Integrating Arguments From Parliamentary Debates Into Args.me
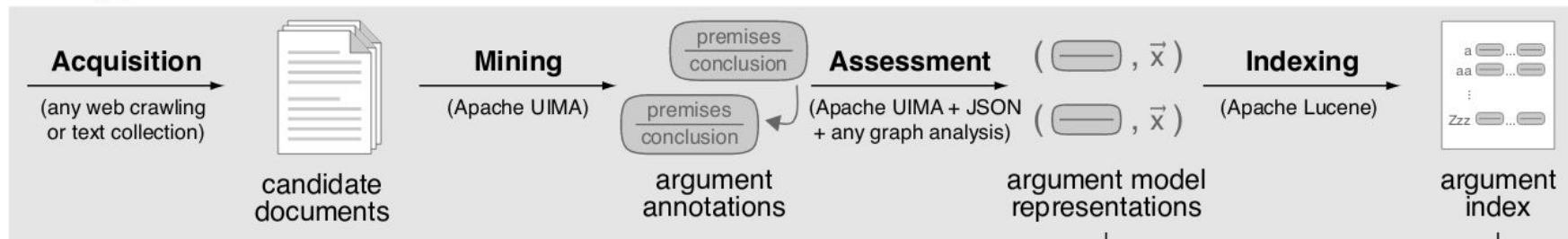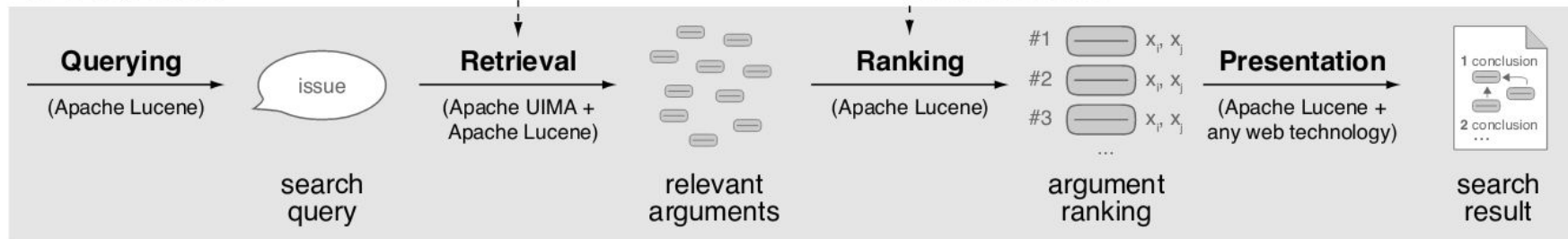
Lukas Trautner

Args - Pipeline

Args - Acquisition

- Acquire data that may contain arguments
- Input:
  - File containing labels for sentences
  - File containing the actual sentences
- Fetches metadata from original corpus and merges it with the labels and sentences based on the "basepk" of a speech
- Outputs a html-file for each speech

```html
<!DOCTYPE html>
<html>
<head>
<title>4128088</title>
<meta charset="UTF-8">
<meta name="speechdate" content="1999-06-09T12:00:00Z"/>
<meta name="role" content="Government"/>
<meta name="speakername" content="Mr. Mac Harb (Ottawa Centre, Lib.)"/>
<meta name="topic" content="Saudi Arabia"/>
<meta name="annotationdate" content="2019-06-21T14:18:41Z"/>
<meta name="stance" content="pro"/>
<meta name="party" content="Liberal"/>
</head>
<body>
<div>
<p class="non-argumentative">this year marks the centennial anniversary of the proclamation of the kingdom of saudi arabia.</p>
<p class="conclusion">the leadership, vision and commitment of the government of saudi arabia and its people have created a society where education, economic development and diversification have become priorities.</p>
<p class="premise">the development of infrastructure and the petrochemical industry have transformed a barren land into one of the most advanced nations of the world.</p>
<p class="premise">he saudi arabian people have built a moderate nation that is dedicated to promoting peace and stability.</p>
<p class="non-argumentative">on behalf of my colleagues and all canadians, i ask his excellency dr. mohammed al-hussaini, the ambassador of saudi arabia, to extend our warmest congratulations to the government of saudi arabia and its people, the people of this exceptional nation.</p>
<p class="non-argumentative">i would like to wish them a happy anniversary.</p>
</div>
</body>
</html>
```

Args - Acquisition - Why html?

All the data of the speeches are merged into html-files, because

- the args-pipeline was designed for working with html
- there already exists a wrapper for html that makes processing of html-files with UIMA very easy
- the processing of e.g csv-files would require a new wrapper for it or some rather complicated processing in the mining-step

Args - Acquisition - Metadata

The metadata that is currently stored consists of

- speakername - the name of the speaker
- speakerparty - the political party of a speaker
- speakerrole - the role of the speaker, e.g. "government" or "opposition"
- speechdate - the date when the speech was held
- topic - the topic of a speech
- stance - whether the stance of the speech is "pro" or "con"
- annotationdate - when the speech was annotated (this is necessary to conform to the acquire() interface of the acquisition step)

Args - Mining

- In the mining-step of the pipeline UIMA-annotations are added
- Makes use of an html-wrapper for the CAS, which adds css-functionality to UIMA, to select the elements that we want to label as "premise" or "conclusion"
- A new collection reader was required to read html-files
- A new analysis engine was required to process the, in the acquisition step, created html-files

Args - Mining - Collection Reader

To process the html-files, that were produced in the acquisition-step, a new collection reader was required, because the args-pipeline could only process warc-files beforehand.

UimaParliamentaryDebateReader is the collection reader for parliamentary debates.

It reads the html-files and converts it into UIMA-format (.xmi) and also converts the metadata in the html-files into "MetadataArgs"-annotations, which are used by the following steps of the pipeline.

This required an extension of the "MetadataArgs"-annotation in the Argumentation-Typesystem.

Args - Mining - Analysis Engine

ParliamentaryDebateAnalysisEngine

The analysis engine used for the parliamentary debates adds the "Argument"-annotations.

Currently each speech is labeled as one argument with the first conclusion in the speech as the conclusion and the whole text as the premise.

The premise consists of the whole speech, because otherwise we are not able to display the full speech, because args is only able to display arguments.

Args - Mining - Output

Html-Wrapper
Metadata
Content + Argument Annotations

```xml
<?xml version="1.0" encoding="UTF-8"?><xmi:XMI xmlns:supertype="http:///de/aitools/ie/uima/type/supertype.ecore" xmlns:tcas="http:///uima/tcas.ecore" xmlns:xmi="http://www.omg.org/XMI" xmlns:cas="http:///uima/cas.ecore" xmlns:warc="http:///de/aitools/ie/uima/type/warc.ecore" xmlns:argumentation="http:///de/aitools/ie/uima/type/argumentation.ecore" xmlns:core="http:///de/aitools/ie/uima/type/core.ecore" xmi:version="2.0"><cas:NULL xmi:id="0"/><tcas:DocumentAnnotation xmi:id="8" sofa="1" begin="0" end="833" language="x-unspecified"/><warc:HtmlElement xmi:id="43" sofa="1" begin="0" end="833" xpath="/html[1]/body[1]/div[1]"/><warc:HtmlElement xmi:id="48" sofa="1" begin="0" end="833" xpath="/html[1]/body[1]"/><warc:HtmlElement xmi:id="13" sofa="1" begin="0" end="94" xpath="/html[1]/body[1]/div[1]/p[1]"/><warc:HtmlElement xmi:id="18" sofa="1" begin="94" end="285" xpath="/html[1]/body[1]/div[1]/p[2]"/><warc:HtmlElement xmi:id="23" sofa="1" begin="285" end="432" xpath="/html[1]/body[1]/div[1]/p[3]"/><warc:HtmlElement xmi:id="28" sofa="1" begin="432" end="536" xpath="/html[1]/body[1]/div[1]/p[4]"/><warc:HtmlElement xmi:id="33" sofa="1" begin="536" end="787" xpath="/html[1]/body[1]/div[1]/p[5]"/><warc:HtmlElement xmi:id="38" sofa="1" begin="787" end="833" xpath="/html[1]/body[1]/div[1]/p[6]"/><warc:HtmlSource xmi:id="53" sofa="1" begin="0" end="0" source="&lt;!doctype html&gt;&#10;&lt;html&gt;&#10;  &lt;head&gt;&#10;    &lt;title&gt;4128088&lt;/title&gt;&#10;    &lt;meta charset=&quot;UTF-8&quot;&gt;&#10;    &lt;meta name=&quot;speechdate&quot; content=&quot;1999-06-09T12:00:00Z&quot;&gt;&#10;    &lt;meta name=&quot;role&quot; content=&quot;Government&quot;&gt;&#10;    &lt;meta name=&quot;speakername&quot; content=&quot;Mr. Mac Harb (Ottawa Centre, Lib.)&quot;&gt;&#10;    &lt;meta name=&quot;topic&quot; content=&quot;Saudi Arabia&quot;&gt;&#10;    &lt;meta name=&quot;annotationdate&quot; content=&quot;2019-06-21T14:18:41Z&quot;&gt;&#10;    &lt;meta name=&quot;stance&quot; content=&quot;pro&quot;&gt;&#10;    &lt;meta name=&quot;party&quot; content=&quot;Liberal&quot;&gt;&#10;  &lt;/head&gt;&#10; &lt;body&gt;&#10;   &lt;div&gt;&#10;    &lt;p class=&quot;non-argumentative&quot;&gt;this year marks the centennial anniversary of the proclamation of the kingdom of saudi arabia.&lt;/p&gt;&#10;    &lt;p class=&quot;conclusion&quot;&gt;the leadership, vision and commitment of the government of saudi arabia and its people have created a society where education, economic development and diversification have become priorities.&lt;/p&gt;&#10;    &lt;p class=&quot;premise&quot;&gt;the development of infrastructure and the petrochemical industry have transformed a barren land into one of the most advanced nations of the world.&lt;/p&gt;&#10;    &lt;p class=&quot;premise&quot;&gt;he saudi arabian people have built a moderate nation that is dedicated to promoting peace and stability.&lt;/p&gt;&#10;    &lt;p class=&quot;non-argumentative&quot;&gt;on behalf of my colleagues and all canadians, i ask his excellency dr. mohammed al-hussaini, the ambassador of saudi arabia, to extend our warmest congratulations to the government of saudi arabia and its people, the people of this exceptional nation.&lt;/p&gt;&#10;    &lt;p class=&quot;non-argumentative&quot;&gt;i would like to wish them a happy anniversary.&lt;/p&gt;&#10;   &lt;/div&gt;&#10; &lt;/body&gt;&#10;&lt;/html&gt;"/><argumentation:MetadataArgs xmi:id="58" sofa="1" begin="0" end="0" title="4128088" crawlingDate="2019-06-21T14:18:41Z" url="http://www.parliamentary-debates.com/basepk-4128088" speakerName="Mr. Mac Harb (Ottawa Centre, Lib.)" speakerRole="Government" topic="Saudi Arabia" speechDate="1999-06-09T12:00:00Z" speakerParty="Liberal"/><argumentation:Argument xmi:id="80" sofa="1" begin="0" end="833" premises="75" conclusion="70"><stances>PRO</stances></argumentation:Argument><cas:Sofa xmi:id="1" sofaNum="1" sofaID="InitialView" mimeType="text" sofaString="this year marks the centennial anniversary of the proclamation of the kingdom of saudi arabia.the leadership, vision and commitment of the government of saudi arabia and its people have created a society where education, economic development and diversification have become priorities.the development of infrastructure and the petrochemical industry have transformed a barren land into one of the most advanced nations of the world.he saudi arabian people have built a moderate nation that is dedicated to promoting peace and stability.on behalf of my colleagues and all canadians, i ask his excellency dr. mohammed al-hussaini, the ambassador of saudi arabia, to extend our warmest congratulations to the government of saudi arabia and its people, the people of this exceptional nation.i would like to wish them a happy anniversary."/><argumentation:ArgumentativeDiscourseUnit xmi:id="75" sofa="1" begin="0" end="833"/><argumentation:ArgumentativeDiscourseUnit xmi:id="70" sofa="1" begin="94" end="285"/><cas:View sofa="1" members="8 43 48 13 18 23 28 33 38 53 58 80"/></xmi:XMI>
```

url is now "https://args.me/argument/<Integer.toHexString(basepk.hashCode())>"

Args - Assessment

In the assessment-step of the args-pipeline arguments can be enriched with additional metadata, but this is currently not done.

Also the UIMA-annotations from the mining step are converted into me.args.Argument's, which are then converted into a json-file for the indexing process.

To convert the additional metadata of the parliamentary debates some new converters were implemented, one for each kind of metadata.

Args - Assessment - Output

```
{
  "arguments": [
    {
      "id": "e1de3515-2019-06-21T14:18:41Z-00000-000",
      "conclusion": "the leadership, vision and commitment of the government of saudi arabia and its people have created a society where education, economic development and diversification have become priorities.",
      "premises": [
        {
          "text": "this year marks the centennial anniversary of the proclamation of the kingdom of saudi arabia.the leadership, vision and commitment of the government of saudi arabia and its people have created a society where education, economic development and diversification have become priorities.the development of infrastructure and the petrochemical industry have transformed a barren land into one of the most advanced nations of the world.he saudi arabian people have built a moderate nation that is dedicated to promoting peace and stability.on behalf of my colleagues and all canadians, i ask his excellency dr. mohammed al-hussaini, the ambassador of saudi arabia, to extend our warmest congratulations to the government of saudi arabia and its people, the people of this exceptional nation.i would like to wish them a happy anniversary.",
          "stance": "PRO"
        }
      ],
      "context": {
        "acquisitionTime": "2019-06-21T14:18:41Z",
        "previousArgumentInSourceId": "",
        "nextArgumentInSourceId": "",
        "sourceId": "e1de3515-2019-06-21T14:18:41Z",
        "sourceUrl": "http://www.parliamentary-debates.com/basepk-4128088",
        "sourceTitle": "4128088",
        "speakerName": "Mr. Mac Harb (Ottawa Centre, Lib.)",
        "speakerParty": "Liberal",
        "speakerRole": "Government",
        "speechDate": "1999-06-09T12:00:00Z",
        "sourceText": "this year marks the centennial anniversary of the proclamation of the kingdom of saudi arabia.the leadership, vision and commitment of the government of saudi arabia and its people have created a society where education, economic development and diversification have become priorities.the development of infrastructure and the petrochemical industry have transformed a barren land into one of the most advanced nations of the world.he saudi arabian people have built a moderate nation that is dedicated to promoting peace and stability.on behalf of my colleagues and all canadians, i ask his excellency dr. mohammed al-hussaini, the ambassador of saudi arabia, to extend our warmest congratulations to the government of saudi arabia and its people, the people of this exceptional nation.i would like to wish them a happy anniversary.",
        "sourcTextConclusionStart": 94,
        "sourceTextConclusionEnd": 285,
        "sourceTextPremiseStart": 0,
        "sourceTextPremiseEnd": 833
      }
    }
  ]
}
```

Args - Indexing

Create an index of all assessed arguments, based on the full text of each argument, by using Apache Lucene.

This did not require any changes.

Arguments are currently indexed by the "sourceText". This may not be ideal, because adding the "sourceText" in the assessment-step, makes its output file unnecessarily large, but it should suffice for the moment.

Args - Retrieval

Based on a query, specified by the user, arguments are retrieved from the index.

The retrieved arguments are then ranked and presented to the user.

Currently the retrieval process can only be executed from the command line.

Also no "real" ranker is implemented and the retrieved arguments are returned to the user as they are retrieved from the index.

Args - Retrieval - Example

The query "arabia" retrieves this argument:

```
Arguments 1 - 1 of 1 for query 'arabia' retrieved in 0.033 seconds
With presentation.offset = 10
With presentation.length = 5
--------------------
1) the leadership, vision and commitment of the government of saudi arabia and its people have created a society where education, economic development and
diversification have become priorities.
---
PRO this year marks the centennial anniversary of the proclamation of the kingdom of saudi arabia.the leadership, vision and commitment of the government of saudi
arabia and its people have created a society where education, economic development and diversification have become priorities.the development of infrastructure and the
petrochemical industry have transformed a barren land into one of the most advanced nations of the world.he saudi arabian people have built a moderate nation that is
dedicated to promoting peace and stability.on behalf of my colleagues and all canadians, i ask his excellency dr. mohammed al-hussaini, the ambassador of saudi arabia,
to extend our warmest congratulations to the government of saudi arabia and its people, the people of this exceptional nation.i would like to wish them a happy
anniversary.
--------------------
```

(Other output formats, with more information about the retrieved arguments, like json or html are also available)

Args - Problems

- Indexing arguments from online debate portals alongside arguments from the parliamentary debate may result in some conflicts
- If we index all arguments (debate-portals + parliamentary debates) it could be hard to find the right query to show them at the "summaery", to solve this a ranker could be implemented that ranks arguments from the parliamentary debates higher than the arguments from the online debate portals for demonstration purposes
- Presentation for arguments from parliamentary debates needs to be adjusted, because they can not be presented as the arguments from online debate portals

Args - Presentation

Pro

If all drugs were legal, addicts would no longer...
http://www.debate.org/debates/Drugs/7/
If all **drugs** were legal, addicts would no longer pay black-market prices to criminals for **drugs** of questionable and dangerous origin. They would get **drugs** produced by legitimate ... ▾ score

Con

Ok, I am against Drugs being legalised; therefore,...
http://www.debate.org/debates/Drugs/7/
Ok, I am against **Drugs** being legalised; therefore, I shall be arguing that **drugs** should be illegal ... of **drugs** in some countries have shown to be effective. Recent goverment ... ▾ score

The arguments retrieved from the parliamentary debates link to the webpage they stem from and the full argument can not be accessed on args.me directly.

Linking to the original webpage is obviously not possible for the arguments for parliamentary debates.

A solution for this would be to dynamically create webpages for the arguments from parliamentary debates and link to those pages, where we then can display the whole speech and the metadata about that speech.