Readings: "Processing Text and Text as Data"

Required reading:

Hase, V. (2023). Automated Content Analysis. In F. Oehmer-Pedrazzi, S. H. Kessler, E. Humprecht, K. Sommer, & L. Castro (Eds.), *Standardisierte Inhaltsanalyse in der Kommunikationswissenschaft – Standardized Content Analysis in Communication Research* (pp. 23–36). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36179-2_3

Additional sources:

Benoit, K. (2020). Text as Data: An Overview. L. Cuirini & R. Franzese (eds.), *Handbook of Research Methods in Political Science and International Relations* (pp. 461–497). Thousand Oaks: Sage.

Chai, C. P. (2023). Comparison of Text Preprocessing Methods. *Natural Language Engineering*, *29*(3), 509–553. https://doi.org/10.1017/S1351324922000213

Denny, M. J., & Spirling, A. (2018). Text Preprocessing For Unsupervised Learning: Why It Matters, When It Misleads, and What to Do about It. *Political Analysis*, *26*(2), 168–189. https://doi.org/10.1017/pan.2017.44

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2022). *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Chapter 6 on "Bag of Words". Princeton University Press.

Van Atteveldt, W., Trilling, D., & Calderón, C. A. (2022). *Computational Analysis of Communication*. Chapter 9 on "Processing Text" and Chapter 10 on "Text as Data". Wiley Blackwell. Link to online chapter 9 and Link to online chapter 10.

Welbers, K., van Atteveldt, W., & Benoit, K. (2017). Text Analysis in R. *Communication Methods and Measures*, *11*(4), 245–265. https://doi.org/10.1080/19312458.2017.1387238