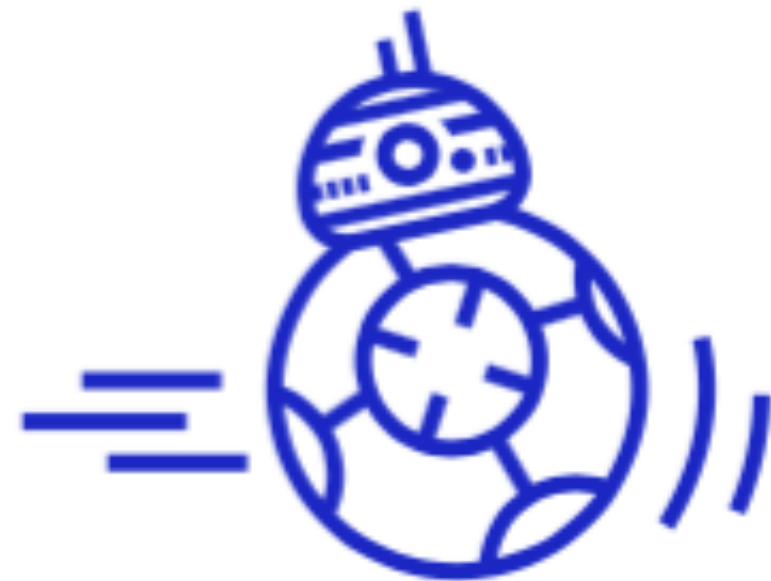


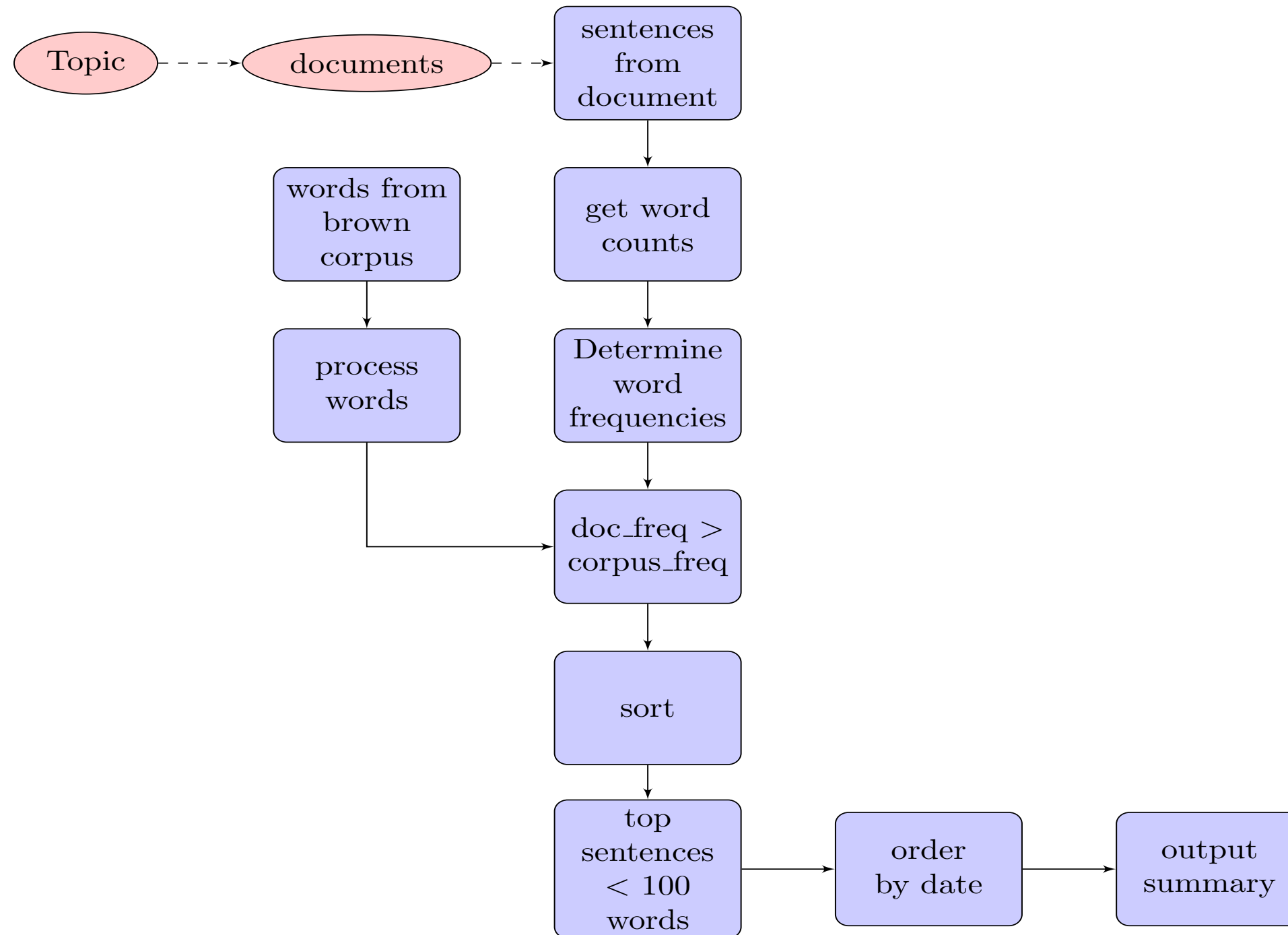
Deliverables 2

Matt Calderwood

Kirk LaBuda

Nick Monaco





System Overview

- Content Selection - after preprocessing, tf*idf comparison with lemmatized Brown corpus.
- Choose sentences with highest score (thematizing sentences)



System Overview

- Information Ordering - Order the sentences according to date of publication. (Ad hoc heuristic.)
- Content Realization - keep summaries under 100 words.



Issues and Successes

- 1 ROUGE-1 Average_R: 0.10987** (95%-conf.int. 0.09229 - 0.12813)
- 1 ROUGE-2 Average_R: 0.01891** (95%-conf.int. 0.01412 - 0.02389)
- 1 ROUGE-3 Average_R: 0.00502** (95%-conf.int. 0.00317 - 0.00720)
- 1 ROUGE-4 Average_R: 0.00129** (95%-conf.int. 0.00039 - 0.00242)

