

RATIO 2024

OASiS: Objective Argument Summarization in Search

Timon Ziegenbein, Shahbaz Syed, Martin Potthast, Henning Wachsmuth

07.06.2024

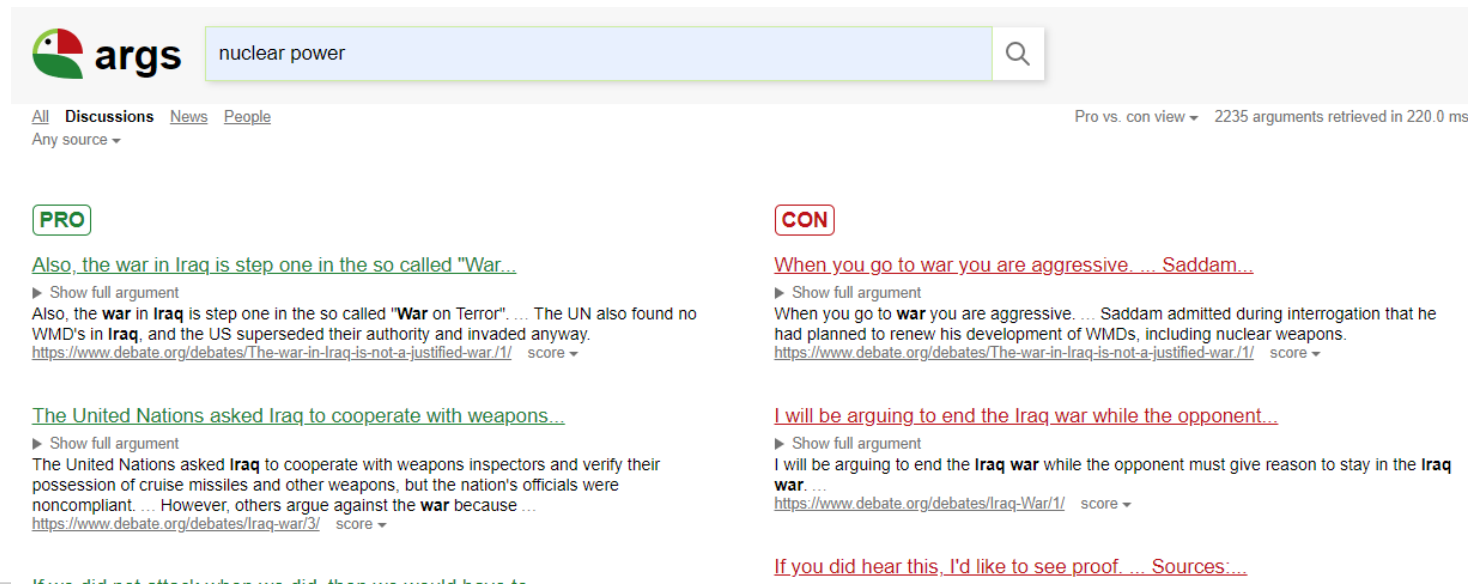


UNIVERSITÄT
LEIPZIG

U N I K A S S E L
V E R S I T Ä T

OASiS: Objective Argument Summarization in Search

- Rendition of arguments may affect user beliefs, and cognitive biases (Wachsmuth et al., 2017b; Kiesel et al., 2020, 2021)
 - Users expect good quality arguments, sensibly organized



The screenshot shows the 'args' search interface. At the top, there is a search bar with the text 'nuclear power' and a magnifying glass icon. Below the search bar, there are navigation links: 'All', 'Discussions', 'News', and 'People'. To the right of these links, it says 'Pro vs. con view' and '2235 arguments retrieved in 220.0 ms'. Below the navigation links, there are two columns of search results. The left column is labeled 'PRO' and contains two results. The right column is labeled 'CON' and contains two results. Each result includes a title, a 'Show full argument' link, a short summary, and a URL with a 'score' dropdown.

args nuclear power

All Discussions News People Pro vs. con view 2235 arguments retrieved in 220.0 ms
Any source ▾

PRO

[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 The UN also found no WMD's in **Iraq**, and the US superseded their authority and invaded anyway.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 ▾

CON

[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:...](#)

OASiS: Objective Argument Summarization in Search

- Rendition of arguments may affect user beliefs, and cognitive biases (Wachsmuth et al., 2017b; Kiesel et al., 2020, 2021)
 - Users expect good quality arguments, sensibly organized
- Personal biases in the arguments can hamper unbiased decisions (Durmus and Cardie, 2018)
 - Most arguments on the web are exchanged outside “professional circles”

The screenshot shows the 'argos' search engine interface. At the top, there is a search bar with the text 'nuclear power' and a magnifying glass icon. Below the search bar, there are navigation links: 'All', 'Discussions', 'News', and 'People'. To the right, it says 'Pro vs. con view' and '2235 arguments retrieved in 220.0 ms'. Below this, there are two columns of results. The left column is labeled 'PRO' and contains two entries. The right column is labeled 'CON' and contains two entries. Each entry includes a title, a 'Show full argument' link, a short summary, and a URL with a 'score' dropdown.

PRO

[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 ▼

CON

[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:...](#)

OASiS: Objective Argument Summarization in Search

- Rendition of arguments may affect user beliefs, and cognitive biases (Wachsmuth et al., 2017b; Kiesel et al., 2020, 2021)
 - Users expect good quality arguments, sensibly organized
- Personal biases in the arguments can hamper unbiased decisions (Durmus and Cardie, 2018)
 - Most arguments on the web are exchanged outside “professional circles”
- Conventional snippet generation is insufficient (Alshomary et al., 2020)
 - Highlighting search query appearances in the arguments is not enough

The screenshot shows the 'args' search engine interface. At the top, there is a search bar with the text 'nuclear power' and a magnifying glass icon. Below the search bar, there are navigation links: 'All', 'Discussions', 'News', and 'People'. To the right of these links, it says 'Pro vs. con view' and '2235 arguments retrieved in 220.0 ms'. Below the navigation links, there are two columns of search results. The left column is labeled 'PRO' and contains two results. The right column is labeled 'CON' and contains two results. Each result includes a snippet of text, a 'Show full argument' link, and a URL. The snippets in the 'PRO' column are green, and the snippets in the 'CON' column are red. The snippets in the 'PRO' column are: 'Also, the war in Iraq is step one in the so called "War...' and 'The United Nations asked Iraq to cooperate with weapons...'. The snippets in the 'CON' column are: 'When you go to war you are aggressive. ... Saddam...' and 'I will be arguing to end the Iraq war while the opponent...'. The search results are displayed in a clean, modern layout with a light gray background.

OASiS: Objective Argument Summarization in Search

What are the important properties of snippets in argument search?

- Hypothesis 1: Snippets should **summarize** the argument
- Hypothesis 2: Snippets should be as “**objective**” as possible



OASiS: Objective Argument Summarization in Search

What are the important properties of snippets in argument search?

- Hypothesis 1: Snippets should **summarize** the argument
- Hypothesis 2: Snippets should be as **“objective”** as possible

What does objective mean?

- \approx **Appropriateness**:
 - An argument has an appropriate style if the used language supports the **creation of credibility** and **emotions** as well as if it is **proportional to the issue**.
(Wachsmuth et al., 2017a)
 - The **minimal required quality** of an argument that makes it **worthy to be considered in a discussion**.
(Ziegenbein et al., 2023)



<https://www.ft.com/content/5637672-d31e-11e0-9ba8-00144feab49a>

Approach Overview

Why not ban teenage sexuality altogether?



Relevant Document

HAHAHA! It's not going to happen.
People are free to make their own decisions in life.
Yeah, maybe it sucks that there are such big consequences to it, but really you sound stupid!

Sex should not be taught as an evil thing. It's not evil or dirty. Maybe you should teach your kids about responsibility and learn to talk to them about it.

Jails are over crowded now. Imagine what it would be like if ppl are thrown into jail because of that.

You can't stop anybody from doing what they want to do. You should know that already from when you were younger.

The more you discourage it the more they are going to want to do it. People have that right. Yes, even kids. That's the only way they learn.

Approach Overview

Why not ban teenage sexuality altogether?



Relevant Document

HAHAHA! It's not going to happen.
People are free to make there own decisions in life.
Yeah, maybe it sucks that there is such big consequences to it, but really you sound stupid!

Sex should not be taught as an evil thing. It's not evil or dirty. Maybe you should teach your kids about responsibility and learn to talk to them about it.

Jails are over crowded now. Imagine what it would be like if ppl are thrown into jail because of that.

You can't stop nobody from doing what they want to do. You should know that already from when you where younger.

The more you discouage it the more they are going to want to do it. People have that right. Yes, even kids. That's the only way they learn.

TextRank

Baseline Snippet

People are free to make there own decisions in life.
Yeah, maybe it sucks that there is such big consequences to it, butreally you sound stupid!. Jails are over crowded now. The more you discouage it the more they are going to want to do it.

Approach Overview

Why not ban teenage sexuality altogether? 🔍

Relevant Document

HAHAHA! It's not going to happen. People are free to make there own decisions in life. Yeah, maybe it sucks that there is such big consequences to it, but really you sound stupid!

Sex should not be taught as an evil thing. It's not evil or dirty. Maybe you should teach your kids about responsiblity and learn to talk to them about it.

Jails are over crowded now. Imagine what it would be like if ppl are thrown into jail because of that.

You can't stop nobody from doing what they want to do. You should know that already from when you where younger.

The more you discouage it the more they are going to want to do it. People have that right. Yes, even kids. That's the only way they learn.

Snippet Generation →

Snippet

Sex should not be taught as an evil thing. It's not evil or dirty. Maybe you should teach your kids about responsiblity and learn to talk to them about it.

TextRank →

Baseline Snippet

People are free to make there own decisions in life. Yeah, maybe it sucks that there is such big consequences to it, butreally you sound stupid!. Jails are over crowded now. The more you discouage it the more they are going to want to do it.

Approach Overview

Why not ban teenage sexuality altogether? 🔍

Relevant Document

HAHAHA! It's not going to happen. People are free to make there own decisions in life. Yeah, maybe it sucks that there is such big consequences to it, but really you sound stupid!

Sex should not be taught as an evil thing. It's not evil or dirty. Maybe you should teach your kids about responsibility and learn to talk to them about it.

Jails are over crowded now. Imagine what it would be like if ppl are thrown into jail because of that.

You can't stop nobody from doing what they want to do. You should know that already from when you where younger.

The more you discourague it the more they are going to want to do it. People have that right. Yes, even kids. That's the only way they learn.

Snippet Generation

Neutralization

TextRank

Snippet

Sex should not be taught as an evil thing. It's not evil or dirty. Maybe you should teach your kids about responsibility and learn to talk to them about it.

Objective Snippet

It is important to provide teens with accurate information about sexuality, while also stressing the importance of responsibility. Open communication between parents and children is essential for teens to understand the risks associated with sexual activity and to make informed decisions.


Baseline Snippet

People are free to make there own decisions in life. Yeah, maybe it sucks that there is such big consequences to it, butreally you sound stupid!. Jails are over crowded now. The more you discourague it the more they are going to want to do it.

Data

The Args.me corpus (Ajour et al., 2019)

- 387,606 arguments from debate portals
- Filtered by length and inappropriateness (Ziegenbein et al., 2023)
- Top 10 most frequent used queries
- Top five pro and con arguments for each query based on the args.me API
 - Total of 99 arguments for manual annotation studies



Snippet generation
+
Argument search

Data

The Args.me corpus (Ajjour et al., 2019)

- 387,606 arguments from debate portals
- Filtered by length and inappropriateness (Ziegenbein et al., 2023)
- Top 10 most frequent used queries
- Top five pro and con arguments for each query based on the args.me API
 - Total of 99 arguments for manual annotation studies

The Appropriateness Corpus (Ziegenbein et al., 2023)

- 2,191 arguments from three genres, manually annotated with 14 flaw dimensions
- Filtered by length and inappropriateness
 - Random sample of 100 arguments for manual annotation studies

Snippet generation
+
Argument search

Argument
neutralization

Snippet Generation

(Supervised) Extractive-Sum. (Alshomary et al., 2020)

- TextRank balancing the centrality and argumentativeness of an argument given a search query
- Extract the main claim and supporting premise as an argument snippet

Snippet Generation

(Supervised) Extractive-Sum. (Alshomary et al., 2020)

- TextRank balancing the centrality and argumentativeness of an argument given a search query
- Extract the main claim and supporting premise as an argument snippet

(Supervised) Abstractive-Sum. (Lewis et al., 2020)

- BART finetuned to the task of abstractive news summarization

Snippet Generation

(Supervised) Extractive-Sum. (Alshomary et al., 2020)

- TextRank balancing the centrality and argumentativeness of an argument given a search query
- Extract the main claim and supporting premise as an argument snippet

(Supervised) Abstractive-Sum. (Lewis et al., 2020)

- BART finetuned to the task of abstractive news summarization

(Unsupervised) Instruction-Sum. (Taori et al., 2020)

- LLaMA finetuned to followed instructions (Alpaca)

Instruction: The following is an argument on the topic "<topic>".
Extract a coherent gist from it that is exactly two sentences long.

Input: <argument>

Response:



Evaluation: Snippet Generation

What defines a “good” snippet?

- Similarity to the original argument (ROUGE, BERTScore)
- Fluency (Perplexity)
- Appropriateness

Approach	(a) Automatic						(b) Manual			
	R1	R2	RL	Sim.	PPL↓	App.↑	#1	#2	#3	Avg.↓
Extractive-Sum.	0.29	0.28	0.29	0.25	67.7	0.21	42	126	327	2.58
Abstractive-Sum.	0.40	0.38	0.38	0.35	50.9	0.31	274	149	72	1.59
Instruction-Sum.	0.24	0.11	0.16	0.13	26.5	0.58	179	220	96	1.83

Evaluation: Snippet Generation

What do we expect from a “good” snippet?

- Similarity to the original argument (ROUGE, BERTScore)
- Fluency (Perplexity)
- Appropriateness

Which approach is the best?

- Ranking the snippet generation approaches (5 annotators, 99 examples, Kendall’s τ 0.22)
- Similarity appears to be most important
- Trade-off between appropriateness and similarity
- Abstractive summarization is preferred

Approach	(a) Automatic						(b) Manual			
	R1	R2	RL	Sim.	PPL↓	App.↑	#1	#2	#3	Avg.↓
Extractive-Sum.	0.29	0.28	0.29	0.25	67.7	0.21	42	126	327	2.58
Abstractive-Sum.	0.40	0.38	0.38	0.35	50.9	0.31	274	149	72	1.59
Instruction-Sum.	0.24	0.11	0.16	0.13	26.5	0.58	179	220	96	1.83

Argument Neutralization

(Unsupervised) Instruction-Neut. (Taori et al., 2020)

- LLaMA finetuned to followed instructions (Alpaca)

Instruction: Rewrite the following argument on the topic of "<topic>" to be more appropriate and make only minimal changes to the original argument.

Input: <argument>

Response:



Argument Neutralization

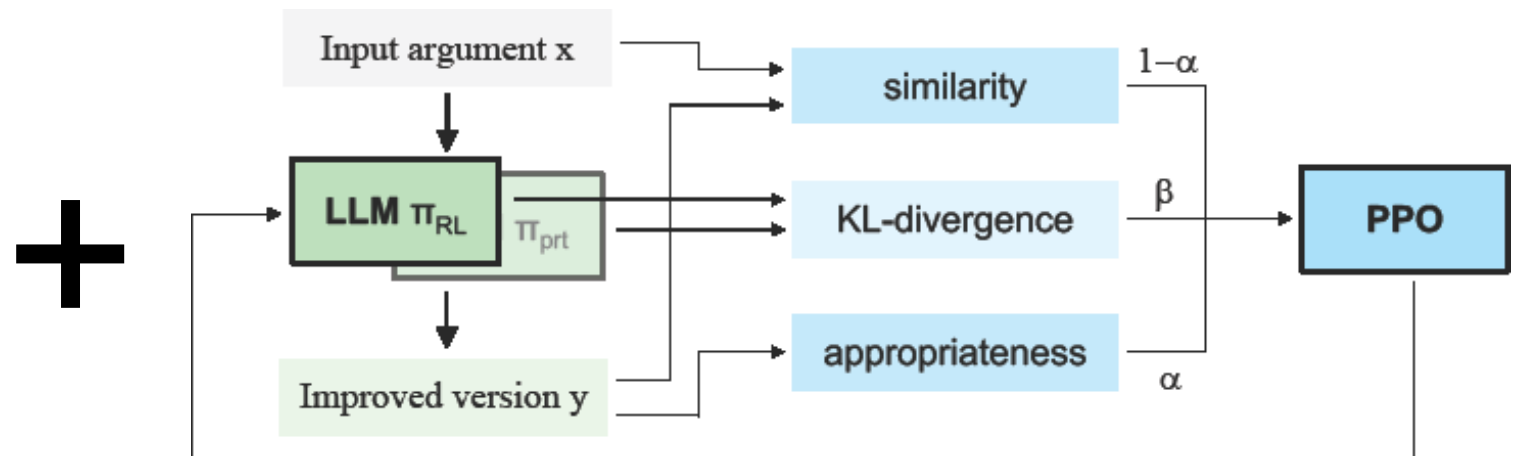
(Unsupervised) Instruction-Neut. (Taori et al., 2020)

- LLaMA finetuned to followed instructions (Alpaca)

Instruction: Rewrite the following argument on the topic of "<topic>" to be more appropriate and make only minimal changes to the original argument.

Input: <argument>

Response:



(Reinforced) Aligned-Neut. (Ziegenbein et al., 2024)

- LLaMA finetuned to followed instructions (Alpaca)
- RLHF inspired alignment balancing appropriateness and semantic similarity

Evaluation: Neutralization

What do we expect from a “neutral” version of an argument?

- Similarity to the original argument (ROUGE, BERTScore)
- Fluency (Perplexity)
- Appropriateness

Approach	(a) Automatic						(b) Manual			
	R1	R2	RL	Sim.	PPL↓	App.↑	#1	#2	#3	Avg.↓
Exact-Copy	1.00	1.00	1.00	1.00	66.1	0.00	10	91	399	2.78
Instruction-Neut.	0.79	0.66	0.73	0.67	29.5	0.40	67	345	88	2.04
Aligned-Neut.	0.41	0.16	0.27	0.18	18.4	0.97	423	64	13	1.18

Evaluation: Neutralization

What do we expect from a “neutral” version of an argument?

- Similarity to the original argument (ROUGE, BERTScore)
- Fluency (Perplexity)
- Appropriateness

Which approach is the best?

- Ranking arguments by their tendency to lead to a productive discussion (5 annotators, 100 examples, Kendall's τ 0.48)
- Appropriateness appears to be most important
- Trade-off between appropriateness and semantic similarity
- RL-aligned neutralization is preferred

Approach	(a) Automatic						(b) Manual			
	R1	R2	RL	Sim.	PPL↓	App.↑	#1	#2	#3	Avg.↓
Exact-Copy	1.00	1.00	1.00	1.00	66.1	0.00	10	91	399	2.78
Instruction-Neut.	0.79	0.66	0.73	0.67	29.5	0.40	67	345	88	2.04
Aligned-Neut.	0.41	0.16	0.27	0.18	18.4	0.97	423	64	13	1.18

Evaluation: Argument Search

What do we expect from a “good” snippet?

- Similarity to the original argument (ROUGE, BERTScore)
- Fluency (Perplexity)
- Appropriateness

Approach	(a) Automatic						(b) Manual	
	R1	R2	RL	Sim.	PPL↓	App.↑	Pref.↑	%↑
Abstractive-Sum.	0.40	0.38	0.38	0.35	50.9	0.31	57	0.11
+ Aligned-Neut.	0.25	0.10	0.17	0.11	20.0	0.87	438	0.89

Evaluation: Argument Search

What do we expect from a “good” snippet?

- Similarity to the original argument (ROUGE, BERTScore)
- Fluency (Perplexity)
- Appropriateness

Which approach is the best for argument search?

- Ranking snippets by their likelihood to be clicked on when searching for a given topic
(5 annotators, 100 examples, Krippendorff’s α 0.29)
- Appropriateness appears to be most important
- Trade-off between appropriateness and semantic similarity
- Abstractive summarization + RL-aligned neutralization is preferred

Approach	(a) Automatic						(b) Manual	
	R1	R2	RL	Sim.	PPL↓	App.↑	Pref.↑	%↑
Abstractive-Sum.	0.40	0.38	0.38	0.35	50.9	0.31	57	0.11
+ Aligned-Neut.	0.25	0.10	0.17	0.11	20.0	0.87	438	0.89

Qualitative Evaluation

What are the influential properties of the evaluated approaches?

- Manually analyzed ~500 comments from our annotation studies

Qualitative Evaluation

What are the influential properties of the evaluated approaches?

- Manually analyzed ~500 comments from our annotation studies

Task	Quality Dimensions (Preferred by Annotators)
Snippet Generation	specificity, clarity, positive/inoffensive language, conciseness, self-containment, informativeness, focus on the issue, avoiding personal attacks, structure and coherence, accuracy/correctness
Neutralization	openness, simple language, absence of profanity, facilitating critical evaluation, seriousness, absence of grammatical/orthographic errors, balanced emotions, well-reasoned, structure and coherence, formal language, non-speculative
Objective Snippets	conciseness, simple language, fluency, balanced emotions, includes quotes/evidence/statistics, specificity, coherence

Limitations

- Stance flipping (not really reflected in similarity measures)
 - Happens rarely, but when it happens it is very bad
- Content addition
 - Hallucination vs. Clarification

Limitations

- Stance flipping (not really reflected in similarity measures)
 - Happens rarely, but when it happens it is very bad
- Content addition
 - Hallucination vs. Clarification

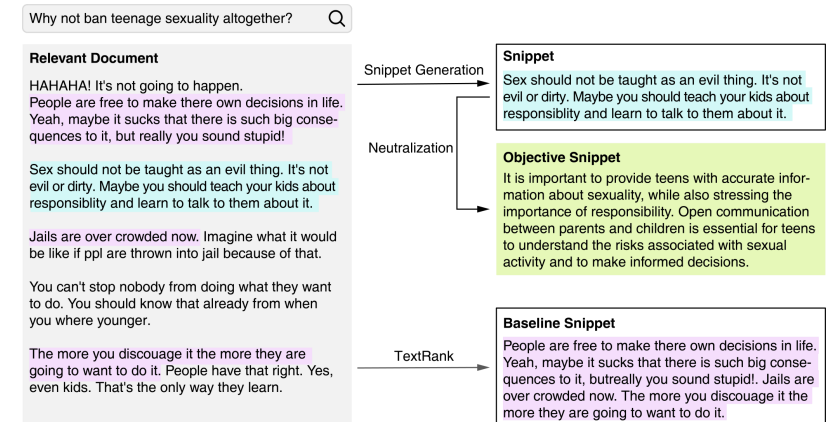
What we did not test for:

- The effect of clicking on a neutralized snippet and then seeing the original argument
 - Should the arguments also be neutralized?
 - First neutralize vs. first summarize?
- Stakeholder disagreement (author vs. reader vs. provider)
 - Readers prefer appropriateness
 - What do authors/providers prefer?

Summary and Outlook

An evaluation of snippet generation approaches

- Similarity appears to be most important
- (Supervised) Abstractive summarization is preferred



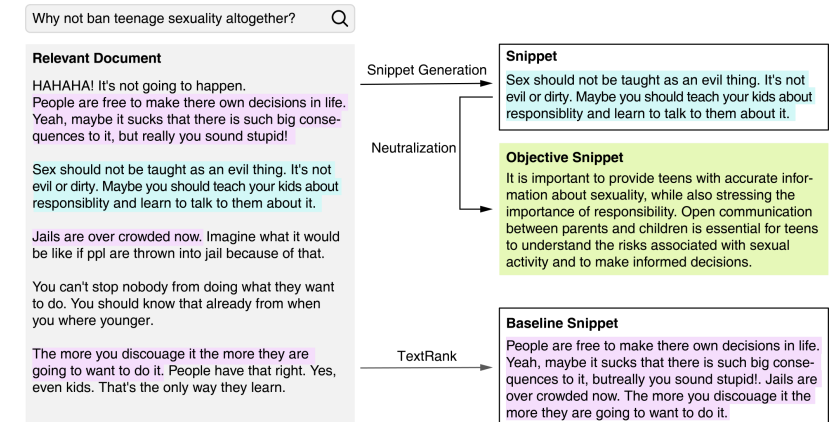
Summary and Outlook

An evaluation of snippet generation approaches

- Similarity appears to be most important
- (Supervised) Abstractive summarization is preferred

An evaluation of neutralization approaches

- Appropriateness appears to be most important
- (Reinforced) Aligned neutralization is preferred



Summary and Outlook

An evaluation of snippet generation approaches

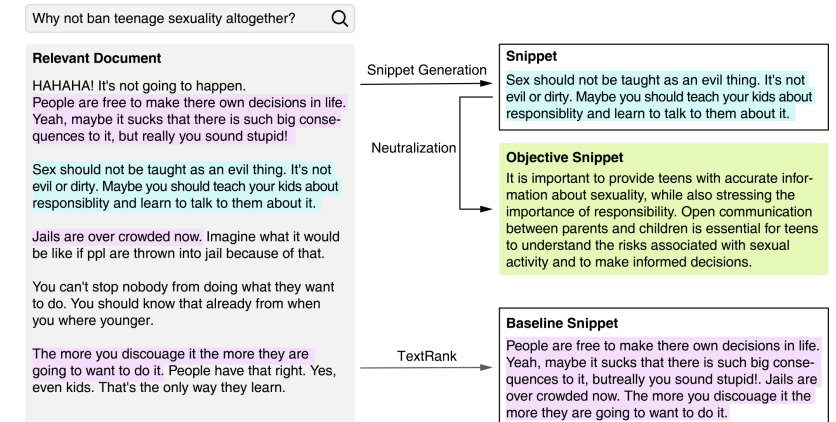
- Similarity appears to be most important
- (Supervised) Abstractive summarization is preferred

An evaluation of neutralization approaches

- Appropriateness appears to be most important
- (Reinforced) Aligned neutralization is preferred

An improved snippet generation approach for argument search

- Abstractive summarization + RL-aligned neutralization
- Appropriateness appears to be most important



Summary and Outlook

An evaluation of snippet generation approaches

- Similarity appears to be most important
- (Supervised) Abstractive summarization is preferred

An evaluation of neutralization approaches

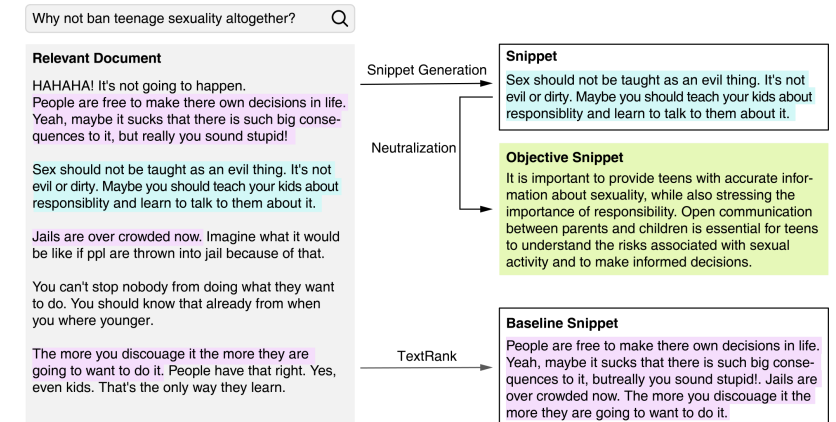
- Appropriateness appears to be most important
- (Reinforced) Aligned neutralization is preferred

An improved snippet generation approach for argument search

- Abstractive summarization + RL-aligned neutralization
- Appropriateness appears to be most important

Takeaways

- Trade-off between appropriateness and semantic similarity
- The effects of using “objective” snippets in a real search setting remain to be investigated



Summary and Outlook

An evaluation of snippet generation approaches

- Similarity appears to be most important
- (Supervised) Abstractive summarization is preferred

An evaluation of neutralization approaches

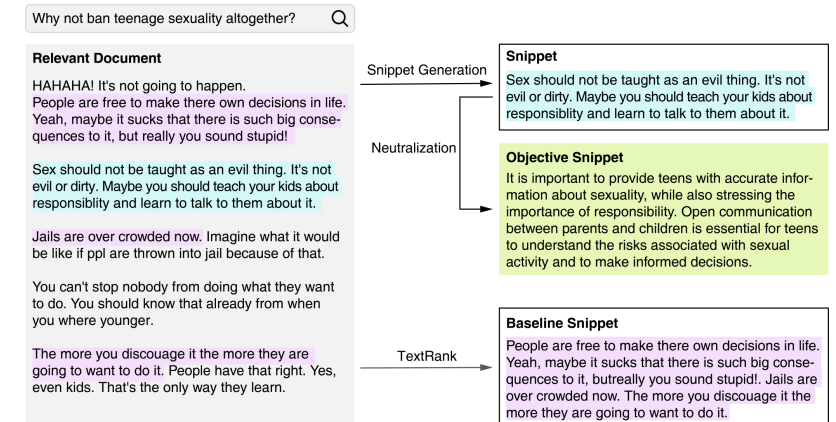
- Appropriateness appears to be most important
- (Reinforced) Aligned neutralization is preferred

An improved snippet generation approach for argument search

- Abstractive summarization + RL-aligned neutralization
- Appropriateness appears to be most important

Takeaways

- Trade-off between appropriateness and semantic similarity
- The effects of using “objective” snippets in a real search setting remain to be investigated



Corpus & Code

<https://github.com/webis-de/RATIO-24>

Summary and Outlook

An evaluation of snippet generation approaches

- Similarity appears to be most important
- (Supervised) Abstractive summarization is preferred

An evaluation of neutralization approaches

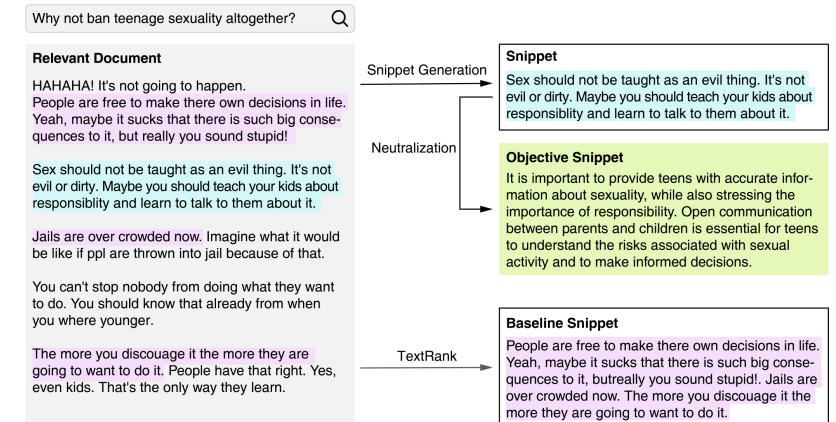
- Appropriateness appears to be most important
- (Reinforced) Aligned neutralization is preferred

An improved snippet generation approach for argument search

- Abstractive summarization + RL-aligned neutralization
- Appropriateness appears to be most important

Takeaways

- Trade-off between appropriateness and semantic similarity
- The effects of using “objective” snippets in a real search setting remain to be investigated



Corpus & Code

<https://github.com/webis-de/RATIO-24>

Thank you!

Bibliography

- **Wachsmuth et al. (2017b)** Wachsmuth, Henning, Benno Stein, and Yamen Ajjour. "PageRank for argument relevance." Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers. 2017.
- **Kiesel et al. (2020)** Kiesel, Johannes, et al. "Investigating expectations for voice-based and conversational argument search on the web." Proceedings of the 2020 Conference on Human Information Interaction and Retrieval. 2020.
- **Alshomary et al. (2020)** Alshomary, Milad, Nick Düsterhus, and Henning Wachsmuth. "Extractive snippet generation for arguments." Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2020.
- **Durmus and Cardie (2018)** Durmus, Esin, and Claire Cardie. "Exploring the Role of Prior Beliefs for Argument Persuasion." Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Vol. 1. 2018.
- **Ziegenbein et al. (2023)**, Timon Ziegenbein, Shahbaz Syed, Felix Lange, Martin Potthast, and Henning Wachsmuth. 2023. Modeling Appropriate Language in Argumentation. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 4344–4363, Toronto, Canada. Association for Computational Linguistics.
- **Wachsmuth et al., (2017a)** Henning Wachsmuth, Nona Naderi, Yufang Hou, Yonatan Bilu, Vinodkumar Prabhakaran, Tim Alberdingk Thijm, Graeme Hirst, Benno Stein. "Computational Argumentation Quality Assessment in Natural Language." Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers. 2017.

Bibliography

- **Taori et al., (2023)** Rohan Taori and Ishaan Gulrajani and Tianyi Zhang and Yann Dubois and Xuechen Li and Carlos Guestrin and Percy Liang and Tatsunori B. Hashimoto . Stanford Alpaca: An Instruction-following LLaMA model.
- **Lewis et al., (2023)** Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 7871–7880, Online. Association for Computational Linguistics.
- **Ajjour et al., (2019)** Yamen Ajjour, Henning Wachsmuth, Johannes Kiesel, Martin Potthast, Matthias Hagen, and Benno Stein. Data Acquisition for Argument Search: The args.me corpus. In Christoph Benzmüller and Heiner Stuckenschmidt, editors, 42nd German Conference on Artificial Intelligence (KI 2019), pages 48–59, September 2019. Springer.
- **Ziegenbein et al. (2024)**, Timon Ziegenbein, Gabriella Skitalinskaya, Alireza Bayat Makou, and Henning Wachsmuth. 2024. Modeling Appropriate Language in Argumentation. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics.