## Touché Task 1 2022

- For the 2022 edition of the Touché Task 1, the goal is to retrieve relevant pairs of sentences for a controversial question.



# Touché Task 1 2022 - Data



Id	
"Sf9294c83-Af186e851"	
Conclusion	
"Saddam Hussein is gone and Iraq is now functioning as one of very few democracies in the Middle East"	
Text	Stance
"His removal provides stability and security not only for Iraq but for the Middle East as a region"	"PRO"
discussionTitle/topic	
"the War in Iraq was Worth the Cost"	
Sentences	
<p>'sent_id': 'Sf9294c83-Af186e851_PREMISE_1'            'sent_text': 'His removal provides stability and security not only for Iraq but for the Middle East as a region'</p> <p>'sent_id': 'Sf9294c83-Af186e851_CONC_1',            'sent_text': 'Saddam Hussein is gone and Iraq is now functioning as one of very few democracies in the Middle East'</p>	

## Touché Task 1 2022 - Approach

### Pair of sentences formation

Creates all the possible pairs from sentences in the same argument;

### Queries and pairs variations

Create two types of queries and two types of pairs;

### Sparse search

With the four different data combinations, Pyserini performs the sparse search.

## Touché Task 1 2022 - Pairs of sentences formation

It's important to be clear that this debate is looking at the results of the Iraq war and, by any definition Iraq is in a much more stable and secure position than it was in 2003 when American, British and other international troops arrived in the country

**Sf9294c83-A9a4e056e\_PREMISE\_1**

Whatever one thinks of the initial justifications for the war there is no doubt that the country, the region and the world are better and safer places without Saddam Hussein[i].

**Sf9294c83-A9a4e056e\_PREMISE\_2**

Saddam Hussein is gone and Iraq is now functioning as one of very few democracies in the Middle East

**Sf9294c83-A9a4e056e\_CONC\_1**

Data  
processin  
g 

It's important to be clear that this debate is looking at the results of the Iraq war and, by any definition Iraq is in a much more stable and secure position than it was in 2003 when American, British and other international troops arrived in the country. Whatever one thinks of the initial justifications for the war there is no doubt that the country, the region and the world are better and safer places without Saddam Hussein[i].

**Sf9294c83-A9a4e056e\_PREMISE\_1, Sf9294c83-A9a4e056e\_PREMISE\_2**

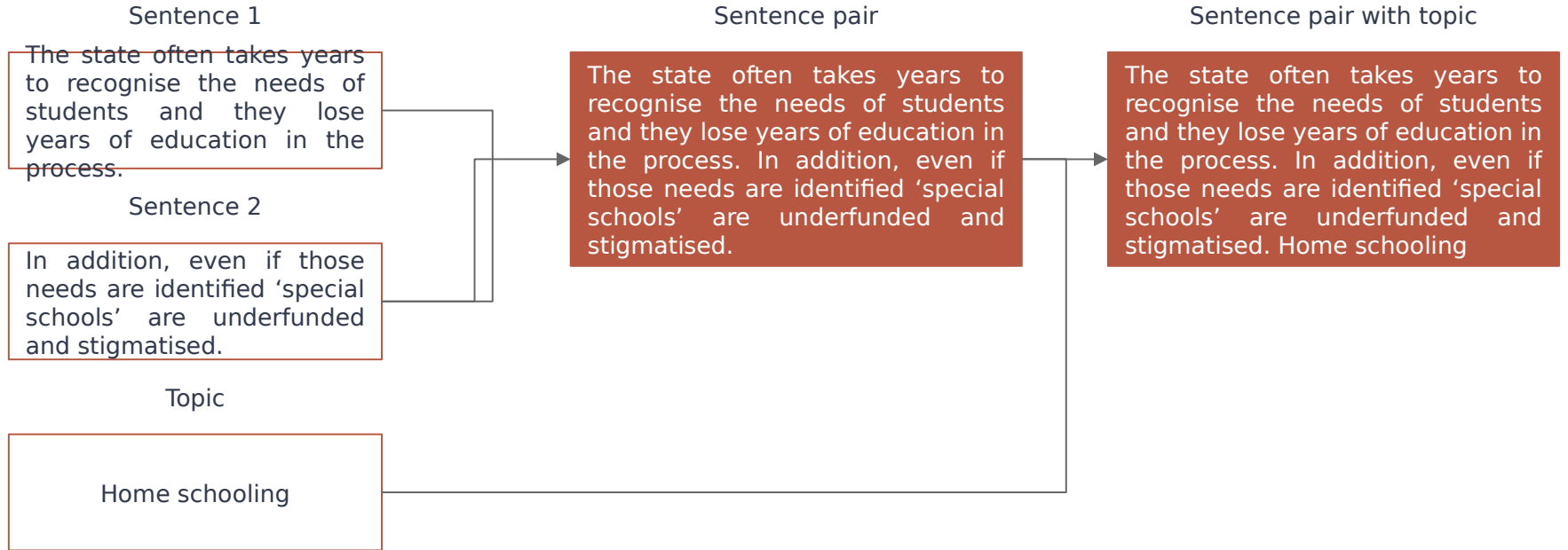
It's important to be clear that this debate is looking at the results of the Iraq war and, by any definition Iraq is in a much more stable and secure position than it was in 2003 when American, British and other international troops arrived in the country. Saddam Hussein is gone and Iraq is now functioning as one of very few democracies in the Middle East

**Sf9294c83-A9a4e056e\_PREMISE\_1, Sf9294c83-A9a4e056e\_CONC\_1**

Whatever one thinks of the initial justifications for the war there is no doubt that the country, the region and the world are better and safer places without Saddam Hussein[i]. Saddam Hussein is gone and Iraq is now functioning as one of very few democracies in the Middle East

**Sf9294c83-A9a4e056e\_PREMISE\_2, Sf9294c83-A9a4e056e\_CONC\_1**

## Touché Task 1 2022 - Pairs of sentences



## Touché Task 1 2022 - Query

<topic>

<number>1</number>

<title>Should teachers get tenure?</title> → Query with title only

<description> A user has heard that some countries do give teachers tenure and others don't. Interested in the reasoning for or against tenure, the user searches for positive and negative arguments. The situation of school teachers vs. university professors is of interest. </description>

<narrative> Highly relevant arguments make a clear statement about tenure for teachers in schools or universities. Relevant arguments consider tenure more generally, not specifically for teachers, or, instead of talking about tenure, consider the situation of teachers' financial independence. </narrative>

</topic>

Query with title, description and narrative concatenated;

## Touché Task 1 2022 - Runs

	Query title	Query title, description, and narrative
Sentence pair	pyserini_sparse_v2	pyserini_sparse_v1
Sentence pair with topic	pyserini_sparse_v4	pyserini_sparse_v3



## Experiments - Touché Task 1 2022

- Touché provided the evaluation files;
- Human assessors label the retrieved documents manually;
- The metrics used for evaluation are:
  - rhetorical **quality** of the document;
  - general topic **relevance**;
  - whether the sentence pair forms a **coherent** text;
- Trec\_eval was used to compare our results with the evaluation files.

## Touché Task 1 2022 - Submitted runs results

Run	Quality NDCG@5	Relevance NDCG@5	Coherenc e NDCG@5
pyserini_sparse_v1	<b>0.772</b>	0.641	<b>0.378</b>
pyserini_sparse_v2	0.701	0.580	0.353
pyserini_sparse_v3	0.760	<b>0.651</b>	0.354
pyserini_sparse_v4	0.709	0.586	0.357

## Touché Task 1 2022 - Classification

Team	Relevance NDCG@5
Porthos	0.742
Daario Naharis	0.683
<b>Bruce Banner</b>	<b>0.651</b>
D Artagnan	0.642
Gamora	0.616
Hit Girl	0.588
Pearl	0.481
Gorgon	0.408
General Grievous	0.403
Swordsman	0.356
Korg	0.252

Team	Quality NDCG@5
Daario Naharis	0.913
Porthos	0.873
Gamora	0.785
Hit Girl	0.776
<b>Bruce Banner</b>	<b>0.772</b>
Gorgon	0.742
D Artagnan	0.733
Pearl	0.678
Swordsman	0.608
General Grievous	0.517
Korg	0.453

Team	Coherence NDCG@5
Daario Naharis	0.458
Porthos	0.429
Pearl	0.398
<b>Bruce Banner</b>	<b>0.378</b>
D Artagnan	0.378
Hit Girl	0.377
Gamora	0.285
Gorgon	0.282
Swordsman	0.248
General Grievous	0.231
Korg	0.168

## Conclusion



For the 2022 edition of the Touché Task 1 the Pyserini sparse search resulted in third, fourth and fifth places in relevance, coherence and quality, respectively;



The results for the 2022 edition could be improved by using a different data arrangement.

## Future work



Different data arrangement to form sentence pairs;



Develop a considerably smaller sentence-pairs dataset to use Pyserini hybrid search.

Thank you for your attention!