

# Building an Argument Search Engine for the Web

**Henning Wachsmuth**, Martin Potthast, Khalid Al-Khatib, Yamen Ajjour, Jana Puschmann, Jiani Qu, Jonas Dorsch, Viorel Morari, Janek Bevendorff, Benno Stein

Bauhaus-Universität Weimar

www.webis.de henning.wachsmuth@uni-weimar.de



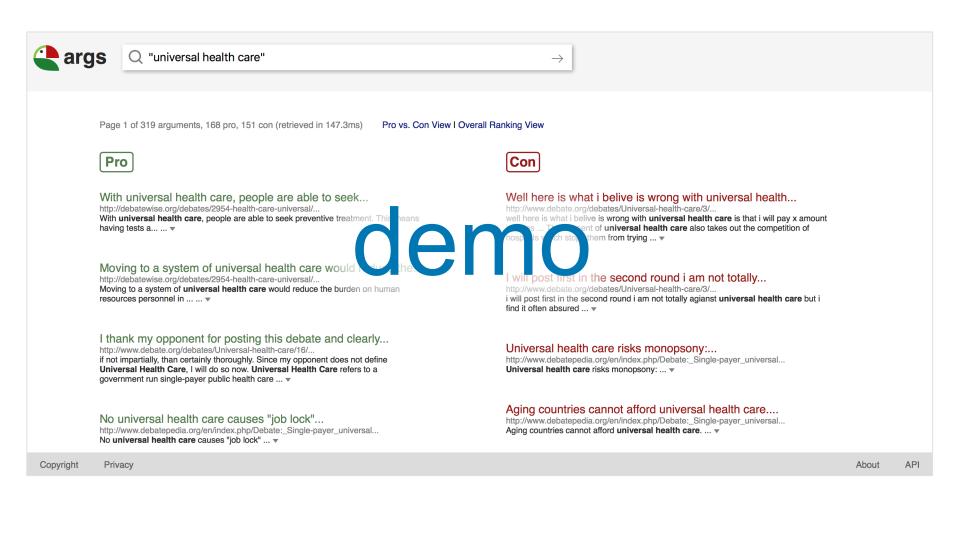








#### args.me



# Vision

- Computational argumentation in future search
  - Support forming opinions on controversial issues
  - Make it easy to find relevant arguments
- Search results should...
  - Rank the *best* arguments highest
  - Cover diverse aspects of an issue
  - Cover heterogeneous sources
  - Be up-to-the-minute
  - Be concise but detailed enough
  - Be traceable and evaluable ... and much more
- Our argument search engine...
  - Is improvable on all these criteria
  - Defines a basis to work towards the vision



No universal health care causes "job lock"... http://www.debitopedia.org/en/index.php?bebate:\_Single-payer\_universal... No universal health care causes "job lock"... v

The well insured also face risks without universal... http://www.debatepedia.org/en/index.php?bebate.\_Single-payer\_universal. The well insured also face risks without universal health care ... +

A payroll tax could help fund universal health... htp://www.debatepolia.org/en/index.php?bobate:\_Single-payer\_universal. A payroll tax could help fund universal health care ... v

Universal health care creates a universal standard... http://www.debatepedia.org/en/index.php?bebate..Single-payer\_universal Universal health care creates a universal standard of care ... v

health care grounds doctors in... antesedia.org/sn/index.php?Debate:\_Single-payer\_universal... Mb care grounds doctors in professionalism, not commercialism. Hi, guys. Tim new to debating, and this is my... http://www.debate.org/debatastur/vicenale/eath-care/21/ hi, guys. Tim new to debating, and this is my first debate ever. Ti be fighting against universal health care in the US while pro will fighting for t just accord in the first ... v

Aging countries cannot afford universal health... http://www.desatepedia.org/en/index.phpDabate\_Single-payer\_universal Aging countries cannot afford universal health care....v

niversal health care creates dependencies on... p/lwww.debatepodia.org/on/index.php/Debate\_Single-payer\_universa iversal health care creates dependencies on government ... v

Universal health care leads to rationing..... http://www.debatepodia.org/nn/index.php/Debater\_\_Single-payer\_universa Universal health care leads to rationing.....v

Sovernment-funded universal health care would... http://www.debatepedia.org/wr/index.php/Debate\_\_Single-payer\_universe Sovernment-funded universal health care would violate patient privacy

Government-funded universal health care is no... http://www.debatepedia.org/en/index.php/Debater\_Single-payer\_univ Government-funded universal health care is no right ... v



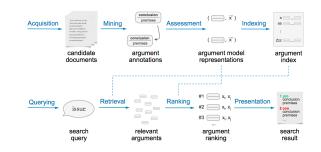
# Contributions

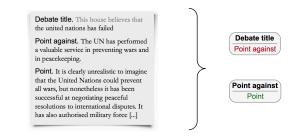
#### 1. An open search framework

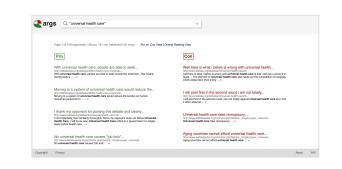
- Common argument model
- Modular indexing and retrieval processes
- Java source code coming soon

#### 2. An initial search index

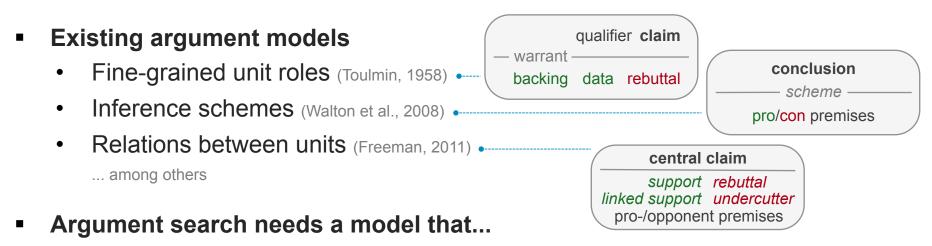
- 300k debate portal arguments
- Mining with distant supervision
- www.args.me/static/args-api.html
- 3. A prototype search engine
  - Based on framework and index
  - Standard retrieval methods
  - args.me







# Argument search framework: Model

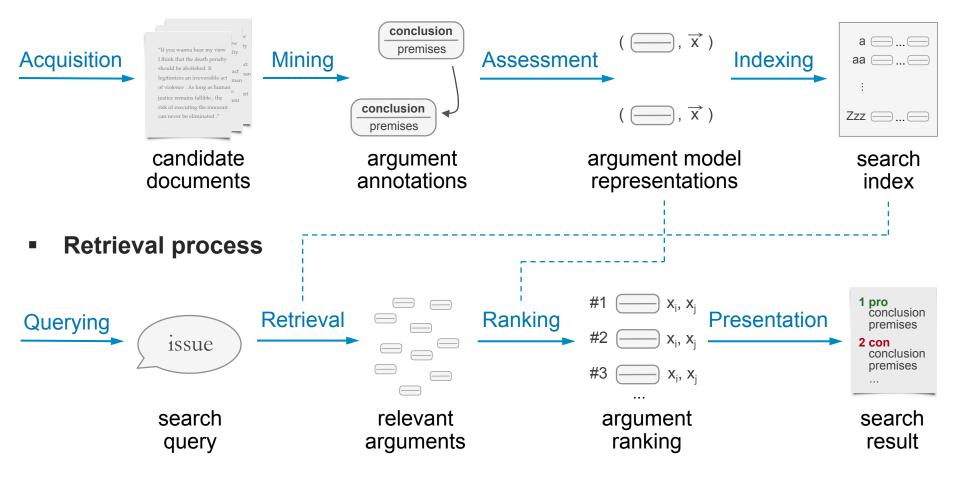


- Is applicable to all arguments
- Allows treating all arguments analogously
- Works with simple mining approaches
- Supports sophisticated approaches
- Our model
  - Argument. Conclusion, *k* premises with stance
  - Context. URL, discussion, argument position
  - Extensions. Roles, schemes, scores, ...



# Argument search framework: Processes

Indexing process



#### (implementation based on Apache UIMA and Apache Lucene)

### Argument search index

#### Crawling of 5 debate portals

- *idebate.org*, *debatepedia.org*, and *debatewise.org*
- debate.org and forandagainst.com

#### • "Mining" with distant supervision

- Portal-specific mappings to model
- Conclusion + 1 premise with stance in all cases
- 376,129 arguments initially
- 291,440 after cleaning
- Coverage of phrase queries
  - 1082 Wikipedia controversial issues
  - 78% match with  $\geq$  1 argument
  - 42% match with  $\geq$  1 conclusion

Debate title. This house believes that the united nations has failed

**Point against**. The UN has performed a valuable service in preventing wars and in peacekeeping.

**Point.** It is clearly unrealistic to imagine that the United Nations could prevent all wars, but nonetheless it has been successful at negotiating peaceful resolutions to international disputes. It has also authorised military force [...]



# Argument search engine

- Our prototype
  - Based on framework and index
  - Content-based argument search
- Capabilities so far
  - Querying. Free text *phrase* and *and* queries
  - Retrieval. Exact matching with conclusion
  - Ranking. BM25F on conclusion, premise, and context (Robertson and Zaragoza, 2009)
  - Presentation. Pro vs. con and overall ranking
- Limitations so far
  - Querying. No argument-specific interpretation
  - Retrieval. Stance taken from premises
  - Ranking. Quality not considered
  - Presentation. Snippets only premise excerpts



arg	js	Q, "universal h	ealth care* $\rightarrow$		
		<mark> args</mark>	Q, "usiversal heath care" →		
	Wi has	Page	1 of 319 arguments (Instands In 33 Bins) Pro vs. Can Vew I Deval Reality Vew		
	M: No No	con	Well here is what I belie is among with notivenal health "Extension belies adjustation interview and the set is full adjustation of the set is		
	1  10  10  10  10	pro	With universal health care, people are able to seek		
	Nc http: No	pro	) Moving to a system of universal health care would reduce the https://docume.rstatematics/surveysite/survesite/surveysite/surveysite/surveysite/surveysite/surveysite/surveysite/surveysite/surveysite/surveysite/surveysite/surveysite/surveysite/surveysite/surveysite/surveysite/surve		
pyright	Pi	con	I will post first in the second record i sam not lotally //spinnession or polyhometal-investant-investi- //spinnession i sam not tasky space aniwmal healt one but i fold taken alavedv		
		Copyright Prive	ag	About	API

0	capitalism ver socialism
against abortion	

# Conclusion

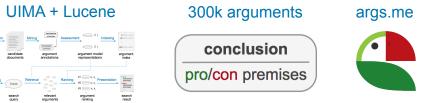
#### Towards argument search on the web

- An open search framework
- An initial search index ٠
- A prototype search engine ٠

#### Environment for collaborative research

- Combining computational argumentation methods •
- Evaluating methods in practice ٠
- Shaping the future of web search •
- Several open questions
  - Mining. How to find arguments robustly on the web? •
  - Querying. How do users search for arguments? ٠
  - Ranking. What arguments should be seen as best? •

... and many more









### References

- **Freeman (2011).** Argument Structure: Representation and Theory. Springer.
- **Robertson and Zaragoza (2009).** Stephen Robertson and Hugo Zaragoza. 2009. The Probabilistic Relevance Framework: BM25 and Beyond. *Foundations and Trends in Information Retrieval* 3(4):333–389.
- **Toulmin (1958).** Stephen E. Toulmin. 1958. *The Uses of Argument*. Cambridge University Press.
- Walton et al. (2008). Douglas Walton, Christopher Reed, and Fabrizio Macagno. 2008. *Argumentation Schemes*. Cambridge University Press.

... many more in the paper