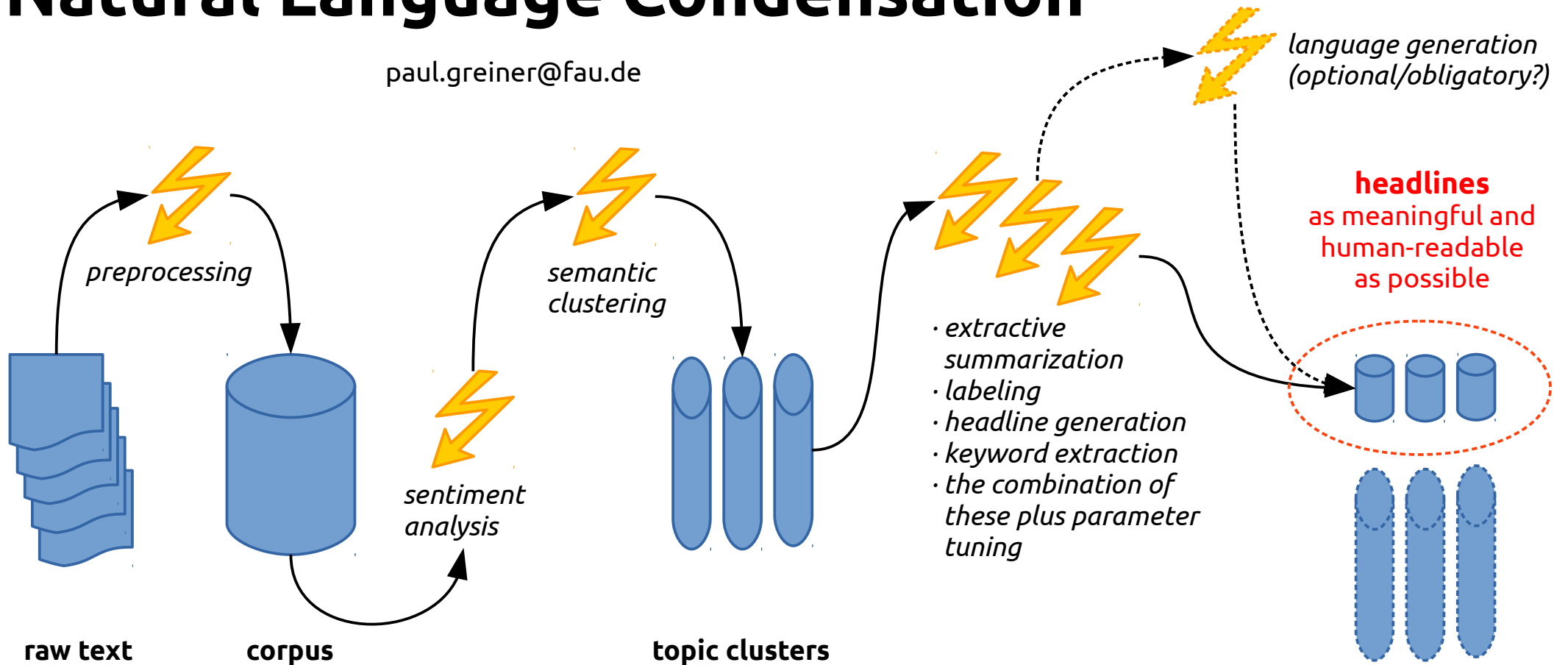


Natural Language Condensation

paul.greiner@fau.de



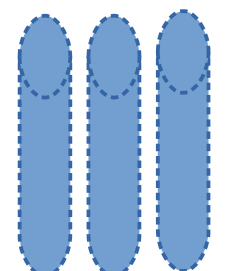
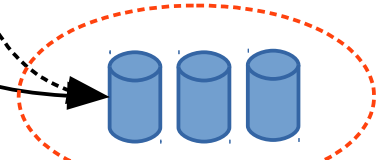
raw text

corpus

topic clusters

headlines
as meaningful and
human-readable
as possible

- extractive summarization
- labeling
- headline generation
- keyword extraction
- the combination of these plus parameter tuning



source

TKE

thesis

goal

(opinion polls,
potentially others)

Miscellaneous notes:

objective (vs. subjective human-generated) · abstractive (idiomatic labels based on extractive techniques, language generation) · discard noise & compress information

Working title:

"A Human Readable Topic Clustering – An Applied Case Study on NLC"

If this shouldn't

work: Add
towards to the
title ;)