

Data Acquisition for Argument Search: The args.me corpus



Yamen
Ajjour
Weimar



Henning
Wachsmuth
Paderborn



Johannes
Kiesel
Weimar



Matthias
Hagen
Halle

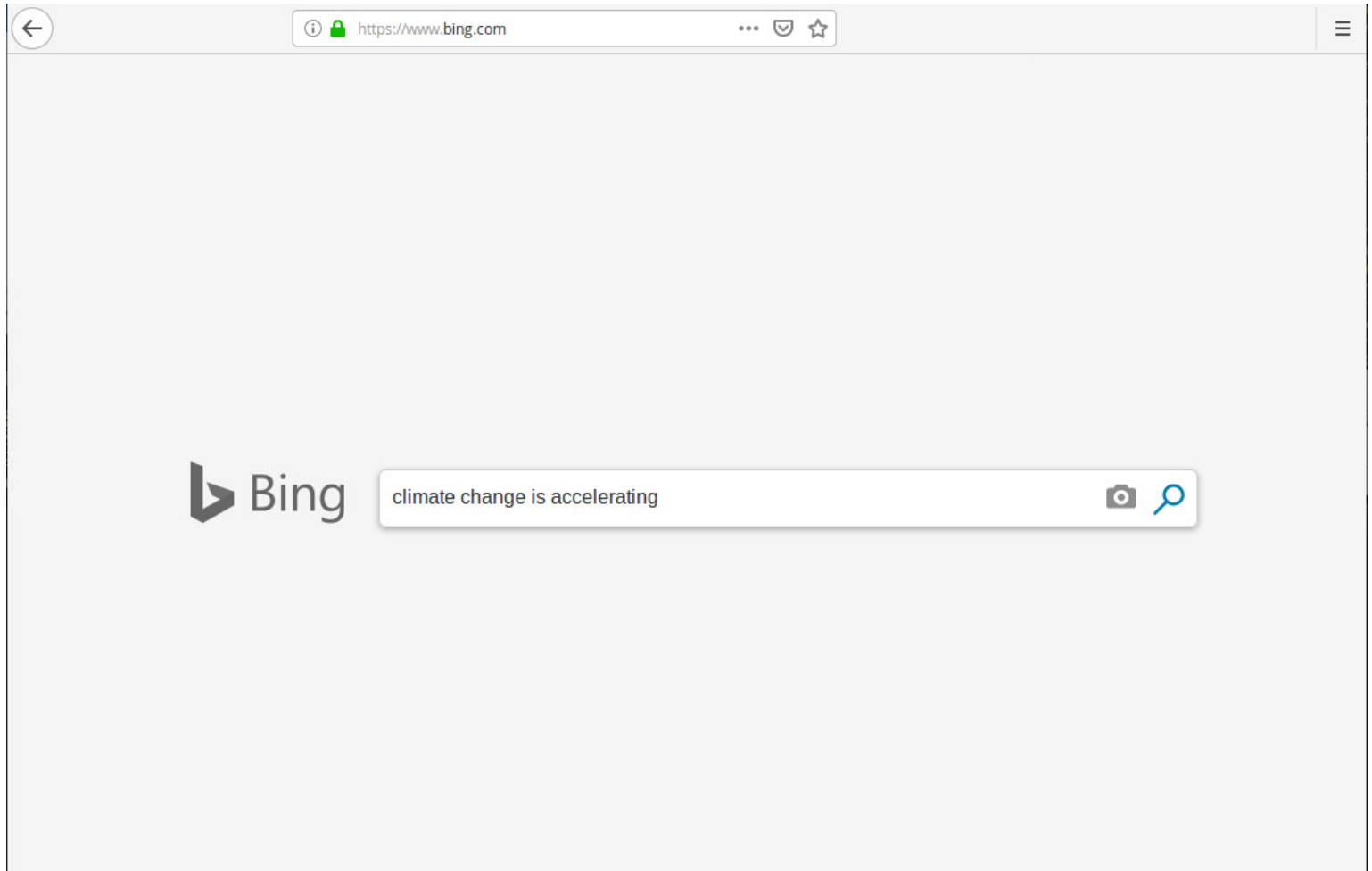


Martin
Potthast
Leipzig

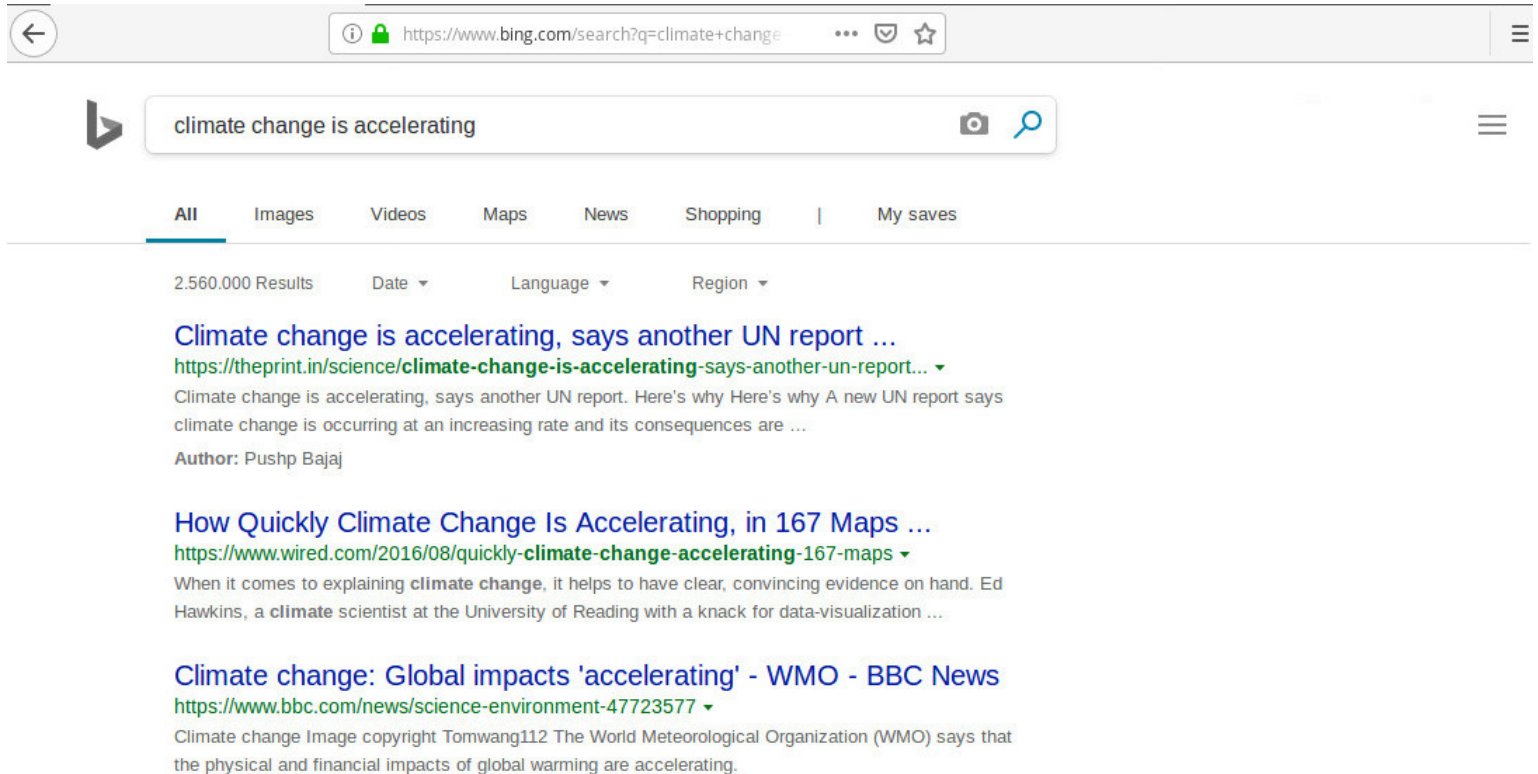


Benno
Stein
Weimar

Limitations of conventational search engines



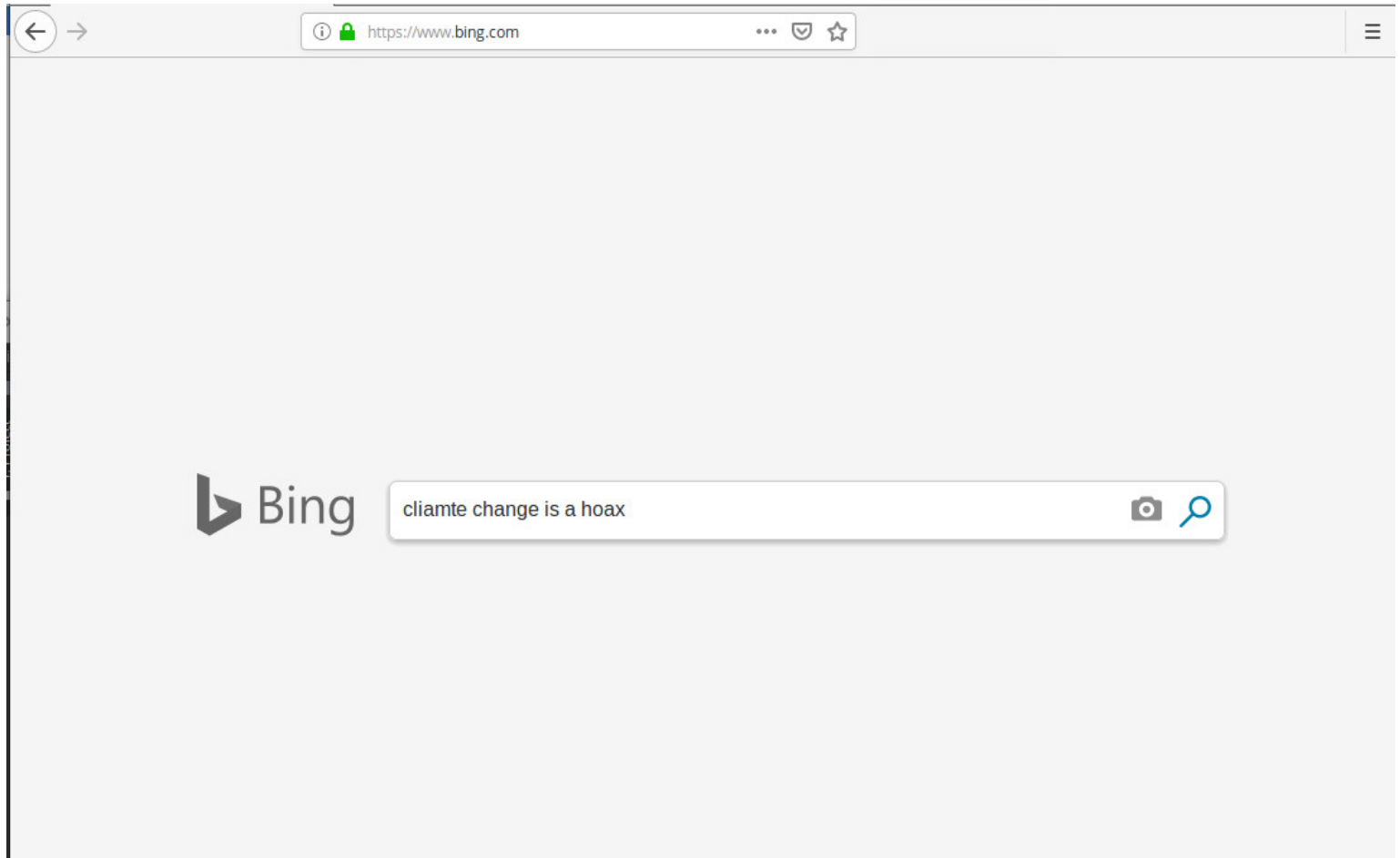
Limitations of conventational search engines



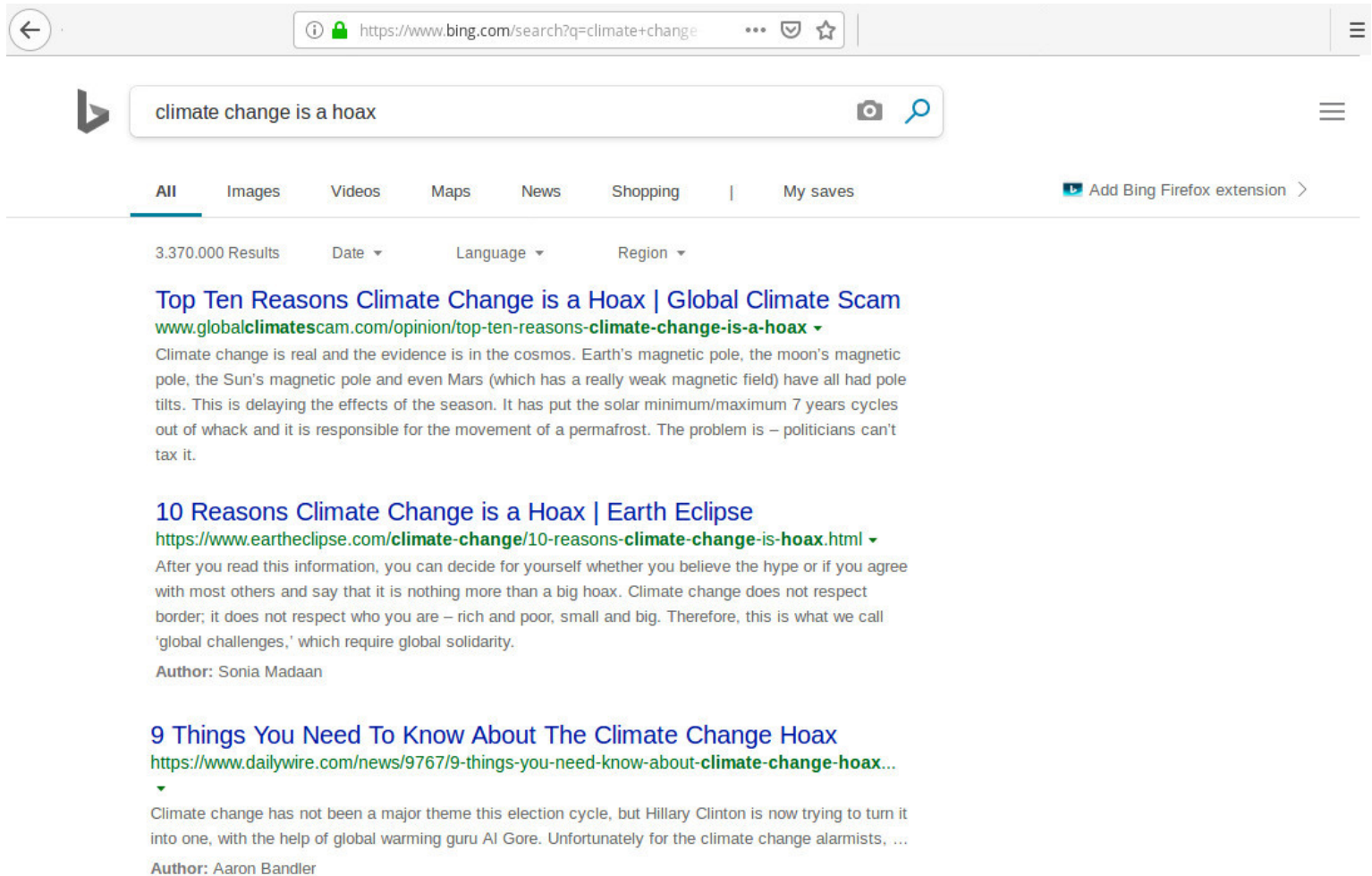
The screenshot shows a Bing search results page. The address bar at the top displays the URL <https://www.bing.com/search?q=climate+change>. The search bar contains the query "climate change is accelerating". Below the search bar, there are tabs for "All", "Images", "Videos", "Maps", "News", "Shopping", and "My saves". The "All" tab is selected. The results section shows "2,560,000 Results" and filters for "Date", "Language", and "Region". Three search results are visible:

- Climate change is accelerating, says another UN report ...**
<https://theprint.in/science/climate-change-is-accelerating-says-another-un-report...>
Climate change is accelerating, says another UN report. Here's why Here's why A new UN report says climate change is occurring at an increasing rate and its consequences are ...
Author: Pushp Bajaj
- How Quickly Climate Change Is Accelerating, in 167 Maps ...**
<https://www.wired.com/2016/08/quickly-climate-change-accelerating-167-maps>
When it comes to explaining climate change, it helps to have clear, convincing evidence on hand. Ed Hawkins, a climate scientist at the University of Reading with a knack for data-visualization ...
- Climate change: Global impacts 'accelerating' - WMO - BBC News**
<https://www.bbc.com/news/science-environment-47723577>
Climate change Image copyright Tomwang112 The World Meteorological Organization (WMO) says that the physical and financial impacts of global warming are accelerating.

Limitations of conventational search engines



Limitations of conventational search engines



The screenshot shows a web browser with the address bar displaying <https://www.bing.com/search?q=climate+change>. The search bar contains the text "climate change is a hoax". Below the search bar, there are tabs for "All", "Images", "Videos", "Maps", "News", "Shopping", and "My saves". The "All" tab is selected. The search results show 3,370,000 results. The first three results are listed below:

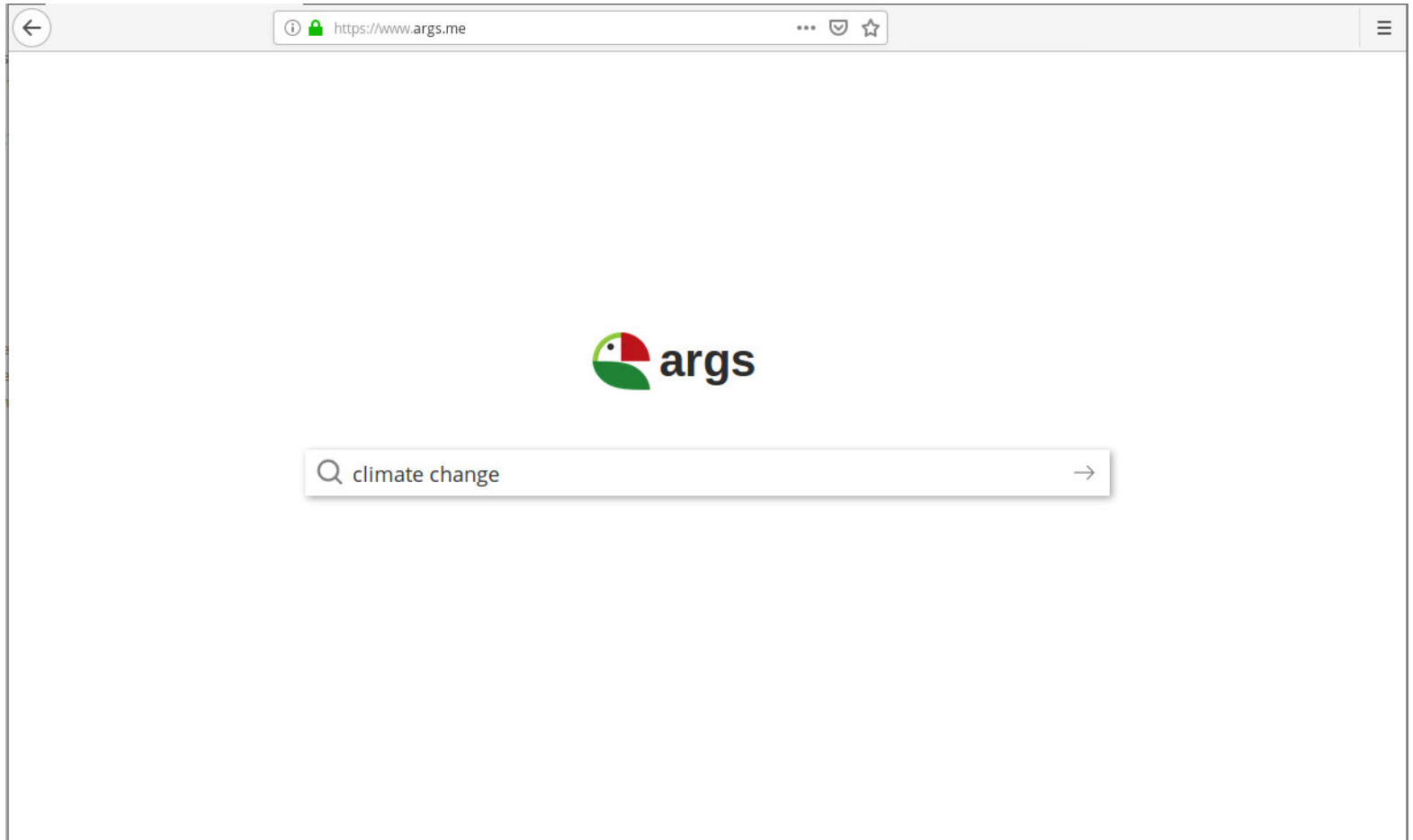
3.370.000 Results Date ▾ Language ▾ Region ▾

Top Ten Reasons Climate Change is a Hoax | Global Climate Scam
www.globalclimatescam.com/opinion/top-ten-reasons-climate-change-is-a-hoax ▾
Climate change is real and the evidence is in the cosmos. Earth's magnetic pole, the moon's magnetic pole, the Sun's magnetic pole and even Mars (which has a really weak magnetic field) have all had pole tilts. This is delaying the effects of the season. It has put the solar minimum/maximum 7 years cycles out of whack and it is responsible for the movement of a permafrost. The problem is – politicians can't tax it.

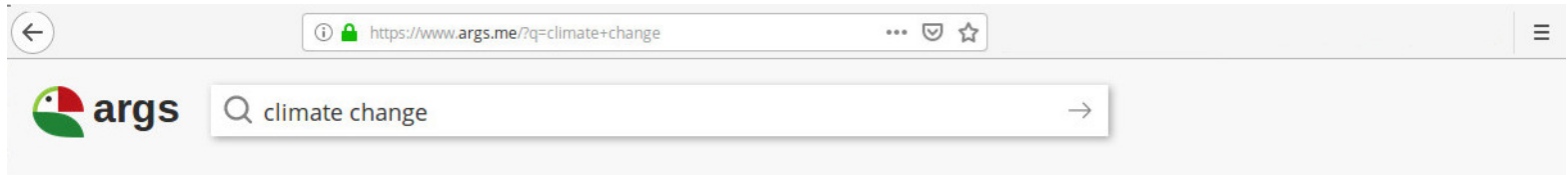
10 Reasons Climate Change is a Hoax | Earth Eclipse
<https://www.eartheclipse.com/climate-change/10-reasons-climate-change-is-hoax.html> ▾
After you read this information, you can decide for yourself whether you believe the hype or if you agree with most others and say that it is nothing more than a big hoax. Climate change does not respect border; it does not respect who you are – rich and poor, small and big. Therefore, this is what we call 'global challenges,' which require global solidarity.
Author: Sonia Madaan

9 Things You Need To Know About The Climate Change Hoax
<https://www.dailywire.com/news/9767/9-things-you-need-know-about-climate-change-hoax...> ▾
Climate change has not been a major theme this election cycle, but Hillary Clinton is now trying to turn it into one, with the help of global warming guru Al Gore. Unfortunately for the climate change alarmists, ...
Author: Aaron Bandler

Argument Search Engine



Argument Search Engine



Page 1 of 883 arguments, 393 pro, 490 con (retrieved in 1236.5ms)

[Pro vs. Con View](#) | [Topic Space View](#)

Pro

My position on climate change is that not only do...

<http://www.debate.org/debates/Climate-Change-is-man-caused/1/>

My position on **climate change** is that not only do I believe in it but I believe it is man caused. I ... that disagrees with me is ignorant on this subject. 97% of scientists believe that ... ▼ [score](#)

There is an overwhelming scientific consensus...

<http://www.debate.org/debates/Jail-climate-change-deniers./1/>

There is an overwhelming scientific consensus **climate change** presents a serious risk to the future of the Earth and to mankind. Many of the threats we face due to **climate** ... ▼ [score](#)

Climate Change is causing the Earth to warm up...

<http://www.debate.org/debates/Climate-Change-is-happening/1/>

Climate Change is causing the Earth to warm up measurably, and there are already signs of disaster ... reviewed scientific papers on this subject, 0% disagreed that **climate** ... ▼ [score](#)

Con

Climate change is quite complicated and requires...

<http://www.debate.org/debates/Climate-Change-is-man-caused/1/>

Climate change is quite complicated and requires more than two minds. **Climate change** is happening ... enough research done on the natural causes of **climate** ... ▼ [score](#)

The effects of climate change will not necessarily...

<http://debatewise.org/debates/200-climate-change-is-the-end-of-the-wor...>

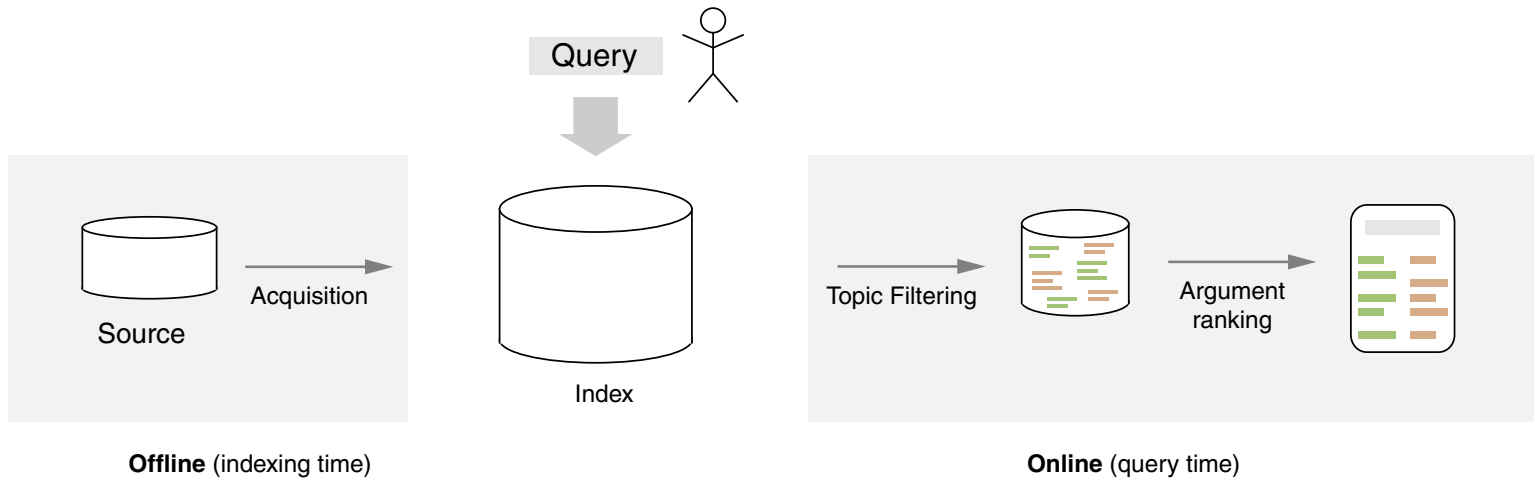
The effects of **climate change** will not necessarily be bad ... ▼ [score](#)

I'm against jailing Climate change deniers. It is...

<http://www.debate.org/debates/Jail-climate-change-deniers./1/>

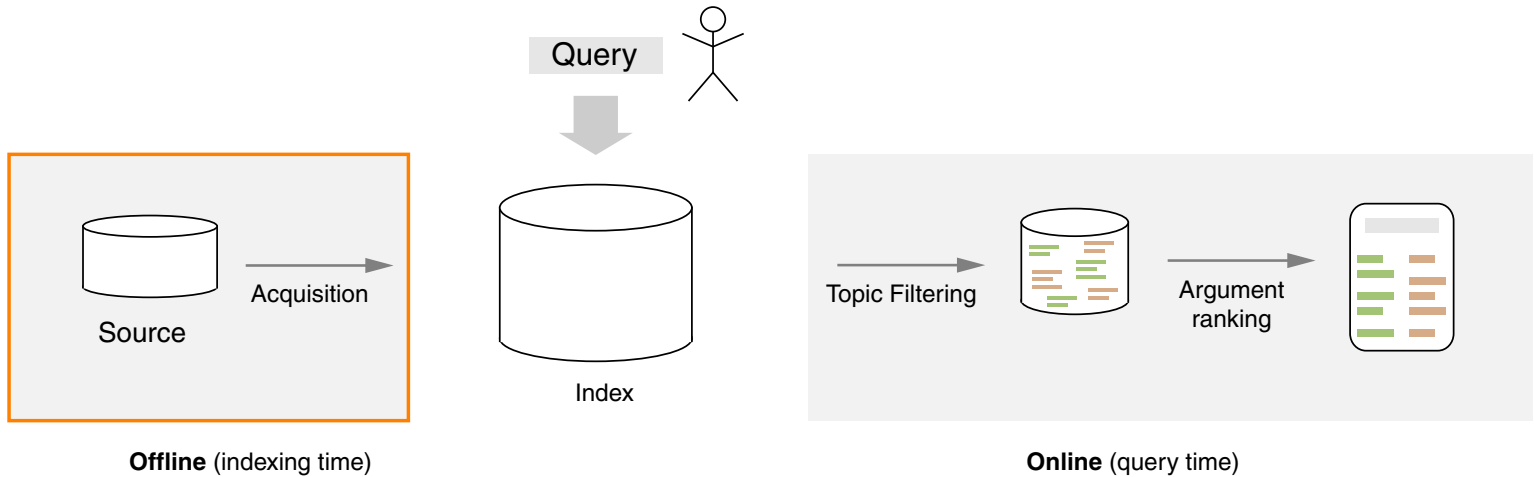
I'm against jailing **Climate change** deniers. It is a violation of free speech. I also think it sets ... Nevertheless some people think we should jail **climate change** deniers. ... ▼ [score](#)

Design of an Argument Search Engine

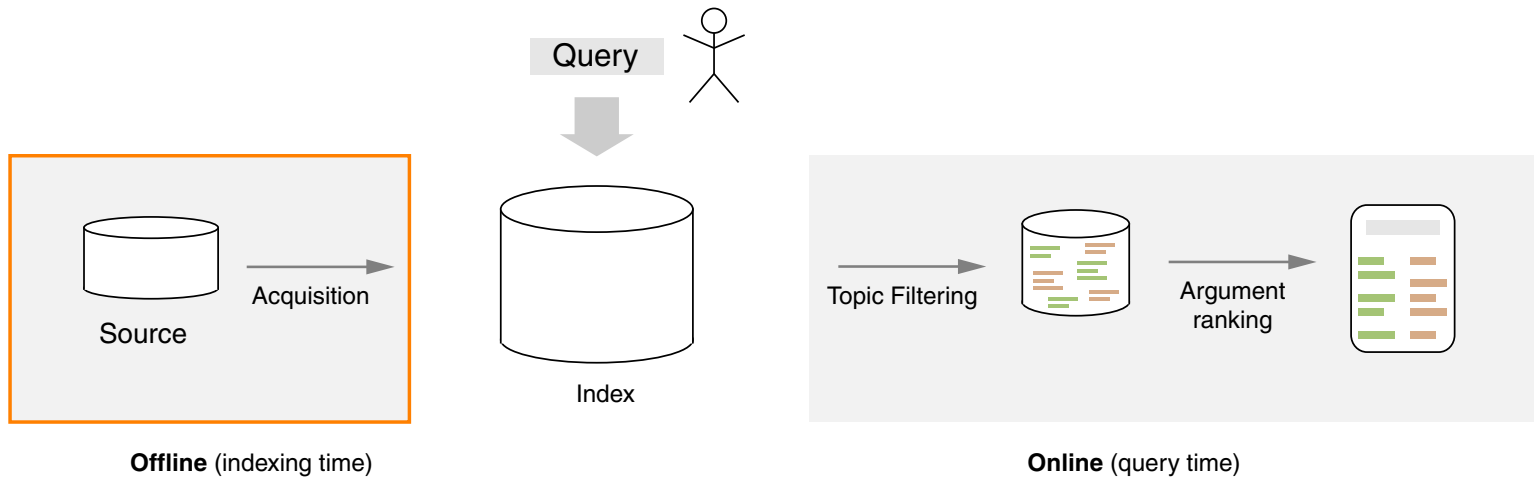


Argument Acquisition

We focus on the first step in designing an argument search engine: argument acquisition.



Argument Acquisition

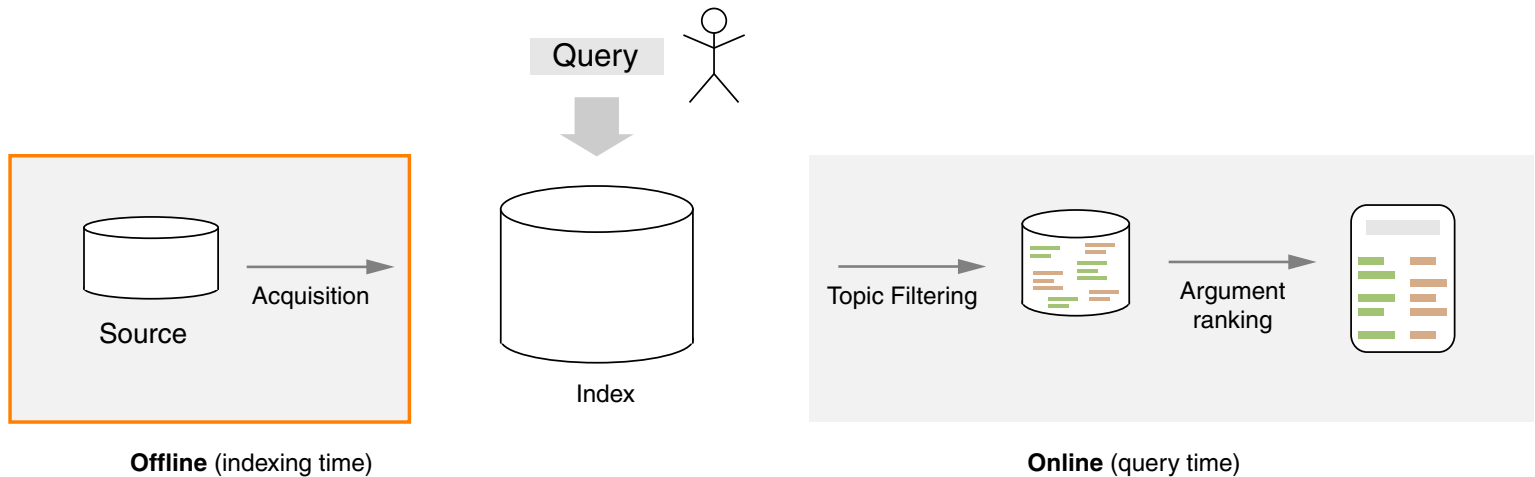


Argument Acquisition is a real challenge because:

- ❑ Annotating arguments is very expensive.
- ❑ Mining arguments from the Web is a hard task.

Approach	F1-score
Baseline (Linear Regression)	0.6
Modiefied BiLSTM (Stab et al., 2018)	0.66
Baseline (Majority)	0.6
Distant-supervision (Al-khatib., 2016)	0.7

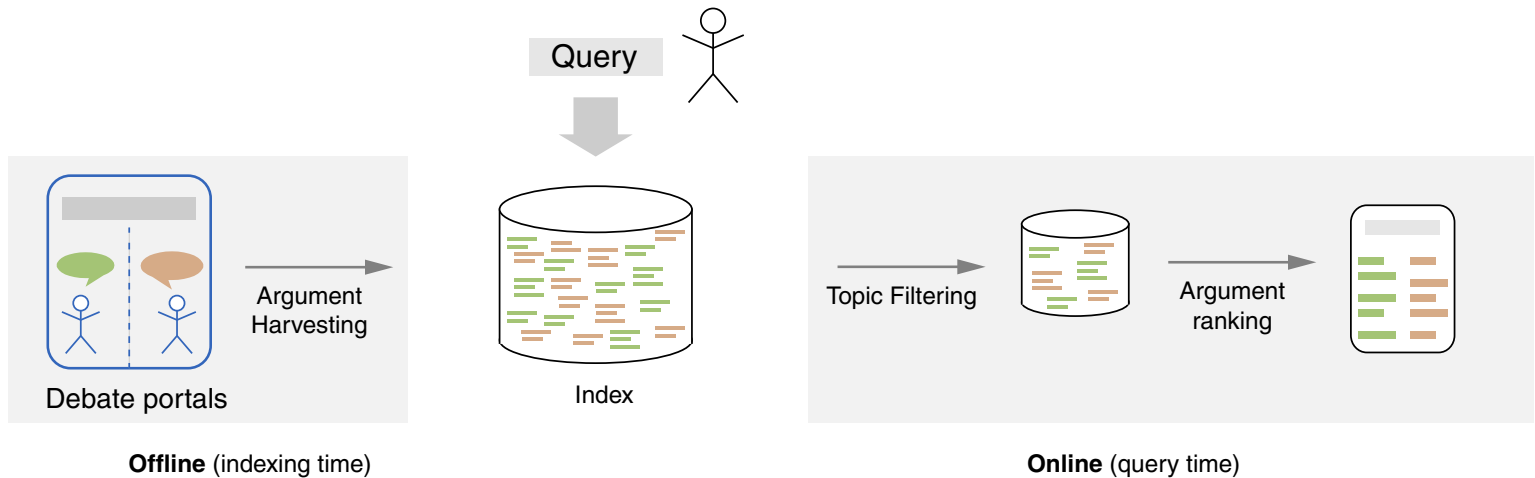
Argument Acquisition Paradigms: Consequences



Consequences of choosing the source and approach to extract arguments from:

- ❑ **Precision vs Recall:** Trade off between the quality of the arguments and the coverage of the retrieved arguments.
- ❑ **Stance balance:** how balanced are the stances of the presented arguments?
- ❑ **Efficiency:** How fast is the argument search engine at retrieving arguments?
- ❑ **Research Focus:** What are the research challenges that the paradigm allows to study?

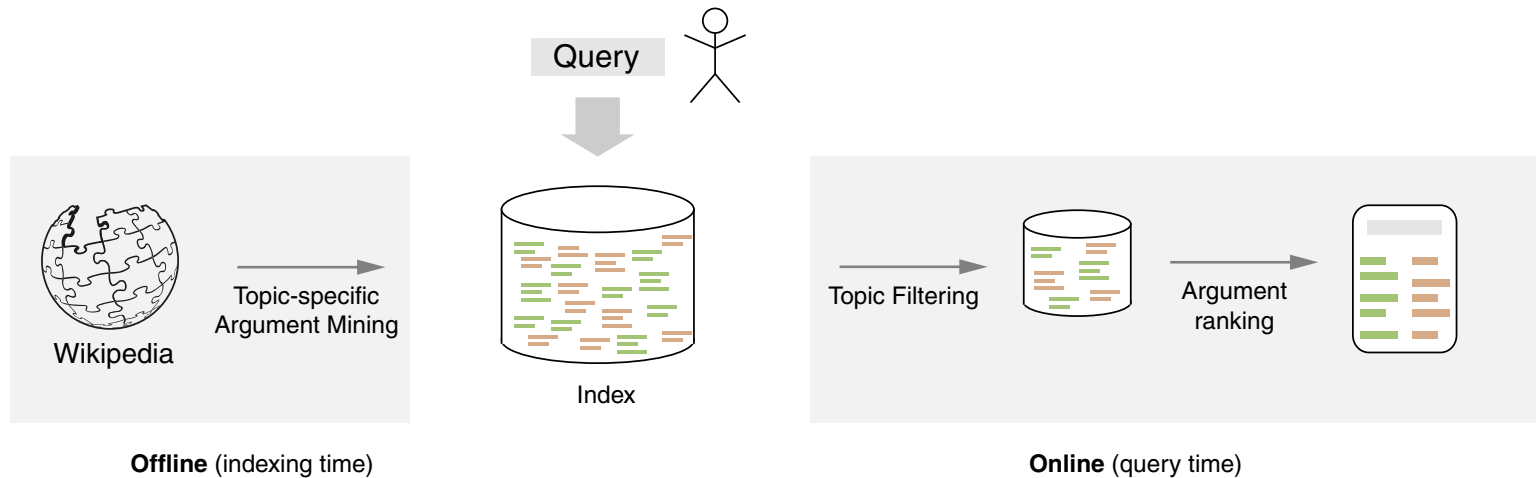
Argument Acquisition Paradigm of args.me



Consequences:

- ❑ High precision, low recall.
- ❑ Stance balance is guaranteed.
- ❑ Very efficient.
- ❑ Research focus on all computational argumentation tasks, e.g, argument generation.

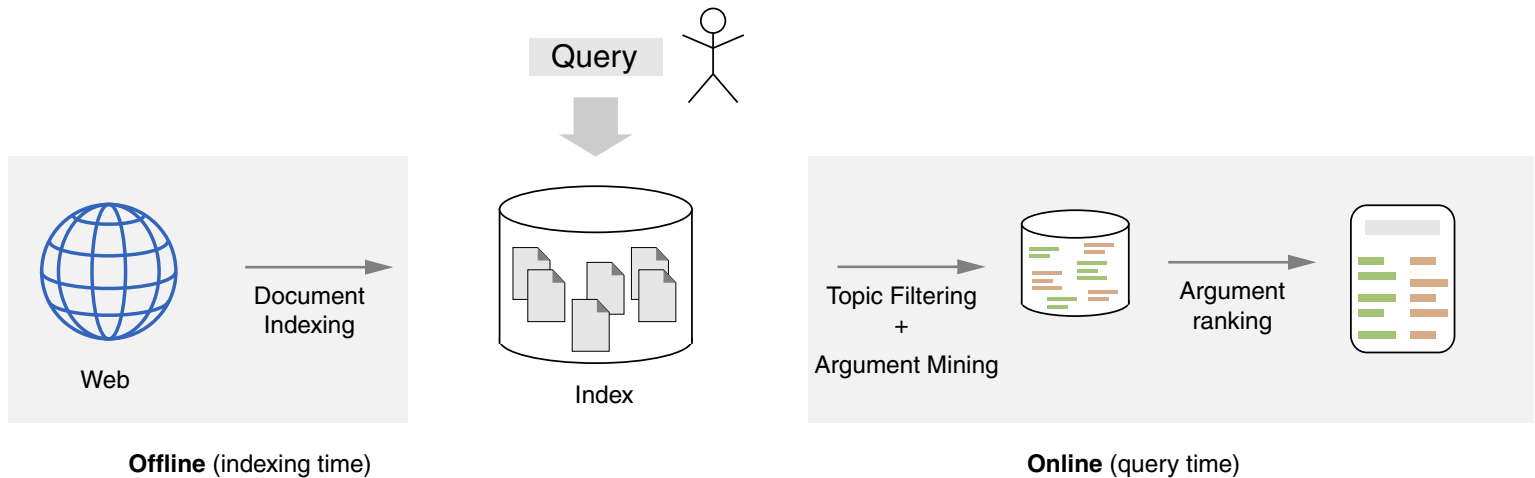
Argument Acquisition Paradigm of IBM Debater



Consequences:

- ❑ High precision, low recall.
- ❑ Stance balance is guaranteed.
- ❑ Very efficient.
- ❑ Research focus on all computational argumentation tasks, e.g, argument ranking.

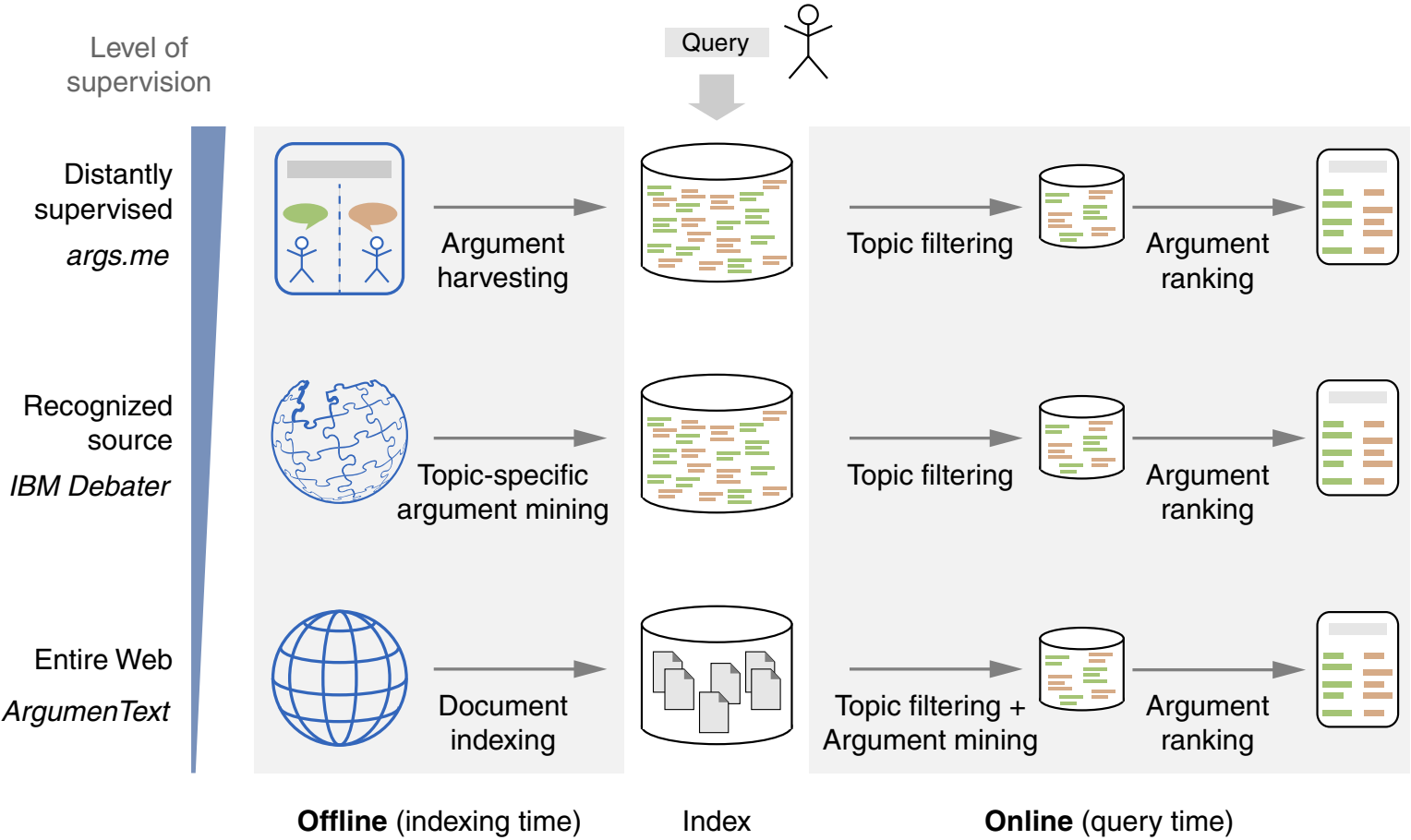
Argument Acquisition Paradigm of ArgumentText



Consequences:

- ❑ Low precision, high recall.
- ❑ Stance balance is not guaranteed.
- ❑ Slow since argument mining is performed online.
- ❑ Research focus on argument mining.

Argument Acquisition Paradigms

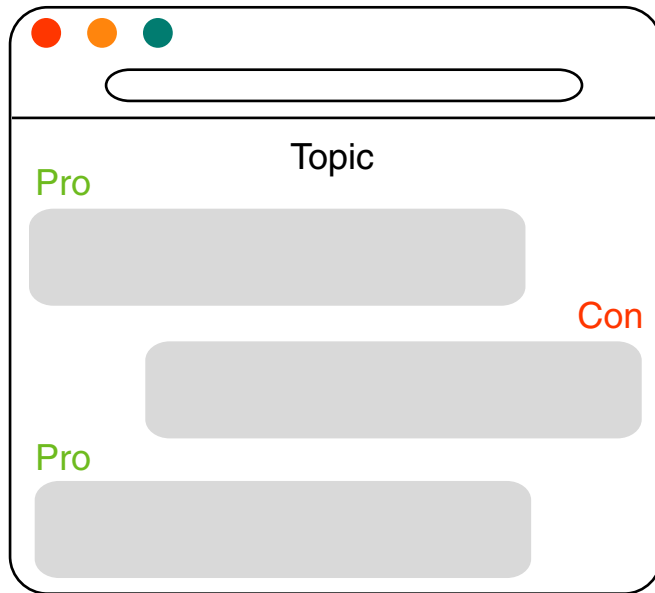


Low Hanging Fruits in Debateportals

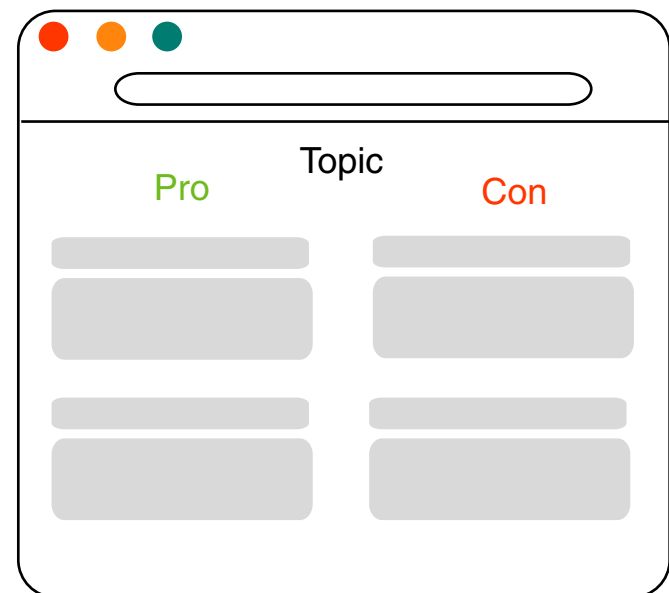
Debate portals are websites where people discuss or list arguments about controversial topic, e.g, climate change.

There are two types of Debateportals:

- ❑ Dialogical: people discuss topics in rounds.
- ❑ Monological: people or admin list arguments.



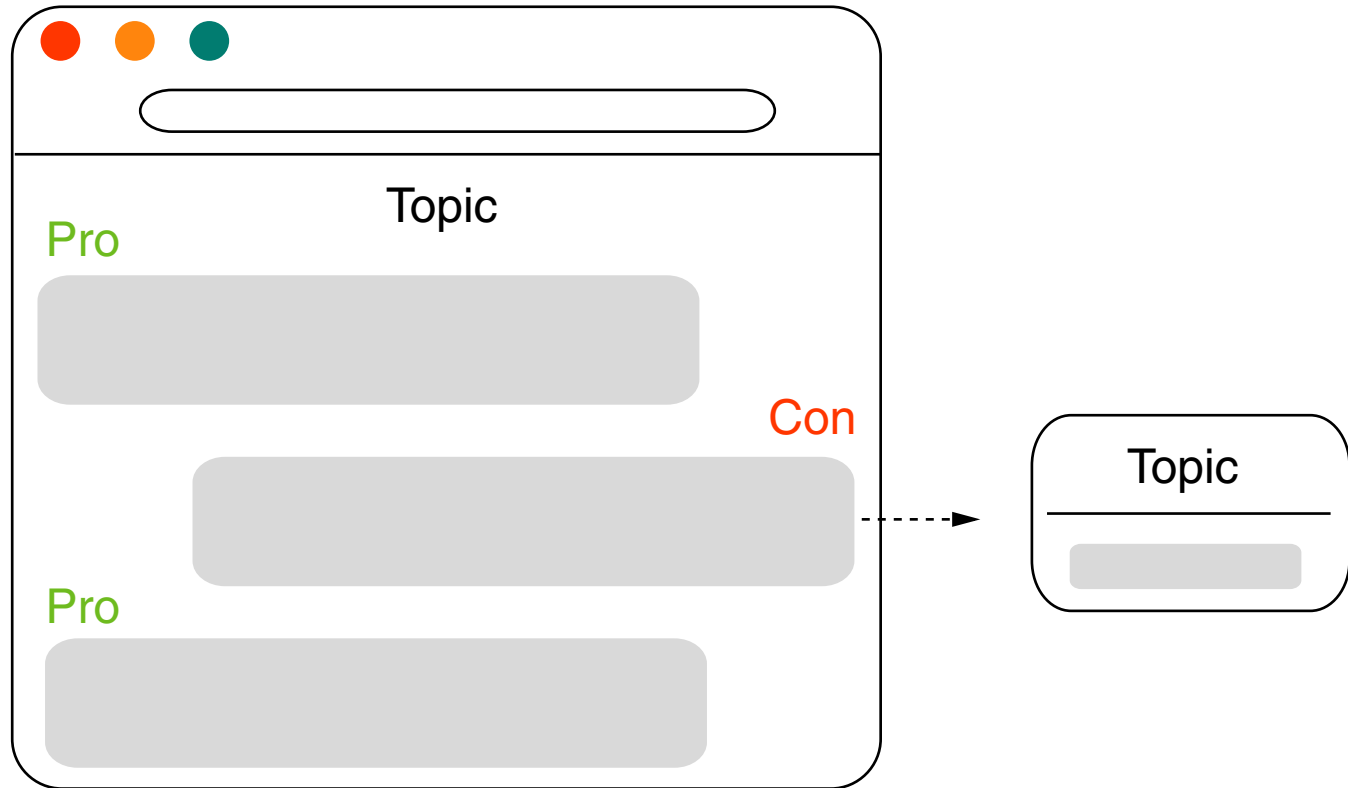
Dialogical Debate portals



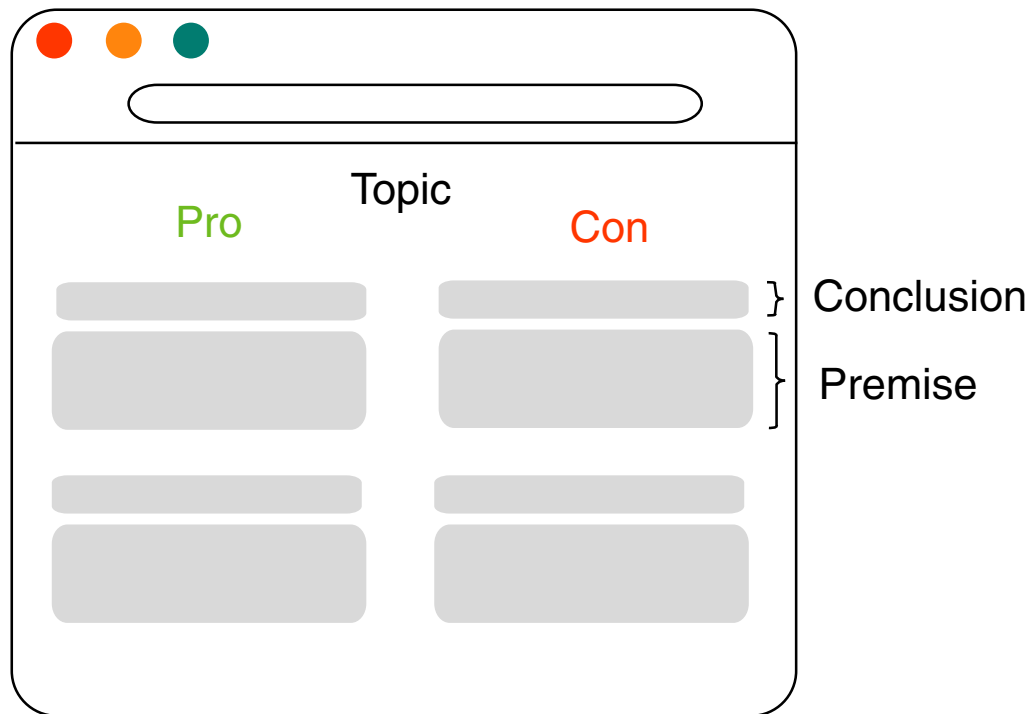
Monological Debate portals

Extraction Heuristics for Dialogical Debateportals

We extract for each post an argument whose conclusion is the topic.



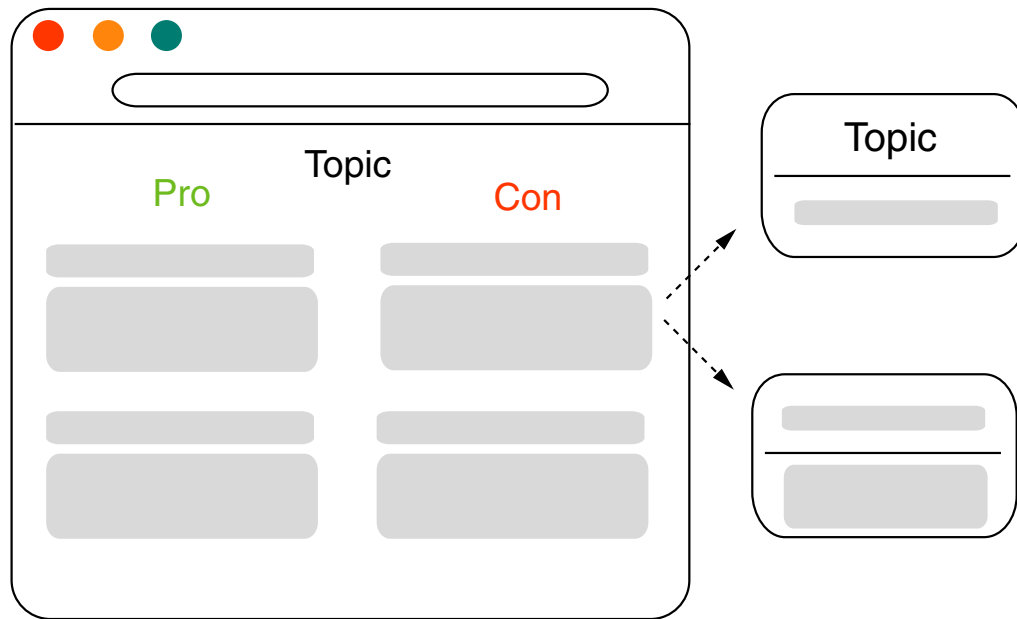
Extraction Heuristics for Monological Debateportals



Extraction Heuristics for Monological Debateportals

We extract for each post two arguments:

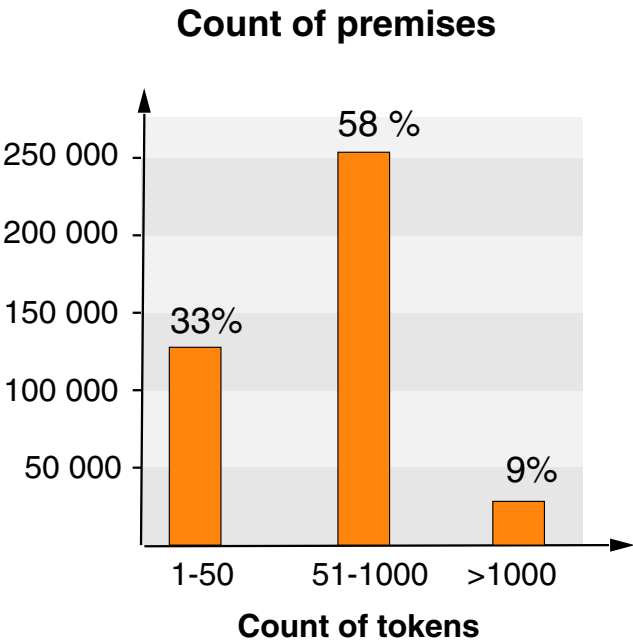
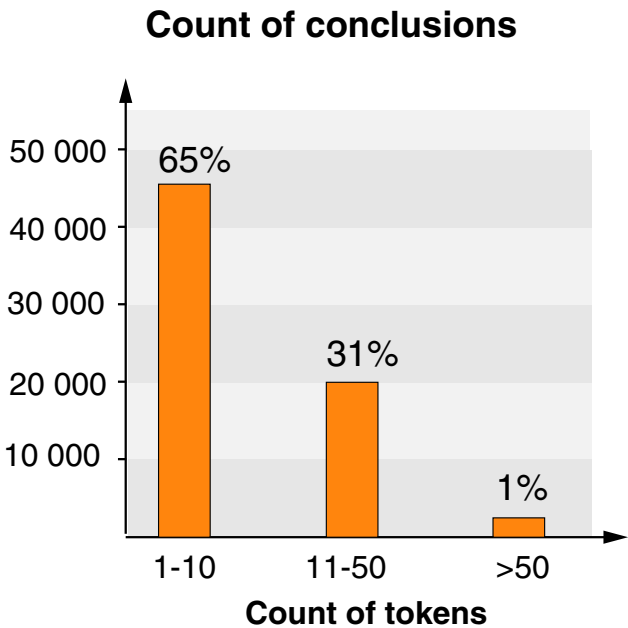
- ❑ The conclusion and premise of the post.
- ❑ An argument whose conclusion is the topic and its premise is the conclusion of the post.



Args.me corpus statistics

Args.me corpus is the largest argument corpus with 387,606 arguments.

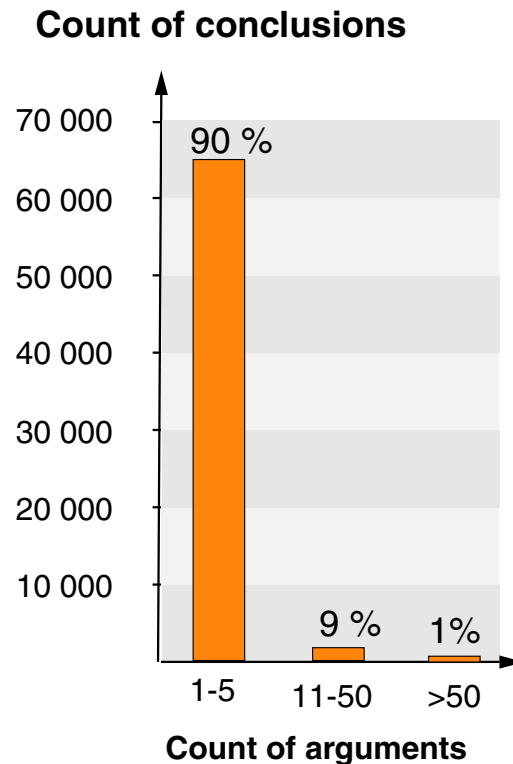
Count of arguments	Count of pro stance	Count of con stance	Count of debates
387 606	200 099	187 507	59 637



Args.me corpus statistics

Most of the conclusions in args.me corpus have 1-5 supporting or attacking arguments.

Some conclusions have more than 50 supporting or attacking arguments.



Argument search tasks

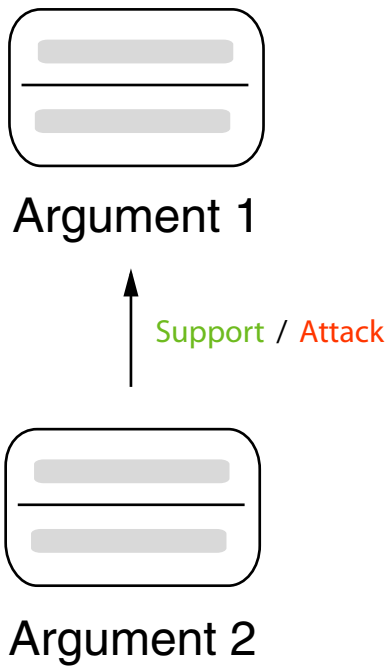
List of tasks that can be performed on the args.me corpus:

- ❑ Argument relation classification
- ❑ Stance classification
- ❑ Same side classification
- ❑ Argument generation
- ❑ ..

Argument Relation Classification

Task Description: Given two arguments detect if one attacks or support the other.

Application: Retrieval of counter arguments for an given argument.

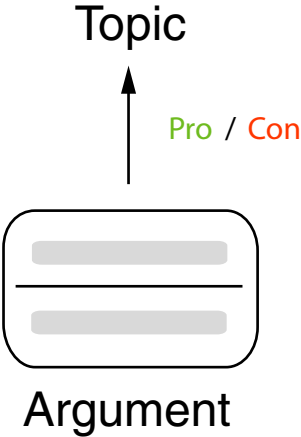


Approach	Accuracy
Baseline (Linear Regression)	0.77
LSTM (Cocarascu et al., 2017)	0.89

Stance Classification

Task Description: Given a topic and an argument, detect whether the argument is pro or con the topic.

Application: Classifying the retrieved arguments for a query into pro or contra.

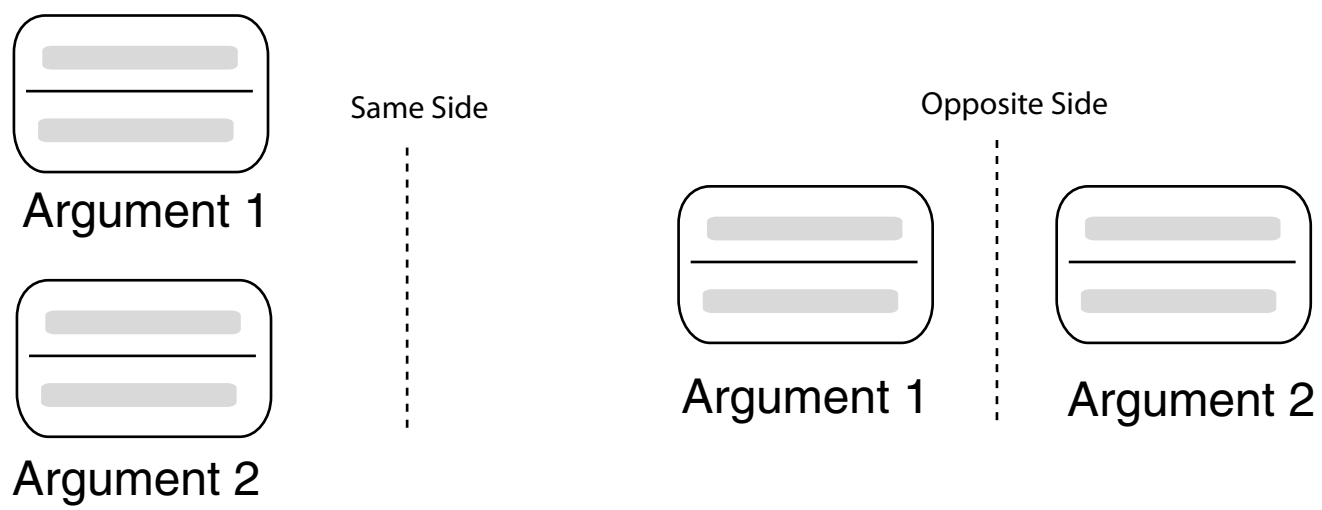


Approach	Accuracy
Baseline (Majority)	0.51
Sentiment-based (Bar-Haim et al., 2017)	0.65

Same Side Classification

Task Description: Given two arguments, detect whether they are on the same or opposite side.

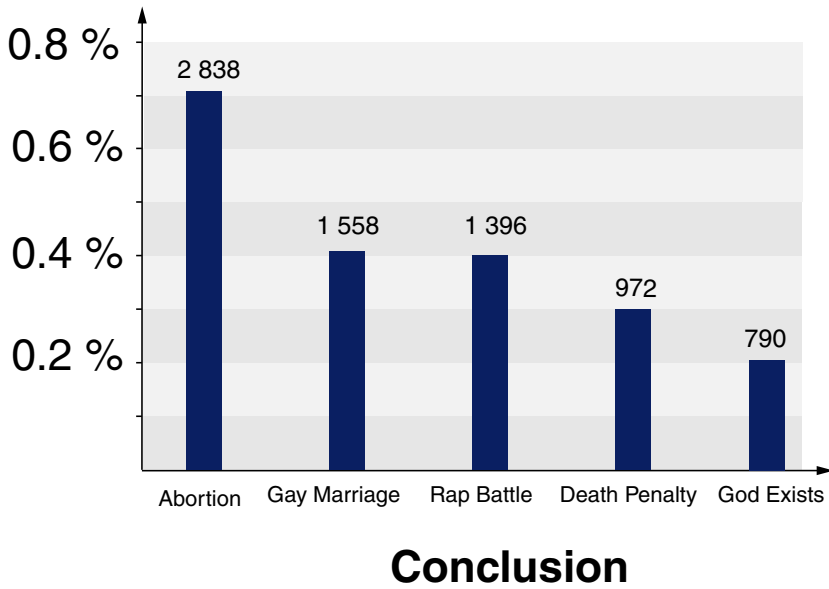
Application: Grouping the retrieved arguments for a query under different claims.



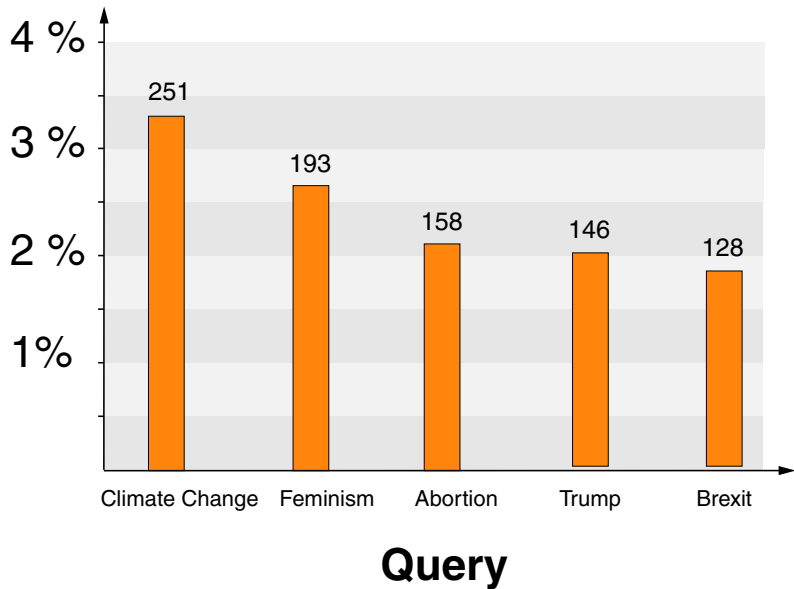
Approach	Accuracy
Majority	0.50
BERT	0.73

First insights from args.me

Count in args.me corpus



Count in args.me logs



Top five conclusions in args.me corpus

Top five queries in args.me query logs

Summary

- ❑ Argument acquisition paradigm: a set of decisions made while designing an argument search engine.
- ❑ Comparison of three major argument search engines under the argument acquisition paradigm.
- ❑ Introduction of the largest argument corpus args.me corpus.
- ❑ Introduction of several tasks that can be performed on args.me corpus.