Handbook

Parliament Browser

Erik Brandes, Roman Shumliakivskyi, Nicholas Getta

Contents

1 Server-Side Instructions

1.1 Before Starting the Program

There are two things that need to be done before starting the program for it to work properly:

1. The user must have a TeX-Distribution (such as TeXLive) installed on their system, otherwise the downloadable PDFs can not be created.

2. The user must run docker and a docker container each for SpaCy, GerVader, and ParlBERT. By default, these programs expect to be running on ports 1000, 1001, and 1002 respectively. Optionally, the docker image for SpaCy can be modified to support GPU usage, which will significantly increase performance. Download https://github.com/texttechnologylab/duui-uima, go to duui-spacy/docker_build_single.sh and change `"de_core_news_sm"` to `"de_core_news_lg"`. Then, execute docker_build_single.sh by calling it in the shell. Once this is done, you can run the image.
   Run the following three commands in a shell:

   ```
   docker run --gpus all --rm -p 1000:9714 docker.texttechnologylab.org/
   textimager-duui-spacy-single-de_core_news_lg:0.1.6

   docker run --rm -p 1001:9714 docker.texttechnologylab.org/gervader_duui:latest

   docker run --rm -p 1002:9714 docker.texttechnologylab.org/parlbert-topic-german:latest
   ```

   Once "Application startup complete." is printed out for all of them, the application can be started.

1.2 Starting the Program

The program has three optional parameters:

1. -db (*path*): add the path to a text file with MongoDB credentials and an encryption key for a JWT token. -db is a mandatory parameter. If -db is not used the program will not connect to any MongoDB by default.

2. -ws: starts the web service. Even f -ws is not used the program can still read protocol data and insert into the database.

3. -data (*path*): add the path to a folder with the members of parliament's metadata in an XML-File. All necessary data (protocols, speeches, comments, metadata etc.) will then be parsed and saved in the database.

1.3 After Starting the Program

Visit http://localhost:4567 via any web browser.

2 Instructions for Users of the Web Service

The application's main feature is the display of Bundestag-speeches, analysed and visualized with different natural language processing points of view in mind. To navigate, enter a speaker's name into the search field on the main site. You will then be redirected to the speaker's metadata-site (this might take a short while) and can there choose from their speeches. On a speech's site you will see:

- Some basic information about the speech.

- The speech's full text. Indicators at the end of each sentence reflect that sentence's sentiment. Named Entities (locations, people, organisations) are colour-coded.
- All comments made during the speech.
- Expandable and collapsable windows containing graphic visualisations about some of the speech's properties.
- A preview of the speech as a PDF, which you can download by clicking the button below.

Additionally, users can log into the site. Usernames must contain only letters and numbers and have between 6 and 24 characters, passwords much have between 8 and 24 characters. Once a user is logged in, they can edit a speech's text. Clicking the edit-button will open a window in which the user can make any changes to the text, confirming them will update the speech permanently. If the edit is successful, the page will reload and reflect any changes made.

New Protocols (should there have been new debates in the Bundestag) can be loaded in from the main page. This might take a while.

In theory, a full text search over all speeches stored in the database can be performed from the main page by entering a search query into the designated field. This feature is however currently not working properly, as the amount of data that needs to be looked through by this operation is simply to large and requires too much memory.

An infographic about the amount of speeches held by each speaker can also be accessen via the main page.

Disclaimer: A frequent reason for an Internal Server Error in the webapp is a MongoSocketOpenException, which is usually fixed by clicking the same link again.

3 Promo-Video

There is a promo video here: https://ppr.gitlab.texttechnologylab.org/ErikBrandes/parliament_browser_3_5/-/blob/aacf825eb3774962f804d18f193c305b533adab2/promo_video.mov.