

# Visit the Dome

Extracting arguments in parliamentary debates

Anh Phuong Le, Lukas Peter Trautner, Ruta Bakhda, Thang Nguyen, Yamen Ajjour, Wei-Fan Chen

## Parliamentary Debates

Citizens care about **societal issues** and await changes by their governments. To discuss those problems, governments frequently have **debates** in their parliament.

Some countries publish those debates, yet they are often **not easy to comprehend**. In this project, we analyze those debates and present the **essential arguments** to internet users.

## Data Preprocessing

We obtain a record of **3,796,847 speeches** from the **Canadian Parliament** and combine several methods to extract only those concerning **controversial issues (21,332 speeches)**.

**Methods used:**

- Cleansing data (e.g. removing duplicates/clichés)
- Finding controversial topics
- Sentiment analysis

## Argument Mining

**Manual Annotation:**  
We set up a tool to manually annotate argumentative texts (**claim/premise**) and **stance** of an entire speech towards a topic.

**Predicting** argumentative texts:  
Using annotated datasets from **other domains**, we **train** several machine learning models with different combination of features and **test** them.

**Statistics** (after cleaning)

- Online Debates (47,427 sentences)
- Student Essays (6,089 sentences)
- Parliamentary Debates (108 sentences sampled)

|          |                       | Training Set   |                |
|----------|-----------------------|----------------|----------------|
|          |                       | Online Debates | Student Essays |
| Test Set | Online Debates        | 0.92 / 0.84    | 0.80 / 0.51    |
|          | Student Essays        | 0.47 / 0.55    | 0.76 / 0.66    |
|          | Parliamentary Debates | 0.72 / 0.17    | 0.79 / 0.15    |

**Features used:** syntactic, semantic and positional

Best Results (**Accuracy** and **F<sub>1</sub> score**) of In-/Cross-domain experiments

## Integrating into args.me

We **integrate** the results into the **argument search engine** args.me, together with arguments from online debate portals. Also, we **highlight** statements along with the details of the speakers.

