Chapter IR:XI

XI. IR Applications

- Web Technology
- Web Graph
- Web Crawling
- Web Archiving
- Web Content Extraction
- Near-duplicate Detection
- Link Analysis

- The Treachery of Answers
- Argument Retrieval Problems
- Argument Ranking I
- Argument Ranking II
- Argumentation-Related Resources
- Argument Search Engines
- Argument Search Evaluation
15 years
Domesticated

How Long Do Cats Live? | petMD
www.petmd.com/blogs/thedailyyvet/.../how_long_do_cats_live-11496
Aug 8, 2011 - This question, typically rephrased as, "How long will my cat (or dog, horse, etc.) live," is something veterinarians hear on a daily basis.

Aging Cats: Changes, Health Problems, Food, and More
pets.webmd.com/cats/guide/aging-cats-qa
WebMD veterinarian experts answer common questions cat owners have ... What else can you expect as your cat ages? ... Q: How long do cats usually live?

What Is the Life Span of the Common Cat? - Cats - About.com
cats.about.com › About Home › Cats
How long is the common cat supposed to live? Questions and answers from the About Guide to Cats.

Ageing - How long do cats live | Adelaide Animal Hospital
Life expectancy depends on many things, including one important factor - whether your cat is an indoor-only cat or an outdoor cat. Indoor cats generally live from 12-18 years of age. Many may live to be in their early 20s. The oldest reported cat lived to be an
15 years
Domesticated

How Long Do Cats Live? | petMD
www.petmd.com/blogs/thedailyvet/.../how_long_do_cats_live-11496
Aug 8, 2011 - This question, typically rephrased as, "How long will my cat (or horse, etc.) live," is something veterinarians hear on a daily basis.

Aging Cats: Changes, Health Problems, Food, and More
pets.webmd.com/cats/guide/aging-cats-qa
WebMD veterinarian experts answer common questions cat owners have ... What can you expect as your cat ages? ... Q: How long do cats usually live?

What Is the Life Span of the Common Cat? - Cats - About
cats.about.com › About Home › Cats
How long is the common cat supposed to live? Questions and answers from Guide to Cats.

Ageing - How long do cats live | Adelaide Animal Hospital
Life expectancy depends on many things, including one important factor - whether your cat is an indoor-only cat or an outdoor cat. Indoor cats generally live from 12-15 years, while outdoor cats can live from 15-20 years. Many may live to be in their early 20s. The oldest reported cat lived to be 33 years old. The average lifespan is 13.4 years.
The Treachery of Answers

Retrieving answers as a retrieval paradigm:

- Users ask questions that concern them.
- Search engines return direct answers from knowledge bases and the web.
The Treachery of Answers

Retrieving answers as a retrieval paradigm:

- Users ask questions that concern them.
- Search engines return direct answers from knowledge bases and the web.

Observations:

- Answers from knowledge bases often lack source reference and justification.
- Are answers chosen with attention to their accuracy and source credibility?
- Direct answers may lead users to believe that there are no other answers.
- Some users expect to learn why an answer is an answer.
The Treachery of Answers

Retrieving answers as a retrieval paradigm:

- Users ask questions that concern them.
- Search engines return direct answers from knowledge bases and the web.

Observations:

- Answers from knowledge bases often lack source reference and justification.
- Are answers chosen with attention to their accuracy and source credibility?
- Direct answers may lead users to believe that there are no other answers.
- Some users expect to learn why an answer is an answer.

The dilemma of the direct answer: [Potthast/Hagen/Stein 2020]

The dilemma of the direct answer is a user’s choice between convenience and diligence when using an information retrieval system.

The impact on society of giving direct answers at scale is not well-understood.
The Treachery of Answers

Hi, how can I help?

How long do cats live?

15 years.

Leci n’est pas une réponse.
Remarks:

- Copyright notice:
  - “La Trahison des réponses” (2020; “The Treachery of Answers”) by the Webis Group is licensed [CC BY-NC 2.0](https://creativecommons.org/licenses/by-nc/2.0/).
  - “La Trahison des réponses” is a derivation from “La Trahison des images” (1929; “The Treachery of Images”) by René Magritte.
  - The canvas and handwriting have been derived from a 2019 public domain reproduction of Magritte’s painting by Thomas Hawk at [publicdelivery.org](http://publicdelivery.org).
  - The image of the cat has been taken from a public domain reproduction of the painting “Sitting Cat” (1815) by Jean Bernard Duvivier at [rawpixel.com](https://www.rawpixel.com).
  - The cat’s image was kindly colorized manually by user [BlueBudgieOne](https://www.reddit.com/user/BlueBudgieOne) on Reddit’s /r/colorizationrequests.
<table>
<thead>
<tr>
<th>Conclusion</th>
<th><em>Mankind will be able to travel to other galaxies.</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Premise 1</td>
<td><em>Photon drives can take you up to relativistic velocities.</em></td>
</tr>
<tr>
<td>Premise 2</td>
<td><em>In August 2019 Lightsail2 demonstrated its functioning.</em></td>
</tr>
<tr>
<td>Premise 3</td>
<td><em>NASA announces progress on torpor (human hibernation).</em></td>
</tr>
</tbody>
</table>
## Argument Retrieval Problems

### Basic Argument Model

<table>
<thead>
<tr>
<th>Conclusion</th>
<th><em>Mankind will be able to travel to other galaxies.</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Premise 1</td>
<td><em>Photon drives can take you up to relativistic velocities.</em></td>
</tr>
<tr>
<td>Premise 2</td>
<td><em>In August 2019 Lightsail2 demonstrated its functioning.</em></td>
</tr>
<tr>
<td>Premise 3</td>
<td><em>NASA announces progress on torpor (human hibernation).</em></td>
</tr>
</tbody>
</table>

**Argument:**

- A conclusion (claim) supported by premises (reasons).  
  Conclusion and premises are considered as propositions.  
  
- Conveys a stance on a controversial topic.  
  Assignment of truth values to the propositions:
  \[
  I(“Mankind will be able to travel to other galaxies.”) = 1, \quad I(“Photon ...”) = 1, \quad \ldots
  \]

- The mechanism (“calculus”, “argumentation type”) to obtain (“derive”) the conclusion from the premises is let implicit and is usually informal.
Argument Retrieval Problems
Basic Argument Model

Conclusion  *Mankind will be able to travel to other galaxies.*

Premise 1  *Photon drives can take you up to relativistic velocities.*
Premise 2  *In August 2019 Lightsail2 demonstrated its functioning.*
Premise 3  *NASA announces progress on torpor (human hibernation).*

Argument:

- A conclusion (claim) supported by premises (reasons).  [Walton et al. 2008]
  Conclusion and premises are considered as propositions.

- Conveys a stance on a controversial topic.  [Freeley and Steinberg, 2009]
  Assignment of truth values to the propositions:
  \[ I("Mankind will be able to travel to other galaxies.") = 1, \quad I("Photon \ldots") = 1, \quad \ldots \]

- The mechanism ("calculus", "argumentation type") to obtain ("derive") the conclusion from the premises is let implicit and is usually informal.
Argument Retrieval Problems
Basic Argument Model

**Conclusion**  *Mankind will be able to travel to other galaxies.*

- **Premise 1**  *Photon drives can take you up to relativistic velocities.*
- **Premise 2**  *In August 2019 Lightsail2 demonstrated its functioning.*
- **Premise 3**  *NASA announces progress on torpor (human hibernation).*

**Argument:**

- A conclusion (claim) supported by premises (reasons).  [Walton et al. 2008]
  
  Conclusion and premises are considered as propositions.

- Conveys a stance on a controversial topic.  [Freeley and Steinberg, 2009]
  
  Assignment of truth values to the propositions:
  
  \[ I(\text{"Mankind will be able to travel to other galaxies."}) = 1, \ I(\text{"Photon \ldots ")} = 1, \ldots \]

- The mechanism ("calculus", "argumentation type") to obtain ("derive") the conclusion from the premises is let implicit and is usually informal.
Argument Retrieval Problems
Basic Argument Model

**Thesis / Major claim** $t$  Human beings will colonize other planets.

$A_{pro}$

$P_1$

$c_1$  Mankind will be able to travel to other galaxies.

$p_1$  Photon drives can take you up to relativistic velocities.

$p_2$  In August 2019 Lightsail2 demonstrated its functioning.

$p_3$  NASA announces progress on torpor (human hibernation).
Argument Retrieval Problems

Basic Argument Model

**Thesis / Major claim** \( t \)  Human beings will colonize other planets.

\[ A_{\text{pro}} \]

\( c_1 \)  Mankind will be able to travel to other galaxies.

\( p_1 \)  Photon drives can take you up to relativistic velocities.

\( p_2 \)  In August 2019 Lightsail2 demonstrated its functioning.

\( p_3 \)  NASA announces progress on torpor (human hibernation).

**Note:**  \( c_1 \preceq t \)

- \( "c_1" \) supports \( t \) (entailment in a cogent, nonobligatory sense)
- \( "t" \) is compatible with \( c_1 \) (but the real argumentation focus)
### Thesis / Major claim

**t**  
*Human beings will colonize other planets.*

### Argument Retrieval Problems

#### Basic Argument Model

---

#### Thesis / Major claim

*Human beings will colonize other planets.*

---

#### Argument for (Pro)

**$A_{pro}$**

- **$c_1$**  
  *Mankind will be able to travel to other galaxies.*

- **$p_1$**  
  *Photon drives can take you up to relativistic velocities.*

- **$p_2$**  
  *In August 2019 Lightsail2 demonstrated its functioning.*

- **$p_3$**  
  *NASA announces progress on torpor (human hibernation).*

---

#### Argument against (Con)

**$A_{con}$**

- **$c_2$**  
  *Mankind will never explore other galaxies.*

- **$p_4$**  
  *Matter cannot pass through wormholes.*

- **$p_5$**  
  *Hawking explained why time travel is impossible.*
### Argument Retrieval Problems

**Basic Argument Model**

#### Thesis / Major claim \( t \)
*Human beings will colonize other planets.*

---

**Argument Pro (A_pro)**

- **Claim** \( c_1 \)
  *Mankind will be able to travel to other galaxies.*
  
  **Premises**
  - **Pro** \( P_1 \)
    - \( p_1 \) *Photon drives can take you up to relativistic velocities.*
    - \( p_2 \) *In August 2019 Lightsail2 demonstrated its functioning.*
    - \( p_3 \) *NASA announces progress on torpor (human hibernation).*

**Argument Con (A_con)**

- **Claim** \( c_2 \)
  *Mankind will never explore other galaxies.*
  
  **Premises**
  - **Con** \( P_2 \)
    - \( p_4 \) *Matter cannot pass through wormholes.*
    - \( p_5 \) *Hawking explained why time travel is impossible.*

---

- The standard interpretation \( I \) of all propositions, \( t, c_i, p_j \), is 1 (true).

**Note:**

- \( c_1 \cong \neg c_2 \) "\( \neg c_2 \) is a paraphrase of \( c_1 \)"
- \( c_2 \) can be expressed as \( c_1 \) with opposite truth assignment, \( I(c_1) = 0, \ I(c_2) = 1 \)
Argument Retrieval Problems

(1) Argument Relevance $\Pi_{rel}$

Query  

*Will human beings colonize other planets?*
Argument Retrieval Problems

(1) Argument Relevance $\Pi_{\text{rel}}$

Query  
*Will human beings colonize other planets?*

**Given in $\Pi_{\text{rel}}$:**

- information need, expressed as query, $q \in Q$
- set of arguments, $A = \{(c_1, P_1), (c_2, P_2), \ldots, (c_n, P_n)\}$
  * (possibly hidden) human selection of the relevant arguments, $A^*_q$, $q \in Q$
Argument Retrieval Problems

(1) Argument Relevance $\Pi_{rel}$

**Query**  
*Will human beings colonize other planets?*

**Given in $\Pi_{rel}$:**
- information need, expressed as query, $q \in Q$
- set of arguments, $A = \{(c_1, P_1), (c_2, P_2), \ldots, (c_n, P_n)\}$
  * (possibly hidden) human selection of the relevant arguments, $A_q^*, q \in Q$

**Sought in $\Pi_{rel}$:**
- a relevance function $\rho : Q \times A \rightarrow \{0, 1\}$, such that ...  
  
  the macro-averaged $F$-measure (precision, recall) regarding $A_q^*, q \in Q$, is maximum
Argument Retrieval Problems
(2) Argument Ranking $\Pi_{\text{rank}}$

**Query**  
*Will human beings colonize other planets?*

Given in $\Pi_{\text{rank}}$:

- information need, expressed as query, $q \in Q$
- set of relevant arguments, $A_q = \{(c_1, P_1), (c_2, P_2), \ldots, (c_m, P_m)\}$
- (possibly hidden) human ranking of the relevant arguments, $\pi^*_q$, $q \in Q$
Argument Retrieval Problems

(2) Argument Ranking $\Pi_{\text{rank}}$

**Query**

*Will human beings colonize other planets?*

Given in $\Pi_{\text{rank}}$:

- information need, expressed as query, $q \in Q$
- set of relevant arguments, $A_q = \{(c_1, P_1), (c_2, P_2), \ldots, (c_m, P_m)\}$
  - (possibly hidden) human ranking of the relevant arguments, $\pi^*_{A_q}$, $q \in Q$

Sought in $\Pi_{\text{rank}}$:

- a ranking function $\sigma : Q \times \mathcal{P}(A) \rightarrow \Pi$, such that...
  - the mean rank correlation $\bar{\tau}$ regarding $\pi^*_{A_q}$, $q \in Q$, is maximum

**Relevant Arguments**

$A_{\text{pro}}$

\[
\begin{align*}
A_{\text{pro}} &= \left\{ \begin{array}{l}
\begin{array}{l}
c_1 \quad \text{Mankind will be able to travel to other galaxies.} \\
p_1 \quad \text{Photon drives can take you up to relativistic velocities.} \\
p_2 \quad \text{In August 2019 Lightsail2 demonstrated its functioning.} \\
p_3 \quad \text{NASA announces progress on torpor (human hibernation).}
\end{array}
\end{array} \right. \\
\end{align*}
\]
Argument Retrieval Problems

(3) – (7) Further Problems

3. \( \Pi_{\text{counter}} \) Retrieve the “best” counterargument
   Given: query \( q \), argument set \( A \), argument \( A \)

4. \( \Pi_{\text{sameside}} \) Retrieve (all) arguments with the same stance
   Given: argument set \( A \), argument \( A \)

5. \( \Pi_{\text{argdoc}} \) Is the document argumentative?
   Given: document \( d \)

6. \( \Pi_{\text{argquery}} \) Is the query argumentative?
   Given: query \( q \)

7. \( \Pi_{\text{argsum}} \) Summarize an argument.
   Given: argument \( A \)
3. $\Pi_{\text{counter}}$ Retrieve the “best” counterargument
   Given: query $q$, argument set $A$, argument $A$

4. $\Pi_{\text{sameside}}$ Retrieve (all) arguments with the same stance
   Given: argument set $A$, argument $A$

5. $\Pi_{\text{argdoc}}$ Is the document argumentative?
   Given: document $d$

6. $\Pi_{\text{argquery}}$ Is the query argumentative?
   Given: query $q$

7. $\Pi_{\text{argsum}}$ Summarize an argument.
   Given: argument $A$

Notes:
- $\Pi_{\text{counter}}$ can be cast as $\Pi_{\text{rank}}$ if the query is negated.
- $\Pi_{\text{argdoc}}$ and $\Pi_{\text{argquery}}$ are decision problems.
- $\Pi_{\text{counter}}$ and $\Pi_{\text{sameside}}$ can be cast as decision problems as well.
- Challenge: development of domain-independent or “topic-agnostic” approaches.
XI. IR Applications

- Web Technology
- Web Graph
- Web Crawling
- Web Archiving
- Web Content Extraction
- Near-duplicate Detection
- Link Analysis

- The Treachery of Answers
- Argument Retrieval Problems
- Argument Ranking I
- Argument Ranking II
- Argumentation-Related Resources
- Argument Search Engines
- Argument Search Evaluation
**Argument Ranking I**

### PRO

- **Abortion is the ending of pregnancy by the removal or...**
  - Show full argument
  - Abortion is the ending of pregnancy by the removal or forcing out from the womb of a fetus or embryo before it is able to survive on its own. An abortion can occur spontaneously, in which ...
  - [https://www.debate.org/debates/abortion/3290](https://www.debate.org/debates/abortion/3290)
  - score:

- **Great, another forfeiter. As someone who has debated...**
  - Show full argument
  - Great, another forfeiter. As someone who has debated abortion before, I will post a link to my original abortion debate right here: [https://www.debate.org](https://www.debate.org). I will be using arguments that...
  - [https://www.debate.org/debates/abortion/3293](https://www.debate.org/debates/abortion/3293)
  - score:

- **This should be fun :) The legalisation of abortion has...**
  - Show full argument
  - This should be fun :) The legalisation of abortion has been a big issue worldwide for a long period of time; not only politically but also on social and religious fronts. Abortion can be ...
  - [https://www.debate.org/debates/abortion/3292](https://www.debate.org/debates/abortion/3292)
  - score:

- **There are many good and bad sides to abortion. But just...**
  - Show full argument
  - There are many good and bad sides to abortion. But just like everything, there is no black and white, just a shade of gray. Abortion is one way that poverty can decrease. Most unplanned ...
  - [https://www.debate.org/debates/abortion/3291](https://www.debate.org/debates/abortion/3291)
  - score:

- **Although I oppose abortion in most cases, I accepted this...**
  - Show full argument
  - Although I oppose abortion in most cases, I accepted this debate because Con's position is that abortion "can never be justified regardless of circumstances." That is the point I want to ...
  - [https://www.debate.org/debates/abortion/3290](https://www.debate.org/debates/abortion/3290)
  - score:

- **Abortion is needed to control the population so that the...**
  - Show full argument
  - Abortion is needed to control the population so that the population doesn't get too excess. By the 22 century, the population estimated to be 11.2 billion people and if abortion were illegal ...
  - [https://www.debate.org/debates/abortion/3273](https://www.debate.org/debates/abortion/3273)
  - score:

### CON

- **In 2011 there were about 730,322 abortions reported to...**
  - Show full argument
  - In 2011 there were about 730,322 abortions reported to the centers for disease control. There are about 1.7% of abortion of women's ages from 15-44 each year. Women who already had abortion ...
  - [https://www.debate.org/debates/abortion/456](https://www.debate.org/debates/abortion/456)
  - score:

- **The greatest destroyer of peace is abortion because if a...**
  - Show full argument
  - *The greatest destroyer of peace is abortion because if a mother can kill her own child, what is left for me to kill you and you to kill me? There is nothing between.* says Walter Torcas.
  - [https://www.debate.org/debates/abortion/507](https://www.debate.org/debates/abortion/507)
  - score:

- **Yes the government has the obligation to protect the...**
  - Show full argument
  - Yes the government has the obligation to protect the rights of people. In general. Women have a right to decide whether and when to become a parent. But not abortion, it's an ending life ...
  - [https://www.debate.org/debates/abortion/393](https://www.debate.org/debates/abortion/393)
  - score:

- **Thank you. Pro. Negative CaseA1. False equivalenc...**
  - Show full argument
  - Thank you. Pro. Negative CaseA1. False equivalency in the ...
  - [https://www.debate.org/debates/abortion/492](https://www.debate.org/debates/abortion/492)
  - score:

- **Abortion is wrong! Abortion Is gross! Abortion is...**
  - Show full argument
  - Abortion is wrong! Abortion is gross! Abortion is MURDER!!
  - [https://www.debate.org/debates/abortion/485](https://www.debate.org/debates/abortion/485)
  - score:

- **Thank you to both the audience and my opponent for yet...**
  - Show full argument
  - Thank you to both the audience and my opponent for yet another debate on abortion. The resolution is simply "Abortion" and my opponent has stated that he supports the affirmative. I shall ...
  - [https://www.debate.org/debates/abortion/437](https://www.debate.org/debates/abortion/437)
  - score:
Argument Ranking I

Argumentative documents

Retrieval

Argument mining

Arguments

Relation detection

Paraphrased argument units

Graph analysis

Argument graph

Centrality: PageRank

Ranking
Query  Reintroduce death penalty?
Query: Reintroduce death penalty?
Query  Reintroduce death penalty?
Query: Reintroduce death penalty?

Premises:

- Death penalty should be abolished.
- It does not prevent people from committing crimes.
- The death penalty doesn't deter people from committing serious violent crimes.
- A survey of the UN on the relation between the death penalty and homicide rates gave no support to the deterrent hypothesis.

Conclusion:

- The death penalty should be abolished.
Query: Reintroduce death penalty?

Premises:
1. Death penalty should be abolished.
2. It does not prevent people from committing crimes.
3. The death penalty doesn't deter people from committing serious violent crimes.
4. A survey of the UN on the relation between the death penalty and homicide rates gave no support to the deterrent hypothesis.

Conclusion: The death penalty should be abolished.
Argument Ranking I

\[ p(d_i) = (1 - \alpha) \cdot \frac{1}{|D|} + \alpha \cdot \sum_j \frac{p(d_j)}{|D_j|} \]

Original PageRank \[\text{[Page et al. 1999]}\]
Argument Ranking I

\[ p(d_i) = (1 - \alpha) \cdot \frac{1}{|D|} + \alpha \cdot \sum_j \frac{p(d_j)}{|D_j|} \]

1. ground relevance + recursive relevance

Original PageRank [Page et al. 1999]

1. ground relevance + recursive relevance
Argument Ranking I

\[ p(d_i) = (1 - \alpha) \cdot \frac{1}{|D|} + \alpha \cdot \sum_j \frac{p(d_j)}{|D_j|} \]

Original PageRank [Page et al. 1999]

1. ground relevance + recursive relevance
2. \( d_j \) links to \( d_i \) \( \sim \) increase \( \text{PageRank}(d_i) \)
Argument Ranking I

\[ p(d_i) = (1 - \alpha) \cdot \frac{1}{|D|} + \alpha \cdot \sum_j \frac{p(d_j)}{|D_j|} \]

Original PageRank [Page et al. 1999]

1. ground relevance + recursive relevance
2. \( d_j \) links to \( d_i \) \( \leadsto \) increase \( \text{PageRank}(d_i) \)
3. reward exclusive links
Argument Ranking I

\[ p(d_i) = (1 - \alpha) \cdot \frac{1}{|D|} + \alpha \cdot \sum_j \frac{p(d_j)}{|D_j|} \]

Original PageRank \[\text{[Page et al. 1999]}\]

1. ground relevance + recursive relevance
2. \(d_j\) links to \(d_i\) \(\sim\) increase PageRank\((d_i)\)
3. reward exclusive links
4. uniform ground relevances (sum to 1)
**Argument Ranking I**

\[
p(d_i) = (1 - \alpha) \cdot \frac{1}{|D|} + \alpha \cdot \sum_j \frac{p(d_j)}{|D_j|}
\]

**Original PageRank**  
[Page et al. 1999]

1. ground relevance + recursive relevance
2. \(d_j\) links to \(d_i\) \(\leadsto\) increase \(\text{PageRank}(d_i)\)
3. reward exclusive links
4. uniform ground relevances (sum to 1)

---

\[
\hat{p}(c_i) = (1 - \alpha) \cdot \frac{p(d_i) \cdot |D|}{|A|} + \alpha \cdot \sum_j \frac{\hat{p}(c_j)}{|P_j|}
\]

**ArgRank**  
[Wachsmuth/Stein 2017]

1. ground strength + recursive relevance
2. \(c_i\) premise for \(c_j\) \(\leadsto\) increase \(\text{ArgRank}(c_i)\)
3. reward exclusive premises
4. ground strength \(\approx\) PageRank
Argument Ranking I

**Original PageRank** [Page et al. 1999]

1. ground relevance + recursive relevance
2. \( d_j \) links to \( d_i \) \( \sim \) increase PageRank(\( d_i \))
3. reward exclusive links
4. uniform ground relevances (sum to 1)

**ArgRank** [Wachsmuth/Stein 2017]

1. ground strength + recursive relevance
Argument Ranking I

\[ p(d_i) = (1 - \alpha) \cdot \frac{1}{|D|} + \alpha \cdot \sum_j \frac{p(d_j)}{|D_j|} \]

Original PageRank  [Page et al. 1999]

1. ground relevance + recursive relevance
2. \( d_j \) links to \( d_i \) \( \rightarrow \) increase PageRank\((d_i)\)
3. reward exclusive links
4. uniform ground relevances (sum to 1)

\[ \hat{p}(c_i) = (1 - \alpha) \cdot \frac{p(d_i) \cdot |D|}{|A|} + \alpha \cdot \sum_j \frac{\hat{p}(c_j)}{|P_j|} \]

ArgRank  [Wachsmuth/Stein 2017]

1. ground strength + recursive relevance
2. \( c_i \) premise for \( c_j \) \( \rightarrow \) increase ArgRank\((c_i)\)
**Argument Ranking I**

$$p(d_i) = (1 - \alpha) \cdot \frac{1}{|D|} + \alpha \cdot \sum_j p(d_j) \cdot \frac{|D_j|}{|D|}$$

**Original PageRank** [Page et al. 1999]

1. ground relevance + recursive relevance
2. $d_j$ links to $d_i \leadsto$ increase $\text{PageRank}(d_i)$
3. reward exclusive links
4. uniform ground relevances (sum to 1)

**ArgRank** [Wachsmuth/Stein 2017]

$$\hat{p}(c_i) = (1 - \alpha) \cdot \frac{p(d_i) \cdot |D|}{|A|} + \alpha \cdot \sum_j \hat{p}(c_j) \cdot \frac{|P_j|}{|P_j|}$$

1. ground strength + recursive relevance
2. $c_i$ premise for $c_j \leadsto$ increase $\text{ArgRank}(c_i)$
3. reward exclusive premises
Argument Ranking I

Original PageRank  [Page et al. 1999]

1. ground relevance + recursive relevance
2. \(d_j\) links to \(d_i\) \(\rightarrow\) increase PageRank\((d_i)\)
3. reward exclusive links
4. uniform ground relevances (sum to 1)

ArgRank  [Wachsmuth/Stein 2017]

1. ground strength + recursive relevance
2. \(c_i\) premise for \(c_j\) \(\rightarrow\) increase ArgRank\((c_i)\)
3. reward exclusive premises
4. ground strength \(\sim\) PageRank
Argument Ranking I

\[ p(d_i) = (1 - \alpha) \cdot \frac{1}{|D|} + \alpha \cdot \sum_j \frac{p(d_j)}{|D_j|} \]

Original PageRank [Page et al. 1999]

1. ground relevance + recursive relevance
2. \( d_j \) links to \( d_i \) \( \xrightarrow{} \) increase \( \text{PageRank}(d_i) \)
3. reward exclusive links
4. uniform ground relevances (sum to 1)

ArgRank [Wachsmuth/Stein 2017]

\[ \hat{p}(c_i) = (1 - \alpha) \cdot \frac{p(d_i) \cdot |D|}{|A|} + \alpha \cdot \sum_j \frac{\hat{p}(c_j)}{|P_j|} \]

1. ground strength + recursive relevance
2. \( c_i \) premise for \( c_j \) \( \xrightarrow{} \) increase \( \text{ArgRank}(c_i) \)
3. reward exclusive premises
4. ground strength \( \sim \) PageRank

“Reversal of Evidence”

PageRank: Author cannot enforce links to her web page.
ArgRank: Author cannot enforce use of her argument.
Argument Ranking I
From Premise Scores to Argument Ranks

Thesis $t$

Premises

Conclusion

$\approx$

$\approx$

$\approx$

$\approx$

$\approx$

$\approx$

$\approx$

$\approx$

$\approx$

$p_1 : 0.01$

$p_2 : 0.12$

$p_3 : 0.03$

How to weigh the premise scores of the matching arguments?
(maximum, average, etc.)
Argument Ranking I
Case Study: Graph Construction

Construction of a raw graph using 57 corpora from the Argument Web:

- 28,875 Argument units, used in . . .
- 17,877 Arguments

Processing steps towards an argument graph:

- 3,113 Conclusions with $\geq 1$ argument, where . . .
  - 498 have multiple premises, from which . . .
  - 70 have a relevant claim, from which . . .
  - 32 are used in 110 intelligible arguments.
Construction of a raw graph using 57 corpora from the Argument Web:

- 28,875 Argument units, used in...
- 17,877 Arguments

Processing steps towards an argument graph:

- 3,113 Conclusions with ≥ 1 argument, where...
  - 498 have multiple premises, from which...
  - 70 have a relevant claim, from which...
  - 32 are used in 110 intelligible arguments.

Acquisition of a ranking ground truth:

- 7 experts from NLP and IR ranked all arguments (110) for each conclusion (32)
- $\tau = 0.59$ as highest agreement between two experts (mean: $\tau = 0.36$)
## Argument Ranking I

### Case Study: Results

<table>
<thead>
<tr>
<th>Ranking approach</th>
<th>Premise score computation</th>
<th>Best</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td>(\tau)</td>
<td>(\tau)</td>
</tr>
<tr>
<td>1. ArgRank</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>2. Frequency</td>
<td>–0.10</td>
<td>–0.03</td>
</tr>
<tr>
<td>3. Similarity</td>
<td>–0.13</td>
<td>–0.05</td>
</tr>
<tr>
<td>4. Sentiment</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>5. Most premises</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. Random</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Approach 1: An argument’s relevance corresponds to the ArgRank of its premises.
### Argument Ranking I

**Case Study: Results**

<table>
<thead>
<tr>
<th>Ranking approach</th>
<th>Premise score computation</th>
<th>Best</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Average</td>
</tr>
<tr>
<td>1. ArgRank</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>2. Frequency</td>
<td>–0.10</td>
<td>–0.03</td>
</tr>
<tr>
<td>3. Similarity</td>
<td>–0.13</td>
<td>–0.05</td>
</tr>
<tr>
<td>4. Sentiment</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>5. Most premises</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. Random</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Approach 2: An argument’s relevance corresponds to the frequency of its premises in the graph.
Argument Ranking I
Case Study: Results

<table>
<thead>
<tr>
<th>Ranking approach</th>
<th>Premise score computation</th>
<th>Best</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td>τ</td>
<td>τ</td>
</tr>
<tr>
<td>1. ArgRank</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>2. Frequency</td>
<td>−0.10</td>
<td>−0.03</td>
</tr>
<tr>
<td>3. Similarity</td>
<td>−0.13</td>
<td>−0.05</td>
</tr>
<tr>
<td>4. Sentiment</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>5. Most premises</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. Random</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Approach 3: An argument’s relevance corresponds to the Jaccard similarity of its premises to its conclusion.
Argument Ranking I
Case Study: Results

<table>
<thead>
<tr>
<th>Ranking approach</th>
<th>Premise score computation</th>
<th>Best</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Average</td>
</tr>
<tr>
<td>1. ArgRank</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>2. Frequency</td>
<td>-0.10</td>
<td>-0.03</td>
</tr>
<tr>
<td>3. Similarity</td>
<td>-0.13</td>
<td>-0.05</td>
</tr>
<tr>
<td>4. Sentiment</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>5. Most premises</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. Random</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Approach 4: An argument’s relevance corresponds to the positivity of its words in the premises according to SentiWordNet.
## Argument Ranking I

### Case Study: Results

<table>
<thead>
<tr>
<th>Ranking approach</th>
<th>Premise score computation</th>
<th>Best</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Average</td>
</tr>
<tr>
<td>τ</td>
<td>τ</td>
<td>τ</td>
</tr>
<tr>
<td>1. ArgRank</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>2. Frequency</td>
<td>–0.10</td>
<td>–0.03</td>
</tr>
<tr>
<td>3. Similarity</td>
<td>–0.13</td>
<td>–0.05</td>
</tr>
<tr>
<td>4. Sentiment</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>5. Most premises</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. Random</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Approach 5: An argument’s relevance corresponds to its number of premises.
## Argument Ranking I

### Case Study: Results

<table>
<thead>
<tr>
<th>Ranking approach</th>
<th>Premise score computation</th>
<th>Best</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Average</td>
</tr>
<tr>
<td>1. ArgRank</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>2. Frequency</td>
<td>-0.10</td>
<td>-0.03</td>
</tr>
<tr>
<td>3. Similarity</td>
<td>-0.13</td>
<td>-0.05</td>
</tr>
<tr>
<td>4. Sentiment</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>5. Most premises</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. Random</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Approach 6: The relevance is decided randomly.
Idea: Given an argument \( A \), the best counterargument \( \overline{A}^* \) employs premises that are similar wrt. topic, but takes the opposite stance.

Consider both similarities to the premises and conclusion [Walton 2009]:
Argument Ranking II

**Idea:** Given an argument $A$, the best counterargument $\overline{A}^*$ employs premises that are similar wrt. topic, but takes the opposite stance.

- Consider both similarities to the premises and conclusion [Walton 2009]:

$$
\begin{array}{c|c|c}
\text{Premises} & \phi_c & \text{Counter-argument} \\
\hline
A & & \overline{A} \\
\end{array}
$$

How to compute these similarities?

How to combine these similarities?

(= What is a sensible hypothesis space of promising model functions?)
Argument Ranking II

Idea: Given an argument \( A \), the best counterargument \( \overline{A}^* \) employs premises that are similar \( \text{wrt. topic} \), but takes the opposite stance.

Consider both similarities to the premises and conclusion \([Walton\ 2009]\):

\[
\begin{align*}
A & \quad \text{Premises} \quad \phi_P \quad \text{Conclusion} \quad \phi_C \\
\overline{A} & \quad \text{Counter-argument}
\end{align*}
\]

Proposed model function to rank counterarguments \([Wachsmuth\ et\ al.,\ 2018]\):

\[
R(A, \overline{A}) = \alpha \cdot \left( \phi_{\text{conclusion}} \circ \phi_{\text{Premises}} \right) \quad \text{max}
- (1 - \alpha) \cdot \left( \phi_{\text{conclusion}} \circ \phi_{\text{Premises}} \right) \quad \text{min}
\]

where

\( \phi \) combines both word and embedding similarities
\( \circ \in \{\min, \max, +, *\} \)
\( \alpha \in [0; 1] \)
## Argument Ranking II

### Corpus and Analysis

<table>
<thead>
<tr>
<th>Theme</th>
<th>Debates</th>
<th>Points</th>
<th>Counters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>46</td>
<td>278</td>
<td>278</td>
</tr>
<tr>
<td>Digital freedoms</td>
<td>48</td>
<td>341</td>
<td>341</td>
</tr>
<tr>
<td>Economy</td>
<td>95</td>
<td>590</td>
<td>588</td>
</tr>
<tr>
<td>Sport</td>
<td>23</td>
<td>130</td>
<td>130</td>
</tr>
<tr>
<td><strong>∑</strong></td>
<td><strong>1069</strong></td>
<td><strong>6779</strong></td>
<td><strong>6753</strong></td>
</tr>
</tbody>
</table>

Corpus:
- based on the iDebate.org portal
- Download: ArguAna Counterargs
Argument Ranking II
Corpus and Analysis

<table>
<thead>
<tr>
<th>Theme</th>
<th>Debates</th>
<th>Points</th>
<th>Counters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>46</td>
<td>278</td>
<td>278</td>
</tr>
<tr>
<td>Digital freedoms</td>
<td>48</td>
<td>341</td>
<td>341</td>
</tr>
<tr>
<td>Economy</td>
<td>95</td>
<td>590</td>
<td>588</td>
</tr>
<tr>
<td>Sport</td>
<td>23</td>
<td>130</td>
<td>130</td>
</tr>
<tr>
<td><strong>∑</strong></td>
<td><strong>1069</strong></td>
<td><strong>6779</strong></td>
<td><strong>6753</strong></td>
</tr>
</tbody>
</table>

Corpus:
- based on the iDebate.org portal
- Download: ArguAna Counterargs

Retrieval experiments (selected results):

<table>
<thead>
<tr>
<th>Find the best counterargument within . . .</th>
<th>True-to-false ratio</th>
<th>Accuracy*</th>
</tr>
</thead>
<tbody>
<tr>
<td>all counters of the same debate</td>
<td>1 : 3</td>
<td>0.75</td>
</tr>
<tr>
<td>all counters of the same theme</td>
<td>1 : 136</td>
<td>0.54</td>
</tr>
<tr>
<td>all arguments of the entire portal</td>
<td>1 : 2 800</td>
<td>0.32</td>
</tr>
</tbody>
</table>

* The parameters for $R(A, \overline{A})$ were determined by a systematic ranking analysis.
Chapter IR:XI

XI. IR Applications

- Web Technology
- Web Graph
- Web Crawling
- Web Archiving
- Web Content Extraction
- Near-duplicate Detection
- Link Analysis

- The Treachery of Answers
- Argument Retrieval Problems
- Argument Ranking I
- Argument Ranking II
  - Argumentation-Related Resources
  - Argument Search Engines
  - Argument Search Evaluation
Argumentation-Related Resources
# Argumentation-Related Resources

<table>
<thead>
<tr>
<th>Leverage effort*</th>
<th>Resource type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>very low</td>
<td>Technology</td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>Corpora</td>
<td></td>
</tr>
<tr>
<td>medium</td>
<td>Debate portals</td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>Discussion pages</td>
<td></td>
</tr>
<tr>
<td>very high</td>
<td>Articles</td>
<td></td>
</tr>
</tbody>
</table>

* Estimated effort / expertise to exploit a resource of the respective type within own research.
## Argumentation-Related Resources

<table>
<thead>
<tr>
<th>Leverage effort*</th>
<th>Resource type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>very low</td>
<td>Technology</td>
<td>Visual inspection Acquisition, Tagging</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Argument Web Truthmapping</td>
</tr>
<tr>
<td>low</td>
<td>Corpora</td>
<td></td>
</tr>
<tr>
<td>medium</td>
<td>Debate portals</td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>Discussion pages</td>
<td></td>
</tr>
<tr>
<td>very high</td>
<td>Articles</td>
<td></td>
</tr>
</tbody>
</table>

* Estimated effort / expertise to exploit a resource of the respective type within own research.
## Argumentation-Related Resources

<table>
<thead>
<tr>
<th>Leverage effort*</th>
<th>Resource type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>very low</strong></td>
<td>Technology</td>
<td>Visual inspection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Acquisition, Tagging</td>
</tr>
<tr>
<td><strong>low</strong></td>
<td>Corpora</td>
<td>Argumentative structure analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Argumentation quality analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stance detection</td>
</tr>
<tr>
<td><strong>medium</strong></td>
<td>Debate portals</td>
<td></td>
</tr>
<tr>
<td><strong>high</strong></td>
<td>Discussion pages</td>
<td></td>
</tr>
<tr>
<td><strong>very high</strong></td>
<td>Articles</td>
<td></td>
</tr>
</tbody>
</table>

* Estimated effort / expertise to exploit a resource of the respective type within own research.
<table>
<thead>
<tr>
<th>Leverage effort*</th>
<th>Resource type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>very low</td>
<td>Technology</td>
<td>Visual inspection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Acquisition, Tagging</td>
</tr>
<tr>
<td>low</td>
<td>Corpora</td>
<td>Argumentative structure analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Argumentation quality analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stance detection</td>
</tr>
<tr>
<td>medium</td>
<td>Debate portals</td>
<td>English</td>
</tr>
<tr>
<td></td>
<td></td>
<td>German</td>
</tr>
<tr>
<td>high</td>
<td>Discussion pages</td>
<td></td>
</tr>
<tr>
<td>very high</td>
<td>Articles</td>
<td></td>
</tr>
</tbody>
</table>

* Estimated effort / expertise to exploit a resource of the respective type within own research.
# Argumentation-Related Resources

<table>
<thead>
<tr>
<th>Leverage effort*</th>
<th>Resource type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>very low</td>
<td>Technology</td>
<td>Visual inspection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Acquisition, Tagging</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Argumentative structure analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Argumentation quality analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stance detection</td>
</tr>
<tr>
<td>low</td>
<td>Corpora</td>
<td>English</td>
</tr>
<tr>
<td></td>
<td></td>
<td>German</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>medium</td>
<td>Debate portals</td>
<td>English</td>
</tr>
<tr>
<td></td>
<td></td>
<td>German</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Focus on persuasion</td>
</tr>
<tr>
<td>high</td>
<td>Discussion pages</td>
<td>Controversial issues</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Focus on deliberation</td>
</tr>
<tr>
<td>very high</td>
<td>Articles</td>
<td></td>
</tr>
</tbody>
</table>

* Estimated effort / expertise to exploit a resource of the respective type within own research.
### Argumentation-Related Resources

<table>
<thead>
<tr>
<th>Leverage effort*</th>
<th>Resource type</th>
<th>Example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>very low</strong></td>
<td>Technology</td>
<td>Visual inspection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Acquisition, Tagging</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>low</strong></td>
<td>Corpora</td>
<td>Argumentative structure analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Argumentation quality analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stance detection</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>medium</strong></td>
<td>Debate portals</td>
<td>English</td>
</tr>
<tr>
<td></td>
<td></td>
<td>German</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>high</strong></td>
<td>Discussion pages</td>
<td>Focus on persuasion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controversial issues</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Focus on deliberation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>very high</strong></td>
<td>Articles</td>
<td>Editorials, Essays</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Legal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Scientific publications</td>
</tr>
</tbody>
</table>

* Estimated effort / expertise to exploit a resource of the respective type within own research.
Argumentation-Related Resources

The Argument Web

AIFdb Corpora
Structured argument data in uniform format

AIFdb Search
Search interface for argument resources

ARG-tech API
Several argument web services
Argumentation-Related Resources

The Argument Web

AIFdb Corpora

Structured argument data in uniform format

AIFdb Search

Search interface for argument resources

ARG-tech API

Several argument web services

Argublogging

Widget for argument annotation in blogs

OVA

Online visualization and analysis of arguments

Arvina

Dialogue platform based on AIFdb
## Argument Search Engines

### Vision of Argument Search

<table>
<thead>
<tr>
<th>Search Query</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is time travel possible</td>
<td>NASA Space Place: Time travel is one of my favorite topics! I wrote some time travel stories in junior high school that used a machine of my own invention to travel backwards in time, ...</td>
</tr>
<tr>
<td>Is time travel backwards possible</td>
<td></td>
</tr>
<tr>
<td>Is time travel a paradox?</td>
<td></td>
</tr>
<tr>
<td>Are wormholes possible?</td>
<td></td>
</tr>
<tr>
<td>Can we travel close to the speed of light?</td>
<td></td>
</tr>
</tbody>
</table>

- **We can build a real time machine** - BBC News - BBC.com
  - Travelling in time might sound like a flight of fancy, but some physicists think it might really be possible. BBC Horizon looked at some of the ... [Link](https://www.bbc.com/news/science-environment-44771942)

- **Is Time Travel Possible?** - Explore - physics.org
  - Travelling forwards in time is surprisingly easy. Einstein’s special theory of relativity, developed in 1905, shows that time passes at different rates for people who ... [Link](https://www.physics.org/article-questions)

- **Is time travel possible?** - Tomorrow Today - The Science ... - DW
  - This week’s viewer question comes from Richard Mack’oloo in Dar es Salaam, Tanzania. [Link](https://www.dw.com/is-time-travel-possible)
Argument Search Engines*

Vision of Argument Search

Arguments in future web search:
- support forming opinions
- make it easy to find relevant arguments
- deliberation: learn about other views
- education: learn to debate

Search results should . . .
- rank the best arguments highest
- cover diverse aspects
- cover reliable and heterogeneous sources
- be up-to-the-minute
- be traceable and evaluable

---

Argument Search Engines
Basic Elements and Process

Crawling  Mining  Cleansing  Indexing  Filtering  Ranking  Presentation

Acquisition  Retrieval
Argument Search Engines

Basic Elements and Process

Crawling  Mining  Cleansing  Indexing  Filtering  Ranking  Presentation

Sources  Candidate documents  Candidate arguments  Model-conform arguments

Acquisition

Index  Relevant arguments  Ranked arguments  Argument map

Retrieval

Result
Argument Search Engines
Basic Elements and Process

$\Pi_{\text{argdoc}}$

Crawling  Mining  Cleansing  Indexing

$\Pi_{\text{rel}}$  $\Pi_{\text{rank}}$  

Query

Sources  Candidate documents  Candidate arguments  Model-conform arguments

Index

Retrieval

Relevant arguments

Ranked arguments  Argument map

Result

$\Pi_{\text{argdoc}}$  $\Pi_{\text{rel}}$  $\Pi_{\text{rank}}$
Argument Search Engines

Basic Elements and Process

Acquisition paradigm [Ajjour et al. 2019]:

- distribution of processing steps regarding offline time and online time
- tradeoff between precision, recall, and topicality
Argument Search Engines

Acquisition Paradigms: (a) args.me

- Research focus: argument ranking
- Supervision level: medium (distantly supervised)
- Effectiveness profile: high precision, low recall
- Stance balance: guaranteed
- Efficiency: high
Argument Search Engines
Acquisition Paradigms: (b) IBM Debater

- Research focus: debating technology
- Supervision level: medium (recognized source)

- Effectiveness profile: high precision, high recall on topic
- Stance balance: guaranteed
- Efficiency: high
Argument Search Engines
Acquisition Paradigms: (c) ArgumenText

- **Research focus:** argument mining
- **Supervision level:** low

- **Effectiveness profile:** low precision, high recall
- **Stance balance:** cannot be guaranteed
- **Efficiency:** low
Designing a ranking algorithm:

- Analyze conclusions, premises, or both?
- Use fulltext or elite terms only?
- Exploit metadata and sentiment?
- Analyze relations between arguments?

...
# Argument Search Engines
## Ranking Paradigms in IR

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>Boolean</td>
<td>VSM</td>
<td>FuzzySet</td>
<td>LSI</td>
<td>Genre</td>
<td>DivRand</td>
<td>CL-ESA</td>
</tr>
<tr>
<td>1970</td>
<td>2-Poisson</td>
<td>GVSM</td>
<td>Inquery</td>
<td>BeliefNet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>ProbabilityIndex</td>
<td>BIM</td>
<td>BII</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Empirical Models**

**Probabilistic Models**

**Language Models**

[Stein et al. 2017]
New research indicates that *Divergence from Randomness* and *Language Models* are the currently most effective retrieval models to address $\Pi_{rank}$.

[Pottast et al. 2019]
Argument Search Engines

More on Args  [args.me]

Argument sources:

<table>
<thead>
<tr>
<th>#</th>
<th>Debate Portal</th>
<th>Argument Units</th>
<th>Arguments</th>
<th>Debates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>idebate.org</td>
<td>16 084</td>
<td>15 384</td>
<td>698</td>
</tr>
<tr>
<td>2</td>
<td>debatepedia.org</td>
<td>34 536</td>
<td>33 684</td>
<td>751</td>
</tr>
<tr>
<td>3</td>
<td>debatewise.org</td>
<td>39 576</td>
<td>33 950</td>
<td>2 252</td>
</tr>
<tr>
<td>4</td>
<td>debate.org</td>
<td>210 340</td>
<td>182 198</td>
<td>28 045</td>
</tr>
<tr>
<td>5</td>
<td>forandagainst.com</td>
<td>29 255</td>
<td>26 224</td>
<td>3 038</td>
</tr>
<tr>
<td>∑</td>
<td></td>
<td>329 791</td>
<td>291 440</td>
<td>34 784</td>
</tr>
</tbody>
</table>

Design decisions:

- **Argument model**: conclusion + 1 premise with stance information
- **Query**: free text phrase, interpreted as AND query
- **Retrieval**: exact matching against conclusion
- **Ranking**: BM25F based on conclusion (1.0), premise (0.5), and debate (0.2)
Argument Search Engines
More on Args [args.me]

Top queries (Sep.’17 – Apr.’19):

<table>
<thead>
<tr>
<th>Query</th>
<th>Absolute</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 climate change</td>
<td>251</td>
<td>3.5%</td>
</tr>
<tr>
<td>2 feminism</td>
<td>193</td>
<td>2.7%</td>
</tr>
<tr>
<td>3 abortion</td>
<td>158</td>
<td>2.2%</td>
</tr>
<tr>
<td>4 trump</td>
<td>146</td>
<td>2.0%</td>
</tr>
<tr>
<td>5 brexit</td>
<td>128</td>
<td>1.8%</td>
</tr>
<tr>
<td>6 death penalty</td>
<td>73</td>
<td>1.0%</td>
</tr>
<tr>
<td>7 google</td>
<td>58</td>
<td>0.8%</td>
</tr>
<tr>
<td>8 vegan</td>
<td>57</td>
<td>0.8%</td>
</tr>
<tr>
<td>9 nuclear energy</td>
<td>56</td>
<td>0.8%</td>
</tr>
<tr>
<td>10 donald trump</td>
<td>47</td>
<td>0.7%</td>
</tr>
</tbody>
</table>

Coverage of 1082 Wikipedia controversial issues:
- 78% match with ≥ 1 argument
- 42% match with ≥ 1 conclusion
Argument Search Engines
Presentation and Analytics

PRO
Abortion is the ending of pregnancy by the removal or forcing out from the womb of a fetus or embryo before it is able to survive on its own. Abortion occurs naturally, in which... 
https://www.debate.org/debates/abortion735/ score =

Great, another forfeiter. As someone who has debated... 
https://www.debate.org/debates/abortion735/ score =

This should be fun :) The legalisation of abortion has... 
https://www.debate.org/debates/abortion735/ score =

There are many good and bad sides to abortion. But just... 
https://www.debate.org/debates/abortion735/ score =

Although I oppose abortion in most cases, I accepted this... 
https://www.debate.org/debates/abortion735/ score =

Abortion is needed to control the population so that... 
https://www.debate.org/debates/abortion735/ score =

CON
In 2011 there were about 736,322 abortions reported to... 
https://www.debate.org/debates/abortion548/ score =

The greatest destroyer of peace is abortion because if a... 
https://www.debate.org/debates/abortion548/ score =

Yes the government has the obligation to protect the... 
https://www.debate.org/debates/abortion548/ score =

Thank you, Pro. Negative Case A1. False equivalencin... 
https://www.debate.org/debates/abortion548/ score =

Abortion is wrong! Abortion is gross! Abortion is... 
https://www.debate.org/debates/abortion548/ score =

Attacks: Abortion... 
https://www.debate.org/debates/abortion548/ score =

Thank you to both the audience and my opponent for... 
https://www.debate.org/debates/abortion548/ score =
Abortion is the ending of pregnancy by the removal or\n
The greatest destroyer of peace is abortion because it a…

Yes the government has the obligation to protect the…

Thank you, Pro. Negative CaseA1: False equivalences in the…

Abortion is wrong! Abortion is moral! Abortion is…

Abortion is needed to control the population so that the…

Argument Search Engines
Presentation and Analytics
Argument Search Engines
Presentation and Analytics

This Friday, Canada will undergo its third universal review at the United Nations. This is an important moment when Canada will be held accountable by other UN member states on human rights issues.

**CON**
It has been 29 years since abortion was legalized in this country. In those days abortion was used only where a mother's life or health was endangered. But today I think it has gone too far...

**PRO**
January 28 was the 20th anniversary of the Supreme Court decision that deincriminalized abortion in Canada. The legal decision was a victory for Canadian women...

Poverty
Truth
Cancer
Other
Incest
God
Rape

Woman
Women
XI. IR Applications

- Web Technology
- Web Graph
- Web Crawling
- Web Archiving
- Web Content Extraction
- Near-duplicate Detection
- Link Analysis

- The Treachery of Answers
- Argument Retrieval Problems
- Argument Ranking I
- Argument Ranking II
- Argumentation-Related Resources
- Argument Search Engines

- Argument Search Evaluation
Task: Given two arguments regarding a certain topic, decide whether or not the two arguments have the same stance.

Topic: “Gay marriage should be legalized.”

<table>
<thead>
<tr>
<th>Argument 1</th>
<th>Argument 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marriage is a commitment to love and care for your spouse till death. This is what is heard in all wedding vows. Gays can clearly qualify for marriage according to these vows, and any definition of marriage deduced from these vows.</td>
<td>Marriage is the institution that forms and upholds for society, its values and symbols are related to procreation. To change the definition of marriage to include same-sex couples would destroy its function.</td>
</tr>
</tbody>
</table>
Argument Search Evaluation
Same Side Stance Classification  [sameside.webis.de]

Task: Given two arguments regarding a certain topic,
decide whether or not the two arguments have the same stance.

Topic: “Gay marriage should be legalized.”

Argument 1
Marriage is a commitment to love and care for your spouse till death. This is what is heard in all wedding vows. Gays can clearly qualify for marriage according to these vows, and any definition of marriage deduced from these vows.

Argument 2
Marriage is the institution that forms and upholds for society, its values and symbols are related to procreation. To change the definition of marriage to include same-sex couples would destroy its function.

Argument 1
Marriage is a commitment to love and care for your spouse till death. This is what is heard in all wedding vows. Gays can clearly qualify for marriage according to these vows, and any definition of marriage deduced from these vows.

Argument 2
Gay marriage should be legalized since denying some people the option to marry is discriminatory and creates a second class of citizens.
**Task:** Given two arguments regarding a certain topic, decide whether or not the two arguments have the same stance.

**Topic:** “Gay marriage should be legalized.”

**Argument 1**
Marriage is a commitment to love and care for your spouse till death. This is what is heard in all wedding vows. Gays can clearly qualify for marriage according to these vows, and any definition of marriage deduced from these vows.

**Argument 2**
Marriage is the institution that forms and upholds for society, its values and symbols are related to procreation. To change the definition of marriage to include same-sex couples would destroy its function.

**Conclusion:** Different side

**Argument 1**
Marriage is a commitment to love and care for your spouse till death. This is what is heard in all wedding vows. Gays can clearly qualify for marriage according to these vows, and any definition of marriage deduced from these vows.

**Argument 2**
Gay marriage should be legalized since denying some people the option to marry is discriminatory and creates a second class of citizens.

**Conclusion:** Same side
Argument Search Evaluation
Same Side Stance Classification: Task Rationale

Same side classification needs not to distinguish topic-specific pro- / con-vocabulary.

→ "Only" argument similarity within a stance needs to be assessed.
→ Same side classification may be solved in a topic-agnostic fashion.

Applications:

- measure the bias strength within argumentation
- structure a discussion
- find out who or what is challenging me in a discussion
- filter wrongly labeled stances in a large argument corpus
- . . .
Argument Search Evaluation

Same Side Stance Classification: Tasks Details

Two topics (domains):

1. Should gay marriage be legalized?
2. Should abortion be legalized?

**Within domain setting:**

Training. Instances from both domains.
Test. Instances from both domains.

**Cross domain setting:**

Training. Instances from abortion.
Test. Instances from gay marriage.
Argument Search Evaluation
Same Side Stance Classification: Tasks Details

Two topics (domains):
1. Should gay marriage be legalized?
2. Should abortion be legalized?

Within domain setting:
Training. Instances from both domains.
Test. Instances from both domains.

Cross domain setting:
Training. Instances from abortion.
Test. Instances from gay marriage.

Form of an instance:
1. Name of the topic (domain) $d$.
2. Argument 1 from $A_d$.
3. Argument 2 from $A_d$.
4. One of \{ $\circ=\circ$, $\circ\neq\circ$ \}.

Timeline:
8.6. 2019: Training data online.
# Argument Search Evaluation

## Same Side Stance Classification: Results “Within Domain”

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
<td>Pre</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.90</td>
<td>0.73</td>
<td>0.83</td>
<td>0.79</td>
</tr>
<tr>
<td>Leipzig University</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
<td>0.78</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.73</td>
<td>0.63</td>
<td>0.70</td>
<td>0.64</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.74</td>
<td>0.56</td>
<td>0.68</td>
<td>0.63</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.76</td>
<td>0.35</td>
<td>0.62</td>
<td>0.65</td>
</tr>
<tr>
<td>LMU</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
<td>0.53</td>
</tr>
</tbody>
</table>
### Argument Search Evaluation

#### Same Side Stance Classification: Results “Within Domain”

| Team               | Gay marriage |            | | Abortion |            | | All |            |
|--------------------|---------------|------------| |          |------------| |     |------------|
|                    | Pre | Rec | Acc | Pre | Rec | Acc | Pre | Rec | Acc |
| Trier University   | 0.90 | 0.73 | 0.83 | 0.79 | 0.59 | 0.71 | 0.85 | 0.66 | 0.77 |
| Leipzig University | 0.80 | 0.78 | 0.79 | 0.78 | 0.68 | 0.75 | 0.79 | 0.73 | 0.77 |
| IBM Research       | 0.73 | 0.63 | 0.70 | 0.64 | 0.54 | 0.62 | 0.69 | 0.59 | 0.66 |
| TU Darmstadt       | 0.74 | 0.56 | 0.68 | 0.63 | 0.48 | 0.60 | 0.68 | 0.52 | 0.64 |
| Düsseldorf University | 0.76 | 0.35 | 0.62 | 0.65 | 0.32 | 0.57 | 0.70 | 0.33 | 0.60 |
| LMU                | 0.53 | 1.00 | 0.55 | 0.53 | 1.00 | 0.55 | 0.53 | 1.00 | 0.55 |
| ...                |      |      |     |      |      |     |      |      |     |
## Argument Search Evaluation

### Same Side Stance Classification: Results “Within Domain”

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage</th>
<th></th>
<th></th>
<th>Abortion</th>
<th></th>
<th></th>
<th>All</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.90</td>
<td>0.73</td>
<td>0.83</td>
<td>0.79</td>
<td>0.59</td>
<td>0.71</td>
<td>0.85</td>
<td>0.66</td>
<td>0.77</td>
</tr>
<tr>
<td>Leipzig University</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
<td>0.78</td>
<td>0.68</td>
<td>0.75</td>
<td>0.79</td>
<td>0.73</td>
<td>0.77</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.73</td>
<td>0.63</td>
<td>0.70</td>
<td>0.64</td>
<td>0.54</td>
<td>0.62</td>
<td>0.69</td>
<td>0.59</td>
<td>0.66</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.74</td>
<td>0.56</td>
<td>0.68</td>
<td>0.63</td>
<td>0.48</td>
<td>0.60</td>
<td>0.68</td>
<td>0.52</td>
<td>0.64</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.76</td>
<td>0.35</td>
<td>0.62</td>
<td>0.65</td>
<td>0.32</td>
<td>0.57</td>
<td>0.70</td>
<td>0.33</td>
<td>0.60</td>
</tr>
<tr>
<td>LMU</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
</tr>
</tbody>
</table>

...
## Argument Search Evaluation

### Same Side Stance Classification: Results “Within Domain”

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage</th>
<th>Abortion</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.90</td>
<td>0.73</td>
<td>0.83</td>
</tr>
<tr>
<td>Leipzig University</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.73</td>
<td>0.63</td>
<td>0.70</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.74</td>
<td>0.56</td>
<td>0.68</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.76</td>
<td>0.35</td>
<td>0.62</td>
</tr>
<tr>
<td>LMU</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
</tr>
</tbody>
</table>

...
## Argument Search Evaluation

### Same Side Stance Classification: Results “Within Domain”

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage</th>
<th></th>
<th>Abortion</th>
<th></th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
<td>Pre</td>
<td>Rec</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.90</td>
<td>0.73</td>
<td>0.83</td>
<td>0.79</td>
<td>0.59</td>
</tr>
<tr>
<td>Leipzig University</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
<td>0.78</td>
<td>0.68</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.73</td>
<td>0.63</td>
<td>0.70</td>
<td>0.64</td>
<td>0.54</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.74</td>
<td>0.56</td>
<td>0.68</td>
<td>0.63</td>
<td>0.48</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.76</td>
<td>0.35</td>
<td>0.62</td>
<td>0.65</td>
<td>0.32</td>
</tr>
<tr>
<td>LMU</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
<td>0.53</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Trier University. BERT (large, uncased, sequence length 512), tuning for 3 epochs.
# Argument Search Evaluation

Same Side Stance Classification: Results “Within Domain”

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage</th>
<th>Abortion</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.90</td>
<td>0.73</td>
<td>0.83</td>
</tr>
<tr>
<td>Leipzig University</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.73</td>
<td>0.63</td>
<td>0.70</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.74</td>
<td>0.56</td>
<td>0.68</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.76</td>
<td>0.35</td>
<td>0.62</td>
</tr>
<tr>
<td>LMU</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Leipzig University. BERT (uncased, sequence length 512, tuning for 5 epochs), loss function: sigmoid_binary_crossentropy.
### Argument Search Evaluation

**Same Side Stance Classification: Results “Within Domain”**

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage</th>
<th></th>
<th></th>
<th></th>
<th>Abortion</th>
<th></th>
<th></th>
<th></th>
<th>All</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
<td>Pre</td>
<td>Rec</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.90</td>
<td>0.73</td>
<td>0.83</td>
<td>0.79</td>
<td>0.59</td>
<td>0.71</td>
<td>0.85</td>
<td>0.66</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leipzig University</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
<td>0.78</td>
<td>0.68</td>
<td>0.75</td>
<td>0.79</td>
<td>0.73</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.73</td>
<td>0.63</td>
<td>0.70</td>
<td>0.64</td>
<td>0.54</td>
<td>0.62</td>
<td>0.69</td>
<td>0.59</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.74</td>
<td>0.56</td>
<td>0.68</td>
<td>0.63</td>
<td>0.48</td>
<td>0.60</td>
<td>0.68</td>
<td>0.52</td>
<td>0.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.76</td>
<td>0.35</td>
<td>0.62</td>
<td>0.65</td>
<td>0.32</td>
<td>0.57</td>
<td>0.70</td>
<td>0.33</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMU</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IBM Research. Two BERT models fine-tuned in cascade starting from the vanilla BERT model.
Argument Search Evaluation

Same Side Stance Classification: Results “Within Domain”

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage</th>
<th>Abortion</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.90</td>
<td>0.73</td>
<td>0.83</td>
</tr>
<tr>
<td>Leipzig University</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.73</td>
<td>0.63</td>
<td>0.70</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.74</td>
<td>0.56</td>
<td>0.68</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.76</td>
<td>0.35</td>
<td>0.62</td>
</tr>
<tr>
<td>LMU</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
</tr>
</tbody>
</table>

TU Darmstadt. Microsoft’s Multi-Task Deep Neural Network mt-dnn. Basis for the mt-dnn is BERT (large). No hyper-parameter tuning, 4 epochs.
### Argument Search Evaluation

**Same Side Stance Classification: Results “Within Domain”**

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage</th>
<th></th>
<th>Abortion</th>
<th></th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
<td>Pre</td>
<td>Rec</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.90</td>
<td>0.73</td>
<td>0.83</td>
<td>0.79</td>
<td>0.59</td>
</tr>
<tr>
<td>Leipzig University</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
<td>0.78</td>
<td>0.68</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.73</td>
<td>0.63</td>
<td>0.70</td>
<td>0.64</td>
<td>0.54</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.74</td>
<td>0.56</td>
<td>0.68</td>
<td>0.63</td>
<td>0.48</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td><strong>0.76</strong></td>
<td>0.35</td>
<td>0.62</td>
<td>0.65</td>
<td>0.32</td>
</tr>
<tr>
<td>LMU</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
<td>0.53</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Düsseldorf University. Manhattan LSTM – a siamese network – which measures the similarity of both arguments. Document embeddings via BERT (base, uncased, not fine-tuned, sequence length 512 tokens).
## Argument Search Evaluation

### Same Side Stance Classification: Results “Within Domain”

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage</th>
<th>Abortion</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.90</td>
<td>0.73</td>
<td>0.83</td>
</tr>
<tr>
<td>Leipzig University</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.73</td>
<td>0.63</td>
<td>0.70</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.74</td>
<td>0.56</td>
<td>0.68</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.76</td>
<td>0.35</td>
<td>0.62</td>
</tr>
<tr>
<td>LMU</td>
<td>0.53</td>
<td>1.00</td>
<td>0.55</td>
</tr>
</tbody>
</table>

LMU. Bert (base). Arguments organized as graph: edges are weighted with the confidence that arguments agree and confidence that they disagree. If known from training set that the arguments agree or disagree the confidence is 0 and 1 or 1 and 0 accordingly.
### Argument Search Evaluation

**Same Side Stance Classification: Results “Cross Domain”**

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage (large)</th>
<th>Gay marriage (small)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
</tr>
<tr>
<td>LMU</td>
<td>0.67</td>
<td>0.53</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.64</td>
<td>0.59</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.62</td>
<td>0.49</td>
</tr>
<tr>
<td>Paderborn University</td>
<td>0.60</td>
<td>0.38</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.69</td>
<td>0.16</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.72</td>
<td>0.53</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Argument Search Evaluation

**Same Side Stance Classification: Results “Cross Domain”**

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage (large)</th>
<th></th>
<th>Gay marriage (small)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>Acc</td>
<td>Pre</td>
</tr>
<tr>
<td>LMU</td>
<td>0.67</td>
<td>0.53</td>
<td>0.63</td>
<td>0.78</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.64</td>
<td>0.59</td>
<td>0.63</td>
<td>0.71</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.62</td>
<td>0.49</td>
<td>0.60</td>
<td>0.74</td>
</tr>
<tr>
<td>Paderborn University</td>
<td>0.60</td>
<td>0.38</td>
<td>0.56</td>
<td>0.79</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.69</td>
<td>0.16</td>
<td>0.54</td>
<td>1.00</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.72</td>
<td>0.53</td>
<td>0.66</td>
<td>0.68</td>
</tr>
</tbody>
</table>

...
## Argument Search Evaluation

**Same Side Stance Classification: Results “Cross Domain”**

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage (large)</th>
<th>Gay marriage (small)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
</tr>
<tr>
<td>LMU</td>
<td>0.67</td>
<td>0.53</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.64</td>
<td>0.59</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.62</td>
<td>0.49</td>
</tr>
<tr>
<td>Paderborn University</td>
<td>0.60</td>
<td>0.38</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.69</td>
<td>0.16</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.72</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

...
## Argument Search Evaluation

### Same Side Stance Classification: Results “Cross Domain”

<table>
<thead>
<tr>
<th>Team</th>
<th>Gay marriage (large)</th>
<th>Gay marriage (small)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
</tr>
<tr>
<td>LMU</td>
<td>0.67</td>
<td>0.53</td>
</tr>
<tr>
<td>TU Darmstadt</td>
<td>0.64</td>
<td>0.59</td>
</tr>
<tr>
<td>IBM Research</td>
<td>0.62</td>
<td>0.49</td>
</tr>
<tr>
<td>Paderborn University</td>
<td>0.60</td>
<td>0.38</td>
</tr>
<tr>
<td>Trier University</td>
<td>0.69</td>
<td>0.16</td>
</tr>
<tr>
<td>Düsseldorf University</td>
<td>0.72</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Most of the submitted classifiers are robust regarding:

- imbalances between domain proportions in training and test
- imbalances between domain proportions within test
- imbalances between same side / different side proportions
Argument Search Evaluation
Argument Retrieval Task @ CLEF 2020 [touche.webis.de]

Task 1: Supporting argumentative conversations
- Scenario: Users search for arguments on controversial topics
- Task: Retrieve “strong” pro/con arguments on the topic
- Data: 300,000 “arguments” (short text passages)

Task 2: Answering comparative questions with arguments
- Scenario: Users face personal decisions from everyday life
- Task: Retrieve arguments for “Is X better than Y for Z?”
- Data: ClueWeb12 or ChatNoir [chatnoir.eu]

- Run submissions similar to “classical” TREC tracks
- Software submissions via TIRA [tira.io]
Touché @ CLEF 2020
1st Shared Task on Argument Retrieval

Alexander Bondarenko
Matthias Hagen
Martin Potthast
Henning Wachsmuth
Meriem Beloucif
Chris Biemann
Alexander Panchenko
Benno Stein

[touche.webis.de]