

## Annotation Goal and Presentations (3/9, 3/11)

As a group you will present your annotation goal for the semester. Each group will have approximately 10 minutes. The presentation should include:

- Planned dataset, or a few possible datasets if you haven't firmly decided on one yet
- What is the sampling technique for the dataset? Do you plan on further subsampling from the dataset?
- Balance of documents
  - Genre, time period, authorship?
- Source
  - Where did the data come from?
- Size of data
  - How big is the dataset?
  - Measured in GBs? Sentences? Tokens?
- Copyright or license of your dataset
  - Is it publicly available?
  - Are there restrictions on redistribution or derivative works?
- Is there already any annotation of your dataset? If so, what kind?
- Planned task
  - Clear statement of annotation task
- What do you hope to achieve by annotating for this task?
- Discussion of 2-3 relevant papers (similar corpus, similar annotation goals, etc).

## Submission:

You will submit your slides for the presentation on LATTE by 3/9/26 @ 9:00am.