

Identifying Argumentative Questions in Web Search Logs



Yamen Ajjour
Leipzig



Pavel Braslavski
Yekaterinburg



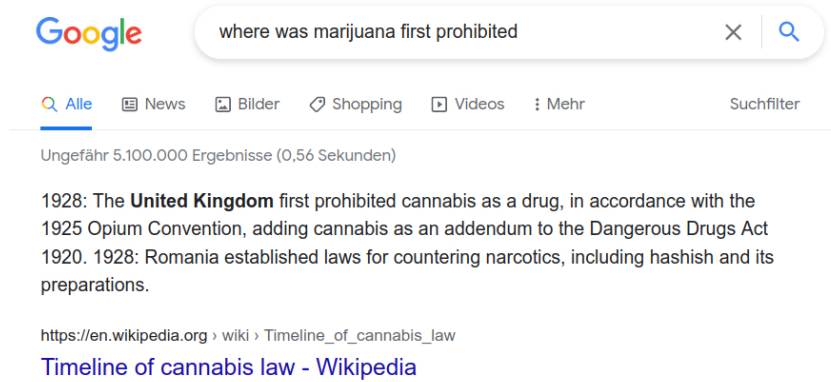
Alexander Bondarenko
Halle



Benno Stein
Weimar

Motivation: Search Engines

Search engines are good at answering factual questions.




The image shows a Google search interface. The search bar contains the text "where was marijuana first prohibited". Below the search bar, there are navigation options: "Alle", "News", "Bilder", "Shopping", "Videos", "Mehr", and "Suchfilter". The search results show "Ungefähr 5.100.000 Ergebnisse (0,56 Sekunden)". The first result is a snippet from Wikipedia: "1928: The **United Kingdom** first prohibited cannabis as a drug, in accordance with the 1925 Opium Convention, adding cannabis as an addendum to the Dangerous Drugs Act 1920. 1928: Romania established laws for countering narcotics, including hashish and its preparations." Below the snippet is the URL "https://en.wikipedia.org › wiki › Timeline_of_cannabis_law" and a blue link "Timeline of cannabis law - Wikipedia".

Motivation: Search Engines


Search engines struggle at delivering all perspectives on argumentative questions.

The image shows a Google search interface with the query "should marijuana stay illegal". The search results are as follows:

- Result 1:** <https://www.drugrehab.us> › news › Diese Seite übersetzen
Why Recreational Marijuana Should Stay Illegal - Drug Rehab
Many people think marijuana isn't harmful enough to be illegal, but this is dangerous - See the reasons why recreational **marijuana should stay illegal** now!
pro
- Result 2:** <https://www.liberty.edu> › 2022/03 › Diese Seite übersetzen
Cons of Legalizing Recreational Marijuana Use - Liberty ...
07.03.2022 — Despite the growing acceptance, decriminalizing **marijuana should be** a federal issue, and it **should** not be legal to use recreationally. Although ...
pro
- Result 3:** <https://www.youtube.com> › watch
3 Arguments Why Marijuana Should Stay Illegal Reviewed

3 Arguments Why **Marijuana Should Stay Illegal** Reviewed ...
We take a fair look at some of the best counter arguments for...
YouTube · Kurzgesagt – In a Nutshell · 10.06.2018
3 wichtige Momente in diesem Video
pro
- Result 4:** <https://www.pewresearch.org> › In-... › Diese Seite übersetzen
Why Americans Support or Oppose Legalizing Marijuana
14.04.2015 — A small share of opponents (7%) say that while the recreational use of **marijuana should be illegal**, they do not object to legalizing medical ...
Survey Report · Current Opinion on Legalizing... · Other Opinions: Federal...
neutral
- Result 5:** <https://www.ncbi.nlm.nih.gov> › pmc · Diese Seite übersetzen
More Reasons States Should Not Legalize Marijuana - NCBI
von ST Wilkinson · 2013 · Zitiert von: 14 — Medical **marijuana should be** subject to the same rigorous approval process as ... A substantial and growing literature documents **legalized...**
Recreational Marijuana · Myth: Marijuana is Not Addictive · Effects on Cognition
pro

Motivation: Argument Retrieval System

Argument retrieval systems retrieve pro and con arguments for a query.

Q

All Discussions News People

Pro vs. con view 1513 arguments retrieved in 1.0 ms

PRO

[This debate will be about whether weed should be...](#)

► Show full argument

This debate will be about whether weed **should** be legalized or **remain illegal**. ... For round 1, when accepting my challenge, introduce yourself, but the arguing will not begin until the second round.

<https://www.debate.org/debates/Marijuana-should-remain-illegal/1/> score ▼

[Well, there isn't anything that especially demonizes...](#)

► Show full argument

Well, there isn't anything that especially demonizes **marijuana**, but what lifting the ban of **marijuana** would do is 1: Increase use, whether it be recreational or medicinal, leaving millions susceptible to falling prey to ...

<https://www.debate.org/debates/Marijuana-should-remain-illegal/1/> score ▼

[This "gift of nature" is a cause of several thousand...](#)

► Show full argument

This "gift of nature" is a cause of several thousand employment terminations, incarcerations, relationship breakups, car accidents, and more. ... The fact that it "grows naturally" doesn't excuse the fact it is an abused drug, and ...

<https://www.debate.org/debates/Marijuana-should-remain-illegal/1/> score ▼

[In my state of CA you will find more dispensaries than...](#)

► Show full argument

In my state of CA you will find more dispensaries than star bucks.. however I think everyone **should** have this choice, and that it shouldn't be only ok for the terminally ill. ... (con) will argue **marijuana should remain** and/or be ...

<https://www.debate.org/debates/Marijuana-should-be-legalized/13/> score ▼

[definitions: marijuana noun 1/2 1/2 Drug smoked or ingested...](#)

► Show full argument

definitions: **marijuana** noun 1/2 1/2 Drug smoked or ingested for euphoric effect, '... definitions: <http://www.ninjawords.com>

<https://www.debate.org/debates/Marijuana-Should-be-Legalized-in-the-United-States-of-America/1/> score ▼

[First off let me say that I believe drugs should be...](#)

► Show full argument

First off let me say that I believe drugs **should** be legalized and regulated but I don't think there is justification for alcohol to **remain** legal while other drugs are **illegal**. This debate **should** be accepted by someone that ...

<https://www.debate.org/debates/As-long-as-drugs-such-as-marijuana-are-illegal-alcohol-should-be-as-well/1/> score ▼

CON

[I am for the legalization of recreational and medical...](#)

► Show full argument

I am for the legalization of recreational and medical uses of **Marijuana**. ... **Marijuana-Cannabis illegal**- In this context, to have the sale, possession, distribution and growth **illegal to be punished by the law**

<https://www.debate.org/debates/Marijuana-should-remain-illegal/1/> score ▼

[At 18 we also can make the choice to enroll in the...](#)

► Show full argument

At 18 we also can make the choice to enroll in the dangerous life of a soldier. ... Yes, but the state **should** not baby it's citizens through making choices in their life. ...

<https://www.debate.org/debates/Marijuana-should-remain-illegal/1/> score ▼

[I don't deny this, however what my opponent does not...](#)

► Show full argument

I don't deny this, however what my opponent does not address in round 2 is if these bans work or not. ... "This "gift of nature" is a cause of several thousand employment terminations, incarcerations, relationship breakups, car ...

<https://www.debate.org/debates/Marijuana-should-remain-illegal/1/> score ▼

[Marijuana is used almost like alcohol, but has been used...](#)

► Show full argument

Marijuana is used almost like alcohol, but has been used in a more abusive way. ... just because liquor is legal doesn't mean we **should** use and abuse it extensively, like we have with **Marijuana**. ...

<https://www.debate.org/debates/Marijuana-should-be-legal/9/> score ▼

[With the legalisation of marijuana, there would be more...](#)

► Show full argument

With the legalisation of **marijuana**, there would be more **marijuana** users, and thus an increase in healthcare costs. ... British Journal of Psychiatry, 178: 574-5 <http://learnaboutsam.org..> <http://www.usnews.com...> ...

<https://www.debate.org/debates/Should-marijuana-be-legal-or-illegal/1/> score ▼

[Thus it is obvious that preserving and promoting public...](#)

► Show full argument

Thus it is obvious that preserving and promoting public health is a duty of the government, and legalising **marijuana** goes against it. ... References: <http://education-portal.com...> <http://www.thefreedictionary.com...>

<https://www.debate.org/debates/Should-marijuana-be-legal-or-illegal/1/> score ▼

Research Question

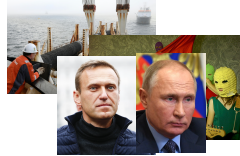
How to identify argumentative questions from other question types in the query stream of a search engine?

Preparing a Russian Questions Dataset

1. Filter from Yandex logs 4.5 million Russian questions on 19 controversial topic.

Local Issues

Putin
Navalny
Pussy Riots
Nord stream
...

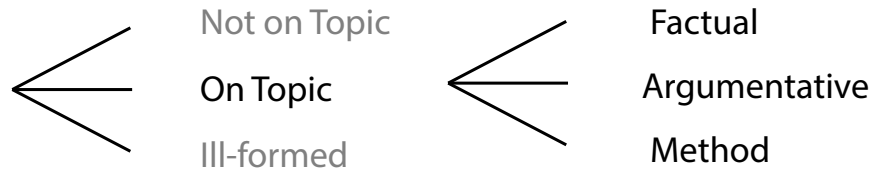


Global Issues

Abortion
Death penalty
Gay marriage
...



2. Sample 54,850 questions and annotate them with the annotation scheme:



1. Is Context Controversial?

2. Which Question Type?

Preparing a Russian Questions Dataset

1. Filter from Yandex logs 4.5 million Russian questions on 19 controversial topic.

Local Issues

- Putin
- Navalny
- Pussy Riots
- Nord stream
- ...

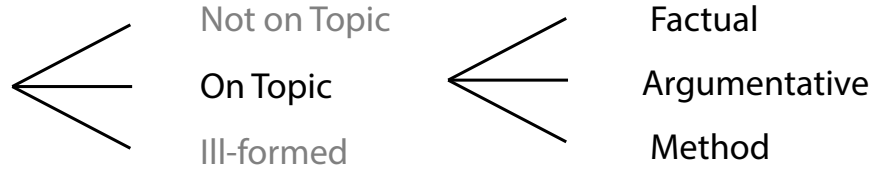


Global Issues

- Abortion
- Death penalty
- Gay marriage
- ...



2. Sample 54,850 questions and annotate them with the annotation scheme:



1. Is Context Controversial?
2. Which Question Type?

Statistics and Examples:

| Question Type | Perc | Count | Example |
|---------------|------|--------|--|
| Factual | 64% | 25,332 | Is marujana legalized in Belgium? |
| Argumentative | 28% | 10,982 | Will the president legailze marijuana? |
| Method | 8% | 3,026 | How to use medical marijuana? |

Analysis of Argumentative Questions

Comparison of argumentative questions with factual and method questions using lexical and syntactical patterns.

| Characteristic | Factual | Argumentative | Method |
|-----------------------|----------------|----------------------|---------------|
| Yes/No | 7.2% | 13.8% | 0.4% |
| Predictions | 3.8% | 8.2% | 0.6% |
| Comparisons | 3.2% | 5.7% | 4.4% |
| Personal Pronouns | 0.3% | 3.8% | 0.4% |
| Wh-words | 67.0% | 62.0% | 94.3% |
| <i>why</i> | 1.3% | 20.7% | 0.0% |

Analysis of Argumentative Questions

Comparison of argumentative questions with factual and method questions using lexical and syntactical patterns.

| Characteristic | Factual | Argumentative | Method |
|-----------------------|----------------|----------------------|---------------|
| Yes/No | 7.2% | 13.8% | 0.4% |
| Predictions | 3.8% | 8.2% | 0.6% |
| Comparisons | 3.2% | 5.7% | 4.4% |
| Personal Pronouns | 0.3% | 3.8% | 0.4% |
| Wh-words | 67.0% | 62.0% | 94.3% |
| <i>why</i> | 1.3% | 20.7% | 0.0% |

Finding: Argumentative questions look for predictions and explicitly for reasons.

Question Type Classification Experiment

Developing classifiers to map questions to argumentative, factual or method.

Experimental setting is leave-one-topic-out: test on one topic after training on remaining topics.

F1-score of the three question types and their macro average.

| Classifier | Factual | Argument. | Method | Macro |
|---------------------|----------------|------------------|---------------|--------------|
| Majority Baseline | 0.78 | 0.00 | 0.00 | 0.26 |
| Logistic Regression | 0.80 | 0.61 | 0.52 | 0.65 |
| RuBERT | 0.85 | 0.74 | 0.74 | 0.78 |

Question Type Classification Experiment

Developing classifiers to map questions to argumentative, factual or method.

Experimental setting is leave-one-topic-out: test on one topic after training on remaining topics.

F1-score of the three question types and their macro average.

| Classifier | Factual | Argument. | Method | Macro |
|---------------------|----------------|------------------|---------------|--------------|
| Majority Baseline | 0.78 | 0.00 | 0.00 | 0.26 |
| Logistic Regression | 0.80 | 0.61 | 0.52 | 0.65 |
| RuBERT | 0.85 | 0.74 | 0.74 | 0.78 |

Finding: Identifying argumentative questions is feasible, even on unseen topics.

Summary

Contributions:

- ❑ Effective approach to label and identify argumentative questions in web search logs
- ❑ Exemplary dataset of 100 questions
- ❑ Analysis of the form and structure of argumentative questions

Findings:

- ❑ High prevalence of argumentative questions on controversial topics in web search logs
- ❑ Argumentative questions look mainly for reasons and predictions