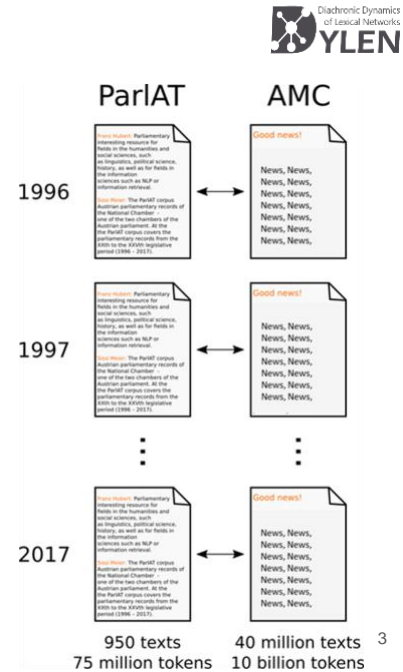


# Supplement W4

## Data abstraction

# Data

- We use two corpora spanning 20 years
  - **ParlAT: Austrian parliamentary records**  
(950 texts, 75 million token)
    - ❖ Speaker (Politicians, Party)
    - ❖ Date (Legislative Period)
    - ❖ POS
  - **AMC: Austrian media corpus**  
(40 million texts, 10 billion tokens)
    - ❖ Name of the medium
    - ❖ Section (e.g. politics, sport)
    - ❖ Date
    - ❖ POS



## Timetable

- 01:00 pm – 01:30 pm: Introduction  
01:30 pm – 03:00 pm: Wishful Thinking  
03:00 pm – 03:20 pm: Break  
03:20 pm – 03:40 pm: Visualization Analogies  
03:40 pm – 05:30 pm: Storyboarding  
05:30 pm – 06:00 pm: Conclusion and Closing