

Quantitative Textanalyse in den Politikwissenschaften: Grundlagen, Methoden und aktuelle Anwendungen

Freie Universität Berlin,
Hauptseminar, SS 2020
Wöchentliche Online Sitzungen

Dr. Christian Rauh
Wissenschaftszentrum Berlin für Sozialforschung
www.christian-rauh.eu
christian.rauh@wzb.eu

Stand: 19.04.2020

***Die FU konnte mich bisher nicht mit Zugangsdaten ausstatten,
bitte prüfen Sie den Syllabus regelmäßig auf Updates!***

A. Kursbeschreibung und Lernziele

Politik findet durch und in Texten statt. Reden und Presserklärungen, Debatten und Parteiprogramme, Gesetze und Positionspapiere oder aber Zeitungsberichte und Blogs verraten viel über inhaltliche Prioritäten, Positionen und die Macht politischer Akteure. Solche Texte liefern die relevante Information aber häufig nur in unstrukturierter und interpretationsbedürftiger Form. Das Seminar bietet den TeilnehmerInnen deshalb einen einführenden Überblick zu Methoden, mit denen sie systematische, vergleichende Informationen zu eigenen Fragestellungen aus größeren Textmengen gewinnen können.

Der Kurs gliedert sich dazu in drei Blöcke. In einem ersten Teil besprechen wir die Grundlagen der Inhalts- und insbesondere der Textanalyse anhand einzelner Lehrbuchkapitel. Themen sind der Aufbau eines typischen Forschungsprojekts, das Sampling relevanter Texte, mögliche Kodierungsverfahren und die Qualitätskriterien inhaltsanalytischer Forschungsdesigns. Im zweiten Block geht es dann um klassische politikwissenschaftliche Anwendungen von Inhaltsanalysen etwa zur Parteienpositionierung, zur Salienz politischer Themen oder zur Gewinnung von Protestdaten aus der Medienberichterstattung. Darauf aufbauend werden im dritten Block moderne Verfahren zur (halb-) automatisierten Textanalyse vorgestellt, mit denen aktuell etwa der Einfluss von Interessengruppen, der Wandel politischer Interpretationsmuster oder aber die Salienz internationaler Politik in nationalen Parlamentsdebatten untersucht wird.

Die TeilnehmerInnen erörtern so die Stärken und Schwächen einzelner Ansätze an konkreten Beispielen und erhalten zudem praktische Tipps zu relevanten Analyseverfahren, Datenquellen und Werkzeugen.

B. Kursorganisation und Leistungsnachweise

Der Kurs richtet sich an fortgeschrittene Bachelorstudenten, die erste Kenntnisse in quantitativen politikwissenschaftlichen Methoden und Forschungsdesigns vertiefen wollen und ggf. eigene Inhalts- oder Textanalyseprojekte anstreben. Teilnahmevoraussetzungen sind die regelmäßige vorbereitende Auseinandersetzung mit der zumeist englischsprachigen *Pflichtliteratur*, die aktive *Mitwirkung* sowie mindestens eine *Präsentation* eines wissenschaftlichen Aufsatzes. Für einen Leistungsschein ist zudem eine Hausarbeit verlangt, in der entweder ein textanalytisches Projekt fachlich kritisiert oder aber ein eigenes textanalytisches Forschungsdesign zu einer selbstgewählten Fragestellung – etwa zur Vorbereitung von Bachelorarbeiten – entwickelt wird.

Da wir uns in diesem Semester nicht physisch treffen können, wird das Seminar in einer digitalen Umgebung durchgeführt. Trotzdem ist der Anspruch der gleiche: Ich möchte (a) Sie mit der gegenwärtigen Forschungsagenda vertraut machen, (b) sicherstellen, dass wir voneinander lernen können, indem es Raum für Ihre Fragen, Kommentare und Kritik gibt, und (c) auch eine faire Bewertung Ihrer Leistungen möglich machen.

Unser primärer „Treffpunkt“ wird die Seminar-Website auf dem FU-Blackboard sein (<https://lms.fu-berlin.de/>). Solange mir die FU keinen Zugang gewährt, finden sie die Kursmaterialien auch unter www.christian-rauh.eu/teaching. Dort wird das Seminar entlang folgender Parameter organisiert:

- Ihr wöchentliches *Selbststudium* entlang der *Pflichtliteratur* für jede Sitzung (siehe Abschnitt C unten)
- Eine wöchentliche kurze *Videovorlesung* (ppt Voiceover, asynchron, jeweils bis spätestens Donnerstag) von mir, die Sie durch die wichtigsten Argumente und Schwachstellen der Pflichtliteratur führt.
- Ein offenes *Message Board*, in dem jeder Teilnehmer Fragen und Beobachtungen teilen kann (und sollte) und gleichzeitig Kommentare zu denen der anderen Teilnehmer liefert. Bitte überprüfen Sie dieses Board mindestens wöchentlich,
- Zwei *Videokonferenzen* (KW23 und KW29), die vermutlich über WebEx organisiert werden
- *Zusammenfassung* von mindestens einer wissenschaftlichen Studie (siehe Abschnitt C unten), der die Präsentation ersetzt. Diese Zusammenfassung sollte die Forschungsfrage und Befunde der Studie vorstellen, diskutieren warum und wie ein inhaltsanalytisches Verfahren gewählt wurde und dieses dann anhand der im Seminar entwickelten Kriterien kritisieren. Analog zu einer Präsentation wird das Ergebnis über Blackboard allen Teilnehmern zur Verfügung gestellt. Ich akzeptiere die folgenden Formate:
 - Selbstaufgezeichnete Videopräsentation (z. B. ppt Voiceover)
 - Wissenschaftliches Poster
 - Schriftliche Zusammenfassung (pdf) von ca. 2000-3000 Wörtern
- *Individuell vereinbarte Kontakte* per E-Mail und / oder Skype

Ich bin gerade dabei, mehr über die an der FU verfügbaren Technologien herauszufinden. Weitere Details finden Sie im Video für die erste Sitzung. Wenn Sie Fragen oder Anregungen haben, wenden Sie sich bitte direkt an mich.

C. Detaillierte Kursübersicht

Pflichttexte für die einzelnen Sitzungen sind in der nachfolgenden Kursübersicht mit (PL) markiert. Die für die Präsentationen vorgesehenen Texte sind mit (PR) gekennzeichnet; gerne könne Sie aber auch selbst gewählte Literatur oder eigene Projekte vorstellen, sofern sie zum Thema der jeweiligen Sitzung passen (bitte nehmen Sie dazu Kontakt auf). Freiwillige weiterführende Informationsquellen sind mit (WF) vermerkt. Als einführende Literatur empfehle ich die folgenden Bücher, stelle Ihnen aber die wichtigsten Kapitel mit der übrigen Literatur auch per Blackboard zur Verfügung:

Krippendorff, Klaus (2004) *Content Analysis: An Introduction to Its Methodology*. 2nd Edition. London: Sage Publications.

Neuendorf, Kimberly (2002) *The Content Analysis Guidebook*. SAGE Publications.

Am Ende der Übersicht finden Sie zudem einige weiterführende Online-Ressourcen zu einzelnen Sitzungen.

Block I: Grundlagen der quantitativen Inhalts- und Textanalyse

1. Sitzung (KW 17):

Kursorganisation und kurze Geschichte der Inhaltsanalyse

- (PL) Krippendorff, Klaus (2004) *Content Analysis: An Introduction to Its Methodology*. London: 2nd Edition. London: Sage Publications: Chapter 1 "History" (pp. 3-17).
- (PL) Neuendorf, Kimberly (2002) *The Content Analysis Guidebook*. SAGE Publications: Chapter 2 "Milestones" (pp. 27-46)

2. Sitzung (KW 18):

Definitionen und typischer Aufbau von Inhaltsanalysen

- (PL) Krippendorff, Klaus (2004) *Content Analysis: An Introduction to Its Methodology*. London: 2nd Edition. London: Sage Publications: Chapter 2 "Conceptual foundation" (pp. 19-43).
- (PL) Neuendorf, Kimberly (2002) *The Content Analysis Guidebook*. SAGE Publications: Chapter 1 "Defining content analysis" (pp. 1-26)
- (PL) Neuendorf, Kimberly (2002) *The Content Analysis Guidebook*. SAGE Publications: Chapter 3 "Beyond description" (pp. 47-52)

3. Sitzung (KW 19):

Festlegung und Auswahl relevanter Texte und Analyseeinheiten

- (PL) Neuendorf, Kimberly (2002) *The Content Analysis Guidebook*. SAGE Publications: Chapter 4 "Message Units" (pp. 71-93)
- (PL) Krippendorff, Klaus (2004) *Content Analysis: An Introduction to Its Methodology*. London: 2nd Edition. London: Sage Publications: Chapter 5 "Unitizing" (pp. 97-110) and Chapter 6 "Sampling" (pp. 111-124).

4. Sitzung (KW 20):

Gütekriterien der Textkodierung – Reliabilität und Validität

- (PL) Neuendorf, Kimberly (2002) *The Content Analysis Guidebook*. SAGE Publications: Chapter 6 “Measurement techniques” (pp. 111-25 & 132-37)
- (PL) Krippendorff, Klaus (2004) *Content Analysis: An Introduction to Its Methodology*. London: 2nd Edition. London: Sage Publications: Chapter 11 “Reliability” (pp. 211-23 & 241-50) and Chapter 13 “Validity” (pp. 313-38).

Block II: Klassische, manuelle Inhaltsanalysen in den Politikwissenschaften

5. Sitzung (KW 21)

Parteipositionen im Wahlkampf und die Verteilung politischer Aufmerksamkeit

Die Comparative Manifestos und Comparative Policy Agendas Projekte

- (PL) Klingemann, Hans-Dieter, Andrea Volkens, Judith Bara, Ian Budge & Michael McDonald (2006) *Mapping Policy Preferences II: Estimates for Parties, Electors, and Governments in Eastern Europe, European Union and OECD 1990-2003*. Oxford: Oxford University Press: Introduction (pp. xvii-xxiii) and Appendices I-II (pp. 150-94)
- (PL) John, P. (2006) ‘The policy agendas project: a review’, *Journal of European Public Policy* 13(7): 975-86.
- (PR) Pappi, Franz Urban; Susumu Shikano (2004) ‘Ideologische Signale in den Wahlprogrammen der deutschen Bundestagsparteien 1980 bis 2002’, *Arbeitspapiere - Mannheimer Zentrum für Europäische Sozialforschung*: 76.
- (PR) Green-Pedersen, C. (2007) ‘The growing importance of issue competition: The changing nature of party competition in Western Europe’, *Political Studies* 55(3): 607-628.
- (PR) Mikhaylov, et al. (2012). ‘Coder Reliability and Misclassification in the Human Coding of Party Manifestos’. *Political Analysis* 20(1):78–91.
- (PR) Mortensen, P. B., C. Green-Pedersen, G. Breeman, L. Chaqués-Bonafont, W. Jennings, P. John, A. M. Palau, and A. Timmermans (2011) ‘Comparing government agendas. Executive Speeches in the Netherlands, United Kingdom, and Denmark’, *Comparative Political Studies* 44(8): 973-1000.
- (PR) Alexandrova, P. and A. Timmermans (2013) ‘National interest versus the common good: The presidency in european council agenda setting’, *European Journal of Political Research* 52(3): 316-38.
- (WF) Baumgartner, F., Jones, B. and MacLeod, M. (1998) ‘Lessons from the trenches: quality, reliability, and usability in a new data source’, *The Political Methodologist* 8(2): 1–10.
- (WF) Alexandrova, P., M. Carammia, S. Princen, and A. Timmermans (2014) ‘Measuring the European Council Agenda: Introducing a new approach and dataset’, *European Union Politics* 15(1): 152-67.
- (WF) Volkens, Andrea, Judith Bara and Ian Budge (2009) ‘Data Quality in Content Analysis. The Case of the Comparative Manifestos Project’, *Historical Social Research* 34(1): 234-51.

6. Sitzung (KW 22)

Politische Signale im öffentlichen Raum

Die Erhebung von Ereignisdaten und politischen Claims aus Medienberichten

- (PL) Earl, J., A. Martin, J. D. McCarthy, and S. A. Soule (2004) 'The use of newspaper data in the study of collective action', *Annual Review of Sociology* 30(1): 65-80.
- (PL) Koopmans, R. & Statham, P. (1999) 'Political Claims Analysis: Integrating Protest Event and Political Discourse Approaches', *Mobilization* 4: 203-221.
- (PR) Uba, Katrin, and Fredrik Uggla (2011) 'Protest Actions against the European Union, 1992-2007', *West European Politics* 34(2): 384-393.
- (PR) Rauh, Christian, and Gerald Schneider (2013) 'There is No such Thing as a Free Open Sky: Financial Markets and the Struggle over European Competences in International Air Transport', *Journal of Common Market Studies* 51(6): 1124-40. (plus Web Appendix A6!)
- (PR) Kriesi, Hanspeter, Edgar Grande, Romain Lachat, Martin Dolezal, Simon Bornschie, and Timotheos Frey (2006) 'Globalization and the transformation of the national political space: Six European countries compared', *European Journal of Political Research* 45(6): 921-956.
- (PR) Nonhoff, Martin, Jennifer Gronau, Frank Nullmeier, and Steffen Schneider (2009) 'Zur Politisierung internationaler Institutionen: Der Fall G8', *Zeitschrift für Internationale Beziehungen* 16(2): 233-63.
- (PR) De Wilde, Pieter (2019) 'The Quality of Representative Claims: Uncovering a Weakness in the Defense of the Liberal World Order', *Political Studies*: Online First.
- (WF) Koopmans, Ruud and Dieter Rucht (2002) 'Protest Event Analysis', In: Bert Klandermans and Suzanne Staggenborg (eds.) *Methods of Social Movement Research*. Minneapolis: University of Minnesota Press: 231-59.
- (WF) Hutter, Swen (2014) *Protesting Culture and Economics in Western Europe: New Cleavages in Left and Right Politics* Minneapolis: University of Minnesota Press: esp. Appendix A (pp. 147-65).
- (WF) Pieter de Wilde, Ruud Koopmans and Michael Zürn (2014) 'The Political Sociology of Cosmopolitanism and Communitarianism: Representative Claims Analysis', Discussion Paper SP IV 2014-102 Wissenschaftszentrum Berlin für Sozialforschung.
- (WF) Marc Helbling and Anke Tresch (2011): 'Measuring Party Positions and Issue Salience from Media Coverage: Discussing and Crossvalidating New Indicators', *Electoral Studies* 30(1): 174-183.
- (WF) Schneider, Steffen (2010) 'Empirische Legitimationsforschung'. In: F. Nullmeier, D. Bigeon, J. Gronau, M. Nonhoff, H. Schmidtke and S. Schneider (eds) *Prekäre Legitimitäten. Rechtfertigung von Herrschaft in der postnationalen Konstellation*. Frankfurt/New York: Campus Verlag: 45-67.

Block III: (Halb-) Automatisierte Verfahren politikwissenschaftlicher Textanalyse

7. Sitzung (KW 23)

Grundlagen automatisierter Textanalyseverfahren

- (PL) Krippendorff, Klaus (2004) *Content Analysis: An Introduction to Its Methodology*. London: 2nd Edition. London: Sage Publications: Chapter 12 "Computer Aids" (pp. 257-312).
- (PL) Grimmer, Justin, and Brandon Stewart (2013) 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts', *Political Analysis* 21(3): 267-297.
- (WF) Lowe, Will (2003) 'The Statistics of Text: New Methods for Content Analysis', Paper presented as the *Midwest Political Science Association Conference* April 2003. Available at SSRN: <http://ssrn.com/abstract=2210808>.

8. Sitzung (KW 24)

Inhaltsanalyse als Klassifikationsproblem – Wörterbuch-basierte Verfahren

- (PL) Neuendorf, Kimberly (2002) *The Content Analysis Guidebook*. SAGE Publications: Chapter 6 "Measurement techniques" (pp. 125-32)
- (PL) Rauh, C. (2018) 'Validating a sentiment dictionary for German political language—a workbench note', *Journal of Information Technology & Politics* 15(4): 319–343.
- (PR) Rauh, C. and De Wilde, P. (2018) 'The opposition deficit in EU accountability: Evidence from over 20 years of plenary debate in four member states', *European Journal of Political Research* 57(1): 194–216.
- (PR) Rauh, C., Bes, B. J. and Schoonvelde, M. (2019) 'Undermining, defusing, or defending European integration? Assessing public communication of European executives in times of EU politicization', *European Journal of Political Research* Online First.
- (PR) Rauh, C. and Zürn, M. (2019) 'Authority, politicization, and alternative justifications: endogenous legitimation dynamics in global economic governance¹', *Review of International Political Economy*: Online First.
- (PR) Proksch, S.-O., Lowe, W., Wäckerle, J. and Soroka, S. (2018) 'Multilingual Sentiment Analysis: A New Approach to Measuring Conflict in Legislative Speeches', *Legislative Studies Quarterly*: Online first.
- (PR) Reinhard, J., J. Biesenbender, and K. Holzinger (2014, May). Do arguments matter? argumentation and negotiation success at the 1997 Amsterdam intergovernmental conference. *European Political Science Review* 6, 283-307.
- (PR) Dodds, Peter Sheridan, and Christopher M Danforth 'Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents', *Journal of Happiness Studies* 11(4)
- (WF) Young, Lori, and Stuart Soroka (2012) 'Affective News: The Automated Coding of Sentiment in Political Texts', *Political Communication* 29(2): 205-31.
- (WF) Mohammad, S. M. and Turney, P. (2013) 'Crowdsourcing a Word-Emotion Association Lexicon', *Computational Intelligence* 29(3): 436–465.

9. Sitzung (KW 25)

Komplexere Klassifikationsverfahren: Topic Modelle und Machine Learning

- (PL) Blei, David (2012) 'Probabilistic topic models', *Commun. ACM* 55(4): 77-84.
- (PL) Quinn, Kevin, Burt Monroe, Michael Colaresi, Michael Crespin, and Dragomir Radev (2010) 'How to Analyze Political Attention with Minimal Assumptions and Costs', *American Journal of Political Science* 54(1): 209-228.
- (PL) Hopkins, Daniel, and Gary King (2010) 'A Method of Automated Nonparametric Content Analysis for Social Science', *American Journal of Political Science* 54(1): 229-247.
- (PR) Ceron, Andrea, Luigi Curini, Stefano Micus, and Giuseppe Porro (2013) 'Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France', *New Media & Society* 16(2): 340-358.
- (PR) Quinn Albaugh, Stuart Soroka, Jeroen Joly, Peter Loewen, Julie Sevenans, and Stefaan Walgrave (2014) 'Comparing and Combining Machine Learning and Dictionary-Based Approaches to Topic Coding', *Unpublished Manuscript*.
- (PR) Grimmer, J. (2010, December). A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases. *Political Analysis* 18(1): 1-35.
- (WF) Cardie, Claire, and John Wilkerson (2008) 'Text Annotation for Political Science Research', *Journal of Information Technology & Politics* 5(1): 1-6.
- (WF) Scharkow, M. (2013) 'Thematic content analysis using supervised machine learning: An empirical evaluation using German online news', *Quality and Quantity* 47(2): 761-773.

10. Sitzung (KW 26)

Die Positionen von politischen Sprechern oder Autoren: Skalierungsverfahren

- (PL) Laver, M., K. Benoit, and J. Garry (2003). Extracting policy positions from political texts using words as data. *American Political Science Review* 97(2): 311-31.
- (PL) Slapin, J. B. and S.-O. Proksch (2008) 'A scaling model for estimating Time-Series party positions from texts', *American Journal of Political Science* 52(3): 705-22.
- (PR) Klüver, H. (2009) 'Measuring interest group influence using quantitative text analysis', *European Union Politics* 10(4): 535-49.
- (PR) Proksch, Sven-Oliver, and Jonathan Slapin (2009) 'How to Avoid Pitfalls in Statistical Analysis of Political Texts: The Case of Germany', *German Politics* 18(3): 323-344.
- (PR) Bernauer, Julian, and Thomas Bräuninger (2009) 'Intra-Party Preference Heterogeneity and Faction Membership in the 15th German Bundestag: A Computational Text Analysis of Parliamentary Speeches', *German Politics* 18(3).
- (WF) Lowe, Will (2008) 'Understanding Wordscores', *Political Analysis* 16(4): 356-71.

11. Sitzung (KW 27)

Semantik und Struktur in der automatisierten Analyse politischer Texte **Grammatikalisches Parsing und Word Embeddings**

- (PL) Schrodtt, Philip A. and Deborah J. Gerner (2012) 'Chapter 2: Fundamentals of Machine Coding', In: Schrodtt, Philip A. and Deborah J. Gerner (eds.) *Analyzing International Event Data: A Handbook of Computer-Based Techniques*. Unpublished book manuscript, available at: <http://eventdata.parusanalytics.com/books.html>
- (PL) Rodman, E. (2020) 'A Timely Intervention: Tracking the Changing Meanings of Political Concepts with Word Vectors', *Political Analysis* 28(1): 87–111.
- (PR) Hammond, Jesse, and Nils B Weidmann (2014) 'Using machine-coded event data for the micro-level study of political violence', *Research & Politics* 1(2):
- (PR) Benoit, K., Munger, K. and Spirling, A. (2019) 'Measuring and Explaining Political Sophistication through Textual Complexity', *American Journal of Political Science* 63(2): 491–508.
- (PR) Rudkowsky, E. et al. (2018) 'More than Bags of Words: Sentiment Analysis with Word Embeddings', *Communication Methods and Measures* 12(2–3): 140–157.
- (PR) Rheault, L. and Cochrane, C. (2020) 'Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora', *Political Analysis* 28(1): 112–133.
- (PR) Spirling, A. and Rodriguez, P. L. (n.d.) 'Word Embeddings What works, what doesn't, and how to tell the difference for applied research', *Unpublished Manuscript*, available at https://github.com/ArthurSpirling/EmbeddingsPaper/blob/master/Paper/Embeddings_SpirlingRodriguez.pdf.
- (WF) Leetaru, Kalev and Philip A: Schrodtt (2013) 'GDELT: Global Data on Events, Location and Tone, 1979-2012', Paper presented at the *International Studies Association meetings*, San Francisco, April 2013.
- (WF) King, G. and W. Lowe (2003, June). An automated information extraction tool for international conflict data with performance as good as human coders: A rare events evaluation design. *International Organization* 57 (03), 617-642.
- (WF) Schories, M. (2020) 'Using Word Embeddings for Journalistic Research', <https://towardsdatascience.com/using-word-embeddings-as-a-method-for-journalistic-research-ae82ffea7a62> (last accessed: 06.04)

12. Sitzung (KW 28)

Praktische Aspekte – Datenquellen, Textaufbereitung und Analysesoftware

- (PL) Rauh, Christian, and Sebastian Bödeker (2013) 'The international trade regime in the public sphere, 1986-2012: Evaluating the social legitimacy of global governance with semi-automated text mining approaches', Paper presented at the 1st *European Workshops on International Studies (EWIS)* Tartu, Estonia, 5-8 June 2013.
- (PL) Alexa, Melina and Cornelia Zuell (2000) "Text Analysis Software: Commonalities, Differences and Limitations: The Results of a Review." *Quality and Quantity* 34(3):299–321.
- (PL) Alonso, Sonia, Andrea Volkens, and Braulio Gómez (2012) *Content-analyzing political texts*. Madrid: CIS. Kapitel 1 (esp. pp. 24-30)
- (WF) Greene, Z., Ceron, A., Schumacher, G. and Fazekas, Z. (2016) 'The Nuts and Bolts of Automated Text Analysis. Comparing Different Document Pre-processing Techniques in Four Countries', *Open Science Framework* November 1 (osf.io/ghxj8).

- (WF) Benoit, K. et al. (2018) 'quanteda: An R package for the quantitative analysis of textual data', *Journal of Open Source Software* 3(30): 774.
- (WF) Jackman, Simon (2006) 'Data from the Web Into R', *The Political Methodologist* 14(2): 11-15
- (WF) Munzert, S., Rubba, C., Meißner, P. and Nyhuis, D. (2015) *Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining*, Wiley.

13. Sitzung (KW 29)

Zusammenfassung und Feedback

D. Weiterführende Online Ressourcen zu einzelnen Sitzungen

5. Sitzung: Comparative Manifestos and Policy agendas

CMP Datensatzbeschreibungen und Daten zum Download:

<https://manifesto-project.wzb.eu/>

Kurzes Anwendungsbeispiel:

<http://democracy.blog.wzb.eu/2013/09/17/and-they-do-differ-an-analysis-of-the-german-parties-election-programs-2013/>

Überblick, Codebooks und weitere Ressourcen zu den Policy Agendas Projekten:

<http://www.comparativeagendas.info/>

<http://www.policyagendas.org/>

6. Sitzung: Ereignisdaten und Claims Analysen

Beschreibung, Publikationen und Datensätze des PRODAT Projekts:

<http://www.wzb.eu/de/forschung/beendete-forschungsprogramme/zivilgesellschaft-und-politische-mobilisierung/projekte/prodat-dokument>

Beschreibung und Kodierverfahren des EUROPUB Projekts: <http://europub.wzb.eu/>

Kodebuch des Projekts 'The Political Sociology of Cosmopolitanism and Communitarianism:

Representative Claims Analysis': <https://bibliothek.wzb.eu/pdf/2014/iv14-102.pdf>

8. Sitzung: Wörterbuch-basierte Verfahren

Will Lowes JFREQ und Yoshikoder:

<http://conjugateprior.org/software/jfreq/>

<http://conjugateprior.org/software/yoshikoder/>

Der von Young/Soroka verwendete Lexicoder einschließlich Wörterbücher:

<http://lexicoder.com/>

Das quanteda dictionaries Paket für R:

<https://github.com/kbenoit/quanteda.dictionaries>

Quanteda Tutorial für ein einfaches Wörterbuch Verfahren:

<https://tutorials.quanteda.io/advanced-operations/targeted-dictionary-analysis/>

10. Sitzung: Skalierungsverfahren

Weiterführende Informationen und Werkzeuge für WORDSCORES:

http://www.tcd.ie/Political_Science/wordscores/

Weiterführende Informationen und Werkzeuge für WORDFISH

www.wordfish.org

11. Sitzung: Automatisierte Ereignisdaten

Alle Informationen rund um KEDS, CAMEO, TABARI:

<http://eventdata.parusanalytics.com/>

Alle Informationen rund um GDELT:

<http://www.gdeltproject.org/>

Anwendungen und kontroverse Diskussionen rund um GDELT:

Twitter: #gdelt

12. Sitzung: Praktische Aspekte

Gute Übersichtsseite mit vielfältigen Ressourcen:

<http://www.textanalysis.info/>

Nützliche Software zum Edit von Text-Dateien (inkl. Regular Expressions, Batch Edits):

<http://www.digitalvolcano.co.uk/textcrawler.html>

<http://notepad-plus-plus.org/>

Grundlegendes zu Text Manipulation in R:

http://en.wikibooks.org/wiki/R_Programming/Text_Processing

Quanteda tutorials:

<https://tutorials.quanteda.io/>

Online Buch mit vielen Tutorials zum Text Mining mit den tidy tools in R:

<https://www.tidytextmining.com/>

Einführung zur Automatisierten Inhaltsanalyse mit R (Cornelius Puschmann, deutsch):

<http://inhaltsanalyse-mit-r.de/index.html>

Weitere R Pakete zur automatischen Sprachverarbeitung:

<http://cran.r-project.org/web/views/NaturalLanguageProcessing.html>

Frei zugängliche Korpora politischer Texte (Asuwahl; Datenquellen zitieren!):

ParlSpeech: Über 6,3 Millionen parlamentarische Reden aus neun Staaten und Zeiträumen von bis zu 20 Jahren: <https://doi.org/10.7910/DVN/L4OAKN>

EUSpeech: Über 17.000 Reden von Staats- und Regierungschefs, EU Kommissaren und MEPs während der Eurokrise:
<https://dataverse.harvard.edu/dataverse/euspeech>

UNGD corpus: Volltexte der Reden in der Generalversammlung der Vereinten Nationen seit 1947:
<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/0TJX8Y>

Manifesto Corpus: Volltexte von über 2500 Parteiprogrammen: <https://manifesto-project.wzb.eu/information/documents/corpus>