

# Schriftliche Bearbeitung von Portfolio Aufgaben

Thomas Haase

10. November 2023

Justus-Liebig-Universität Gießen

Matrikelnummer - 6033199

Blockseminar:

Prof. Dr. Elmar Schlüter - *Digitale Methoden der empirischen Sozialforschung*

---

## Inhaltsverzeichnis

<b>1</b>	<b>Aufgabe 1 - Bibliographie textbasierter Social Media Analysen</b>	<b>1</b>
<b>2</b>	<b>Aufgabe 2 - Vorstellung und Einordnung eines Artikels</b>	<b>2</b>
2.1	Einordnung . . . . .	3
<b>3</b>	<b>Aufgabe 3 - Vorteile und Schwierigkeiten bei und mit Netlytic</b>	<b>4</b>
3.1	Vorteile . . . . .	4
3.2	Schwierigkeiten im Erlernen . . . . .	4
<b>4</b>	<b>Aufgabe 4 - Diskussion zu Qualitative/ Quantitative Inhaltsanalysen</b>	<b>4</b>
<b>5</b>	<b>Aufgabe 5 - Viele Coder für eine Analyse?</b>	<b>5</b>
<b>6</b>	<b>Aufgabe 6 - Cohen's Kappa</b>	<b>6</b>
6.1	Beispiel . . . . .	6
<b>7</b>	<b>Literatur</b>	<b>7</b>

---

# 1 Aufgabe 1 - Bibliographie textbasierter Social Media Analysen

## Literaturliste für Aufgabe 1

1. Sältzer, M. & Stier, S. Die Bundestagswahl 2021 auf Twitter. *easy social sciences* **67**, 30–38. ISSN: 2199-9082 (2022).
2. Czymara, C., Dochow-Sondershaus, S., Drouhot, L., Simsek, M. & Spörlein, C. Catalyst of hate? Ethnic insulting on YouTube in the aftermath of terror attacks in France, Germany and the United Kingdom 2014-2017. *Journal of Ethnic and Migration Studies* (Aug. 2022).
3. Abanoz, E. The Reactions to Muslim Identity Building through Social Media: User Comments on YouTube Street Interview Videos. *Religions* **13**. ISSN: 2077-1444. <https://www.mdpi.com/2077-1444/13/6/498> (2022).
4. Karami, A., Lundy, M., Webb, F. & Dwivedi, Y. K. Twitter and Research: A Systematic Literature Review Through Text Mining. *IEEE Access* **8**, 67698–67717 (2020).
5. Lazard, A. J., Scheinfeld, E., Bernhardt, J. M., Wilcox, G. B. & Suran, M. Detecting themes of public concern: A text mining analysis of the Centers for Disease Control and Prevention’s Ebola live Twitter chat. *American Journal of Infection Control* **43**, 1109–1111. ISSN: 0196-6553. <https://www.sciencedirect.com/science/article/pii/S0196655315006148> (2015).
6. Turenne, N. Net activism and whistleblowing on YouTube: a text mining analysis. *Multimedia Tools and Applications* **82**, 9201–9221 (März 2023).
7. Machuca, C. R., Gallardo, C. & Toasa, R. M. Twitter Sentiment Analysis on Coronavirus: Machine Learning Approach. *Journal of Physics: Conference Series* **1828**, 012104. <https://dx.doi.org/10.1088/1742-6596/1828/1/012104> (Feb. 2021).
8. Stier, S., Bleier, A., Lietz, H. & Strohmaier, M. Election Campaigning on Social Media: Politicians, Audiences, and the Mediation of Political Communication on Facebook and Twitter. *Political Communication* **35**, 50–74. <https://doi.org/10.1080/10584609.2017.1334728> (2018).
9. Stier, S., Posch, L., Bleier, A. & Strohmaier, M. When populists become popular: comparing Facebook use by the right-wing movement Pegida and German political parties. *Information, Communication & Society* **20**, 1365–1388. <https://doi.org/10.1080/1369118X.2017.1328519> (2017).
10. GRÜNDER-FAHRER, S., SCHLAF, A., WIEDEMANN, G. & HEYER, G. Topics and topical phases in German social media communication during a disaster. *Natural Language Engineering* **24**, 221–264 (2018).

## 2 Aufgabe 2 - Vorstellung und Einordnung eines Artikels

In dem Artikel *Finding the bird's wings: Dimensions of factional conflict on Twitter* von Marius Sältzer[1] wird die Fragestellung untersucht ob Angehörige von verschiedenen Parteifractionen ihre unterschiedlichen innerparteilichen, ideologischen Positionen äußern. Fractionen als Art der innerparteilichen Organisation werden genutzt um innerparteiliche Konflikte auszutragen [1, S. 1], das lässt die Frage leicht plausibel erscheinen.

Dieser Frage ging Sältzer mittels einer Korrespondenzanalyse von 800.000 Tweets von 498 Abgeordneten des deutschen Bundestages auf den Grund gegangen. Die Art der Analyse wurde beim säubern des Datensatzes mitbeachtet, indem unter anderem Wörter, die nur 100 Accounts nutzten, entfernt wurden. Eben jene Wörter können bei Clusteringverfahren störend wirken, da der genutzte Algorithmus, der später kurz erläutert werden wird, diese fälschlicherweise als artifizielle Zentren erkennen würde - diese sind aber eben nicht die Zentren, die gemessen werden sollen.

Es gibt bei scaling Modellen die Möglichkeit a priori Zentren festzulegen um die die Daten gruppiert werden sollen, wie in etwa bei *wordscores* als scaling Modell. Alternativ gibt es die Möglichkeit mit scaling Modellen wie *Wordfish*[3] ohne Referenztexte unsupervised auf Basis der beobachteten Worthäufigkeiten zu scalen [2]. In diesem Fall wurde der wordfish-Ansatz genutzt. Dies wurde mittels einer Studie die dieses Modell bereits für Twitterdaten nutzte und mit Salienztheorie begründet [1, S. 3].

Die Idee ist, zuerst die Tweets entlang verschiedener Dimensionen darzustellen um zu sehen welche der getweeteten Begriffe die Politiker auf welchen Achsen unterscheiden. Im Anschluss kann darauf aufbauend der Effekt der Fraktionszugehörigkeit auf die Positionen auf den Achsen getestet werden [1, S. 5].

Es wurde sich dagegen entschieden die Dimensionen theoriegeleitet zu bestimmen, da die Daten in sehr großer Menge und stark unstrukturiert vorlagen. Um die Dimensionen trotzdem zu reduzieren wurde ein Screeplot genutzt. Screeplots sind eine Darstellung des Ergebnis einer Hauptkomponentenanalyse. Bei dieser werden die Eigenwerte einer diagonalisierten Korrelationsmatrix berechnet. Die Eigenwerte der Korrelationsmatrix beschreiben die Transformation des Raumes, durch die unkorrelierte (normalverteilte, "whitenoise") Daten transformiert werden müssten um die gemessenen Rohdaten zu erhalten. Aus dieser Transformation können die sich am stärksten auf die Transformation auswirkenden Eigenwerte/Faktoren als Balkendiagramm grafisch dargestellt werden [4]. Dieser Plot wird Screeplot genannt.

Die ausschlaggebendste Komponente ist die Betonung von Migration, die zweit wichtigste Komponente die links-rechts Dimension und als dritt wichtigste Komponente wurde die liberale Betonung von nicht-Migration festgestellt. Um die ökonomische und kulturellen Dimensionen zu konstruieren wurden eben diese Begriffe zu Kultur und Migration entfernt und hinzu-addiert.

In der Interpretation des Biplots fallen insbesondere auf, dass die CDU besonders a-politische Begriffe nutzt und die AFD als stark separates Cluster auffällt. Die AFD hebt sich nur auf einer Achse - der kulturellen - von den anderen Parteien ab. Das Modell hat Schwierigkeiten die AFD ökonomisch und links-rechts einzuordnen, weil die AFD insbesondere nur Begriffe der kulturellen Achse nutzt [1, S. 7].

Um im zweiten, finalen Schritt des Forschungsdesigns die Individuelle Ebene der Fraktionszugehörigkeit einzubeziehen wurden 3 OLS Modelle mit unterschiedlich vielen Kontrollvariablen gerechnet. Das erste Modell enthielt nur den Haupteffekt, das zweite einmal die die kulturelle Orientierung der Fraktion, einandermal die ökonomische und das dritte Modell den Mandatentyp als Kontrollvariable, weil die Fraktion teilweise vom Mandatentyp abhängig ist [1, S. 7].

Es ergibt sich, dass die Parteizugehörigkeit alle anderen Faktoren stark überwiegt, die Fraktionszugehörigkeit ist sehr niedrig aber signifikant. Daraus lässt sich schließen, dass die inner-

parteiliche Fraktion eine kleine aber signifikante Rolle bei der Äußerung der Positionen eines Abgeordneten spielt [1, S. 8].

Die Methode des Artikels ist innovativ, insbesondere weil mehrere komplexe Analyseverfahren intelligent miteinander kombiniert werden. Die Ergebnisse sowie die Nutzung und Art der Kombination der unterschiedlichen Methoden ist zudem gut begründet. Zu kritisieren ist jedoch, dass die genaueren Spezifikationen des Scalingmodels nicht genauer aufgeführt wurden. Auch in dem im Artikel verlinkten Github Repository ist nur der Datensatz und der R-Code für den zweiten Teil der Analyse zu finden. Den Code oder ein paar Zeilen im Artikel bezüglich der genaueren Umsetzung des Scalings mittels Wordfish wären hier von Relevanz gewesen.

Obwohl im Artikel behauptet wird der Screeplot sei im Onlineappendix zu finden fehlt von diesem jede Spur (bis auf eine Überschrift ist nichts zu finden). Mit welcher `library()` oder mit welchem Algorithmus (rekursiv?,...) die PCA in R gerechnet wurde, wie dabei mit NAs verfahren wird und ob die Daten standardisiert wurden und kann extreme Unterschiede in den Ergebnissen hervorrufen [4, S. 9–10]. Obwohl der Fokus des Artikels auf der Kombination der Methoden lag, stützt sich das gesamte Ergebnis auf dem Ergebnis der PCA. Daher wäre es unglaublich relevant gewesen hier für Klarheit zu sorgen.

## 2.1 Einordnung

Bei dem dargestellten Artikel handelt es sich um eine quantitativ explorative Inhaltsanalyse. Nach Jünger et al. unterscheiden sich Text- und Inhaltsanalyse nach ihrem Interesse, wobei Inhaltsanalyse die Suche nach den "datengenerierenden Prinzipien hinter einem Text"[5, S. 357] statt rein linguistische Ziele ohne nähere Betrachtung der Wortbedeutung verfolgt. Marius Sältzer konnte den Anspruch hunderttausende Tweets zu analysieren nur mittels automatisierter Analyse der Tweetinhalte durchführen. Mit 500 Kandidaten konnte er einen überwiegenden Großteil der Abgeordneten, die auch das Ziel seiner Analyse waren, abbilden. Twitter eignet sich gut für diese Umsetzung, weil die Abgeordneten in Interviews bei Fragen bzgl. dem Verhältnis ihrer Positionierung zu Fraktion und Partei in einen Rollenkonflikt geraten, wodurch Verzerrungen wahrscheinlicher gemacht werden würden. Auf Twitter gibt es zwar auch Verzerrungen, da Twitter als Situation der Meinungsäußerung durch Bestätigung und den Algorithmus auch die Meinungsäußerung verzerren kann (negative Sentiments korrelieren mit mehr Reichweite). Diese Verzerrungen hat Sältzer allerdings schon 2 Jahre zuvor untersucht und festgestellt, dass die Abgeordneten **nicht** so auf das Feedback auf Twitter reagieren, dass sie ihre Äußerungen anpassen [6]. Der Nachteil quantitativer Inhaltsanalyse, dass der Kontext der Entstehung der Tweets unbekannt ist greift bei der spezifischen Zielgruppe Abgeordneter nicht, weil für Abgeordnete Twitter ein Instrument ihres Berufs darstellt, kurz: auch Tweets die in ungünstigen Situationen verfasst werden sind ein Ausdruck der Positionierung im politischen Feld. Bei quantitativen explorativen Analysen muss sehr vorsichtig gearbeitet werden, da Datensätze Kontexte nicht (gut) abbilden können. Die Datenmenge, Unterschiedlichkeit der Daten, die beschriebene Begründung der Validität durch die Auswahl der Daten und die Begründung der Methoden können das Verfahren aber nachvollziehbar und gut rechtfertigen.

## 3 Aufgabe 3 - Vorteile und Schwierigkeiten bei und mit Netlytic

### 3.1 Vorteile

Ein großer Vorteil von Netlytic ist die einfache Handhabung. Diese wird ermöglicht durch den Onlinezugriff, da es sich um eine Website handelt, aber auch, dass man quasi "automatisch" durch die einzelnen Schritte der Analyse geleitet wird. Es gibt wenig direkt zu entscheiden, wodurch ein Datensatz sehr schnell analysiert werden kann. Die Geschwindigkeit der Analyse wird auch dadurch begünstigt, dass die Website insgesamt sehr effizient aufgebaut ist und keine Animationen oder Grafiken Ladezeit verbrauchen.

Die Arbeit mit APIs und Scraping ist mühselig. Netlytic nimmt dem Studierenden so viel dieser Arbeit ab, dass es kaum wie ein relevanter Teil des Projekts wirkt. So kann sich auf die Analyseentscheidungen und Interpretation der Daten fokussiert werden.

### 3.2 Schwierigkeiten im Erlernen

Wenn man sich innerhalb des Menüs eines Datensatzes befindet wird man gut angeleitet, wenn man sich allerdings in der Übersicht der Datensätze befindet eröffnen sich jedoch Einstellungen, die man innerhalb der Datenanalyse nicht gegeben sind wie das Subsetzen oder das Bearbeiten eines Datensatzes.

Zu Schwierigkeiten führt auch die lange Ladezeit der Analysen, es stellen sich dann Fragen, ob das Browserfenster geöffnet bleiben muss oder in den Hintergrund verschoben werden kann.

Auch die Einstellungen innerhalb der Analysen sind teilweise inkonsequent designed. Beim bearbeiten eines Dictionaires kann man löschen, disablen und einen neuen Term zur Kategorie hinzufügen. Wenn allerdings das Dictionaire als .csv, mit der Absicht es zu bearbeiten, heruntergeladen wird eröffnen sich die Spalten "score" und "status" die nicht weiter erklärt werden, sich aber nicht automatisch erschließen (z.B. Warum beträgt der Score bei manchen Wörtern -1?). Auch auf der Website unter dem Reiter "help" findet sich dazu keine Erklärung.

## 4 Aufgabe 4 - Diskussion zu Qualitative/ Quantitative Inhaltsanalysen

Qualitative und Quantitative Inhaltsanalyse unterscheiden sich im Kern (wie jede Qualitative und Quantitative Methode) durch ihren Grad der Standardisierung.

Die Ursache für diesen Unterschied sind die unterschiedlichen Ziele qualitativer und quantitativer Forschung. Qualitative Forschung hat den Anspruch Hypothesen generierend zu wirken und komplexe Handlungssituationen zu rekonstruieren um verallgemeinernd(e) Theorie zu bilden. Mithilfe von standardisierter Forschung wird versucht repräsentative Ergebnisse durch das Überprüfen von aus Theorie abgeleitete Hypothesen zu erzielen [7, S. 22–26].

Da angewandte Forschung wie die empirische Soziologie auch eine Frage von Ressourcen ist, kann nicht ausgelassen werden, dass ein wichtiger Vorteil von quantitativen Analysen die Möglichkeit der Analyse und Abbildung großer Datenmengen ist.

Wirklich verallgemeinern kann man die harte Trennung allerdings nicht. Quantitative Textanalyse wird heutzutage auch oft explorativ genutzt [9]. Kritik an dieser Vorgehensweise ist jedoch, dass ohne theoretischen Ansatz der reine Blick auf die Daten nicht genug Kontext liefert um eines der Ziele von Soziologie als Wissenschaft zu erfüllen: Erklärungen zu liefern [8, S. 15]. Ohne Theorie vorzugehen birgt implizit auch immer das Ziel in sich Theorie produzieren

zu wollen. Auch die, eigentlich vielversprechenden, unsupervised KI Modelle funktionieren letztendlich wie eine Blackbox, deren Ergebnisse viel zusätzliche Validierungen, Tests und weitere Forschung benötigen.

Nicht nur über das Teilgebiet der Computational Social Science zerfließen diese Grenzen, sondern auch über Mixed-Methods.

An den Übergängen zwischen qualitativer und quantitativer Forschung stellt sich ganz besonders die Frage nach Vor- und Nachteilen der Methoden ganz und gar neu. Es lassen sich nicht wie bei klassischen Methoden die allgemeinen Vorteile des quali- oder quantitativen Forschungsansatzes für die jeweils angewandte Methode übernehmen, stattdessen muss im Detail ausgearbeitet werden wo Vor- und Nachteile der einzelnen Forschungsansätze auftauchen und wie diese zusammen mit dem Forschungsgegenstand wirken. Sind diese Fragen geklärt kann man versuchen die Methode zu präzisieren und erst dann ist es auch möglich die Methode deutlich zu begründen.

## 5 Aufgabe 5 - Viele Coder für eine Analyse?

Um zu erläutern warum es bei qualitativen und quantitativen Inhaltsanalysen erstrebenswert ist dieselben Items von mehreren CoderInnen codieren zu lassen muss auf die Gütekriterien von empirischer Forschung zurückgegriffen werden. Diese sind nicht nur bei quantitativer Forschung von Relevanz, sondern für qualitative Forschung ebenso von Bedeutung.

Das grundlegendste Kriterium für Forschung allgemein ist Objektivität. Objektivität kann als Intersubjektivität umformuliert werden [10, S. 277]. Es ist nur möglich den Grad der Intersubjektivität zu kontrollieren und festzustellen, wenn diese getestet wird. Bei der Bewertung der eigenen Forschung wird die eigene Arbeit zum neuen Forschungs- bzw. Bewertungsobjekt um die darauf bezogene Frage zu beantworten, ob sie wissenschaftlichen Standards entspricht. Das hinzuziehen von mehreren CoderInnen für dieselben Items ist eine Methode um sicherzustellen, dass sich Intersubjektivität/Objektivität angenähert wird. Obwohl bei qualitativer Forschung im Vergleich zu quantitativer Forschung die "Objektivität gerade aus der Subjektivität der Interaktionsrelationen" emergieren soll [10, S. 277] ist es sinnvoll diese Größe zur Einschätzung für andere Forscher, oder Leser im allgemeinen, im Wert der Interraterreliabilität festzuhalten.

Durch die Möglichkeit der Berechnung der Interraterübereinstimmung kann nicht nur die Objektivität, sondern auch die Reliabilität getestet und dadurch im Forschungsverlauf verbessert werden. Reliable Ergebnisse liefert das Instrument, also das der/die CoderIn mit dem Kategorienschema selbst.

Nicht nur Intersubjektivität sondern auch die Fehleranfälligkeit kann somit ausgeschlossen werden [10, S. 278]. Vor allem beim qualitativen Codieren ist leicht vorstellbar, dass eine zweite Person bezüglich derselben Passage anderweitige Handlungsmotive aus dem Material zu erkennen behauptet. Diese Unterschiede können auf unterschiedliche Verzerrungen der CodiererInnen zurückgeführt werden. Vorher habe ich die Beurteilung der Forschung als kleine Forschung über die eigene Forschung behandelt. So wie sich in etwa auf ein Interview das Setting auswirken kann, kann sich auch auf die Analyse das Setting, die Erfahrung und Hintergründe des Coders/der Coderin auswirken. Je weniger Störfaktoren und Verzerrungen, desto klarer kann der Forscher analysieren und je klarer analysiert werden kann, desto eher kann die Emergenz der hinter dem Material (bei Inhaltsanalysen sind es oft Texte) verborgenen Motive und/oder Interessen erfolgen. Bei quantitativen standardisierten Inhaltsanalysen kann es sein, dass das a priori gestellte Kategorienschema bei einer Auswahl von Fällen an seine Grenzen gerät. Um solche Fälle zu identifizieren können zusätzliche CoderInnen durch eventuell abweichende Codierungen Differenzen zwischen dem Ziel des Kategorienschemas und dem empirischen Befund

aufzeigen.

## 6 Aufgabe 6 - Cohen's Kappa

”Man versteht darunter die Höhe der Übereinstimmungen der Einschätzungsergebnisse unterschiedlicher Beobachter bzw. unterschiedlicher »Testanwender« (Rater).” [10, S. 24].

Reliabilität als allgemeines Gütekriterium bezeichnet wie zuverlässig ein Instrument gleiche oder ähnliche Ergebnisse liefert. Wenn mehrere Personen zum Beispiel bei einer Textanalyse Textpassagen kategorisieren sollen ist die ”Messung” also die Einschätzung der Rater reliabel, wenn diese zu gleichen oder ähnlichen Ergebnissen kommen. Je höher die Interraterreliabilität ist, desto reliabler ist oder wird auch die ”Messung” (bzw. der Codierprozess) verlaufen. Um die Höhe der Übereinstimmung einschätzen zu können wurden im Laufe der Zeit verschiedene Messgrößen entwickelt, die populärste ist Cohen's Kappa ( $\kappa$ ) [12]. Es gibt viele unterschiedliche Messgrößen, die versuchen einige Nachteile von Cohens  $\kappa$  auszugleichen, diese führen aber in den meisten Fällen durch ihre Veränderung eigene, oft schwerwiegendere, Nachteile ein [11]. Bei mehr als 2 Ratern müsste Cohen's Kappa für alle Raterkombinationen dargestellt werden, oder man bedient sich eben einer alternativen Größe (z.B. Fleiss  $\kappa$ ). Diese birgt aber wie gerade besprochen selbst wieder neue Limitationen. Daher wird im folgenden Beispiel Cohen's Kappa als Messgröße von Interraterreliabilität exemplarisch berechnet werden:

### 6.1 Beispiel

Zu Beginn muss eine Übereinstimmungsmatrix des Codierverhaltens konstruiert werden

		Rater 1				Summe
		K1	K2	K3	K4	
Rater 2	K1	23	6	4	7	40
	K2	7	18	4	6	35
	K3	5	6	12	2	25
	K4	1	2	1	1	5
	Summe	36	32	21	16	105

Die Formel für Cohen's Kappa ist

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (1)$$

Zu Beginn muss  $P_o$  und  $P_e$  berechnet werden, wobei  $P_o$  die Übereinstimmungen der Rater sind und  $P_e$  der Anteil der Fälle ist, bei denen zufällig eine Übereinstimmung erwartet wird.

$$P_o = \frac{23 + 18 + 12 + 1}{105} = 0.5143 \quad (2)$$

$$P_e = \left( \frac{36}{105} \cdot \frac{40}{105} \right) + \left( \frac{32}{105} \cdot \frac{35}{105} \right) + \left( \frac{21}{105} \cdot \frac{25}{105} \right) + \left( \frac{16}{105} \cdot \frac{5}{105} \right) = 0.3476 \quad (3)$$

Die Ergebnisse können nun in die Formel eingesetzt werden.

$$\kappa = \frac{P_o - P_e}{1 - P_e} = \frac{0.5143 - 0.3476}{1 - 0.3476} = 0.255 \quad (4)$$

Das Ergebnis könnte nun schon interpretiert werden, in seinem Text macht Cohen selbst einen Signifikanztest [12, S. 43], dafür muss zunächst der Standardfehler berechnet werden. Danach kann ein Signifikanztest mit einem Signifikanzniveau von 95% durchgeführt werden. Die Formeln finden sich in Cohen's Artikel als Gleichung 7 und 10 [12].

$$\sigma_{\kappa} = \sqrt{\frac{p_o(1-p_o)}{N(1-p_e)^2}} = \sqrt{\frac{0.5143(1-0.5143)}{105(1-0.3476)^2}} \approx 0.0748 \quad (5)$$

$$\sigma_{\kappa_0} = \sqrt{\frac{p_e}{N(1-p_e)}} = \sqrt{\frac{0.3476}{105(1-0.3476)}} \approx 0.0712 \quad (6)$$

$$95\% \text{Konfidenzintervall} = \kappa \pm 1.96\sigma_{\kappa} = 0.255 \pm 1.96 \cdot 0.0748 \approx (0.1084, 0.4016) \quad (7)$$

Die Übereinstimmung der Rater ist statistisch signifikant.

$$Z = \frac{\kappa}{\sigma_{\kappa}} = \frac{0.255}{0.0748} \approx 3.41 \quad (8)$$

$$\text{p-Wert} = 2 * (1 - 0.9997) \approx 0.0006 \quad (9)$$

Das Ergebnis ist sogar auf einem Niveau von  $p < 0.001$  signifikant.

## 7 Literatur

### Literaturliste ab Aufgabe 2

1. Sältzer, M. & Stier, S. Die Bundestagswahl 2021 auf Twitter. *easy social sciences* **67**, 30–38. ISSN: 2199-9082 (2022).
2. *quanteda tutorial* letzter Zugriff: 24.7.2023 (2018). <https://tutorials.quanteda.io/machine-learning/wordfish/>.
3. Slapin, J. B. & Proksch, S.-O. A Scaling Model for Estimating Time-Series Party Positions from Texts. *American Journal of Political Science* **52**, 705–722. ISSN: 00925853, 15405907. <http://www.jstor.org/stable/25193842> (2023) (2008).
4. Carsi González, A. *Unleashing the Power of Principal Component Analysis: Diagnose of OA degenerative joint disease* unpublished. Juli 2023.
5. Jünger, J. & Gärtner, C. *Computational Methods für die Sozial- und Geisteswissenschaften* <https://link.springer.com/10.1007/978-3-658-37747-2> (2023) (Springer Fachmedien Wiesbaden, Wiesbaden, 2023).
6. Mueller, S. D. & Saeltzer, M. Twitter made me do it! Twitter's tonal platform incentive and its effect on online campaigning. *Information, Communication & Society* **25**, 1247–1272. <https://doi.org/10.1080/1369118X.2020.1850841> (2020).
7. Flick, U. *Sozialforschung : Methoden und Anwendungen ; ein Überblick für die BA-Studiengänge*. (Rowohlt, Hamburg, 2009).
8. Weber, M. *Wirtschaft und Gesellschaft* (J.C.B. Mohr (Paul Siebeck), 1922).
9. Geissler, D., Bär, D., Pröllochs, N. & Feuerriegel, S. *Russian propaganda on social media during the 2022 invasion of Ukraine* 2023. arXiv: [2211.04154 \[cs.SI\]](https://arxiv.org/abs/2211.04154).

10. Hussy, W., Schreier, M. & Echterhoff, G. *Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor* 2. Aufl. (Springer, Juni 2013).
11. Hsu, L. M. & Field, R. Interrater Agreement Measures: Comments on Kappan, Cohen's Kappa, Scott's pi, and Aikin's alpha. *Understanding Statistics* **2**, 205–219 (2003).
12. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **20**, 37–46. <https://doi.org/10.1177/001316446002000104> (1960).