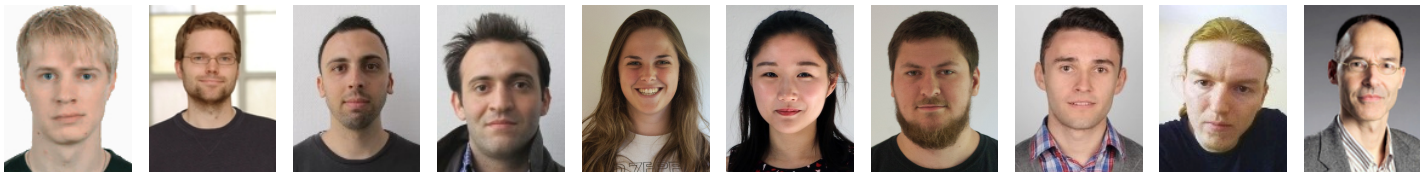**args.me**

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

Contributions

1. An open search framework

- Common argument model

- Modular indexing and retrieval processes

- Java source code coming soon
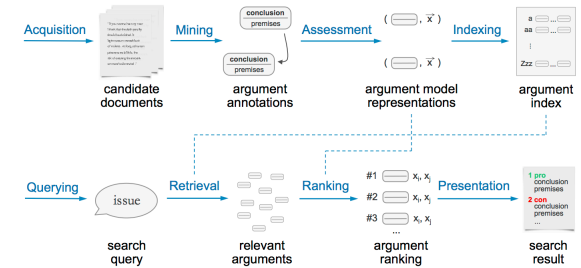
2. An initial search index

- 300k debate portal arguments

- Mining with distant supervision

- www.args.me/static/args-api.html
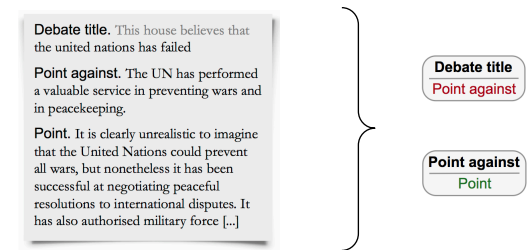
3. A prototype search engine

- Based on framework and index

- Standard retrieval methods

- args.me
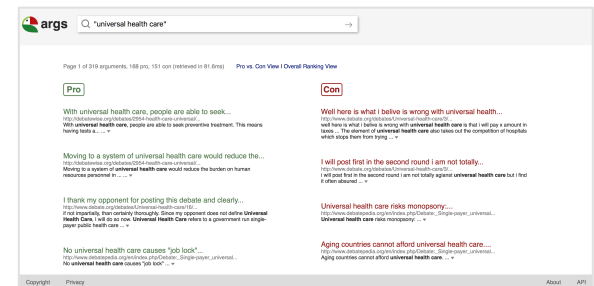
Argument search framework: Model

- **Existing argument models**

  - Fine-grained unit roles (Toulmin, 1958)

  - Inference schemes (Walton et al., 2008)

  - Relations between units (Freeman, 2011)

    ... among others
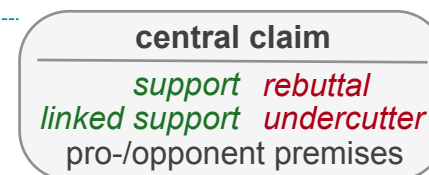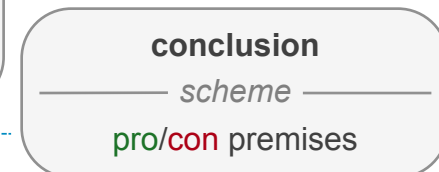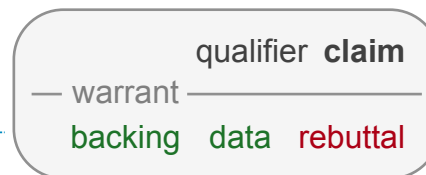
- **Argument search needs a model that...**

  - 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, ...
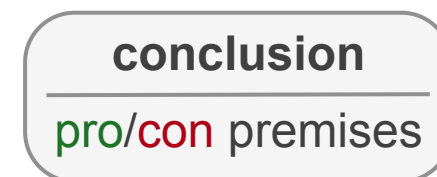
Argument search framework: Processes

- **Indexing process**

Acquisition → candidate documents → Mining → argument annotations → Assessment → argument model representations → Indexing → search index

The argument annotations show blocks labeled **conclusion** / premises.

The argument model representations show $(\boxed{\quad}, \vec{x})$ pairs.

The search index shows entries: a, aa, ⋮, Zzz.

- **Retrieval process**

Querying → search query (issue) → Retrieval → relevant arguments → Ranking → argument ranking → Presentation → search result

The argument ranking shows:
#1 $\boxed{\quad}$ x_i, x_j
#2 $\boxed{\quad}$ x_i, x_j
#3 $\boxed{\quad}$ x_i, x_j
...

The search result shows:
1 pro conclusion premises
2 con conclusion premises
...

(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 ≥1 argument
  - 42% match with ≥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 [...]

Debate title
Point against

Point against
Point

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
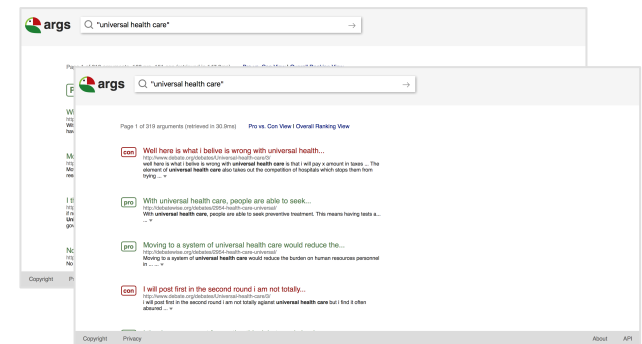
Conclusion

- **Towards argument search on the web**
  - An open search framework
  - An initial search index
  - A prototype search engine

  UIMA + Lucene   300k arguments   args.me
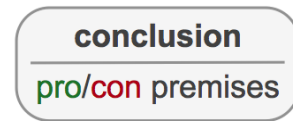
- **Environment for collaborative research**
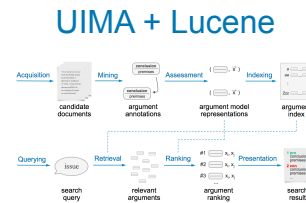 - Combining computational argumentation methods
  - Evaluating methods in practice
  - Shaping the future of web search

- **Several open questions**
  - Mining. How to find arguments robustly on the web?
  - Querying. How do users search for arguments?
  - Ranking. What arguments should be seen as best?
    ... and many more

  arguments for args

References

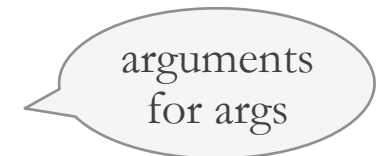- **Freeman (2011).** *Argument Structure: Representation and Theory*. Springer.

- **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