

1) Progress made thus far

- We have obtained the data, cleaned and filtered only articles containing “Gore” or “Bush” to start with
- We have mostly understood the whole paper and identified the approach to implement the PLSA with Dirichlet prior
- We have split up the remaining work

2) Remaining tasks

- Further clean the data into only paragraphs containing “Gore” and “Bush” and remove all non-alphanumeric characters
- Work on the granger causality test combining the time series data with the word counts of document data
- Calculated the prior according the formula specified in the paper and update the maximization step of PLSA with the given μ and prior previously calculated

3) Any challenges/issues being faced

- For now we are not running into any execution issues yet