# Scenario

In this scenario we want to build an application that provides insights on newspaper and poll results. We scrape data from the respective website, process it so that we can, for instance, analyze how often a politician's name was mentioned in the news during one day/week/month. There are several approaches of analyzing the data. As mentioned before, the mentions are counted. Furthermore, the sentiment towards politicians/parties is measured. Mentions and sentiment are correlated with current poll results. It is analyzed in what context specific politicians/parties are mentioned most often and what topics they are associated with. The final results are made accessible via an API but also in form of a regular report.



The statistics team is responsible for the quantitative part of application. They take the poll results and the newspaper article and count/measure the mentions/sentiment for each politician/party

### Scraper

The scraper team is responsible for gathering the data. Their sources are newspaper websites and the API of dawum (poll results).



#### **External Sources**

External sources do not directly belong to the application. They are the newspaper websites and dawum (website that provides the poll results).

#### Semantics

The semantics team is responsible for the qualitative analysis of the data. For instance, what are the words that appear most often with some politician or what is the context in which a politician is mentioned. Their output could be a wordcloud or a topic model.

## Interface

The interface team is the one who is responsible for delivering the results to the audience. This can happen by making the results available via an API or by producing a regular report that appears on a regular basis.

