



universität
wien

MAGISTERARBEIT / MASTER'S THESIS

Titel der Magisterarbeit / Title of the Master's Thesis

Populism? Computer says no!

Über die Verwendung von Supervised Machine Learning in der Erkennung von
Populismus in Medientexten

verfasst von / submitted by

Gregor Rettenegger, Bakk.phil.

gemeinsam mit / in collaboration with

Maximilian Haselsteiner, Bakk.phil.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magister der Philosophie (Mag. phil.)

Wien, 2019 / Vienna, 2019

Studienkennzahl lt. Studienblatt /

A 066 841

degree programme code as it appears on
the student record sheet:

Studienrichtung lt. Studienblatt /

Magisterstudium Publizistik- und

degree programme as it appears on the
student record sheet:

Kommunikationswissenschaft

Betreut von / Supervisor:

Univ.-Prof. Hajo Boomgaarden, PhD

Inhaltsverzeichnis

Abbildungsverzeichnis.....	IV
Anmerkungen zur Autorenschaft dieser Arbeit	V
1. Einführung	1
1.1. Einleitung	1
<i>Autor: Maximilian Haselsteiner</i>	
1.2. Erkenntnisinteresse.....	3
<i>Autor: Maximilian Haselsteiner</i>	
2. Theorie	4
2.1. Politische Kommunikation.....	4
<i>Autor: Maximilian Haselsteiner</i>	
2.2. Definition von Populismus.....	5
<i>Autor: Maximilian Haselsteiner</i>	
2.2.1. Der Bedarf nach Populisten.....	8
<i>Autor: Maximilian Haselsteiner</i>	
2.2.2. Populismus und das Volk	11
<i>Autor: Maximilian Haselsteiner</i>	
2.2.3. Populismus und die Eliten	13
<i>Autor: Maximilian Haselsteiner</i>	
2.3. Die Rolle der Medien	17
<i>Autor: Gregor Rettenegger</i>	
3. Messen von Populismus.....	19
<i>Autor: Gregor Rettenegger</i>	
3.1. Manuelle Ansätze	20
<i>Autoren: Maximilian Haselsteine (S. 20-22) & Gregor Rettenegger (S. 23-25)</i>	
3.2. Automatisierte Ansätze	26
<i>Autor: Gregor Rettenegger</i>	
3.2.1. Wörterbuch Ansatz.....	26

<i>Autor: Gregor Rettenegger</i>	
3.2.2. Supervised Machine Learning (SML)	29
<i>Autor: Gregor Rettenegger</i>	
3.2.3. Wahl des Werkzeuges	36
<i>Autor: Maximilian Haselsteiner</i>	
3.3. Validierung automatisierter Verfahren.....	40
<i>Autor: Maximilian Haselsteiner</i>	
4. Methode	45
<i>Autor: Gregor Rettenegger</i>	
4.1. Forschungsfragen	45
<i>Autor: Gregor Rettenegger</i>	
4.2. Untersuchungsmaterial.....	47
<i>Autor: Gregor Rettenegger</i>	
Exkurs: Die österreichische Nationalratswahl 2017.....	48
<i>Autor: Gregor Rettenegger</i>	
4.3. Manuelle Inhaltsanalyse, Erstellen des Trainings-Sets	50
4.3.1. Voruntersuchung	50
<i>Autor: Maximilian Haselsteiner</i>	
4.3.2. Operationalisierung und Codebuch	52
<i>Autor: Maximilian Haselsteiner</i>	
4.3.3. Durchführung und Nachbearbeitung	56
<i>Autor: Gregor Rettenegger</i>	
5. Ergebnisse	59
<i>Autor: Gregor Rettenegger</i>	
6. Ergebnis Diskussion.....	66
6.1. Faktoren für die Qualität eines Trainingssets.....	66
<i>Autor: Gregor Retteneggerr</i>	
6.2 Beispiele für False Positives und False Negatives	68
<i>Autor: Maximilian Haselsteiner</i>	

7. Conclusio	72
<i>Autor: Maximilian Haselsteiner</i>	
8. Literaturverzeichnis	79
9. Appendix.....	87
9.1. Abstract, deutsche Version.....	87
9.2. Abstract, english version	88
9.3. R-Script	89
9.4. Codebuch.....	93

Abbildungsverzeichnis

Abbildung 1: "Encoding references" aus Jagers et. al. 2007: 339 (Appendix).....	21
Abbildung 2: Auszug aus dem "Populist Speech Rubric" in Hawkins et al. 2009:1063f, Appendix	22
Abbildung 3: Classification Accuracy, aus Burscher et. al., 2014: 201	31
Abbildung 4: Zusammenhang zwischen Genauigkeit der Klassifikation und Anzahl der Trainingsdokumente für die verschiedenen Frames, entnommen aus Burscher et.al. 2014: 201.....	33
Abbildung 5: Interkoder Reabilität für manuelle Inhaltsanalyse, aus Scharkov, 2013: 767.....	34
Abbildung 6: Klassifikations- Qualität für SML, aus Scharkov, 2013: 768.....	34
Abbildung 7: Zusammenhang zwischen Interkoder Reliabilität und Klassifikations- Qualität, ebd.....	35
Abbildung 8: Vereinfachte bildliche Darstellung von Regessions-Trees.....	39
Abbildung 9: "Parallel measurement of populist communication and attitude" aus Wirth et al. 2016: 52.	53
Abbildung 10: : Ein Auszug aus "Conceptualization and Operationalization of Populist Communication" (Ernst et al. 2017: 258f).....	53
Abbildung 11: Variabel- und Kategorieschema für Anti- Elitismus aus dem NCCR Codebuch	54
Abbildung 12: Verwendete Artikel nach Medium, Erstellen des Trainingssets.....	57
Abbildung 13: Populismus ab Wahrscheinlichkeit >50%, N/P im Trainingsset = Anzahl an Positives/Negatives.....	60
Abbildung 14: Populismus ab Wahrscheinlichkeit >25%, N/P im Trainingsset = Anzahl an Positives/Negatives.....	61
Abbildung 15: Entwicklung von Precision und Recall über sechs Iterationen.....	61
Abbildung 16: Verschiedene Experimente in den Iterationen	63
Abbildung 17: Darstellung der wichtigsten "Features" in den Entscheidungsbäumen	64
Abbildung 18: Exemplarischer Descision Tree I.....	65
Abbildung 19: Exemplarischer Descision Tree II	65

Anmerkungen zur Autorenschaft dieser Arbeit

Diese Arbeit wurde gemeinsam verfasst. Sowohl der theoretische, als auch der methodologische Aufbau wurden gemeinsam geplant, diskutiert, durchgeführt und evaluiert. Dementsprechend schwer ist es, eine klare Autorenschaft für einzelne Kapitel festzusetzen. Bei jedem Kapitel gilt der jeweils andere Autor dieses Werks also als Co-Autor.

1. Einführung

1.1. Einleitung

Autor: Maximilian Haselsteiner

*Jegliche geschlechterspezifische Begriffe in diesem Dokument
umfassen Frauen und Männer gleichermaßen.*

Üblicherweise würde man in einer Problembenennung einer Abhandlung zum Thema Populismus an dieser Stelle mit einem leicht warnenden Unterton von einem Aufstieg des Populismus sprechen. Ionescu und Geller eröffnen 1969 ihre viel beachtete Abhandlung über Populismus mit den Worten „A Spectre is haunting the world – Populism“. Doch von einem „Aufstieg“ kann heute beinahe nicht mehr die Rede sein. Überall auf der Welt finden sich populistische Parteien und Akteure in Machtpositionen. Dabei wird das Label „Populist“ verschiedenen politischen Strömungen zugeordnet. Von dem linkspopulistischen Regime in Venezuela, bis hin zu den bürgerlich-populistischen „Cuidadanos“ in Spanien, über die „Syriza“ in Griechenland, bis zu den Rechtspopulisten in Europa und den Vereinigten Staaten – Populisten haben sich adaptiert, und weder der flüchtige und bloß episodenhafte „Populist Zeitgeist“ (Mudde, 2004), noch die prophezeite Selbstzerstörung von Innen scheinen sich zu bewahrheiten. Der Populismus ist ein globaler und sehr erfolgreicher Diskurs (Manucci, 2017: 466).

Über soziale Medien erreichen Politiker wie Donald Trump ihre Zielgruppen besser, als es sich das sozialistische Zentralorgan die „Arbeiterzeitung“ oder „Das kleine Volksblatt“ der ÖVP je erträumen konnten. Vor allem neuen politischen Akteuren und Herausforderern die in traditionellen Medien weniger Beachtung erfahren bieten die neuen Kanäle eine Möglichkeit die Gatekeeper zu neutralisieren (Skovsgaard & Van Dalen, 2013). Die Flut an Informationen wird immer schwieriger zu Ordnen und so sind auch Forscher dazu gezwungen neue Methoden zu entwickeln um Populismus zu systematisieren.

In der jungen Vergangenheit gab es deshalb auch eine Vielzahl an Publikationen zum Thema Populismus. Zwar gibt es immer noch hunderte verschiedene Definitionen für Populismus, und noch immer beginnen Abhandlungen zu Populismus mit einer eigenen Definition, doch kehren mehrere Punkte regelmäßig wieder, und Cas Muddes „Ideational

Approach“, eine Minimaldefinition, hat sich bereits in internationalen Vergleichen bewährt (vgl. z.B. Pauwells, 2017) und findet zumindest – selbst in philosophischen Betrachtungen - immer Erwähnung. Dabei wird Populismus auf seine Kernelemente reduziert: Auf die Kritik an sozialen, ökonomischen und politischen Eliten die als diametraler Gegenspieler zu einem monolithischen, reinen Volk gezeichnet werden. Der Populist versteht sich als einziger Repräsentant dieses Volkes und will ihm seine Souveränität zurückgeben (siehe Kapitel 2.2.).

Es gab in letzter Zeit verschiedenste Versuche Populismus in seinen vielen Facetten valide messbar zu machen (siehe Kapitel 3) und es gibt auch viele positive Beispiele, wie das herausragende und äußerst umfangreiche Codebuch “The Appeal of Populist Ideas and Messages” des Schweizer Instituts “National Centre of Competence in Research (NCCR)” rund um Werner Wirth, Martin Wettstein, Lucca Manucci und anderen. Doch für die Durchführung der Studie mussten außer dem ohnehin schon großen Forscherteam weitere 76 Kodierer trainiert, evaluiert und bezahlt werden (ebd., 2017: 1). Vor allem für kleinere Forscherteams ist es so beinahe unmöglich, repräsentative Samples zu erstellen und große Datenmengen zu verarbeiten. Mit neuen Methoden kann Anti-Elitismus durch computergestützte Verfahren aus einer enormen Masse an Botschaften und Informationen erhoben werden und über die Grenzen von Politischen Lagern und Nationalstaaten generalisiert werden (z.B. Rooduijn und Pauwels, 2011). Doch hat dieses Verfahren, allen voran der so genannte „Dictionary-Approach“, welcher Textmengen anhand von Nennungen von Wörtern kategorisiert, Schwierigkeiten dabei, komplexere sprachliche Elemente zu erkennen. Und da es sich bei Populismus um ein sehr komplexes Konstrukt handelt und schwer an einzelnen Wortzählungen dingfest gemacht werden kann, ist es auch bisher kaum möglich, den zweiten Pfeiler des Populismus, die ständige Referenz an das Volk, durch automatisierte Verfahren zu erheben.

Zu diesem Problem will die vorliegende Arbeit einen Beitrag leisten. Mit Hilfe des rasant wachsenden Forschungsstandes über “Advanced Machine Learning” Modelle und Algorithmen (insbesondere den Supervised Learning Methods), soll ein Instrument erstellt werden, welches die computergestützte Erkennung von Populismus nicht nur umfangreicher, sondern auch präziser machen soll. Durch selbstlernende Programme, welche den aktuellen Forschungsstand entsprechen (siehe Kapitel 3.2.), soll es Forschern künftig möglich sein, signifikant weniger Zeit mit der manuellen Kodierung von Inputs aller Art aufwenden zu müssen. Diese Arbeit will dabei nicht zu tief in die komplexen

Vorgänge und Hintergründe des Advanced Machine Learnings eintauchen, sondern in aller Verständlichkeit die damit verbundenen Möglichkeiten und Risiken darstellen. Als Ergebnis dieser Arbeit sollte ein Instrument mit Hilfe der Programmbibliothek “XGBoost” (Chen & Guestrin, 2015) erstellt werden, welches Zeit und Ressourcen für Teams oder einzelne Forscher sparen kann und auch ohne einen Abschluss in Statistik und Informatik eingesetzt werden kann.

1.2. Erkenntnisinteresse

Autor: Maximilian Haselsteiner

In dieser Arbeit soll eine reproduzierbare Methode entwickelt werden um Populismus in Medientexten nicht nur Anhand eines seiner Elemente, sondern durch Texterkennungs-Algorithmen in seiner Gesamtheit erkannt werden kann. Das Instrument sollte leicht verwendbar und mit geringen statistischen und informationstechnischen Wissen einsetzbar sein.

Dazu wird in einem ersten Schritt eine aktuelle Definition von Populismus gefunden und erklärt werden. In einem zweiten Schritt sollte gezeigt werden, welche verschiedenen Methoden zur Erhebung es gibt und welche Vorteile computergestützte bzw. automatisierte Instrumente mit sich bringen. Drittens wird ein Trainingsset für das statistische Modell erstellt. Diese werden manuell auf Grundlage - jedoch stark vereinfacht - des NCCR Codebuchs kodiert und aufgeteilt in ein Lern- und Validierungsset in das algorithmische Modell gefüttert. Viertens werden mehrere Iterationen vorgenommen, um das Instrument zu präzisieren. Fünftens soll der Prozess ausführlich dargestellt werden, um das Instrument auch für andere Anlässe reproduzierbar zu machen. In einem letzten Schritt werden die Möglichkeiten und Stärken für zukünftige Forscher, aber auch die Schwierigkeiten im Prozess und Risiken benannt werden.

“Supervised Machine Learning” Methoden wurden in Bezug auf Populismus bisher kaum erforscht und in deutscher Sprache stellt diese Arbeit ein absolutes Novum dar.

2. Theorie

2.1. Politische Kommunikation

Autor: Maximilian Haselsteiner

Seit den 1970er Jahren boomt die Forschungsdisziplin „Politische Kommunikation“. Dabei umfasst dieses Feld sowohl jene Kommunikation welche von politischen Akteuren ausgeübt, an sie gerichtet oder sich auf ihre Aktivitäten bezieht (siehe etwa Definition Schulz, 2011:16), als auch „Formulierung, Aggregation, Herstellung und Durchsetzung kollektiv bindender Entscheidungen“ (Jarren & Dongens, 2006: 22). Politische Kommunikation ist also selbst Politik. Diese Arbeit wird sich vor allem mit Zweiterem beschäftigen.

Das Verhältnis zwischen Medien und Politik wurde über die Zeit unter verschiedenen Gesichtspunkten diskutiert, derzeit findet sich meist die Annahme, dass es sich um ein wechselseitiges Penetrationsverhältnis mit wechselseitigen Abhängigkeiten und Anpassungsprozessen handelt: Politiker versorgen die Medien mit Material und diese wiederum verschaffen den Politikern Publizität (vgl. Schulz, 2011: 51).

In der Wahrnehmung von Medien an sich wird oft vergessen, dass es sich bei Medien selbst um Unternehmen handelt. Sie müssen Publizität an Werbetreibende verkaufen um profitabel zu bleiben. Selbst die meisten öffentlich-rechtlichen TV-Sender brauchen Werbung um Kosten zu decken. Laut dem wichtigen linken Intellektuellen Noam Chomsky sind Medien damit auch gezwungen, auf die Interessen der Werbetreibenden und der Konsumenten einzugehen:

(Die Medien) "are corporations 'selling' privileged audiences to other businesses. It would hardly come as a surprise if the picture of the world they present were to reflect the perspectives and interests of the sellers, the buyers, and the product..." Chomsky, 1989:8.

Medien müssen ihre Logik also so anpassen, dass das Interesse des Publikums so lange wie möglich gewährleistet wird. Im Generellen unterscheidet man zwischen der Selektionslogik, also der Auswahl der Nachrichten, und der Präsentationslogik. Ersteres Orientiert sich unter anderen an z.B. von Schulz (1967) definierten Nachrichtenfaktoren, diese umfassen Zeit, Nähe, Status, Dynamik, Valenz (etwa Konflikte, Erfolge oder Schäden) und Identifikation. Die Präsentationslogik umfasst alle Formen der Inszenierung

und Darstellung der Nachrichten (vgl. Meyer, 2006: 82). Unterwirft sich nun ein Politiker dieser Logiken vollkommen, kann man damit ein unglaubliches Medienecho erzeugen. Zeitgemäß muss hier der Name Donald Trump genannt werden. Er ist eine Mischung aus Entertainer und Politiker und transformierte sich von einer Figur der Popkultur und einem Markennamen zu einem “Populist Hero”, der es beherrscht wie kein anderer die Nachrichten Agenda zu setzen (vgl. Wells et al., 2016: 669). Und diese Taktik funktionierte. Entgegen allen Expertenmeinungen, Leitmedien und selbst Politikern aus seiner eigenen Partei schaffte er es, eine nie dagewesene Publizität zu erreichen. Die New York Times schrieb etwa, dass der Wert dieser Präsenz beinahe zwei Milliarden US-Dollar erreicht (vgl. NY-Times; Confessore & Yourish, 2016). Eine nähere Befassung mit dem Thema findet sich weiter unten (Kapitel 2.7.: Die Rolle der Medien).

2.2. Definition von Populismus

Autor: Maximilian Haselsteiner

Die Begriffsverwirrung um den Populismus ist beinahe so alt wie das Konzept selbst. Als Ionescu & Geller 1969 eine noch heute oftmals zitierte Arbeit über Populismus schreiben sind die Rezensionen oft niederschlagend: „The inability of the writers to come up with an acceptable definition of populism simply confounds the reader“ (Saloutos, 1970: 328). Auch heute noch müssten sich “eine verstörend hohe Anzahl“ (Mudde, 2017: 27) von Wissenschaftler vorwerfen lassen, das Konzept zu verwenden ohne es hinreichend definiert zu haben. Selbst jene, welche eine hinreichend begründete Definition verwenden, kommen teilweise auf sehr unterschiedliche Ergebnisse. So wird Populismus von Backes und Jesse (1998) beispielsweise als reines „Machteroberungskonzept“ beschrieben, was sich empirisch jedoch nicht begründen lassen konnte. Jagers und Walgrave (2007) definieren Populismus als Kommunikationsstrategie bzw. Stil. Je nach Disziplin wird Populismus auch als politischer Rahmen (Lee, 2006), als Identifikationsmodus (Panizza, 2005), oder gar als Sprache (Kazin, 2005) definiert.

Der ideenbildende, oder ideologische Ansatz an Populismus war schon immer Teil der Diskussion. Schon 1977 versuchte Ernesto Laclau die Definitionen zu systematisieren und kommt dabei zu dem Schluss, dass von vier grundlegenden Ansätzen, drei Populismus als Bewegung (movement) und Ideologie sehen, während der vierte Populismus lediglich als Ideologie ansieht (vgl. Laclau, 1977: 144). Gerade durch komparatistische Studien hat der

„ideational approach to populism“ in letzter Zeit große Bekanntheit erlangt. Bedeutend weiterentwickelt wurde dieser von Cas Mudde, welcher Populismus als „dünne Ideologie“ beschreibt und die Definition auf Punkte reduziert, welche in beinahe allen Definitionen vorkommen:

„Populism is a thin-centered ideology that considers society to be ultimately separated into two homogenous and antagonistic groups; ‘the pure people’ versus the ‘corrupt elite’, and which argues that politicians should be an expression of the volonté générale of the people.”

(Mudde, 2004: 543)

Als “thin-centered ideology” meint Cas Mudde dabei, dass sie im Gegensatz zu so genannten „vollen Ideologien“ wie Liberalismus, Konservatismus etc. nicht für sich selbst existiert, sondern sich an jene (üblicherweise Neo-Liberalismus, Sozialismus bis hin zu Rechtsradikalismus) anhängt und mehr oder weniger einen limitierten, grundsätzlicheren Ansatz oder eine Perspektive auf die Politik darstellt (vgl. Hawkins, 2012: 4). Neben der Ideologie beinhaltet die Definition drei weitere Kernkonzepte: „the people“ welches sowohl als das Volk als auch als „der kleine Mann“ verstanden werden kann, „the elites“ welche die diametralen Gegenspieler des Volkes bzw. der Leute sind und dem „volonté Générale“ welcher von deutschsprachigen Akteuren meist als “gesunder Hausverstand” oder auch “Wille des Volkes” angesprochen wird. Der Ansatz vereint vier wichtige Stärken: Unterscheidbarkeit, Kategorisierbarkeit, Vergleichbarkeit („travelability“) und Vielseitigkeit („versatility“).

Obwohl es vor allem in den Vereinigten Königreich, in Skandinavien und den Benelux Staaten einen breiten Konsens über Muddes Konzeption von Populismus gibt, hat der Ansatz auch Schwächen. Zwar ist es möglich, Populisten klar von Nicht-Populisten zu unterscheiden, allerdings fällt es schwer verschiedene Formen von Populismus zu unterscheiden und zu gewichten (vgl. Heinisch und Mazzoleni, 2017: 106f). Weiters kritisiert Heinisch und Mazzoleni, dass wichtige Aspekte, wie die Variabilität – etwa, dass Populismus keine universellen Leitwerte haben und stattdessen, chamäleonartig, Attribute ihrer Umwelt aufsaugen (vgl. Taggart, 2000: 5) - und die Ambivalenz - eine kalkulierte Ambivalenz, welches es Populisten erlaubt verschiedene Zielgruppen gleichzeitig anzusprechen (Wodak, 2015:46) - in der Konzeption fehlen. Auch der Ideologiebegriff an sich ist kritikwürdig, so distanziert sich Michael Freeden, der den Begriff der “dünnen Ideologie” eingeführt hat (1966), von dem Konzept, unter anderem auch, wie Aslandis

(2015: 7) anmerkt, wegen der Gefahr Populismus normativ oder pathologisch zu erfassen. Heinisch und Mazzoleni (2017: 117) versuchen diese Probleme zu lösen, indem sie Frames und Akteure einführen und vernachlässigte Komponenten einführen:

“Populism can be defined as a frame containing intrinsically ambivalent claim(s) diffused by individual and collective actors in order to challenge the status-quo in favor of people’s empowerment and elite change.”

Dabei definieren sie verschiedene Frames, wie den “Master-Frame”, die Verteidigung des tugendhaften Volkes gegen die Maschinerie der Eliten und das Versprechen von politische Veränderungen um die Macht dem Volk zurückzugeben (ebd.: 119). Doch mit dieser Definition gehen alte Probleme der Konzeption von Populismus einher. So sind etwa gewisse Teile wie die “intrinsically ambivalent claims” oder auch die Rolle der Akteure wohl kaum universell anwendbar. Wichtige Komponenten, über die ein beachtlicher Konsens besteht, werden dafür vernachlässigt.

Das Forscherteam um Werner Wirth (2016)¹ versucht sich an einem integrativen Rahmenwerk für Populismusforschung. Dabei bauen sie zwar auf Muddes (2004) ideologischen Definition auf, versuchen aber Populismus als Strategie und als Stil systematisch in Verbindung zu bringen, ohne die Konzepte jedoch zu vermischen (Wirth et al., 2016: 39). Dabei beschreiben sie Ideologie sehr ähnlich wie Mudde, als der Gedankenwelt der Populisten (und eventuell auch ihren Wählern) und die Kommunikation als Folge dieser Ideologie.

Erstens, das gute und homogene Volk führt zu Kommunikation welche ein monolithisches Volk direkt andeuten oder durch das Inkludieren und Exkludieren anderer Gruppen und Individuen definieren. Die Tugendhaftigkeit und Reinheit des Volkes wird ebenso hervorgehoben, wie die Tatsache, dass der populistische Kommunikator selbst dieser Gruppe angehört oder sehr nahe ist. Zweitens führt der ideologische Anti-Elitismus zu Kommunikation über negative Eigenschaften der Eliten, negative Folgen ihrer Entscheidungen und der Sicht auf sie als Problem oder Last für die Gesellschaft. Drittens, wird der Volkswille kommuniziert, in dem man mehr Souveränität für das Volk fordert und die Legitimität oder die Macht der Eliten verneint. Ebenfalls wird ein Anti-Populismus beschrieben, welcher etwa Pluralismus befürwortet, dem Volk schlechte Eigenschaften

¹ Ebenfalls genannt werden Esser, Wettstein, Engesser, Wirz, Schulz, Ernst, Büchel, Caramani, Manucci, Steenbergen, Bernhard, Weber, Hänggli, Dalmus, Schemer und Müller.

zuschreibt und vice versa die Eliten lobt. Dabei wird populistische Kommunikation durch eine Kombination von mehreren dieser Elemente und dem Fehlen von Anti-Populismus definiert (vgl. ebd.: 41f).

Als dritten Punkt abstrahieren sie populistischen Stil von den Kernelementen. Aus People-Centrism folgt etwa eine einfache und emotionale Sprache, Simplifizierung und einem lockeren Auftreten. Aus der manichäischen Aufteilung zwischen dem “tugendhaften Volk” und den “korrupten Eliten” ergibt sich ein schwarz-weiß Denken und die Verwendung von Absoluten. Mit Dramatisierung und Skandalisierung werden die schlechten Entwicklungen durch die Eliten weiter betont und schließlich verweist der “gesunde Hausverstand” und die “Alltagsschläue” auf die Souveränität des Volkes (vgl. ebd.).

Auch aus dieser Darstellung von Populismus ergeben sich Probleme, etwa bezüglich der Unterscheidbarkeit und der Vergleichbarkeit. So kann aus dem “populistischen Stil” keineswegs auf eine populistische Ideologie geschlossen werden. Eher verweisen sie auf allgemeine Eigenschaften der Interdependenzen zwischen Medien und politischer Kommunikation (siehe u.a. Kapitel 2.7.).

Auch wenn sich diese Ansätze unterscheiden mögen, so merkt Mudde (2016: 41) selbst an, dass sich die Forschung auf die vielen Gemeinsamkeiten in den verschiedenen Definitionen konzentrieren soll und weniger auf die wenigen Unterschiede. Van Kessel (2015: 8) ergänzt, dass verschiedene Interpretation und Deutungen aus einer empirischen Sicht kein Problem sind, solange nur ein Konsens über die Kernattribute herrscht.

2.2.1. Der Bedarf nach Populisten

Autor: Maximilian Haselsteiner

Es gibt nicht nur ein Angebot, sondern auch einen Bedarf an Populistischer Politik. Ohne diesen Bedarf wäre es nicht möglich, jene Massen zu mobilisieren, welche populistischen Parteien Erfolge ermöglicht hat, wie beispielsweise einer Regierungsbeteiligung in Österreich, Dänemark, Italien, Griechenland, dem Einzug und Erhalt von einem Drittel der Stimmen in der französischen Stichwahl für die Präsidentschaft, oder den Oberbefehl über

das mächtigste Militär der Welt². Eine von vielen Autoren unterstützte These ist, dass sich die demokratische Repräsentation seit dem Beginn des 21sten Jahrhunderts in einer Krise befindet (vgl. z.B. Mastropaolo, 2017: 59). Die großen Lager Parteien sind in den meisten Nationen zusammengebrochen, da die Rolle als Repräsentant für eine bestimmte soziale Schicht oder Gruppe meist nicht mehr gegeben ist. Im Gegenteil haben sich politische Parteien nach Katz und Mairs “cartel party theory” (1995) in Vermittler verschiedener Interessen gewandelt. Die “Alt-”Parteien identifizieren sich mit der Regierung und müssen daher oft auch Politiken unterstützen, welche eigentlich gegen ihre Kernwählerschaft gehen. Sie haben ihre Rolle als Advokaten oder Sprecher der Bürger verloren und egal wie sehr sie sich an der Macht klammern, die allgemeine Unzufriedenheit wird eher größer. Hinzu kommt, dass etwa christliche Volksparteien durch die wachsende Säkularisierung traditionelle Wähler verlieren, wohingegen sozialistische und sozialdemokratische Parteien durch den wachsenden Wohlstand ehemaliger Arbeiter und sich verändernde Industrien einen Teil ihrer Kernwählerschaft verliert.

Laclau (2005) versteht den Populismus u.a. als Reparaturmechanismus für diese Krise. Er erinnert die Demokratie an sich selbst und re-politisiert diese. Die Realität und Komplexität der Repräsentativen Demokratie schafft Frustration und Zorn bei vielen Menschen, sie fühlen sich nicht mehr repräsentiert durch etablierte Eliten, welche die Interessen verschiedenster Akteure aushandeln (vgl. Wodak, 2015: 28). Genau dort setzen Populisten an: Bei der Kritik an eben jenen Eliten, bei der Reduzierung der Komplexität auf einfache Dichotomien - wir gegen die - auf einen kompromisslosen Umgang mit Minderheiten und der Huldigung des ‘gesunden Hausverstandes’, der es schon irgendwie richten wird. Populismus versucht den Desillusionierten eine Stimme zu geben und rüttelt an dem System. Alte Parteien werden als Establishment, Eliten oder “die da oben” verteufelt, man selbst steht neben diesem Parteiensystem und stilisiert sich zum einzigen Beschützer der Nation, der Leute oder des Volkes.

Als konkrete Gründe für einen Bedarf an Populismus nennt Butterwege (1996: 64-67) - allerdings mit einem Fokus auf Deutschland - Reaktionen auf die ökonomischen Krisen, auf die Zuwanderung, immer größere Konkurrenz, Globalisierung und die unüberschaubare Risikogesellschaft. Aber es gibt auch sozioökonomische Aspekte, welche Bedarf an

² Gemeint ist Donald Trump, welcher kürzlich von Historiker Timothy Snyder als “Sado-Populist” betitelt wurde.

Populisten schaffen. Einige der häufigsten Erklärungsansätze für die Wahl rechtsradikaler Parteien sind und bleiben verschiedene Deprivationsthese. Eine generelle Version der Deprivationsthese liefert etwa Lipset (1962: 110). Er sieht niedriges Einkommen, hohe Einkommensunsicherheit und Arbeitslosigkeit als wichtige Faktoren für die Unterstützung eben jener Parteien. Neuere Studien ergänzen die Deprivationsforschung um geringe Arbeitsplatzsicherheit, prekäre Arbeitsverhältnisse (vgl. De Weerd et.al., 2004), aber auch Missständen in den allgemeinen Lebensbedingungen wie soziale Kontakte und Wohnverhältnisse (vgl. Stöss, 1989: 236). Obwohl diese Arbeiten meist nicht Populistische, sondern Parteien am rechten Außenrand behandeln, so gleicht das Untersuchungsmaterial sehr stark jenem, welches Populismus Forscher in Europa untersuchen. Spier (2010) untersuchte jene Komponenten in Bezug auf europäische Wähler von rechtspopulistischen Parteien und kommt zu dem Schluss, dass nicht die tatsächliche Deprivation ausschlaggebend ist, sondern die subjektive, im Vergleich zu seinem Umfeld wahrgenommene Deprivation:

„Statusdeprivation ist ein durchaus sichtbares Merkmal rechtspopulistischer Elektorate. [...] sie bleiben also in ihrem sozioökonomischen Status hinter den Erwartungen zurück, die sie aufgrund ihres höchsten Bildungszertifikates haben konnten.“ (Spier 2010: 263).

Wodak (2015: 26) fasst diese Theorien als die 1994 von Hans-Georg Betz benannten “Modernisierungsverlierer-Theorien” zusammen. Auch wenn der Begriff in der Zwischenzeit stark kritisiert wurde, wurde er in jüngster Zeit, vor allem in subjektiver Form nachgewiesen (vgl. Minkenberg und Perrineau, 2007: 32-42).

Als weitere Gründe für den Erfolg populistischer Parteien wurden unter anderem die mediale Themenlage (Blumer, 2003; Boomgaarden & Vliegenthart, 2007), deren oft charismatische Spitzenkandidaten (Plasser & Ulram, 2002; Stewart et al., 2003) oder sich verändernde politischen System in denen Partei-Treue immer unwichtiger wird, gepaart mit dem Versagen traditioneller Parteien aktuell relevante Themen zu besetzen (Van Kessel, 2011), identifiziert. Schließen lässt sich dieses Kapitel also mit einem Zitat aus einem Vortrag von Dr. Stijn van Kessel im Juli an der Universität Wien: “The rise of populist parties has only partly to do with populism itself...”

2.2.2. Populismus und das Volk

Autor: Maximilian Haselsteiner

Wie schon im Kapitel “Definition von Populismus” erwähnt, ist eine Kernkomponente von Populismus “People Centrist” (Mudde, 2014). Der Begriff “People” kann in die deutsche Sprache sowohl als “das Volk” als auch als “die einfachen Leute” oder oft “der kleine Mann” übersetzt werden. Erschwerend kommt hinzu, dass der Begriff “Volk” im deutschen historisch schwer belastet ist. So finden sich beispielsweise in den Parteimanifesten der von Experten klar als populistisch eingeschätzten “Alternative für Deutschland” kaum Referenzen an das Volk an sich (vgl. Pauwels, 2017: 128ff).

Da dem Begriff in dieser Arbeit eine zentrale Rolle zukommt, soll hier etwas ausgeholt werden. Laut Online-Duden ist eine Demokratie ein “politisches Prinzip, nach dem das Volk durch freie Wahlen an der Machtausübung im Staat teilhat”. Doch wer ist eigentlich dieses Volk - und vor allem, wer ist es nicht? Die Frage wer alle Staats- und Partizipationsrechte hat, ist nach wie vor aktuell; dürfen Gruppen wie etwa Kinder und geistig Behinderte³ einfach ausgegrenzt werden? Ist es legitim, dass Häftlinge nicht wählen dürfen⁴? Äußerst schwierig wird es, wenn es um Migranten, Asylsuchende oder ihre Kinder geht. Oft werden diese Fragen, wie im Fall von österreichischen Häftlingen, rechtlich gelöst, doch der eigentliche Entscheidungsträger, die “höhere Macht”, in einer Demokratie ist immer das Volk. Dieses Paradox, dass das Volk selbst definieren müsste wer das Volk ist, ist unter vielen Namen bekannt (z.B. “the democratic paradox” n. Mouffe 2000 oder “the problem of constituting the demos” n. Godin 2007), da in diesem Absatz allgemein Paulina Ochoa Espejos Argumentation gefolgt wird (2017: 608-612), folgen wir auch ihrer Nomenklatur. Sie bezeichnet das Problem als “the boundary problem” (vgl. Whelan, 1983). Um diese philosophische Diskussion abzukürzen, muss man aufhören, Demokratie als perfekt zu sehen. Ganz nach dem berühmten Zitat von Winston Churchill bei der Rede vor dem Unterhaus 1947: *“No one pretends that democracy is perfect or all-wise. Indeed, it has been said that democracy is the worst form of government except all those other forms that have been tried from time to time.”* Demokratie ist eigentlich eine repräsentative Form des Regierens, man wählt Leute, die zwar einen gewissen Auftrag haben, aber prinzipiell

³ Wobei auch hier diskutiert werden muss - und auch wird - ab welchem Grad der Behinderung ein Wahlrecht entzogen werden kann oder sollte.

⁴ In Österreich dürfen Häftlinge erst nach einer Klage vor dem Gericht für Menschenrechte grundsätzlich wählen. Ein Richter kann dieses Wahlrecht jedoch immer noch aufheben.

frei handeln. Das Volk ist also nicht der regierende, sondern nach Tuck (2016) ein “schlafender Souverän” (vgl. Ochoa Espejo, 2017: 611).

Die politische Philosophie kennt drei Erklärungsansätze; der hypothetische Ansatz geht von einem Volk nicht als Summe von Individuen, sondern einem Volk als normativer Gesetzgeber aus. Es fußt auf den Gesellschaftsvertrag von Kant und Rousseau und wird beschrieben als ein “kontrafaktischer Idealtyp welcher uns erlaubt die Legitimität von Rechtsnormen zu evaluieren (ebd.: 614). Der historische Ansatz sieht das Volk, vereinfacht, als eine durch geschichtliche Ereignisse wie etwa Konflikte, Verhandlungen und Revolutionen erzeugte Gruppe. Diese Gruppe wird retrospektiv zu einem Staatsvolk zusammengefasst. Eine Synthese der beiden Ansätze zum Verständnis von Volk bietet eine prozedurale Ansicht des Volkes:

“This view mediates between the “reason” of the hypothetical account, and the “will” of the historical account. “The People”, [...] is not a collection of individuals, but a procedure of decision-making, by which individuals interact with each other mediated by legal institutions that channel popular demands and force representatives to adopt views and make decisions.” (ebd.: 615).

Das Volk entwickelt sich also dynamisch wie ein Prozess. Es ist nicht starr, sondern kann sich verändern. Es kann dadurch auch nicht homogen sein und nicht mit einer geeinten Stimme sprechen, aber es kann dennoch Normen anfechten. Es kann durch historische Ereignisse, aber auch alltägliche Geschehnisse verändert werden.

Im Gegensatz zu der dynamischen Sicht des Volkes, wird der Volksbegriff bei Populisten konstruiert. Wie bereits in Kapitel 2.2. erwähnt sehen Populisten das Volk als einen monolithischen (vgl. Jagers und Walgrave, 2007: 322), daher homogenen (vgl. Stanley, 2008: 102) und geeinten Block. Diese Gruppe kann objektiv nicht beschrieben werden, vielmehr ist es ein künstlich geschaffenes Produkt. Wenn Populisten im Namen des Volkes sprechen, ist es gar nicht möglich alle Bürger inkludieren, vielmehr erschaffen sie ein Volk welches sie zu repräsentieren behaupten (vgl. Moffit und Tormey, 2014: 389). Auch Mudde bestätigt, dass das Volk oder “the people” wie es Populisten verwenden, *“neither real nor all-inclusive, but in fact a mythical and constructed subset of the whole population”* (2004: 546) ist.

Van Kessel (2011) definiert basierende unter anderem auf den Theorien von Mudde, drei Merkmale von populistischen Parteien. Jede dieser Merkmale weist eine Referenz auf den

Umgang von populistischen Parteien mit dem Begriff “Volk” hier als “ordinary people” auf. Demnach ist eine Partei populistisch, wenn sie folgende Punkte erfüllt:

- (1) Delineate an exclusive community of ‘ordinary people’,*
- (2) Appeal to these ‘ordinary’ people, whose interests and opinions should be central in making political decisions*
- (3) [They] are fundamentally hostile towards the (political) establishment, which allegedly does not act in the interest of the ordinary people” (ebd. S. 2 f)*

Darüber hinaus, wird dieses künstliche Volk mit verschiedenen positiven Eigenschaften in Verbindung gebracht. Es wird beschrieben als tugendhaft (“inherent virtuousness” vgl. Jansen, 2011: 84) und von Natur aus gut (Albertazzi und McDonnell 2008: 6). Sie sind erste-Klasse Bürger, folgen einem gesunden Menschenverstand (als Alltags- oder Hausverstand) und werden oft als eine schweigende Mehrheit (“silent majority”) bezeichnet (Taggard, 2000: 91). Oft wird dazu im Gegensatz das Bild einer elitären oder herrschenden lauten, zeternden und jammernden Minderheit gezeichnet. Erst die Populisten geben dieser schweigenden Mehrheit dann eine Stimme.

Kriesi (2013 3ff, zitiert durch Wirth et al., 2016: 10) unterscheidet zwischen drei verschiedenen Interpretationen des Volksbegriffs. Der politische Begriff, bei dem ein geeintes Staatsvolk der absolute Souverän ist. Eine kulturelle Deutung, in der das Volk in einer ethnischen Dimension begriffen wird. Die Rede ist von “unseren Leuten”, dem Volk als Nation und der klaren Exklusion von Elementen welche nicht zu diesem Volk gehören (Migranten, Ausländer und andere Minderheiten etc.). Drittens, gibt es noch eine ökonomische Sichtweise auf das Volk, in der das Volk beinahe als Klasse gesehen wird. Dabei werden die ökonomischen Bedürfnisse der “kleinen Leute” (Canovan, 1999:5 spricht von “ordinary people”, Kriesi, 2013: 4 von “the little guy”) denen der privilegierten, kosmopolitischen Eliten, welche ihre Taschen mit unverdienten Boni gefüllt haben, gegenübergestellt.

2.2.3. Populismus und die Eliten

Autor: Maximilian Haselsteiner

Das Oxford Dictionary bietet zwei Definitionen für den Begriff „Elite“ an: Allgemein handelt es sich um “A select group that is superior in terms of ability or qualities to the rest of a group or society.” Also jemand mit herausragenden Befähigungen und Qualitäten,

genauso wichtig ist aber die zweite Definition: “A group or class of people seen as having the most power and influence in a society, especially on account of their wealth or privilege.” Eine Gruppe von Personen, welche als mächtig *wahrgenommen* werden. Der Begriff wurde bereits in der Spät-Renaissance im Zusammenhang mit besonders guten Waren verwendet, erst während der französischen Revolution wurde der Begriff *élite* für Personen gebraucht, welche es ohne Zugehörigkeit zu Adel oder Klerus zu Wohlstand brachten. Baruch de Spinoza entwarf „Eliten“ als Funktionsbegriff; nur eine kleine Gruppe an Menschen sei es möglich, durch Intellekt, aber auch moralische Integrität und anderen Tugenden die Komplexität der Welt überschauen zu können und so die Menschheit anzuführen (vgl. Papcke, 2001: 464).

Während der industriellen Revolution wurde der Begriff dann als Abgrenzung zwischen dem besitzenden Bürgertum und den „ungebildeten Massen“ verwendet und bekam dadurch eine wertende Dimension. Damit einher ging auch eine philosophisch-kritische Auseinandersetzung mit den Eliten. Der Politikwissenschaftler Heinrich von Treitschke stellte zur Jahrhundertwende jene „Tausende welche forschen, malen und dichten können“ den „Millionen (welche) ackern, schmieden und hobeln (müssen)“ gegenüber und zeichnet den Elitismus als Rechtfertigung für die Ungleichverteilung von Reichtum (vgl. Tscheke, 1918; zitiert durch Papcke, 2001: 467). Durch die Weiterentwicklung der Thesen des italienischen Philosophen und Politikwissenschaftler Vilfredo Pareto („Die Geschichte ist der Friedhof der Aristokratien“) wird der Begriff dann in den Marxismus eingespeist und geht über, in die „politischen Klasse“ (Gaetano Moca), „Oligarchie“ (Robert Michels) oder gar den „Militärische-industriellen Komplex“ (Charles Wright Mills) in den ideengeschichtlichen Kanon ein.

Elitismus wurde jedoch nicht nur negativ betrachtet, einer der ersten und bekanntesten „Elitisten“ war der griechische Philosoph Platon mit seiner Forderung nach einer „Philosophenherrschaft“. Friedrich Nietzsche steigert Platons Utopie ins Absolute. Er sieht Egalitarismus als Übel, der die Menschen zu eigennützigen (Herden-)Vieh und Pöbel degradieren würde. Selbst Max Horkheimer, einer der führenden Denker der Frankfurter Schule, kommt zu dem Schluss, dass es ein sozialpsychologisches oder funktionales Bedürfnis nach Eliten – ein „Schutzbedürfnis“ – gibt (vgl. Papcke, 2001: 525). Auch heute, in Zeiten wachsender Komplexität, besteht ein Bedürfnis an Eliten als eine Gruppe von Menschen, welche über materielle und immaterielle Ressourcen (Kapital, aber auch bestimmte Befähigungen) verfügen und damit gesellschaftliche Funktionen ausführen.

Die Rolle der Eliten in der Gesellschaft zu erörtern ist allerdings nicht Ziel dieses Kapitels. In dieser Arbeit interessiert man sich jedoch eher für die anfangs als zweite genannte Definition von Eliten, jene Menschen die als Eliten wahrgenommen werden (wobei sich das natürlich nicht ausschließt). Populistische Kommunikatoren haben Eliten zu einem bevorzugten Ziel ihrer Rhetorik und Politik gemacht. Beinahe alle Definitionen von Populismus kommen darin überein, dass die Kritik an den Eliten, Anti-Elitismus, ein zwingendes Merkmal von Populismus ist. Eliten werden dabei als diamantales Gegenteil zum (monolithischen/einheitlichen) Volk gezeichnet. Sie sind „der Klassenfeind“ der Populisten. Schon in der ersten Blüte des lateinamerikanischen Populismus in den 1940er und 50er Jahren galt es, beispielsweise für die argentinischen Perónistas, als einfacher „Diener des Volkes“ die Interessen dieses, gegen jene der „Etablierten“, durzusetzen. Die Eliten hätten den Bezug zur Realität des Volkes verloren und dienen nur noch ihren eigenen Interessen, der populistische Akteur dagegen als selbstloser Diener des Volkes nimmt den Eliten diese Macht und bringt dem Volk ihre Souveränität zurück (Jagers & Walgrave, 2007: 323f). Es können „die da oben“, welche aus einem Porzellanturm verächtlich auf uns herabsehen sein, genauso wie jene, welche von einem hohen moralischen Ross auf uns herabblicken und uns erklären wollen wie wir zu leben haben und jene, welche daumendrehend ihre Millionen verdienen während wir, das arme Volk, uns in ihren Fabriken zu Tode schuften. Wie volatil das Konzept ist, zeigt sich wenn es Populisten schaffen an die Macht zu kommen, denn, wenn man selbst an der Macht ist müssen andere Feinde her, im Falle von Hugo Chavez war es etwa der „imperialistische Teufel USA“, der im absoluten - „manicheanischen“ – Widerspruch zum guten Volk steht. Ein prototypisches Beispiel für Populismus zeigt eine Rede zur Wiederwahl:

„Let no one forget that we are confronting the Devil himself. Sunday, 3 December at the ballot box we will confront the imperialist government of the United States of North America [sic]—that is our real adversary, not these has-beens here, these lackeys of imperialism... Long live Christ, the first great revolutionary of our time! Martyr of the peoples, Christ Redeemer, today is his day, the day of Christ the King [...] You the people are the giant that awoke, I your humble soldier will only do what you say. I am at your orders to continue clearing the way to the greater Fatherland. . . . Because you are not going to reelect Chávez really, you are going to reelect yourselves, the people will reelect the people. Chávez is nothing but an instrument of the people.”

Hugo Chávez 2006, Zitiert durch Hawkings (2009: 1041)

Wie kaum irgendwo sieht man einen so klaren manicheanischen Vergleich: der amerikanische Teufel auf der einen Seite, Volk und Jesus auf der anderen. Die Elite ist in diesen Fall die Elite-Nation und Weltmacht USA und die politische Opposition, welche bloße Vasallen der Supermacht sind. In den USA wiederum müssen (unter anderem) die Medien- und Meinungseliten als Sündenböcke herhalten, aber auch die vorige Regierung oder aber auch Führungspersonen von Institutionen und kleineren Verwaltungseinheiten.

Wie Eliten in Arbeiten zu Populismus definiert werden variiert, meist wird eine klare Definition übersprungen. Van Kessel (vgl. 2011: 3) sieht Eliten als eine von mehreren „Out-Groups“. Migranten und Minderheiten sind genauso Ziel von exklusivistischer Kommunikation, Xenophobie sei im Gegensatz zu Anti-Elitismus jedoch kein zwingendes Element von Populismus. Sie bindet sich bloß an die dünne Ideologie von Populismus, weshalb man, vor allem in Europa und Nordamerika von „*Rechtspopulismus*“ spricht. Stanley (2008: 102) sieht Anti-Elitismus als das Kernelement des Populismus, Mudde (2004) nimmt den theologischen Begriff „Manicheanismus“ um den absoluten, schwarz-weiß-wertenden Unterschied zwischen den korrupten Eliten und dem reinen Volk auszudrücken und Albertazzi & McDonnell (2008: 4) beschreiben die Eliten aus Sicht der populistischen Kommunikation als „Feinde des Volkes“: *„The corrupt elites have occupied, distorted and exploited democracy, sometimes to the extent of its total degeneration“*.

Wer die Eliten sind wurde weiter oben schon erwähnt. Es geht um ökonomische Eliten wie Manager, Banken und Gewerkschaftsführer, Kulturelle Eliten, Meinungsmacher und allen voran die Medien, Intellektuelle wie Publizisten, Schriftsteller und Wissenschaftler und legale Eliten aus der Administration, Gerichten und anderen Institutionen (vgl. Jagers & Walgrave, 2007: 324). Das NCCR Codebuch erweitert diese Liste des Weiteren um Eliten aus den Sicherheitsbereich und dem Militär, Legislative Eliten wie Parlamentarier, Unternehmen, „die oberen Zehntausend“, jene „hinter den Kulissen“ welche am „core of the Power“ sitzen (Wirth et al. 2017: 152f). Je diffuser und pauschalisierender der Anti-Elitismus, desto radikaler und leidenschaftlicher ist dieser (Jagers & Walgrave, 2007: ebd.). Natürlich muss man jedoch aufpassen, dass nicht jede Form von Negative-Campaigning sofort auch als Anti-Elitismus gebrandmarkt wird.

Zusammenfassend zeigt sich also ein eher vages Bild der Eliten – und dieser Umstand macht sie auch zu einem perfekten Feindbild für Populisten. Die Elite ist keine speziell definierte Gruppe, es gibt keine Einkommensgrenzen, keine Mindestzahl an Erwähnungen

in der Klatschpresse, Publikationen in Wissenschaftsmagazinen oder gewisse Zahl an Twitter Followern ab denen man dazugehört. Es ist eher ein vages Konzept und wird solchen Leuten zugeschrieben, welche zumindest zum Teil von der Bevölkerung – um auf die zweite genannte Definition zurückzukommen - als solche wahrnehmbar sind. Präziser; es sind jene Menschen, welche durch kommunikative Maßnahmen *als Eliten wahrnehmbar gemacht werden können*.

2.3. Die Rolle der Medien

Autor: Gregor Rettenegger

Verschiedene Studien haben auf die wichtige Rolle der Medien sowie der Themenlage für den Erfolg rechts-populistischer Parteien aufgezeigt (u.a.: Blumer, 2003; Plasser & Ulram, 2002; Stewart et al., 2003; Boomgarden & Vliegenthart, 2007; Van Kessel, 2011). Am Beispiel der FPÖ in Österreich verweisen Plasser und Ulram (2003) darauf hin, dass rechts-populistische Parteien mehr als andere darauf angewiesen sind, dass für sie wichtige Themen wie z.B. Immigration in den Medien im Sinne der Parteilinie geframed werden. Daher sind diese Parteien auch besonders bemüht die Medien Agenda aktiv zu leiten und zu beeinflussen.

Walgrave und Swert (2004) untersuchten den Zusammenhang zwischen medialer Berichterstattung und dem Erfolg der Partei „Vlaams Blok“ in Belgien. Aus den Theorien von Agenda Setting und Issue Ownership leiten sie ab, dass davon auszugehen sei, dass Medien einen Einfluss auf das Wahlverhalten haben. Die Agenda Setting Theorie geht davon aus, dass die Öffentlichkeit jene Themen als wichtig erachtet, die von den Medien als solche hervorgehoben werden (Cohen, 1963). Diese Effekte wurden erstmals von McCombs und Shaw (1972) in ihrer berühmt gewordenen Chapel Hill Studie nachgewiesen. Seither wurde die Agenda Setting Theorie in zahlreichen Publikationen weiterentwickelt und die Einflüsse von Medieninhalten auf die Meinung der Rezipienten wieder und wieder nachgewiesen (u.a. Lopez-Escobar et al., 1998; McCombs, 2004; Walgrave and Van Aelst, 2006; Vargo and Guo, 2017). Während Wahlkämpfen sind diese Vorgänge von besonderer Bedeutung, weil Parteien stark von einer Medien-Agenda profitieren können, die aus Themen besteht die diesen Parteien zugerechnet werden. Parteien, denen keine Kompetenz zu den bestimmenden Themen des Wahlkampfs zugerechnet werden, haben wiederum schlechtere Karten Wahlen zu gewinnen (Hayes, 2008). Es handelt sich dabei um ein Phänomen, welches von Petrocik (1996) unter dem

Namen „Issue Ownership“ beschrieben wurde und seither rege in der Literatur diskutiert wird (für einen detaillierten Überblick siehe: Belanger & Meguid, 2008).

Den Medien wurde auch unterstellt in einer Komplizenschaft mit populistischen Parteien zu stehen. Dadurch das Medien die provokanten Inhalte populistischer Parteien überproportional verbreiten, haben sie dazu beigetragen die Themen, Schlagworte und kommunikativen Stile der Populisten zu legitimieren (Mazzoleni, 2008). ‘Underdog leaders’ waren dadurch oft erfolgreich die Präferenz der Medien für alles was aus der politischen Routine ausschert, auszunutzen. Medien wiederum erhoffen sich durch die Berichterstattung über markante populistische Aussagen mehr Kunden zu gewinnen. Diese Angebot-und-Nachfrage Dynamik erhöht die Sichtbarkeit populistischer Botschaften und Medien könnten dadurch - ob gewollt oder nicht - zu mächtigen Mobilisierungsinstrumenten für populistische Bewegungen werden (ebd. S. 50).

Diesen Gedanken folgend erscheint es hochrelevant Populismus in medialer Berichterstattung messbar zu machen. Werden populistische Frames zu einem Thema wie Migration oder Anti-Establishment in Medien wiedergegeben, entsteht ein Vorteil für populistischen Parteien, die traditionell mit diesen Themen assoziiert werden. Dabei ist es irrelevant ob die Themen aktiv von diesen Parteien in die Medien gepusht wurden, oder die Themenlage durch äußere Faktoren besonders günstig ist. Ebenso wenig spielt es eine Rolle ob das populistische Framing dieser Themen von Medien nur zitiert oder in journalistischer Eigenregie übernommen wurde.

3. Messen von Populismus

Autor: Gregor Rettenegger

Um Aussagen über das Vorkommen von Populismus in Texten treffen zu können, müssen Entscheidungen getroffen werden, wie Populismus quantifiziert werden soll. Die verschiedenen Ansätze die in der Literatur dazu zu finden sind, sollen in diesem Kapitel vorgestellt und deren Stärken und Schwächen evaluiert werden.

Ziel der Quantifizierung jedes zu messenden Konzeptes in Text Dokumenten muss es sein:

- 1.) Die Reliability und Validität des Messinstrumentes zu maximieren, und dadurch Researcher- Bias zu reduzieren.
- 2.) Mit möglichst geringem Aufwand einen möglichst großen Datensatz zu generieren, der in Folge die Grundlage für statistische und kausale Untersuchungen bietet.

In der Literatur wird auch auf den Mehrwert einer graduellen Messung (Intervall Skala) gegenüber einer dichotomen Messung hingewiesen, da sich aus einer solchen zahlreiche zusätzliche Fragen, die spezifischen Ausprägungen eines Konstrukts besser verstehen zu können, ergeben (Aslandis, 2017). Der Vollständigkeit halber soll dieser Punkt ebenfalls in die Ziele einer Quantifizierung von Populismus aufgenommen werden.

- 3.) Abhängige Variablen als Intervall Skalen (e.g. mehr/ weniger populistisch) und nicht als dichotome Variablen (ja/nein) zu kodieren (Aslandis, 2017)

Während Punkt 1 und 2 für jegliche Quantifizierung essentiell ist, scheint Punkt 3 vor allem für komplexe Untersuchungsgegenstände, deren qualitative Untersuchung noch unzureichend sind, von Bedeutung zu sein. Für das hier untersuchte Phänomen den “Populistischen Diskurses”, können damit beispielsweise das Verhältnis zwischen People-Centrism und Anti-Elitismus beleuchtet werden, oder die kontextspezifischen Eliten besser sichtbar gemacht werden.

3.1. Manuelle Ansätze

Autoren: Maximilian Haselsteiner (S. 20-22) & Gregor Rettenegger (S. 23-25)

Der Messung von Populismus liegen mehrere Probleme zu Grunde. Bereits erwähnt wurde, dass es eine Vielzahl von Definitionen gibt, welche sich auch auf die Messbarkeit auswirken. In “Populist Radical Right Parties in Europe” wendet Mudde (2007) seine eigene Definition an und klassifiziert damit zahlreiche europäische Parteien, basierend auf deren Manifestos, in die Kategorien “Populist Radical-Right”, “Neo-liberal Populist” und “Social-populist” Parties. Er stellt dabei vier Fragen um festzustellen ob eine Partei populistisch ist:

1. Wird das Volk homogen und rein dargestellt?
2. Sind dagegen die Eliten corrupt und schlecht (und ebenfalls homogen)?
3. Sind diese zwei Gruppen zwei antagonistische und ausschließliche Gruppen?
4. Und ist es die hehre Aufgabe der untersuchten Partei die Macht den korrupten Eliten zu entreißen und es dem zu beschützenden Volk zurückzugeben?

Auch wenn sich eine Vielzahl an Forschenden, vor allem aus Nordeuropäischen Staaten und Skandinavien, auf die Definition von Populismus als „dünne Ideologie“ geeinigt haben (vgl. Heinisch & Mazzolezi, 2017: 105f), sollte hier auch ein zweiter Ansatz vorgestellt werden. So wurde bei einem der ersten empirischen Durchbrüchen in der Messung von Populismus durch Jagers und Walgrave (2007) dieser als „Stil“ definiert. Dennoch unterscheidet sich deren Herangehensweise in dieser Arbeit kaum von aktuelleren Arbeiten (siehe weiter unten) (vgl. Pauwels, 2017: 123). Ein wichtiger Punkt ist, dass Populismus neben den bereits erwähnten Operationalisierungen “Anti-Elitismus” und “People-Centrism” auch mit einer dritten Kategorie “Exclusionism” gemessen wurde. Dies beschreibt die Konstruktion einer gefährlichen Außengruppe, vor welcher Populisten das homogene Volk zu schützen geloben. Im europäischen Kontext sind dies vor allem Migranten, während in Lateinamerika eher Intellektuelle als Feindbild konstruiert werden. Ein weiterer Durchbruch dieser Arbeit war, dass ein Codebuch für einen dicken (“thick”) und einen dünnen (“thin”) Populismus erstellt wurde. Wurden beispielsweise nur die “kleinen Leute” oder das Volk angesprochen, wurde der Artikel als “dünn” kategorisiert, wurde dagegen auch Eliten, die Medien oder Außengruppen wie Immigranten negativ

angesprochen, wurde der Artikel als “thick” kategorisiert. Durch eine Inhaltsanalyse wurden diese Kategorien dann von geschulten Codern ausgewertet. Abbildung 1 zeigt das Kodierungsschema für den dünnen Populismus. Als Datenmaterial wurden so genannte “political party broadcasts” von belgischen Parteien verwendet. Diese 10-Minütigen Sendungen werden auf öffentlichen Sendeanstalten ausgestrahlt, allerdings alleinig von den politischen Parteien erstellt und verantwortet. Die Auswertung zeigte, wie vermutet, dass “Vlaams Blok” geradezu ein Lehrbuch Modell für dicken und dünnen Populismus darstellt (vgl. Jagers et. al.: 333).

- 2.1.1 Based on the verbatim term, the reference was put in the corresponding category to avoid a never ending list of referential terms. Content takes priority over verbatim terms.
- 2.1.2 The terms referring to the people could be split up in two kinds: those that refer to the population (group) as an undividable unity mostly proceeded by a definite article (e.g., *the voter*, *the people*, *the consumer*) and those that do not. Since populists are inclined to mobilise as broadly as possible, the first kind of referential terms can be regarded as the most solid indicator of populism.
- 2.1.3 Terms like ‘public opinion’, ‘(political) participation’ and ‘democracy’ are also noted as (indirect) references to the people if they allude to ‘the will of the people’.
- 2.1.4 Nouns used as adjectives to point at a population category were not selected as a reference (e.g., ‘The *Belgian* Flor Van Noppen is 42 years old and self-employed entrepreneur’).
- 2.1.5 The selected references had to reach beyond the private life of the speaker. The reason is the limited mobilising effect of terms that do not go beyond the private life (e.g., ‘My father is a doctor’ versus ‘Doctors want to be respected by their patients’).

Abbildung 1: "Encoding references" aus Jagers et. al. 2007: 339 (Appendix)

Eine weitere Abhandlung ist Hawkings (et. al. 2009) Versuch, empirisch zu testen ob, und wie populistisch der venezolanische Präsident Hugo Chavez ist. Erwähnenswert ist dies unter anderem deshalb, weil hier eine holistische Graduierung verwendet wurde. Es wurden also nicht nur zuvor kodierte Wörter und Paragraphen gezählt, sondern Texte ganzheitlich interpretiert und in drei Kategorien eingeteilt: “extremely populist”, “includes strong, clearly populist elements but either does not use them consistently or tempers them by including nonpopulist elements.”, “few, if any populist elements” (vgl. Hawkings et. al., 2009: 1062). Als Gegenkategorie zu Populismus wurde “Pluralismus” verwendet, was auch in aktuellen Arbeiten immer wieder als Spiegel zu Populismus konzipiert wird (z.B. Mudde, 2017: 34f). Einige Anweisung an die Kodierer werden in Abbildung 2 gezeigt.

Populist	Pluralist
It conveys a Manichaeian vision of the world, that is, one that is moral (every issue has a strong moral dimension) and dualistic (everything is in one category or the other, “right” or “wrong,” “good” or “evil”). The implication—or even the stated idea—is that there can be nothing in between, no fence sitting, no shades of gray. This leads to the use of highly charged, even bellicose language.	The discourse does not frame issues in moral terms or paint them in black and white. Instead, there is a strong tendency to focus on narrow, particular issues . The discourse will emphasize or at least not eliminate the possibility of natural, justifiable differences of opinion.
The moral significance of the items mentioned in the speech is heightened by ascribing cosmic proportions to them, that is, by claiming that they affect people everywhere (possibly but not necessarily across the world) and across time. Especially in this last regard, frequent references may be made to a reified notion of “history.” At the same time, the speaker will justify the moral significance of his or her ideas by tying them to national and religious leaders that are generally revered.	The discourse will probably not refer to any reified notion of history or use any cosmic proportions. References to the spatial and temporal consequences of issues will be limited to the material reality rather than any mystical connections.

Abbildung 2: Auszug aus dem “Populist Speech Rubric” in Hawkins et al. 2009:1063f, Appendix

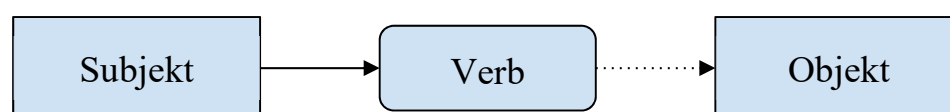
Die Ergebnisse zeigten, dass Hugo Chavez und Evo Morales, wie erwartet, hohe Werte erzielten und auch die Reliabilität der Erfassung hoch war. Diese Herangehensweise ist somit eine gute Ergänzung des Forschungsstandes, hat jedoch zwei starke Nachteile: Erstens sind, wie Hawkins selbst sagt (ebd.: 1049), die Kosten und der Zeitaufwand für diese Methodik sehr hoch, zweitens steht und fällt diese Methode auch mit der Qualität der Ausbildung der Kodierer.

Kurze Erwähnung finden sollen hier auch noch qualitative Methoden. Das Experteninterview gilt allgemein als ökonomische Alternative, allerdings sind Validitätschecks oft schwer durchzuführen. Die Zahl der Experten ist oft limitiert und nicht selten sind diese mit bestimmten politischen Strömungen assoziiert. Hervorzuheben ist eine Arbeit von Wieschomeiner (2016), in welcher zwei verschiedene Ansätze von Experteninterviews zu Populismus in Argentinien und Brasilien angewendet und verglichen werden. Eine interessante Perspektive findet sich auch bei einer Studie von Anup Kumar (2014). Dort untersucht er den so genannten “blauen Populismus” während der Wahlkampagnen von Barack Obama. Es basiert auf dem Konzept von Ernesto Laclau (2005) das Populisten zur Verwendung von “Empty Signifiers” tendieren, also Worte,

deren Interpretation dem Rezipienten selbst überlassen werden. In seiner Studie kommt Kumar zu dem Schluss, dass Journalisten jene Empty Signifiers nicht mit Sinn füllen und mehr oder weniger uninterpretiert an Rezipienten weitergeben (Kumar, 2014: 18).

Die oben beschriebenen Ansätze sind allesamt versuche Populismus holistisch zu messen, was Hawkings (2010) “whole-text Analysis” nennt. Während holistische Ansätze verglichen mit anderen Ansätzen relativ valide Ergebnisse liefern, haben sie doch zwei große Nachteile. Zum einen ist die Interkoderreliabilität oft unzureichend, so berichtet Hawkings (2009) Cohen’s kappa Werte zwischen akzeptablen 0.68 auf der ersten Kodierebene und nur 0.44 auf der zweiten Ebene. Andere Autoren kodieren ihre Daten auch selbst und können deshalb keine Aussagen über Reliabilität treffen (Bernhard et. al., 2015). Ein weiteres Problem dieser Methode ist der beträchtliche Arbeitsaufwand den alle manuellen Methoden gemein haben. Sämtliche Texte des zu untersuchenden Samples müssen von mindestens einem Kodierer aufmerksam gelesen und in Folge möglichst reproduzierbar kodiert werden. Dies kann in einer Zeit immer einfacher zugänglichen, großen Textmengen erhebliche Kosten verursachen.

Zumindest für das erste Problem der unzureichenden Reliabilität der Kodierung gibt es manuelle methodologische Ansätze die Besserung versprechen. So untersuchte Aslanidis (2017) Manifestos von Grassroots Bewegungen auf Basis von “Semantic clauses”. In einem ersten Schritt werden, durch das Anwenden von klaren grammatische Regeln, Texte in sogenannte “semantic triplets” (Franzosi, 2004) umgewandelt. Diese bestehen aus einem *Subjekt*, *Verb* und (optional) einem *Objekt*.



Der Wert diese ursprünglich aus der Linguistik stammende Methode wurde wiederholt für die Sozialwissenschaft aufgezeigt (z.B. Roberts, 1997). Der große Vorteil von Semantic Triplets als Kodiereinheit besteht in ihrer inhärenten Reliabilität. Ihre Erstellung folgt objektiven grammatischen Regeln. Diese strengen Regeln garantieren rigorose und systematische Handhabung von Texten und liefern vergleichbare Kodier-Einheiten in jeder Sprache (Aslanidis, 2017: S. 1251). Aus einem Satz wie “Die Politiker haben die Menschen verraten”, würde das Triplet “Politiker” (Subjekt), “verraten” (Verb) und “Die Menschen”

(Objekt) erstellt werden. Anhand der einzelnen Bausteine lassen sich in Folge granulare Aussagen über den Inhalt machen. So liegt ein “Full Populist Frame” in der Studie von Aslanidis (2017) dann vor, wenn sowohl ein “populist ‘we’” (= People Centrism) als auch ein “populist ‘them’” (= Anti Elitism) vorliegt. Es können aber auch Fälle in denen nur eine dieser zwei Kategorien vorkommt unterschieden werden, oder beispielsweise die häufigsten Vorwürfe (Verb) gegenüber den Eliten erhoben werden.

Aslanidis (2017) berichtet mit dieser Methode Interkoder-Reliabilitätswerte auf den verschiedenen Ebenen zwischen 0.629 und 0.948 (Krippendorff’s alpha). Während das Problem schlechter Interkoderreliabilität mit diesem Ansatz vermieden werden konnte wird das Problem des Aufwandes sogar weiter verschärft. Bevor Kodierer nämlich beginnen können Semantische Triplets zu kodieren müssen diese erst erstellt werden. In der hier erwähnten Studie wurden diese vom Verfasser selbst erstellt. Abgesehen von dem erheblichen Aufwand eines solchen Prozesses besteht auch ein gewisses Objektivitätsproblem. Auch wenn hier rigoros nach strengen Regeln vorgegangen wird bleibt doch eine gewisse Unsicherheit ob nicht bei der Erstellung der Triplets ein Forscher-Bias gegeben war, da an dieser Stelle keine zweite Instanz die Entscheidungen überprüft.

Eine andere sehr umfangreiche Studie die eine ähnliche Methode verwendete, wurde von der NCCR Democracy, einem Forschungsprojekt auf der Universität Zürich durchgeführt. Basierend auf der theoretischen Arbeit von Wirth et. al. (2016) wurde ein umfangreiches Kodierschema erstellt und mit diesem über 57600 politische Texte (= Texte in denen länderspezifisch Relevante politische Akteure vorkommen) von 76 trainierten Kodierern kodiert. Ähnlich dem oben erklärten Ansatz wurden aus dem Text sogenannte “Statements” extrahiert, welche in Folge kodiert werden und die Basis für die Bewertung eines Textes bilden. “Statements” in diesem Sinnen sind Aussagen von politischen Akteuren (z.B. Politiker, Journalisten) und Organisationen (typischerweise Parteien) andere Akteure oder Organisationen betreffend. Dem theoretischen Unterbau folgend werden sie als Anti-Elitismus gewertet, wenn a.) Eliten etwas vorgeworfen wird, b.) Eliten für eine Fehlentwicklung verantwortlich gemacht oder c.) Eliten von den Menschen exkludiert werden. People-Centrism liegt vor wenn a.) die Tugenden der Menschen, b.) die Errungenschaften der Menschen hervorgehoben werden, c.) Nähe zu den Menschen postuliert wird oder d.) die Menschen als monolithisch, homogene Einheit dargestellt werden. Eine detailliertere Beschreibung findet sich im Kapitel 4.3.3. Zusätzlich gibt es noch zwei weitere Indikatoren die der Kategorie “popular sovereignty” zugeordnet sind:

a.) Souveränität für die Menschen wird gefordert, b.) die Eliten sollten weniger Macht haben (Bernhard, 2016). Aus diesen drei Kategorien werden einzeln gewertet und in weiterer Folge dann aufsummiert um einen Populismus- Wert zwischen 0 - 3 zu erhalten. Aus den 57000 Texten wurden so mehr als 160000 einzelne Aussagen extrahiert und kodiert. Beim Erstellen dieser Aussagen hatten die Kodierer eine 90.8% Übereinstimmung (Cohen's Kappa = 0.574) mit ihrem Projektleiter und eine Interkoder- Übereinstimmung von 85.1%; (Kappa=0.597) (Wirth et. al., 2016). Das Kodebuch der NCCR Studie lässt durch seine Vielzahl an Variablen einen sehr detaillierten Blick auf das Phänomen populistischer Kommunikation zu. So gibt es zahlreiche mehr oder weniger klar abgegrenzte Unterkategorien von “die Menschen” oder “Eliten”, aber auch für “Eliten etwas vorwerfen”.

So werden beispielsweise unter der Kategorie “Attribution of suspected impact and responsibility” folgende Unterkategorien kodiert, die in Folge wieder zahlreiche Unterkategorien haben:

Suffix	Content	Type
_thre	Posing a threat	Negative
_burd	Posing a burden	Negative
_enri	Posing an enrichment	Positive
_aneg	Accountable for negative development/situation	Negative
_apos	Accountable for positive development/situation	Positive
_abil	Ability to alleviate/solve a problem	Positive

Ignoriert man den immensen Aufwand dieser Methode und die damit einhergehenden Kosten, scheint hier eine valide Methode gefunden worden zu sein die sowohl einen detaillierten Blick auf die Nuancen populistischer Kommunikation erlaubt, als diesen auch reliabel zu quantifizieren. Tut man dies jedoch nicht, bleibt evident, dass diese Methode nur eine Option für gut finanzierte und große Forschungsgruppen ist und in dieser umfangreichen Form wohl nur selten Anwendung finden wird.

3.2. Automatisierte Ansätze

Autor: Gregor Rettenegger

Im Gegensatz zu den großen Kosten und Reliabilitätsproblemen der oben erwähnten Methoden, liegt es in der Natur der Sache das automatisierte (computer-basierende) Verfahren hier deutliche Vorzüge haben. Egal welcher Ansatz verwendet wird ein Ziel ist es immer den Aufwand menschlicher Kodierung möglichst gering zu halten. Auch die Frage nach Interkoderreliabilität stellt sich, zumindest für den letzten Schritt der Untersuchung, nicht mehr. Die Nachteile liegen in einer meist schlechteren Validität (wird wirklich gemessen was gemessen werden soll?) der Untersuchung sowie der Schwierigkeit komplexe sprachliche Konzepte für den Computer nachvollziehbar zu machen. Das folgende Kapitel soll einen Überblick über die automatisierten oder zumindest computerunterstützten Verfahren, die verwendet wurden um Populismus zu messen, geben.

3.2.1. Wörterbuch Ansatz

Autor: Gregor Rettenegger

Eines der ersten Forscherteams, welche versuchten Populismus mit computergestützten Verfahren zu erheben, waren Matthijs Rooduijn und Teun Pauwels. In ihrer Abhandlung “Measuring Populism: Comparing Two Methods of Content Analysis” (2011) stellen sie einer manuellen eine automatisierte Inhaltsanalyse gegenüber. Sie halten sich dabei stark an die ideologische Definition von Populismus nach Mudde (z.B. 2004: 543, 2007, 2017...) und versuchen, Populismus als eine dünne Ideologie, bestehend aus den Komponenten “Anti-Elitism” und “People-Centrism” (Auch „Reference to the People“) in einer komparativen Studie zu erheben. Als Datenmaterial ziehen sie Partei-Manifeste von 13 politischen Parteien aus verschiedenen Wahlkampf Perioden aus den Niederlanden, Vereinigten Königreich, Deutschland und Italien heran. In der manuellen Inhaltsanalyse wurden Kodierer dazu angehalten, die Kategorie in einen breiten Kontext zu interpretieren. Das Wort “Wir” zum Beispiel, sollte nur kategorisiert werden, wenn es im Kontext “Wir, das Volk, ... “ gebraucht wird. Weiteres wurden folgende Worte (und ihre Übersetzungen) verwendet:

people, citizen(s), community, society, public, population, nation(al), all of us, each of us, everyone, our, we, voter(s), electorate, referenda, direct democracy, public

opinion, country. And words such as: United Kingdom, Britons, Netherlands, Dutch, Italians, Germany, etc. (depending of course on the country under analysis.)

elit konsens* undemokratisch* referend* korrupt* propagand* politiker* täusch* betrüg* betrug* verrat* tradi* scham* schäm* schand* skandal* wahrheit* unfair* unehrlich* establishm* herrsch* kapitul* kaste* leugen* lüge**

(Rooduijin und Pauwels, 2011: 1280, Notes)

Die Kategorie Anti-Elitismus wurde mit Fragen, wie Beispielsweise “Do the authors of the manifesto criticise elites?” erhoben. Bei der computergestützten Inhaltsanalyse wurde nur mehr die zweite Komponente, Anti-Elitism, erhoben, mit folgender Begründung:

“After a first explorative analysis, it turned out that a measurement of people-centrism by means of individual words only is nearly impossible. In many instances, ‘the people’ is referred to by the words ‘we’ and ‘our’ (e.g. ‘we [the people] need to raise our voice’). Yet not every mention of the words ‘our’ or ‘we’ is a reference to the people. Often, these words refer to the political party instead of the people (e.g. ‘we [the party] propose our plans in the next chapter’).”

(ebd.: 1275)

Zum Einsatz kam bei dieser Studie die so genannte “Wörterbuch” (oder “Dictionary”-) Methode. Dabei wird eine Liste mit Wörtern aus der Literatur oder aus qualitativen Verfahren abstrahiert und von einem Computerprogramm gemessen, wie oft diese Wörter vorkommen (vgl. Laver und Garry, 2000). Die Ergebnisse beider Methoden sind ähnlich, jedoch sind alle Validitätswerte für die manuelle Inhaltsanalyse höher. Trotzdem sprechen Rooduijin und Pauwels von guten Neuigkeiten für komparative Forscher (vgl. Rooduijin und Pauwels, 2011: 1278).

Teun Pauwels hat sich seither noch öfter (z.B. 2011b, 2014, 2017) mit der Erhebung von Populismus durch computergestützte Verfahren beschäftigt. Gerade kürzlich erschien ein Beitrag zur computergestützten Erhebung von Populismus in Deutschland. Dabei wendet er eine klassische (holistische) Inhaltsanalyse und eine automatisierte mit und ohne manuelle Nachbearbeitung zur Erkennung von “False Positives” (Treffer des Wörterbuchs die keine sein sollten) an. Die Ergebnisse der Studie zeichnen ein eher problematisches Bild, da Ergebnisse, selbst der manuellen Inhaltsanalyse, nicht nur voneinander, sondern auch von den Annahmen von Experten abweichen. Lediglich bei der “Der Linken” konnten alle drei Methoden Populismus erkennen (Pauwels, 2017: 130). Auch das

Zustandekommen vieler in der Literatur verwendeter Dictionaries muss problematisiert werden. Rooduijin und Pauwels etwa sprechen davon, dass Sie sich von vorigen empirischen Arbeiten “inspirieren” lassen haben und durch theoretische Überlegungen Wörter eliminiert oder hinzugefügt haben. Sie übersetzen das Wörterbuch auch mit nur wenigen Änderungen in vier weitere Sprachen auf Grundlage der Vermutung, dass Anti-Elitismus überall gleich ist (vgl. Rooduijin und Pauwels, 2011: 1275f).

Wie sich zeigt kann eine Untersuchung nur valide Ergebnisse liefern, wenn erst ein möglichst gutes Dictionary um Populismus zu messen, erstellt wurde. Leider sind Aussagen über die ideale Menge an Wörtern nicht zu treffen, Pauwels (2011) verwendet 28 Wörter (davon je nach Sprache ca. 7 spezifische), Harrison and Bruter (2011) ca. 230. Weiters können intuitiven Wörter, wie Beispielsweise “people” oder “elites” oft nicht ins Wörterbuch aufgenommen werden, da diese zu viele falsche Treffer produzieren. Weiteres sind die so erstellten Wörterbücher keineswegs einfach in anderer Sprachen übersetzbar bzw. muss auf regionale Themen und Sprachvariationen Rücksicht genommen werden.

Es gibt zwar Methoden um die Erstellung von Dictionaries zu unterstützen, diese werden jedoch selten angewandt und bergen ihre eigenen Probleme. Als Beispiel soll hier etwa die FAC (Fully Automated Clustering) erwähnt werden. Dabei versucht ein Algorithmus Kategorien aus den Texten zu ziehen und diese dann in eben jene Kategorien einzuteilen. Die Literatur zu FAC Methoden ist sehr umfangreich (einen Überblick geben etwa Manning, Raghavan & Schütze, 2008), die Grundprinzipien sind jedoch ähnlich: Es gibt eine Definition von Abweichung bzw. Nähe zum Cluster-Zentrum, eine Zielfunktion, welche ein ideales Clustering operationalisiert und einen Optimierungsalgorithmus, doch trotz der Masse an Literatur gibt es kaum Anleitungen welche Ähnlichkeitsmaße, objektive Funktionen oder Optimierungsalgorithmen zu guten Ergebnissen führt ohne umfangreiches Wissen über das Datenmaterial zu haben (vgl. Grimmer & Stewart, 2013: 282f). Man spricht bei automatisierten Verfahren, welche nicht von einem Kategorienschema angeleitet werden, vom so genannten „unsupervised machine learning“.

Aus der Arbeit von Pauwels (2011) können wertvolle Lektionen gelernt werden: Allen voran, dass sich Partei-Manifeste vor allem in Deutschland kaum zur Erhebung von Populismus eignen. Das liegt daran, dass vor allem rechte Parteien - eventuell aus historischen Gründen - in Manifesten Wörter wie “Volk” tunlichst vermeiden (vgl. Arzheimer, 2015). Auch versuchen viele Parteien ihren Populismus, welcher in Deutschland sehr negativ konnotiert ist, zu kaschieren. Das ist gut erkennbar an dem

Beispiel “Alternative für Deutschland”. Obwohl Experten diese Partei als die am meisten populistische ansehen, weisen nicht bloß die drei oben genannten Methoden, sondern auch eine qualitative Analyse den Wert Null auf der Skala auf.

Aslanidis (2017) fasst den derzeitigen Stand der Versuche Populismus mit einem Wörterbuch Ansatz treffend zusammen: “Dictionaries carry a broad brush when identifying populism and pose hurdles to comparative research across different languages and political cultures.” (ebd.:1247). Andere Autoren gehen mit dieser Methode noch deutlich strenger ins Gericht. So bezeichnet Bauer (2000) die Idee komplexe Konzepte wie Populismus an Hand des Zählens einzelner Wörter zu messen schlicht als absurd. Obwohl Wörterbuch Ansätze bisher nicht in der Lage waren Populismus valide messbar zu machen, können sie doch eine hilfreiche Rolle bei der Vorauswahl von Material haben (Pauwels, 2014). Aslanidis (2017: 1248) hält weiter fest: “Future developments in information theory may soon improve automated methods, overcoming some of the reservations noted above”. Die vorliegende Studie versucht einen solchen Ansatz vorzustellen.

3.2.2. Supervised Machine Learning (SML)

Autor: Gregor Rettenegger

Machine Learning Ansätze in der Text Analyse basieren darauf den Computer selbst Regeln erstellen zu lassen um Texte in Folge richtig zu klassifizieren. Bei Supervised Machine Learning (SML) wird anders als beim Wörterbuch-Ansatz von den Forschern keine Liste von Wörtern oder sonstigen Regeln erstellt. Dem “learner” des Algorithmus wird ein Trainings-Set, bestehend aus vorab klassifizierten Texten zur Verfügung gestellt anhand dessen dieser automatisch einen “classifier” bildet, mit dem neue Texte automatisch klassifiziert werden können (Sebastiani, 2002). Je nach verwendetem SML-Algorithmus kann die Komplexität eines solchen Classifiers weit über reines Zählen von Wörtern hinausgehen.

Ein kurzer historischer Überblick über die Entwicklung von Machine Learning Methoden findet sich bei Sebastiani (2002). Demnach waren in den 80er Jahren sogenannte “knowledge engineering” Techniken die beliebtesten Methoden zur automatisierten Text-Klassifikation. Dabei wurde, ähnlich zu heutigen Dictionary Ansätzen, von Experten ein Set an Regeln definiert mit dessen Hilfe dann ein Computer Texte entsprechend

klassifizieren konnte. Der Nachteil eines solchen Ansatzes sei der Flaschenhals der “knowledge acquisition” (ebd. S. 9), was meint, dass die Regeln von Experten auf dem jeweiligen Gebiet erstellt und gewartet werden müssen. Ändere sich etwas am zu bewertenden Material oder den Regeln müssten erneut Experten in Anspruch genommen werden. Mit diesen Ansätzen seien Teils sehr hohe Validitätswerte erreicht worden, diese konnten jedoch nie überprüft werden und schienen aus heutiger Sicht unwahrscheinlich. In den frühen 90 Jahren begannen Machine Learning Methoden in der Text Klassifikation populärer zu werden. Wie in der Einführung dieses Kapitels beschrieben, wurden Texte vorab von Experten manuell klassifiziert. Aus diesem Set von Texten baut ein “general inductive process” (auch “learner” genannt) dann einen “classifier”. Von diesem Classifier leitet der Learner ab welche Eigenschaften ein neuer Text haben sollte, um ebenfalls gleich klassifiziert zu werden. In Machine Learning Terminologie wird dieser Prozess “supervised learning” genannt, woraus sich der Name “Supervised Machine Learning” (SML) ableitet. Die Vorteile dieser Vorgehensweise seien evident. Die Programmierarbeit fließe hier in die Erstellung eines vielseitig einsetzbaren Learners, anstatt eines einmaligen Classifiers. Ist dieser erst einmal programmiert, seien bei SML Ansätzen die manuell kodierten Texte die wichtigste Ressource. Es sei einfacher Texte zu klassifizieren als es sei Regeln dafür zu erstellen, da es einfacher sei “to characterize a concept extensionally (i.e. to select instances of it) than intensionally (i.e. to describe the concept in words, or to describe a procedure for recognizing its instances).” (ebd.: 10)

In der Literatur gibt es bisher keine Versuche Populismus mit Hilfe von Machine Learning messbar zu machen, in ihrer Studie “Teaching the computer to code frames in news: Comparing two supervised machine learning approaches to frame analysis.” haben Burscher und Kollegen (Burscher et al., 2014) jedoch verschiedene SLM Algorithmen auf ihre Fähigkeit Frames in der Berichterstattung valide zu erkennen, untersucht. Dazu wurden drei verschiedene Algorithmen zuerst einzeln untersucht und dann zu einem sogenannten “Ensemble” kombiniert. Dabei wurden mit einer Technik die sich “stacked generalization” nennt (wobei unter anderem ein Learner trainiert die Treffsicherheit der Algorithmen im Ensemble zu maximieren) folgende Arten von Machine Learning Algorithmen kombiniert: zwei “Linear Support Vector Machines (SVM)” (Joachims, 1998), ein “Polynomial SVM classifier” (Chang et. al., 2010) und ein “Perceptron Algorithmus” (Lippmann, 1987). (Burscher et. al., 2014: 195).

Untersucht wurden die Titelseiten von holländischen Tageszeitungen zwischen 1995 und 2011. Ähnlich zur vorliegenden Studie wurden nur Artikel berücksichtigt, die in einer vorherigen manuellen Inhaltsanalyse einen Bezug zur Politik aufwiesen. Die Ergebnisse ließen die Autoren Schlussfolgern, dass SML ein funktionierendes Instrument zum Messen von Frames in der politischen Berichterstattung darstellt (ebd.: 197). Auf die verschiedenen untersuchten Medien aufgeteilt zeigte sich, dass die Treffsicherheit (Accuracy) der Learner, Texte nach den vier untersuchten Frames zu klassifizieren, zwischen 0.74 und 0.89 liegt (vgl. Abbildung 3).

TABLE 4
Classification Accuracy of Frames in Sources Outside the Training Set

	<i>VK/NRC</i> → <i>Tel</i>	<i>VK/TEL</i> → <i>NRC</i>	<i>NRC/TEL</i> → <i>VK</i>
Conflict	.69	.74	.75
Economic Cons.	.88	.86	.86
Human Interest	.69	.71	.67
Morality	.97	.90	.89

Note. VK = Volkskrant, NRC = NRC/Handelsblad, TEL = Telegraaf

Abbildung 3: Classification Accuracy, aus Burscher et. al., 2014: 201

Hier sei nur kurz erwähnt, dass Accuracy sowie die ebenfalls angegebenen Kennzahlen ROC Kurve und AUC, in der Literatur nicht undiskutiert sind (mehr dazu in Kapitel 3.3.) und gewisse Qualitätsmerkmale automatisierter Klassifizierung nicht berücksichtigen. Auch die manuelle Kodierung der Dokumente, die für das Trainieren des Learners verwendet wurden, war offensichtlich nicht unproblematisch. So werden für die einzelnen Indikator Fragen, die verwendet wurden um das Vorkommen eines der untersuchten Frames zu bestimmen, Krippendorff's Alpha Werte zwischen .16 und .61⁵ bzw. prozentuale Übereinstimmungen zwischen .57 und .92, berichtet (ebd.: 197).

Trotz dieser nicht unproblematischen Qualitätsmerkmale lassen sich doch interessante Erkenntnisse aus der Arbeit von Burscher et. al. ziehen. So wurde beispielsweise untersucht ob ein holistischer oder ein sogenannter "Indicator Based" Ansatz die besseren Ergebnisse liefert. Bei ersterem werden alle Dokumente die einen Frame enthalten für das Trainieren des SML Learner verwendet werden. Bei Zweitem werden für jede der Indikator Fragen

⁵ Keine einzige der 10 Fragen lag über dem üblicherweise als akzeptabel geltenden Minimalwert von 0.68.

einzelne Learner trainiert, wenn einer dieser Learner ein Dokument positiv klassifizierte zählte dieses für den gesamten Frame. Es zeigte sich, dass der holistische Ansatz bei allen vier untersuchten Frames die besseren Ergebnisse lieferte (ebd.: 199). Ein durchaus überraschendes Ergebnis, würde man doch davon ausgehen, dass je weniger komplex ein zu klassifizierendes Konzept ist, desto einfacher es für einen Learner sein sollte, dieses korrekt zu klassifizieren. Wirft man einen genauen Blick auf die Indikator Fragen zeigt sich jedoch, dass diese nicht dazu geeignet sind einen der Frames in sprachlich klar voneinander abzugrenzende Komponenten aufzuteilen, weshalb die Klassifikation für den Algorithmus dadurch nicht leichter wird. Als Beispiel sollen hier die Indikator Fragen für den “Economic Consequences” Frame dienen. Die erste lautet “Is there a reference to the financial costs/degree of expense involved, or to financial losses or gains?”, die zweite, “Is there a reference to economic consequences of pursuing or not pursuing a course of action?”. Ein SML Learner klassifiziert Dokumente, stark vereinfacht gesagt, auf Grund von Vorkommen und Fehlen gewisser Wörter oder Wortgruppen. Die hier genannten komplexen Fragen haben für den Learner keinerlei Bedeutung, es ist aber davon auszugehen, dass Dokumente, welche viele Begriffe aus der Wirtschaftswelt beinhalten positiv mit beiden dieser Fragen korrelieren.

Eine weitere interessante Erkenntnis betrifft die benötigte Anzahl an Trainings-Dokumenten, die nötig sind um einen Learner soweit zu trainieren, dass er brauchbare Ergebnisse liefert. In einem Experiment wurde die Anzahl der Trainings-Dokumenten in den folgenden Schritten erhöht: 100, 200, 500, 1000, 2000, 3000, und 4000. Dabei zeigte sich, dass die Accuracy der Learner mit zusätzlichen Dokumenten im Trainingsset bei allen vier untersuchten Frames stieg, die Steigerung aber ab einer relativ geringen Anzahl bereits stark abnahm (Siehe Abb. 4). Bei den für den Algorithmus offenbar am leichtesten zu klassifizierenden Frame “Economic Consequences”, stagnierte die Accuracy bereits bei ca. 250 Trainingsdokumenten beinahe. Diese Unterschiede gehen vermutlich auf die unterschiedliche sprachliche Komplexität der verschiedenen Frames zurück, wobei es naheliegend scheint das der “Economic Consequence”- Frame am einfachsten anhand bestimmter Wörter (aus dem Bereich der Wirtschaft) klassifiziert werden kann. Die Autoren nennen die Häufigkeit mit der ein Frame im zu untersuchenden Set vorkommt als einen Grund für die unterschiedlichen Ergebnisse. In dieser Studie wurde aber erstmals gezeigt, dass bereits mit deutlich weniger Dokumenten als in früheren Studien verwendet wurden (z.B. Hillard, 2007), erfolgreich ein SML Learner trainiert werden kann.

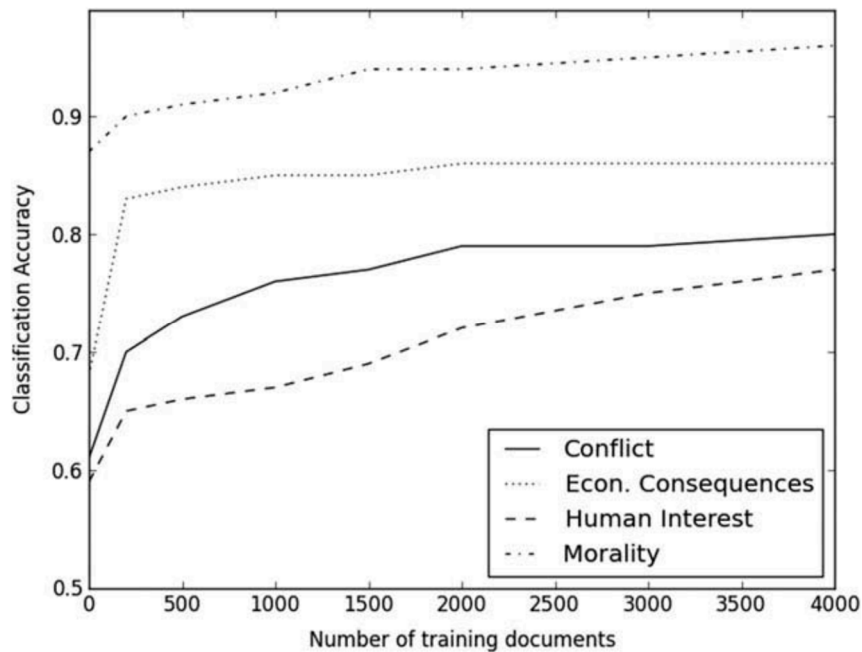


Abbildung 4: Zusammenhang zwischen Genauigkeit der Klassifikation und Anzahl der Trainingsdokumente für die verschiedenen Frames, entnommen aus Burscher et.al. 2014: 201.

In einer weiteren Studie die hier erwähnt werden soll, wurde mit Hilfe von SML versucht News Values und verschiedene Themen richtig zu klassifizieren (Scharkow, 2013). Untersuchungsgegenstand waren sämtliche Artikel aus 12 deutschen online Zeitungen, zwischen Mai 2018 und Juni 2019. Aus über 200.000 Artikeln wurde ein zufälliges Sample von 1000 Artikeln gezogen, die in Folge von 8 trainierten Kodieren vor-kodiert wurden. Jene Texte die für das Trainings-Set verwendet wurden, wurden vorab durch verschiedene übliche Verfahren vorbereitet. So wurden die 1000 häufigsten deutschen Stop-Wörter entfernt, und Word-Stemming, ein Verfahren bei dem Wörter auf ihren Wortstamm reduziert werden, indem Affixe, wie Konjunktionen und andere morphologische Formen von Wörtern entfernt werden, durchgeführt (ebd.: 768).

Auch in dieser Studie wurde gezeigt, dass SML ein valider Ansatz ist um Texte nach traditionellen kommunikationswissenschaftlichen Konzepten zu klassifizieren. Augenmerk wurde unter anderem auf die Qualität des Trainings-Sets gelegt, das für die verschiedenen gemessenen Variablen relativ stark unterschiedliche Interkoderreliabilität, aber auch Häufigkeit im Set aufwies (siehe Abbildung 5).

Table 1 Inter-coder reliability for the manual content analysis

Variable	Percent agr.	CI_{Acc}	Kripp. α	CI_{α}	n_{Art}
National politics	0.90	0.87–0.92	0.69	0.60–0.76	373
International politics	0.93	0.90–0.95	0.76	0.65–0.82	373
Pol. economics	0.93	0.91–0.95	0.74	0.65–0.83	373
Sports	0.99	0.98–1.0	0.98	0.93–0.99	395
Disasters, accidents	0.95	0.93–0.97	0.67	0.54–0.80	373
Crime	0.92	0.89–0.95	0.67	0.56–0.77	373
Controversy	0.69	0.65–0.74	0.49	0.40–0.56	373
Prominence	0.71	0.66–0.75	0.72	0.66–0.77	373

Abbildung 5: Interkoder Reabilität für manuelle Inhaltsanalyse, aus Scharkov, 2013: 767

Vergleicht man diese Tabelle mit den Ergebnissen der SML Learner (siehe Abbildung 6) zeigt sich, dass diese im Schnitt eine ungefähr 5% schlechtere Accuracy aufweisen (Spalte CR_a - CR_m), was bei großen Datenmengen immer noch akzeptable Werte liefert. Die “chance corrected” Accuracy (Spalte α_a) zeigt jedoch, dass “Controversy” und “Crime” nur unzureichend, und “Disasters” gar nicht klassifiziert werden konnten (ebd.: 769).

Table 2 Classification quality for supervised learning

Variable	CR_a	$CR_a - CR_m$	α_a	$\alpha_a - \alpha_m$	Precision	Recall
National politics	0.86	−0.04	0.55	−0.14	0.65	0.63
International politics	0.89	−0.04	0.61	−0.15	0.77	0.60
Pol. economics	0.90	−0.03	0.61	−0.13	0.65	0.69
Sports	0.96	−0.03	0.84	−0.14	0.94	0.80
Disasters/accidents	0.93	−0.02	0.17	−0.50	0.78	0.12
Crime	0.86	−0.06	0.36	−0.31	0.66	0.32
Controversy	0.62	−0.07	0.30	−0.19	0.62	0.52
Prominence	0.60	−0.11	0.45	−0.27	0.73	0.63

Subscripts a and m indicate manual and automatic coding. Precision and recall only apply to automatic coding

Abbildung 6: Klassifikations- Qualität für SML, aus Scharkov, 2013: 768

Warum es zu den schlechten Werten bei diesen Variablen gekommen ist analysiert Scharkow (2013) wie folgt. Vergleicht man die die Accuracy der Coder (bzw. den Krippendorff Alpha Wert, Abb. 7, rechte Grafik) mit der Accuracy der SML Learner zeigt sich, dass bei einer einfachen prozentualen Übereinstimmung (Accuracy), ohne Berücksichtigung der Häufigkeit von positiven Fällen (linke Grafik), ein fast perfekt linearer Zusammenhang zwischen den beiden Kennzahlen besteht. Verwendet man jedoch

einen Krippendorff's Alpha, welcher die Häufigkeit von positiven Fällen mitberücksichtigt, zeigt sich, dass die Kennzahlen stärker vom linearen Verlauf abweichen. Daraus schlussfolgert Scharkow (2013: 769), dass die geringe Anzahl von Dokumenten im Trainingsset, stärker für die schlechte Performance der SML Learner verantwortlich ist, als die Qualität der Kodierung (Interkoderreliabilität). Wir ergänzen hier noch, dass es vor allem das Fehlen von positiven Fällen im Trainingsset ist, das zu den schlechten Ergebnissen bei seltenen Konzepten wie "Disasters" führt.

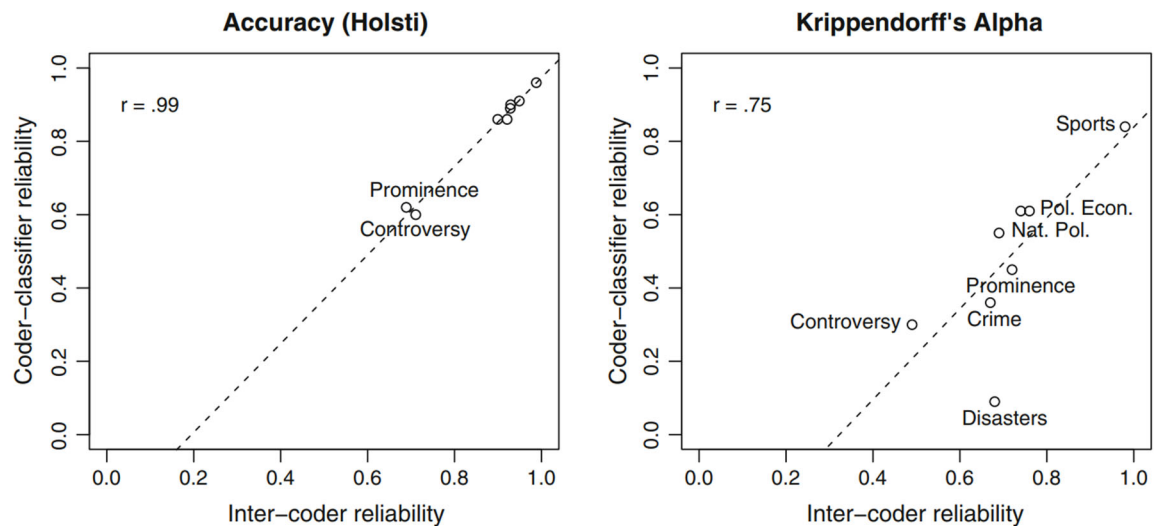


Fig. 1 Relationship between inter-coder reliability and classification quality

Abbildung 7: Zusammenhang zwischen Interkoder Reliabilität und Klassifikations- Qualität, ebd.

Die oben erwähnten Studien zeichnen ein durchwegs positives Bild von SML Methoden. Diese scheinen aus mehreren Gründen gut geeignet um kommunikationswissenschaftlich relevante Konzepte in Texten zu identifizieren. Im Gegensatz zu manuellen Methoden, mit denen man entweder nur eine kleine Anzahl an Dokumenten klassifizieren kann oder einen massiven Kostenfaktor darstellen, ist es bei SML - Ansätzen, sobald ein Learner trainiert wurde, egal ob 100 oder 100.000 Dokumente klassifiziert werden sollen. Ein Vorteil der in Zeiten von immer größeren Textmengen, die Forschern zur Verfügung stehen eine erhebliche Limitation von manuellen Ansätzen darstellt.

Im Gegensatz zu Wörterbuchansätzen, mit welchen bisher keine valide und reliable Klassifizierung von Populismus gelungen ist, scheinen sie deutliche Vorteile in der Qualität der Ergebnisse zu haben. Das mag wie oben erwähnt vor allem daran liegen, dass Populismus und hier vor allem die Komponente "People Centrism" sehr schwer anhand einzelner Wörter zu messen ist. Wörter wie "wir" oder Bi-Gramms wie "die_anderen"

können nicht ohne Kontext interpretiert werden, auf der anderen Seite sind diese jedoch essentiell und fehlen, wenn sie nicht Teil eines Wörterbuchs sind.

Trotz ihrer Probleme haben jedoch auch beide Ansätze ihre Stärken und sind damit bestens geeignet einen SML Ansatz zu unterstützen. So ist eine reliable manuelle Kodierung der Dokumente im Trainingsset essentiell für den Erfolg einen Learner zu trainieren. Vorab können Wörterbuchansätze dabei helfen, Texte zu selektieren in denen das zu untersuchende Konzept tendenziell öfter vorkommt. Vor allem bei selten auftretenden Konzepten kann dies eine große Zeitersparnis darstellen. Somit gilt es aus jedem Bereich die bewährtesten Ansätze zu nützen um das beste Ergebnis zu erreichen. In der manuellen Kodierung scheint dies der Ansatz der “semantic clauses” zu sein, bei Wörterbüchern eine Kombination der bewährtesten Wörterbücher. Im nächsten Kapitel wird auf die Wahl des besten SML Learners näher eingegangen.

3.2.3. Wahl des Werkzeuges

Autor: Maximilian Haselsteiner

Hat man sich erstmal für eine supervised machine-learning Methode entschieden, muss man aus einer Vielzahl von verschiedenen Modellen und Algorithmen wählen. Das Ziel des Modells ist eine Prognose welche sich gut generalisieren lässt, also auch auf andere Daten Sets angewendet werden kann. Schon hier findet sich ein großer Vorteil gegenüber Dictionary Modellen. Loughran und McDonalds (2011) kritisieren etwa die immer häufiger werdende Verwendung von gebrauchsfertigen Dictionaries in einem betriebswissenschaftlichen Kontext. Sie zeigen, dass viele Wörter je nach Kontext verschiedene Konnotationen haben können und viel zu selten validiert werden (können). So müsste eigentlich für jede Fragestellung ein eigenes Wörterbuch angelegt werden und eingehend validiert werden, was einen enormen Aufwand bedeutet. In Kapitel 3.2.1. wurde bereits kritisiert, dass Wörterbücher auch im Hinblick auf Populismus oft mit einer gewissen Willkür erstellt werden und nur selten international generalisierbar sind. Supervised Machine Learning Methoden können dieses Problem beheben, da der Algorithmus aus jedem Daten Set lernt und die Validierung um ein vielfaches einfacher ist (vgl. Kapitel 3.2.2., Grimmer & Steward, 2013: 275).

Es gibt eine Vielzahl konkurrierender Algorithmen und Modelle. SML Methoden teilen sich grundsätzlich in erklärende und prognostizierende Modelle auf, während erklärende Modelle kausale Zusammenhänge suchen (und für diese Arbeit nicht relevant sind), ist das

Ziel von zweitem die Prognose einer abhängigen Variable (oder input variable) mittels unterschiedlicher Kovariate (diese werden in der Literatur oft als Predictors, Einflussgrößen, Input Variable oder unabhängige Variablen benannt). Das Ergebnis jedes prognostizierten Modells ist eine “prediction function” und die Genauigkeit dieser wird durch die “loss function”, welche die Diskrepanz zwischen den vorhergesagten und tatsächlichen Resultaten misst (vgl. Nielsen, 2016: 8), festgestellt. Stark vereinfacht dargestellt, basieren SML Modelle auf einer Risiko-Minimierungs-Funktion (welche sich aus der loss-function berechnet) und einem “Learning-Algorithmus”, welcher in seiner simpelsten Form etwa ein konstantes oder lineares Modell sein kann. Von “adaptiven Basis-Funktionsmodellen” spricht man, wenn die Lernfunktion nicht von vornherein feststeht, sondern jedes Mal gelernt werden. In dieser Kategorie finden sich auch Neuronale Netzwerke sowie Boosting- und (Decision bzw. Regression-) Tree- Modelle. Ein Modell sollte so wenig Vermutungen wie möglich machen müssen und flexibel genug sein, damit alle Strukturen von Interesse in das Datenmaterial passen. Ist das Modell jedoch zu flexibel, kommt es zum sogenannten “overfitting”. Das heißt die Funktion ist zu komplex und kann nicht gut generalisiert werden. Umgekehrt kommt es zum “underfitting”, wenn die Funktion zu simpel sind, das heißt es können keine interessanten Strukturen in die Daten eingebaut werden. In der Machine-Learning Literatur spricht man vom sogenannten “bias-variance tradeoff” (vgl. Nielsen, 2016: 24).

Die Literatur zu SML Methoden für die Analyse von Texten ist ein extrem schnell wachsendes und unübersichtliches Forschungsfeld. Auch wenn sich die Grundlagen (training set erstellen - modell berechnen und ausführen - validieren) bei den meisten Algorithmen gleichen, ist man dennoch auf Hilfe bei der Auswahl angewiesen. Diese findet sich am besten in der Literatur (zum Beispiel Athey, 2017) oder bei Machine Learning Wettbewerben. Der größte Austragungsort dieser Wettbewerbe ist “Kaggle” (<https://www.kaggle.com/competitions>) und schon seit einigen Jahren ist mit MART (“Multiple Additive Regression Trees”), ein sogenannter Gradient-Tree-Boosting Algorithmus (eigentlich eher eine Sammlung oder “library” an Algorithmen) immer auf den ersten Plätzen. Seit 2014 jedoch dominiert ein anderer Algorithmus-Sammlung die kompetitive Machine-Learning Szene: XGBoost. XGBoost steht für “eXtreme Gradient Boosting” und ist eine vollkommen frei erhältliche Software-Library des an der University of Washington (Seattle) praktizierenden PHD Studenten Tianqi Chen, die mit zahlreichen Interfaces wie C++, Python oder R verwendet werden kann. Nicht nur aufgrund der

Leistungsfähigkeit ist XGBoost derzeit eines der meist verwendeten Produkte, sondern auch weil es ressourcenschonend und schnell ist. Unter den 29 “Kaggle Competition” Siegern seit 2015, benutzten 17 XGBoost. Bei einem weiteren Wettbewerb, dem KDD (Conference on Knowledge Discovery and Data Mining)- Cup (<https://www.kdd.org/kdd-cup>) 2015 nutzten sogar die zehn besten Lösungen die XGBoost Library⁶ (vgl. Chen & Guestrin, 2015: 785).

Ein “Tree Model” ist ein Modell, welches aus unabhängigen Variablen subsets bildet und aus den Ausprägungen dieser wiederum subsets bildet. Durch die zahlreichen Verästelungen ergibt sich das Bild eines umgekehrten Baumes, welcher dem Modell auch seinen Namen gibt (Konkrete Beispiele finden sich in Kapitel 5). Als Ergebnis kann man auf den Decision Tree ablesen wie hoch die Wahrscheinlichkeit ist, das eine Ausprägung unter gewissen Umständen (= andere Variablen bzw. Ausprägungen) eintritt. Diese Tree Models haben für sich eher eine geringe Fähigkeit zur Prognose, kombiniert mit anderen Methoden wie “Random Forests” oder in Kombination mit “Boosting” Algorithmen sind sie für maschinelle Lernmethoden unabdingbar (vgl. Nielsen, 2016: 47). Der Prozess, mehr Informationen in ein Modell zu füttern ohne overfitting zu erreichen und Fehler zu minimieren nennt sich “Regularization”. Eine Möglichkeit overfitting zu verhindern, ist, ein größeres Trainingsset zu verwenden, eine weitere, wie bei “Random Forests” (Breiman, 2001) eine Randomisierung einzuführen. Auch beim “Bagging” (Breiman, 1996) funktioniert durch die Erstellung immer wieder neuer randomisierter Trainingssets, man erreicht dadurch eine beliebige Anzahl an Trainingssets, bei der sich einige Features immer wiederholen. Werden diese Features auch noch gewichtet, spricht man von “Boosting”.

Boosting an sich ist eine sequentielle Methode, in welcher zuerst sogenannte “early learners” gewonnen werden. Diese sind oft sehr ungenau (auch “weak learners”) doch werden die Daten auf Fehler untersucht, welche Probleme oder gewisse Spezifika des Untersuchungsgegenstands aufzeigen. Jede weitere Sequenz konzentriert sich nun auf diese Fehler und versucht sie zu gewichten und daraus “predictors” zu erstellen. So wird aus viele “weak learners” ein komplexes Prognosemodell.

Gradient Boosting ist eine von Friedman (2001) erstellte Weiterentwicklung dieser Technik, dabei werden Regressionsmodelle aus einer Sammlung von Regressoren

⁶ Das Wort “Library” bezeichnet im Machine Learning Kontext eine Sammlung von Routinen, Unterprogrammen bzw. Modulen.

errechnet. Eine mögliche und oft verwendete Basis dabei sind die bereits erwähnten Bäume. Es handelt sich jedoch nicht um Decision-Trees, sondern um Regression-Trees, bei welchen jedes “Blatt” (diese werden meist als “End-Nodes” bezeichnet) eine fortlaufende Score beinhaltet (vgl. Chen & Guestrin, 2015: 786). Abbildung 8 zeigt ein gängiges Anschauungsbeispiel für einen Regressions- oder Classification- Tree. Es geht hierbei darum, wie wahrscheinlich es ist, dass jemand Computerspiele spielt. Das Beispiel geht der Einfachheit halber davon aus, dass Menschen unter 15 Jahren eher spielen und es noch viel wahrscheinlicher ist, wenn dieser männlich ist. Daraus ergeben sich wenig aussagekräftige “prediction scores” welche jedoch durch das Boosting mit weiteren Trees gute Vorhersagen ergeben können. In diesem konkreten Beispiel etwa würde man weitere Trees erstellen etwa mit der Frage ob man gerne Computerspiele spielt, ob man ihn täglich verwendet und so weiter.

Regression Tree (CART)

- regression tree (also known as classification and regression tree):
 - Decision rules same as in decision tree
 - Contains one score in each leaf value

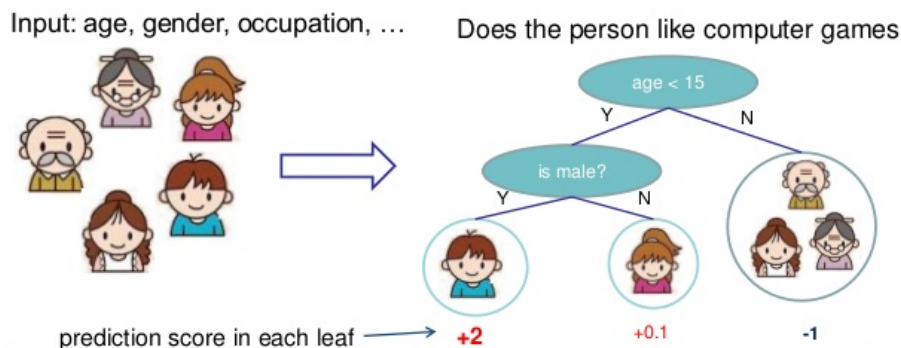


Abbildung 8: Vereinfachte bildliche Darstellung von Regessions-Trees

Ein tieferer Einblick in fortschrittliches Maschinen Lernen würde an dieser Stelle den Rahmen sprengen. Es wurde ein Einblick in die Technik des Gradient Boosting gegeben und gezeigt, dass die Programmbibliothek XGBoost derzeit die state-of-the-art Methode ist um ressourcenschonend, schnell und mit hoher Prognosekraft verschiedenste Daten verarbeiten kann. In der Architektur des Modells sind mehrere subtile Mechanismen für einen optimalen Variance-Bias Tradeoff integriert und er kann sowohl auf kleinere, als

auch enorm große Daten Sets angewendet werden. Erwähnt werden muss dennoch, dass sogenannte Ensembles, also die Verwendung mehrerer Vorhersage-Tools meist noch bessere Ergebnisse liefern kann, XGBoost jedoch, wenn es gut konfiguriert wurde, mit bedeutend weniger Aufwand, nur knapp dahinterliegt:

“Something dramatic happened in Machine Learning over the past couple of years. It is called XGBoost – a package implementing Gradient Boosted Decision Trees that works wonders in data classification. I have tried it on my benchmark dataset – it shows 35% improvement over an SVM. Apparently, every winning team used XGBoost, mostly in ensembles with other classifiers. Sometimes an ensemble of (parametrized) XGBoosts was used – without other classifiers 😊. Most surprisingly, the winning teams report very minor improvements that ensembles bring over a single well-configured XGBoost. After all, in contrary to common belief, one might not need to build an ensemble to win a Data Science competition.”

Ron Bekkerman, 2015.

3.3. Validierung automatisierter Verfahren

Autor: Maximilian Haselsteiner

An dieser Stelle muss ein kurzer Exkurs bezüglich der Validierung von Modellen gemacht werden. Üblicherweise werden Modelle im Machine Learning mit den Werten zu Accuracy Precision und Recall evaluiert. Diese Werte müssen kurz definiert und diskutiert werden. Accuracy (Genauigkeit) meint wie oft ein Dokument richtig klassifiziert worden ist, unabhängig davon, ob es sich dabei um Dokumente handelt die das zu klassifizierende Konzept beinhalten oder nicht. Bei automatisierten Analysen werden Dokumente die das zu messende Konzept enthalten “Positives” genannt, solche die es nicht enthalten “Negatives”. Lässt man Dokumente nun automatisiert klassifizieren, dabei ist es irrelevant ob mit einem Wörterbuch-Ansatz oder einem SML Algorithmus, kann man die klassifizierten Dokumente in vier Kategorien einteilen:

- 1.) True Positives: Dokumente die automatisch als “Positive” klassifiziert wurden, und tatsächlich solche sind.
- 2.) False Positives: Dokumente die automatisch als “Positive” klassifiziert wurden, jedoch in Wahrheit keine sind.
- 3.) True Negatives: Dokumente die automatisch als “Negative” klassifiziert wurden, und tatsächlich solche sind.

- 4.) False Negatives: Dokumente die automatisch als “Negative” klassifiziert wurden, jedoch in Wahrheit keine sind.

Strengere und auch aussagekräftigere Kennzahlen für automatisierte Text Klassifizierungen, die, diese unterschiedlichen Ergebnistypen berücksichtigen, sind “Precision”, “Recall” (und der daraus resultierende “F-Score”) sowie die Spezifität.

- Precision beschreibt die Fähigkeit der Methode Positives richtig zu klassifizieren, bzw. die Wahrscheinlichkeit, dass ein klassifiziertes Positive tatsächlich ein True Positive ist. Bei einer Precision von 1 ist jedes positiv klassifizierte Dokument ein True Positive, bei 0 keines. Die Formel dafür lautet: $TP/TP+FP$
- Recall (auch “sensitivity”) beschreibt die Fähigkeit der Methode Negatives korrekt zu klassifizieren, bzw. den prozentualen Anteil an True Positives in allen untersuchten Dokumenten, die von der Methode auch als Positives klassifiziert wurden. Bei einem Recall von 1 wurden sämtliche Positives aus allen Dokumenten gefunden, bei 0 keines. Die Formel dafür lautet: $TP/TP+FN$
- Der F-Score ist eine Kombination aus beiden, die Formel dazu lautet: $(2 * Precision * Recall) / (Precision + Recall)$
- Schließlich gibt es noch die Spezifität (“specificity” oder “selectivity”) welches prinzipiell die True-Negative Rate misst, also wie oft ein Artikel, welcher das untersuchte Element nicht enthält, korrekt negativ klassifiziert werden.

Precision ist relativ einfach zu messen, da dazu nur die positiv klassifizierten Dokumente untersucht werden müssen, um festzustellen wie viele davon True Positives sind. Recall ist deutlich schwieriger zu messen da hierfür ein ausreichend großes Sample aller Texte untersucht werden muss um festzustellen ob tatsächlich alle negativ klassifizierten Dokument das gesuchte Konzept nicht enthalten. Vor allem für selten vorkommende Konzepte ist das eine Herausforderung, da mitunter sehr viele Dokumente händisch kodiert werden müssen um valide Aussagen treffen zu können. Stryker et al. (2006) haben berechnet, wie viele Dokumente manuell untersucht werden müssen um valide Precision und Recall Werte zu erhalten. Basierend auf anzunehmender Precision/Recall und gewünschtem “Confidence Interval”⁷, wird eine Zahl für die zu untersuchenden Dokumente angegeben, die dann mit der Häufigkeit mit der das zu klassifizierende Konzept

⁷ “Confidence Interval” beschreibt wie sicher das Ergebnis ist und wie groß die Abweichung sein kann. So kann man bei einer Precision von 80% und einen “Confidence Interval” von 5%, zu 95% sicher sein, dass die Precision nicht unter 75% liegt.

in allen untersuchten Dokumente vorkommt, multipliziert werden muss. Wird beispielsweise von einem Precision von .70 ausgegangen und 5% “Confidence Interval” gewählt, ist diese Zahl 548. Tritt ein Konzept nur in 5% aller Dokumente auf, müsste man diese Zahl mit 20 multiplizieren und erhält 10960 Dokumente die man manuell kodieren müsste, um valide Precision und Recall Werte zu bekommen. Bei komplexen Konstrukten mit sprachlich unterschiedlichen Ausprägungen kann ein geringer Recall wert etwa darauf hindeuten, dass nur eine spezielle Art von dem zu untersuchenden Konzept gemessen wird. Diese Werte werden jedoch interdisziplinär in der Literatur vielfach kritisiert (z.B. Chawla, 2005 und Powers, 2007). Accuracy und Specificity etwa müssen in der Machine Learning Literatur mit Vorsicht genossen werden, aus den einfachen Grund, dass die Modelle in nur extrem seltenen Fällen vollkommen händisch geprüft werden. So werden meist nur Precision und Recall oder der daraus resultierende F*-Score angegeben, welcher jedoch die True-Negative Rate (“Selectivity”) vollkommen ignoriert. Außerdem werden existierende Verzerrungen und Vorurteile bestärkt und es wird nicht auf Überzufälligkeit geprüft (Powers, 2007: 1). Besonders schlagend wird das Problem der fehlenden Selectivity bei stark unbalancierten Daten, etwa, weil bei einer unter 5% Positive Rate ein einziger Fehler unverhältnismäßige Auswirkungen auf die Precision hat. Der derzeitige Forschungsstand ist sich noch nicht klar darüber welche Kennwerte am besten sind. Während in der Informatik - wie bereits erwähnt - Accuracy, Recall und Precision angegeben werden, wird in Sozial- und Geisteswissenschaften meist Sensitivity (= Recall) und Specificity angegeben (vgl. ebd.) und oft als sogenannte ROC-Grafik (“Receiver Operating Characteristics”) dargestellt. Diese Graphen werden aus der Konfusionsmatrix aus True Positives, False Positives, True Negatives und False Negatives errechnet, als Vergleichswert gilt die AUC (“Area Under the ROC Curve”). Dieser ist zwar aussagekräftiger für die Messung eines Classifiers (vgl. Fawett, 2005: 873), werden jedoch im Machine Learning in Bezug auf unbalancierte Daten kritisiert, weil der Fall-Out (= False-Positive Rate = $FP / FP + TN$) nicht verhältnismäßig hoch sinkt wenn die Zahl der True Negatives enorm ist (vgl. Pozzolo et al., 2015), weshalb gewichtete Precision-Recall am häufigsten verwendet werden. Am weitesten Fortgeschritten ist der Forschungsstand zur Evaluierung in der Medizin, vor allem deswegen weil die Folgen von schlecht evaluierten Modellen fatal sein können. Dort wurden spezielle Kennzahlen entwickelt, wie Informedness und Markedness. Die Definitionen der Begriffe sind wie folgt:

*“**Informedness** quantifies how informed a predictor is for the specified condition, and specifies the probability that a prediction is informed in relation to the condition (versus chance). **Markedness** quantifies how marked a condition is for the specified predictor, and specifies the probability that a condition is marked by the predictor (versus chance).” - Powers, 2007: 5*

Ergo; Informedness (auch “Bookmaker- Informedness”) errechnet sich aus $[\text{Recall} + \text{Selectivity}] - 1$ und Markedness (in den Geisteswissenschaften v.a. Psychologie meist “DeltaP” genannt) aus $[\text{Precision} + \text{Negative Predictive Value} (= \text{“inverted precision”})] - 1$. In dichotomen Daten wird die Informedness, vor allem in der älteren Literatur auch “Youden’s J Statistic” oder “Youden Index” genannt.

Letztlich ist für die Wahl der Evaluierungs-Kennzahlen von großer Bedeutung, welche Informationen man daraus ziehen will. Im konkreten Beispiel dieser Arbeit etwa wird ein extrem seltenes Phänomen untersucht. Bei der Validierung wurden in einem Validierungsset von 900 Absätzen, - obwohl 4/9 dieses Sets nur semi-randomisiert war (mehr dazu Kapitel 4.3.3. und 5.) - in der manuellen Analyse nur etwa 4% der Absätze als Populismus kodiert. Würde man nun zwei Modelle A und B nur unter Verwendung der “F-Score”, Precision und Recall vergleichen, würde man die exakt selben Werte bekommen, obwohl Modell A richtigerweise nur 100 Absätze ausschließt (=True Negative) während Modell B 800 Absätze korrekt ausschließt. Da auf jeden Fall ein Teilziel dieser Arbeit eine Voruntersuchung des Datenmaterials ist um - ganz konkret - weniger Artikel für zukünftige Forschungsprojekte lesen und kodieren zu müssen, wäre ein solcher Vergleich in diesem Fall äußerst irreführend. Das liegt daran, dass Precision und Recall (und die daraus resultierende F-Score) nur auf den positiven Werten beruhen. Es ist für diese Arbeit also fundamental, dass in der Validierung komplexere Kennzahlen wie der Youden Index (Informedness) und Markedness angewendet werden. Auch die ROC Graphen (und die daraus folgende AUC) haben, wie bereits erwähnt, in unbalancierten Konfusionsmatrizen meist Verzerrungen, der Informatiker Thomas Krafft kommt in seiner Masterarbeit⁸ über den Vergleich verschiedener Qualitätsmaße in der Kriminalprognostik sogar zu folgendem Schluss: “Die hohe Nutzungsquote der AUC bei binären Klassifikatoren für Rückfälligkeitsvorhersagen erscheint willkürlich und erklärt sich

⁸ Natürlich muss man bei der Zitierung von Masterarbeiten vorsichtig sein, allerdings sind seine Schlussfolgerungen konform mit Kopas, 1992 und Urbaniok, 2008 und wurde später in verkürzter Form in einer Fachzeitschrift publiziert.

lediglich durch ihre verbreitete Anwendung in der Historie des maschinellen Lernens” (ebd., 2017: 49). Weitere Möglichkeiten wären eine Gewichtung oder Normalisierung der Daten aus der TP/TN/FP/FN Konfusionsmatrix oder die Verwendung eines angepassten Qualitätsmaßes wie dem von Maratea, Petrosino und Manzo vorgeschlagenen “Adjusted F - Measure” der das F-Maß mit dem “inverted F-Measure” verbindet:

$$\text{Adjusted F - Measure (AGF): } \sqrt{F(2) * F(0,5) \text{ inverted}}$$

wobei (2) und (0,5) für eine Gewichtung steht, um den False Negatives einen höheren Wert zu geben. F(2) berechnet sich also aus $F(2) = (1+2^2) * [(Precision * Recall) / (2^2 * Precision * Recall)]$ (vgl. Maratea et al., 2013: 335). Problematisch an solchen Kennzahlen - und die gibt es zahlreich, vor allem in der Medizin - ist, dass sich anscheinend keine Kennzahl als Standard durchsetzen konnte und somit die Vergleichbarkeit mit anderen Modellen zu wünschen übrig lässt.

4. Methode

4.1. Forschungsfragen

Autor: Gregor Rettenegger

Ambitioniertes Ziel der vorliegenden Arbeit war es das komplexe Phänomen populistischer Kommunikation in der medialen Berichterstattung messbar zu machen, und Forschern damit ein Werkzeug zur Verfügung zu stellen um dieses genauer zu untersuchen. Dazu wurde ein mixed-methods Ansatz gewählt, der die bisher vielversprechendsten Versuche Populismus zu quantifizieren kombiniert und sinnvoll zu erweitert. Die Forschungsfragen die uns in diesem Prozess leitete war:

FF 1: Kann Populismus in der politischen Berichterstattung mit Hilfe von Machine Learning quantifiziert werden?

Weiterführend stellt sich auch die Frage welche Faktoren auf das Ergebnis den größten Einfluss nehmen. Dabei sind verschiedene Faktoren, welche dem Algorithmus zuzuschreiben sind, wie das gewählte Eta, die maximale Tiefe (bei Tree-Boosting Algorithmen) oder überhaupt der statistische Ansatz des verwendeten Algorithmus, zu unterscheiden. Während die ersten beiden, sowie die Wahl des Algorithmus für einfache Nutzer variierbar ist, kann der Algorithmus selbst nur von Experten mit profunden Programmierkenntnissen verändert werden. Am relevantesten für den Erfolg einer SML Klassifizierung scheint jedoch die Zusammensetzung des Trainingssets zu sein (Hillard, 2008). Hier haben Forscher die meiste Kontrolle. Das Trainings-Set bildet sozusagen die Schnittstelle zwischen SML Learner und Forscher. Über die optimale Beschaffenheit und Zusammensetzung eines Trainingssets ist in der Literatur bisher wenig Konkretes zu finden. So zeigte Scharkov (2013) zwar unterschiedliche Ergebnisse je nachdem wie hoch die Interkoder-Reliabilität der Dokumente im Set war, und schlussfolgerte, dass die Anzahl der Dokumente wichtiger sei als deren Qualität. Es wurden in dieser Studie aber weder Anzahl noch Qualität der Trainings-Dokumente variiert, sondern nur die Ergebnisse unterschiedlicher Klassifizierungen verglichen.

Burscher et. al. (2016) wiederum variierten die Anzahl der Trainingsdokumente und zeigten, dass die Ergebnisse bereits bei relativ wenigen Dokumenten nur mehr marginal mit deren Anzahl besser wurden. Weiteres wurden Trainingssets die einen holistischen bzw. Indikator basierenden Ansatz wählten verglichen. Über die genaue Zusammensetzung bzw. Unterschiede oder Gemeinsamkeiten der jeweiligen Indikator-Sets wurde aber auch hier wenig berichtet. Die Frage nach der Zusammensetzung des Trainingssets scheint weiter relevant zu sein, deshalb die folgende Forschungsfrage:

FF 2: Welche Faktoren beeinflussen die Qualität eines SML Trainingssets?

Ein erster Schritt die Qualität eines Trainingssets positiv zu beeinflussen, stellt die manuelle Kodierung der Dokumente dar. Je genauer ein Konzept abgegrenzt ist und je homogener die Dokumente, umso leichter sollte es für den Algorithmus sein Regeln für die automatisierte Klassifizierung daraus abzuleiten. Das Erstellen eines Kodierschemas sowie das Training der Kodierer stellt somit die erste Herausforderung und den ersten Schritt zu einem hochwertigen Trainingsset dar. Wichtigstes Qualitätskriterium hierfür ist eine hohe Interkoderreliabilität (Krippendorff's Alpha) bzw. die prozentuale Übereinstimmung der Kodierer.

Eine weitere Herausforderung sowohl an die manuelle Kodierung, als in Folge auch an die Ergebnisse des SML Algorithmus, scheinen selten vorkommende Konzepte zu sein. Sowohl Scharrow (2013), als auch Burscher et. al. (2016) zeigten, dass das Klassifizieren von seltenen Konzepten schwieriger ist. In den erwähnten Studien wurde dieselbe Anzahl an Trainingsdokumenten für verschiedene Konzepte verwendet, unabhängig davon wie viele positive Vorkommen eines Konzepts beinhaltet waren. Diese geringe Anzahl an "Positives" ist problematisch, da ein SML Learner neben Dokumenten die eine zu messendes Konzept nicht enthalten, auch ausreichend viele Dokumente in denen es vorkommt, benötigt. Demnach wäre eine Aufteilung der Dokumente im Trainingsset von 50% "Positives" und 50% "Negatives" ideal.

An True Positives eines sehr selten vorkommenden Konzepts - wie es auch Populismus in der Berichterstattung ist - zu kommen, stellt jedoch eine erhebliche Schwierigkeit dar. Wie unsere Voruntersuchungen gezeigt haben, müssen mitunter an die hundert Dokumente gelesen werden um einen einzigen positiven Fall zu finden. Will man die Anzahl an

Positives im Trainingsset erhöhen, bietet es sich deshalb an, den ersten Learner zu benutzen, um mit dessen Hilfe schneller an zusätzliche Positives zu kommen. In einem iterativen Prozess kann der trainierte Learner auf einen bisher nicht genutzten Teil des zu untersuchenden Sets an Dokumenten angewandt werden. Sortiert man die Ergebnisse nach der Wahrscheinlichkeit, ob Populismus in einem Dokument vorkommt und kodiert nur die wahrscheinlichsten Treffer, erhält man deutlich schneller zusätzliche Positive Treffer. Für das erste Trainingsset müssen zwar weiterhin mitunter sehr viele Dokumente untersucht werden bis ausreichend Fälle gefunden wurden, hat man jedoch genügend gefunden, könnte man damit einen SML Learner trainieren und diesen benutzen um sich für alle weiteren Dokumente eine Wahrscheinlichkeit, dass jene Absätze Populismus enthalten, ausgeben zu lassen. Untersucht man dann nur die Dokumente die eine sehr hohe Wahrscheinlichkeit aufweisen, sollte man deutlich schneller an zusätzliche Treffer kommen. Mit diesen zusätzlichen Treffern kann das Trainingsset um wertvolle Positives erweitert werden, was zu einer noch besseren Klassifizierung führt, womit wiederum noch besser Positives gefunden werden könnten. Ein derartiger iterativer Prozess könnte helfen auch mit relativ geringem Aufwand die Anzahl an Positives im Trainingsset, und somit dessen Qualität und in Folge die Ergebnisse des Learners, zu verbessern.

4.2. Untersuchungsmaterial

Autor: Gregor Rettenegger

Untersuchungsgegenstand der vorliegenden Studie sind unter anderem Zeitungsartikel der Print- und Onlineausgaben der Tageszeitungen Der Standard, Die Presse, Salzburger Nachrichten, Kurier, Kleine Zeitung, Heute, Österreich und Kronen Zeitung die während der letzten 6 Wochen vor der Nationalratswahl 2017 in Österreich gesammelt wurden. Auch Transkripte bedeutender audiovisueller Wahlkampfauftritte wurden in das Untersuchungsmaterial aufgenommen, so finden sich darin etwa Transkripte der “Zeit in Bild” (ORF) oder dem “Puls 4” Wahldebatten. Vom Sample ausgenommen wurde ein Großteil der digitalen Wahlkampfkommunikation. Theorie und empirische Untersuchungen legen zwar nahe, dass Populismus - auch in seiner Minimaldefinition - besonders stark in Sozialen Medien vorkommt, allerdings ist die Kommunikation in sozialen Medien zu direkt (im Sinne von ungefiltert), die benutzte Sprache darin zu unterschiedlich (im Sinne von Vereinfachung) und die Diskussion über soziale Medien zwischen Wunderwaffe der Partizipation auf der einen Seite und der Trivialisierung, dem Sieg der schrillen Worte über

das rationale Argument, auf der anderen Seite, zu umfangreich (siehe z.B. Kneuer 2017: 45ff). Über die Österreichische Wahlstudie (AUTNES) haben die Autoren Zugang zu einer umfangreichen Datenbank von ca. 50000 Artikeln, die zumindest einen national relevanten politischen Akteur beinhalten. Printmedien sind in Österreich in Wahlkampfzeiten, sowie allgemein, nach wie vor die wichtigste Quelle zur Meinungsbildung. Sie beeinflussen durch Intermedia-Agenda-Setting Effekte auch stark die Themen in Social-Media-Kanälen (Vonbun et al. 2016). Frühere Studien zu Populismus als “dünne Ideologie” beschäftigten sich traditionell mit Partei-Manifestos (Pauwels 2017) aber auch Printmedien haben in letzter Zeit mehr Aufmerksamkeit erfahren (Heinesch 2017). Der Zugang zu einer großen Datenmenge, sowie die Hohe Relevanz von Printmedien für die Meinungsbildung haben uns schlussendlich zu dieser Auswahl geführt.

Das Projekt AUTNES (Austrian National Election Study) ist eine umfangreiche sozialwissenschaftliche Analyse österreichischer Nationalratswahlen, die es sich zum Ziel gesetzt hat, nicht nur das Verständnis der Demokratie in Österreich zu verbessern, sondern auch qualitativ hochwertige Daten für die Öffentlichkeit und die Wissenschaftsgemeinschaft bereit zu stellen (Siehe <https://www.autnes.at/>). So sind bereits eine Vielzahl von Publikationen, von Grundlagenforschung, über Kampagnenforschung, Politische Einstellungen, Methodologie und vieles mehr im Spannungsfeld zwischen Angebot, Nachfrage und den Media-Perspektive, entstanden.

Exkurs: Die österreichische Nationalratswahl 2017

Autor: Gregor Rettenegger

Am 15. Oktober, 2017 fand in Österreich die Nationalratswahl statt. Vor dem Auseinanderbrechen der großen Koalition von Sozial- und Christdemokraten war für lange Zeit die rechts-populistische FPÖ in allen Umfragen auf dem ersten Platz. In einem international viel beachteten Wahlkampf gelang es jedoch der ÖVP unter dem jetzigen Bundeskanzler, Sebastian Kurz das Ruder zu Gunsten seiner Partei zu wenden. Doch wie war dieses Aufholen von über 10% Punkten möglich und welche anderen Akteure gilt es zu beachten? Für die vorliegende Studie ist diese Wahl besonders gut geeignet um sich die mediale Berichterstattung zu untersuchen, da populistische Parteien aus den

unterschiedlichsten Kategorien große Erfolge feiern konnten. In Folge ein kurzer Überblick über die in diesem Sinne wichtigsten Akteure.

Der Wahlkampf wurde stark vom Thema Migration dominiert, welches traditionell von der rechts-populistischen Freiheitlichen Partei Österreichs (FPÖ) besetzt ist (Meyer & Müller, 2013), bei dieser Wahl teilten sich jedoch ÖVP und FPÖ die Themenführerschaft im Bereich *Immigration* (AUTNES Multi-Mode Panel Study 2017). Mit den Freiheitlichen (FPÖ) stand eine typische “Populist Radical Right” nach der Definition von Mudde (2007) zur Wahl, die sich in der Vergangenheit immer wieder mit markanten Sprüchen und grenzwertigen Plakaten Aufmerksamkeit zu verschaffen wusste.

Am anderen Rand des politischen Spektrums trat mit der Liste Pilz (LP), des Ex-Grünen Peter Pilz, eine Partei zum ersten Mal zu einer Wahl an. Peter Pilz trennte sich nicht zuletzt wegen Meinungsverschiedenheiten bezüglich der Parteilinie zum Thema Migration, von der er schon länger einen populistischen Schwenk forderte, von der Grünen Partei. Pilz wurde in Folge von den Medien das Label eines “Links- Populisten” umgehängt, was von österreichischen Experten zumindest nicht dementiert wurde ([Kurier, 27.06.2017](#)). Auch Pilz führte einen Wahlkampf der sehr stark auf ihn als Spitzenkandidaten zugeschnitten war.

Aber auch der Österreichischen Volkspartei (ÖVP), unter der Leitung ihres neuen Parteiobmanns und jetzigem Bundeskanzler Sebastian Kurz, fiel durch einen hoch personalisierten, sowie auf das Migrationsthema fokussierten Wahlkampf auf. Ihr gelang es, sich durch eine Umfärbung der Partei von Schwarz auf Türkis, eines untypisch starken Fokus auf das Thema Migration, und populistischer Rhetorik als “neue Kraft” zu positionieren. Damit konnte die ÖVP viele Wähler der FPÖ für sich und schließlich auch die Wahl zu gewinnen (Plasser & Sommer, 2018). Ein “neuer Stil in der Politik” sei gefragt, man müsse das alte "System aufbrechen", um hier nur zwei Beispiele Anti-Elitärer Aussagen während des Wahlkampfs zu nennen ([DerStandard, 29.09.2017](#)). Mit dieser Mischung aus dem Versprechen am besten geeignet zu sein das Volk gegenüber den Flüchtlingsströmen zu schützen und der markanten Anti-Establishment Rhetorik sind auch bei der Liste Kurz alle Kriterien einer populistischen Partei erfüllt (Van Kessel, 2011), die nach Učeň (2007) der ”New Centrists Populism” Kategorie zuzurechnen wäre.

Theoretische Überlegungen warum Personalisierung und eine populistische Kommunikation besonders für neue Parteien von hoher Relevanz sind, wurden von Van der Brug und de Vreese (2011) zusammengefasst. So seien neue Parteien, da sie als Parteiorganisation noch wenig bekannt sind besonders darauf angewiesen die

Aufmerksamkeit der Medien zu erlangen (McNair, 1995). Weiteres sind die “starken”, “charismatischen” Spitzenkandidaten von populistischen Parteien in der Lage das Image ihrer Partei stärker zu beeinflussen, als die Spitzen traditioneller Parteien. Der populistische Stil dieser Parteien spiegelt wichtige Nachrichtenwerte wie Emotionalisierung, Anti-Establishment Rhetorik oder Personalisierung wieder (Plasser & Ulram, 2003) und wird deshalb auch überproportional von den Medien aufgegriffen (Mudde, 2007). Es ist daher nicht untypisch für neue Parteien sich durch Kommunikationen die dem Populismus zuzuordnen sind Gehör zu verschaffen. Dass es der Traditionspartei ÖVP die seit 1986 ununterbrochen Teil der Regierung war, gelungen ist mit Schlagworten wie “System aufbrechen” zu punkten, kann wohl nur durch das kommunikative Talent und den perfekt inszenierten Auftritten ihres Spitzenkandidaten erklärt werden.

Somit bat der Wahlkampf zur österreichischen Nationalratswahl ein breites Spektrum populistischer Akteure von beiden Rändern des politischen Spektrums als auch einer traditionell in der rechten Mitte angesiedelten Partei, deren Kommunikation populistische Züge trugen.

4.3. Manuelle Inhaltsanalyse, Erstellen des Trainings-Sets

4.3.1. Voruntersuchung

Autor: Gregor Rettenegger

Um den Aufwand bei den manuellen Inhaltsanalysen so gering wie möglich zu halten wurde im Vorhinein eine Selektierung des Untersuchungsmaterials durchgeführt. Diese Voruntersuchung wurde mit einem vergleichsweise einfachen “Bag of Words” durchgeführt. Als Grundlage gelten zwei bereits geprüfte Wörterbücher. Erstens, das von Rooduijn und Pauwels (2011) für die Untersuchung in Deutschland eingesetzte Wörterbuch und zweitens aus einer Untersuchung von Heinisch (2018).

Beide Wörterbücher teilen sich ähnliche Probleme, welche bereits in Kapitel 2.3. erwähnt wurden. Um die Präzision weiter zu erhöhen, wurden gewisse Begriffe mit Einschränkungen versehen. Der Begriff “Oligarch*” etwa erzeugte eine große Anzahl von “False Positives” und wurde daher mit der Einschränkung “NOT (ukrain* OR russ*)” versehen. Der Versuch, Signalwörter für “People Centrist” in den Search String zu verarbeiten, hat die aus der Theorie gewonnen Erkenntnisse bestätigt und das Ergebnis der Voruntersuchung negativ beeinflusst. Der hier verwendete Searchstring wurde möglichst

breit gefasst, um alle unterschiedlichen Arten von Populismus zu erfassen da es das Ziel der vorliegenden Untersuchung war Populismus in all seinen Ausprägungen messbar zu machen. Der verwendete Searchstring beinhaltete, in Anlehnung an Rooduijn und Pauwels (2011), folgende Wörter:

Gier* Grosskonzern* Imerialismus* Imperialistisch* Internationalismus* Kapitalisten* Lakai* Monopol* (Oligarch* NOT (ukrain* OR russ*)) Plutokratie* abgehoben* anti-basisdemokratisch* anti-demokratisch* antibasisdemokratisch* antidemokratisch* aritsokrat* aufhals* aufzwing* ausbeuter* autokrati* elite* elitär* eurokraten* eurokratie* geldadel* herrschend* internationalistisch* kooptier* korrupt* kumpanen* plünder* propagand* technokrat* ungewählt* unterjochen* „politi* kaste“~5 „die da oben“ "das volk" "als volk" "dem volk" („die bürger“ NOT (madrid kataloni* spani*)) „die wähler“

Der verwendete Search-String schloss in etwa die Hälfte der 59.000 Artikel aus, was insofern ein gutes Zeichen ist, da es bei so vielen Artikeln wahrscheinlich scheint, dass verschiedenste Ausprägungen von Populismus enthalten sind. Damit war die erste Vorauswahl der Artikel abgeschlossen.

Als nächster Schritt wurde ein Codebuch erarbeitet, mit dessen Hilfe Populismus auf Absatzebene kodiert werden sollte. Bei computerunterstützten, sowie rein manuellen Inhaltsanalysen von Texten, stellt sich die Frage nach der zu kodierenden Einheit. Üblicherweise wird entweder auf Satz-, Absatz-, oder gesamter Text-Ebene analysiert. Werden sehr lange Texte wie Partei-Manifestos untersucht (z.B. Rooduijn & Pauwels, 2011) wird üblicherweise auf Absatz-Ebene kodiert. Bei kürzeren Texten wie Zeitungsartikel kann auch der gesamte Artikel hergezogen werden (Heinesch, 2018). Es wurde entschieden, Absätze als Texteinheit in dieser Arbeit zu verwenden. Bei Sätzen besteht die Gefahr, dass nach Stemming und Entfernen der Stop-Wörter nicht mehr genug Features für den Algorithmus bereitstehen um ein Modell zu berechnen. Gegen die Articlebene wurde entschieden, da bei dieser sehr viele Wörter oder Textpassagen enthalten sind, die nichts mit der Klassifizierung von Populismus zu tun haben, was das Erstellen eines validen Modells für den SML Learner möglicherweise erschwert hätte. Außerdem liegt die Vermutung nahe, dass sich im Laufe eines Artikels, das Ziel und der Kommunikator von Statements und Aussagen verändert, je länger dieser ist was ebenfalls zu intervenierenden Effekten führt.

4.3.2. Operationalisierung und Codebuch

Autor: Maximilian Haselsteiner

Ein großes Problem bei der computergestützten Untersuchung ist, dass sich vor allem die Referenz an das Volk, welches eigentlich als die zentrale Säule des Populismus ist (Taggart 2000: 91), kaum durch “einfachere” computergestützte Methoden, zu erkennen ist. Dies liegt vor allem daran, dass die Referenz an das Volk oft durch einfache, mehrdeutige Wörter wie “wir”, “uns” dargestellt werden. Durch eine Methode wie dem Wörterbuch-Ansatz, bei dem Samples aufgrund der Nennung bestimmter Wörter kodiert werden, ist es kaum möglich zu unterscheiden, ob “wir” als “wir, das Volk”, oder aber als “wir, die Partei” oder zur Gänze anders konnotiert ist.

Einige Ansätze zur Erhebung von People Centrism wurden bereits im Kapitel 2.2.2. und 3.1. gezeigt. Etwa Rooduijn und Pauwels (2011) Wörterliste und Pauwels (2017) adaptierte Liste, welche sich jedoch wenig erfolgreich für den deutschsprachigen Raum gezeigt hat. Besser zeigen sich hier die Ansätze verschiedener Forscherteams für das Schweizer “National Center of Competence in Research” (NCCR). In einem Versuch Jager und Walgraves (2007) Auffassung von Populismus als politischen Kommunikationsstil mit der ideologischen Ansicht von Mudde (2004) zu versöhnen, wird dort “People-Centrism” in vier verschiedene “Key Messages” unterschieden:

		Communication	Attitude
Conflictive	Anti-Elitism	• Discrediting the Elite • Blaming the Elite • Excluding the Elite from the People	• Negative Attitude toward the Elite • Blaming the Elite
	Restoring Sovereignty	• Denying Power to the Elite	• Desire for More Power for the People at the Expense of the Elite
• Demanding Power for the People			
Advocative	People-Centrism	• Stressing Virtues of the People • Praising Achievements of the People • Stating a Monolithic People • Demonstrating Closeness to the People	• Positive Attitude toward the People • Perception of a Homogeneous / Monolithic People

Abbildung 9: “Parallel measurement of populist communication and attitude” aus Wirth et al. 2016: 52.

Diese Aufstellung von Populismus allgemein zeigt die drei Säulen nach Mudde (2004): Anti-Elitismus, People Centrism und “Restoring Sovereignty to the People”. Die Komponenten der Ideologie werden weiter auf verschiedene “Key Messages” heruntergebrochen. Noch detaillierter wird dieser Dekonstruktion gezeigt bei Ernst et al. (2017), ebenfalls Forscherin für das Projekt “Challenges to Democracy in the 21st Century” des Schweizer NCCR.

People-Centrism	Stressing the people's virtues	The people are virtuous.	The people are bestowed with morality, charisma, credibility, intelligence, competence, consistency, etc. The people are exempt from being malevolent, criminal, lazy, stupid, extremist, racist, undemocratic, etc.
	Praising the people's achievements	The people are beneficial.	The people are described as being an enrichment or responsible for a positive development/situation. The people are described as not being a threat/burden, not being responsible for negative developments/situations, nor as having committed mistakes or crimes.
	Stating a monolithic people	The people are homogenous.	People are described as sharing common feelings, desires, or opinions.
	Demonstrating closeness to the people	The populist represents the people.	The speaker describes himself as belonging to the people, being close to the people, knowing the people, speaking for the people, caring for the people, agreeing with the people, or performing everyday actions. The speaker claims to represent or embody the people.

Abbildung 10: : Ein Auszug aus “Conceptualization and Operationalization of Populist Communication” (Ernst et al. 2017: 258f)

Diese Operationalisierung zielt im Gegensatz zu den meisten anderen, auf die Erhebung von *medial vermittelter populistischer Kommunikation*. Die dünne Ideologie wird dabei auf insgesamt neun “Key Messages” (siehe Abb. 10) heruntergebrochen. Diese Key Messages wiederum zeigen auf die drei grundlegenden Elemente des Populismus in seiner Minimaldefinition.

Populismus mit einer hohen Reliabilität manuell zu kodieren ist keine triviale Aufgabe, wie zahlreiche Studien gezeigt haben (siehe Kapitel 3.1.). Der vielversprechendste Ansatz schien die die Arbeit des NCCR Democracy, einem Forschungsprojekt auf der Universität Zürich, basierend auf der theoretischen Arbeit von Wirth et. al. (2016), zu bieten. Mit Hilfe

von “semantic clauses” werden Aussagen nach strengen grammatikalischen Regeln extrahiert um die Reliabilität möglichst hoch zu halten. Dieses sehr umfangreiche Codebuch, das auf zahlreiche Facetten von Populismus sehr detailliert eingeht, wurde für die vorliegende Arbeit reduziert und adaptiert. Jene Teile die beschreiben wie People-Centrism, Anti-Elitismus und Sovereignty zu kodieren sind, wurden mehr oder weniger eins zu eins übernommen, jedoch stark reduziert. Das NCCR Codebuch weist eine enorme Anzahl an Kategorien und Unterkategorien auf, um Populismus und seine Elemente so weit wie möglich zu abstrahieren und damit die Interkoderreliabilität so hoch wie möglich zu halten. Aufgrund der hohen Ambitionen die sich das Projekt gestellt hat, musste dies auch gemacht werden; insgesamt wurden beinahe 55.000 Artikel von 87 trainierten Kodierern manuell bearbeitet. Mit einem Cohen’s Kappa von 0.597 (vgl. Wirth et al. 2017: 222) sind diese Ergebnisse zwar nicht berauschend, für eine solche Dimension allerdings mehr als akzeptabel. Die Kategorie Anti-Elitismus wird etwa durch 18 Variablen, siehe Abbildung 11, erhoben.

Category 47: Attribution of negative personality traits

Malevolence, viciousness, bad character (Att_Neg_malev)
 Criminal, corrupt, unlawful (Att_Neg_crim)
 Laziness, inactivity, inertia (Att_Neg_lazy)
 Stupidity, lack of understanding, being wrong (Att_Neg_stu)
 Extremism (Att_Neg_ext)
 Racism (Att_Neg_raci)
 Undemocratic (Att_Neg_unde)

Category 1 (48): Attribution of suspected impact and responsibility

Posing a Threat to the people, identity, nation, or society in general (Att_Impact_thre)
 Posing a burden (Att_Impact_burd)
 Posing an Enrichment (Att_Impact_enri)
 Accountability/Responsibility for negative development/situation (Att_Impact_aneg)
 Accountability/Responsibility for positive development/situation (Att_Impact_apos)
 Ability to solve a problem (Att_Impact_abil)

Category 3 (51): Attribution/Description of committed actions

Performing Everyday actions (Att_Act_every)
 Wrong, stupid action or mistake (Att_Act_mist)
 Immoral/Indecent Behavior (Att_Act_immo)
 Criminal or unlawful behavior (Att_Act_crim)
 Important, successful, right action (Att_Act_right)

Abbildung 11: Variabel- und Kategorieschema für Anti- Elitismus aus dem NCCR Codebuch

Da im Falle dieser Arbeit nur zwei Kodierer tätig waren und diese über ein großes Wissen über Populismus verfügen, wurden diese Kategorien auf die drei Kategorievariablen verkürzt, die kleineren Variablen wurden bei Unsicherheit oder Zweideutigkeit konsultiert. Da keinerlei statistische Auswertungen oder Aussagen mit den Ergebnissen der Inhaltsanalyse getätigt werden sollten, wurde sich außerdem die Freiheit genommen, bei komplizierten und mehrdeutigen Statements gemeinsam zu diskutieren. Weiteres wurden die Kategorien zu “Relation to the People” und “Demanding Sovereignty to the People” aus dem Rahmenwerk dieses Codebuchs deduziert. Ersteres wird durch das Attribuieren allgemein, sowie der Darstellung als Repräsentant (“embodiment”) verstanden, zweiteres durch die Attribuierung bzw. Zuschreibung von Macht beschrieben. Als wichtiges Grundelement des NCCR Codebuchs wurde auch die Herangehensweise, nur Statements zu kodieren, übernommen. Dabei werden nur jene Textteile kodiert, welche eindeutig ein Statement sind, d.h. es einen “Speaker” (Kommunikator) und ein “Target” (Ziel) gibt. Zur genaueren Erklärung folgen hier zwei Auszüge aus dem Codebuch, Wirth et al. 2017: Seite 92 bzw. 147:

Definition Speaker: A speaker is an actor which is quoted in the text either directly or indirectly to utter any statement on an issue covered in the article or another actor. For each speaker and its wholesome statement within the text, the content is coded in detail. This includes the Identity of the speaker, his position toward the main issue of the statement, the argumentation and sources provided. In addition to the content, the style of the statement of each speaker is assessed.

A speaker has to fulfil minimum requirements to be coded: First, a speaker has to be cited with a (at least partial) statement. Vague references to what an actor may have said are not valid statements. The statement of the speaker may be any interpretation or evaluation of an issue or an evaluation or description of any target actor.

Definition Target Actor: The utterance of a speaker in any text may contain one or several statements on target actors. Since populist communication strategies may become evident in the way a politician talks about actors such other politicians, supranational organizations, elites, or the people, these statements are crucial to the analysis of populist communication in the media.

A target therefore can be politicians, parties or coalitions, the people, any kind of elites, supranational institutions, foreign countries or other actors and organizations. Not to be coded as targets are mere mentioning of how persons/organizations are affected by something, mere mentioning of the feelings of persons/institutions or mere connection of a person/institution to an issue without indicating a position or elaboration.

Die aufwendige Unterscheidung und Festhaltung der verschiedenen Issues der Statements, sowie abschwächende Ko-Variablen wie Mediatisierung und Dirty Campaigning wurde in unserer Version aus forschungsökonomischen Gründen vollkommen ausgelassen. Auch bei den formalen Kategorien wurden die meisten ausgespart. Hier wurden zwar Elemente wie “Medium” behalten, wirklich essentiell waren für die Erstellung des Trainingssets allerdings nur die Artikel ID für die Reproduzierbarkeit, die dazugehörige Absatz ID sowie die Anzahl der Absätze für die Berechnung von Interkoderreliabilität. Weitere von AmCAT vorgeschlagene formale Kategorien (z.B. Ressort, Datum etc.) wurden zwar erfasst, für die Weiterverwendung jedoch verworfen. Die für diese Arbeit bedeutenden Auszüge des NCCR Codebuchs sind am Anhang zu finden, das vollständige Codebuch ist auf der Website der NCCR zu finden.

4.3.3. Durchführung und Nachbearbeitung

Autor: Gregor Rettenegger

Nachdem ein Pretest mit 100 Artikeln durchgeführt und das verwendete Codebuch weiter angepasst wurde, wurde anhand von 73 Artikeln bzw. 969 Absätzen die Interkoderreliabilität der zwei Koder ermittelt. Auf Artikelebene wurde ein solider Krippendorffs Alpha Wert von .80 erreicht, auf Absatzebene von .68, was bei einem derart seltenen Phänomen ebenfalls als guter Wert angesehen werden kann. Nachdem eine zufriedenstellende Interkoderreliabilität sichergestellt war, begannen zwei Kodierer zufällig Artikel, aus dem vorselektierten Set, auf Absatz Ebene zu kodieren. Ziel war es 200 Absätze die populistische Kommunikation nach der beschriebenen Definition enthielten zu generieren um damit ein erstes Trainingsset von 200 positiven und 200 negativen Absätzen zu erstellen. Burscher et. al. (2016, siehe ebenfalls Abbildung 4, Seite 34) haben in ihrer bereits erwähnten Studie gezeigt, dass bereits mit 200 Trainingsdokumenten Ergebnisse mit hoher Accuracy erreicht werden können, weshalb 400 eine angemessene Zahl zu sein scheint. Aus dem vorselektierten Set wurden insgesamt 644 Artikel kodiert bis 200 Absätze gefunden wurden die populistische Kommunikation enthielten. Abbildung A1 zeigt die verwendeten Artikel nach Medium.

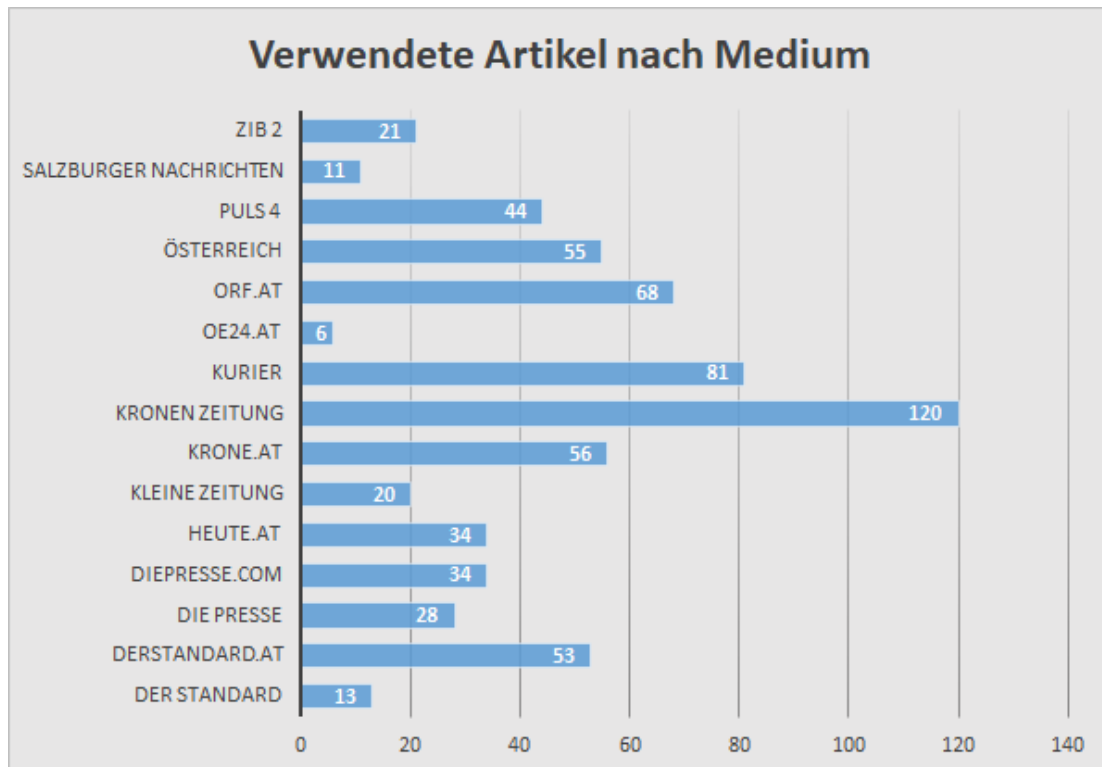


Abbildung 12: Verwendete Artikel nach Medium, Erstellen des Trainingssets

Nachdem genügend Absätze gefunden waren, wurden einige gängige automatisierte Verfahren zur Text Vorbereitung durchgeführt. So wurde mit Hilfe des R-Packages “Quanteda” Word-Stemming durchgeführt, ein Verfahren, bei dem Wörter auf ihren Wortstamm reduziert werden bzw. verschiedene Wortendungen entfernt werden. Aus “Bürger”, “Bürgern” oder “Bürgerinnen” wird “Burg” womit sichergestellt ist, dass der Learner all diese Formen als das an sich selbe Wort erkennt. Weiteres wurden deutsche Bindewörter wie z.B. “der”, “die”, “das”, “und”, “dann” usw. entfernt. Hier sei noch kurz erwähnt, dass jegliche Stop-Wörter, also die gesammelten entfernten Wörter in ihrer Gesamtheit (siehe weiter unten), in Bi-Gramms weiter erhalten blieben, so war beispielsweise “den_Menschen” weiter ein relevantes Feature des finalen Modells. In späteren Tests wurde bestätigt, dass sich diese beiden Schritte positiv auf die Validität des Modells auswirken. Schließlich wurden im Zuge der verschiedenen Iterationen kontextspezifische Wörter ausgeschlossen, da diesen eine zu hohe Gewichtung im Modell verliehen wurde. So wurden unter anderem Partei und Politikernamen vorab entfernt, und andere nachvollziehbare Fehler des Modells so korrigiert. Die vollständige Liste an kontext-spezifischen Wörtern die ausgeschlossen wurden, ist folgende:

"heinz-christian", "hc", "heinz", "christian", "strache", "hc_strache", "christian_strache",
 "heinz-christian_strache", "strache_heinz", "strache_heinz-christian", "strache_hc",
 "c_strache", "strache_h", "peter", "pilz", "peter_pilz", "pilz_peter", "kern",
 "christian_kern", "kern_christian", "sebastian", "kurz", "sebastian_kurz", "kurz_sebastian",
 "ulrike", "lunacek", "ulrike_lunacek", "lunacek_ulrike", "matthias", "strolz",
 "matthias_strolz", "strolz_matthias", "matthias_neos", "neos_das", "corinna", "milborn",
 "corinna_milborn", "milborn_corinna", "armin", "wolf", "armin_wolf",
 "wolf_armin", "dass", "kickl", "spö", "spo", "spoe",
 "sozialdemokrat*", "die_spo", "der_spo", "fpö", "fpo", "fpoe", "freiheitlich*", "die_fpo",
 "övp", "oevp", "ovp", "liste kurz", "liste_kurz", "neue volkspartei", "neue_volkspartei",
 "gruene", "die gruenen", "die_gruenen", "gruen", "gruenen", "grüne", "die grünen",
 "die_grünen", "grün", "grünen", "neos", "liste pilz", "liste_pilz", "behinderung_menschen",
 "behinderte_menschen", "beeinträchtigte_menschen", "beeinträchtigung_menschen",
 "blinde_menschen", "bürgermeister"

Mit dem so erarbeiteten Set wurden in Folge verschiedenen SML Algorithmen trainiert und dabei die In-Set Accuracy verglichen. Mithilfe des Interface „R-Script“, einer freien Programmiersprache bzw. einem Programmierinterface wurde ein Skript geschrieben das zum ersten aus einer gewissen Anzahl von Artikel lernt, zweitens das gelernte auf einen anderen Korpus anwendet und drittens die Ergebnisse aufbereitet und als .CSV Datei auswirft. Das hier verwendete Skript findet sich im Appendix dieser Arbeit, ebenso wie das von uns verwendete Trainingsset um die Ergebnisse sofort reproduzierbar zu machen. Wie bereits erwähnt wurde aus mehreren Gründen die Programmbibliothek XGBoost verwendet, siehe dazu Kapitel 3.2.3.

5. Ergebnisse

Autor: Gregor Rettenegger

Um die erste Forschungsfrage: “Kann Populismus in der politischen Berichterstattung mit Hilfe von Machine Learning quantifiziert werden?” zu untersuchen, wurde mit anfänglich 400 Absätzen ein Learner trainiert. Für die Absätze in diesem ersten Trainingsset lag die Interkoderreliabilität bei einem Krippendorff’s Alpha Wert von .68 ($n = 969$). Dieser erste Learner wurde verwendet um die Absätze von 100 neuen Artikeln zu klassifizieren. Um die Qualität dieser Klassifizierung festzustellen, wurden alle Absätze überprüft für die eine Wahrscheinlichkeit von $>50\%$ Populismus zu enthalten ausgegeben wurde. Diese Absätze wurden in Folge manuell mit “0 - keine populistische Kommunikation” und “1 - populistische Kommunikation” kodiert und dem Trainingsset hinzugefügt. Mit dem damit erweiterten Trainingsset wurde dieser Schritt wiederholt solange bis bei der manuellen Überprüfung keine Verbesserung mehr feststellbar war. Insgesamt wurden auf diese Weise sechs Iterationen durchgeführt. Für die finale Validierung wurden weitere 900 Absätze manuell kodiert und die Learner der unterschiedlichen Iterationsschritte in Folge anhand dieser Absätze überprüft. Um die Anzahl an Positives in diesen Validierungsabsätzen zu erhöhen wurden 10000 Artikeln, die bisher nicht für das Trainieren des Learners verwendet wurden, zufällig ausgewählt. Aus diesen wurden in einem weiteren Schritt 100 Artikeln ausgewählt, die vom Algorithmus als am wahrscheinlichsten klassifiziert wurden Populistische Absätze zu enthalten. Aus diesen 100 Artikeln wurden dann wiederum zufällig 400 Absätze entnommen die von zwei Kodern manuell klassifiziert wurden. Absätzen die schwer zu bestimmen waren wurden von beiden Kodern diskutiert bevor sie final klassifiziert wurden.

Die Ergebnisse für die einzelnen Learner sowie die Anzahl an Positives und Negatives im jeweiligen Trainingsset sind in den Abbildungen 13 und 14 zusammengefasst. Abb. 13 zeigt die Ergebnisse, wenn alle Absätze die eine Wahrscheinlichkeit von größer $>50\%$ aufweisen als populistische Kommunikation gewertet werden, Abbildung 14 zeigt die Ergebnisse, wenn bereits Absätze mit einer Wahrscheinlichkeit von $>25\%$ als populistisch angenommen werden.

	<i>Iteration 1</i>	<i>Iteration 2</i>	<i>Iteration 3</i>	<i>Iteration 4</i>	<i>Iteration 5</i>	<i>Iteration 6</i>
True Positives	24	24	20	16	18	15
True Negatives	711	764	783	796	813	811
False Positives	154	101	82	69	52	54
False Negatives	11	11	15	19	17	20
Precision	.13	.19	.20	.19	.26	.22
Recall	.69	.69	.57	.46	.51	.43
F-Score	.23	.30	.29	.27	.34	.29
Selectivity	.98	.99	.98	.98	.98	.98
Accuracy	.82	.88	.89	.90	.92	.92
Youden's Index	.67	.67	.55	.43	.49	.40
N im Trainingsset	208	326	387	442	459	492
P im Trainingsset	192	219	231	246	254	275

Abbildung 13: Populismus ab Wahrscheinlichkeit >50%, N/P im Trainingsset = Anzahl an Positives/Negatives

	<i>Iteration 1</i>	<i>Iteration 2</i>	<i>Iteration 3</i>	<i>Iteration 4</i>	<i>Iteration 5</i>	<i>Iteration 6</i>
True Positives	30	29	23	19	22	20
True Negatives	630	716	753	765	777	773
False Positives	235	149	112	100	88	92
False Negatives	5	6	12	16	13	15

Precision	.11	.16	.17	.16	.20	.18
Recall	.86	.83	.66	.54	.63	.57
F-Score	.20	.27	.27	.25	.30	.27
Selectivity	.99	.99	.98	.98	.98	.98
Accuracy	.73	.83	.86	.87	.89	.88
Youden's Index	.85	.82	.64	.52	.61	.55
N im Trainingsset	208	326	387	442	459	492
P im Trainingsset	192	219	231	246	254	275

Abbildung 14: Populismus ab Wahrscheinlichkeit >25%, N/P im Trainingsset = Anzahl an Positives/Negatives

Abbildung 15 zeigt die Entwicklung von Precision und Recall über die sechs Iterationen hinweg, wenn Absätze ab einer Wahrscheinlichkeit >.50 als populistisch gewertet werden

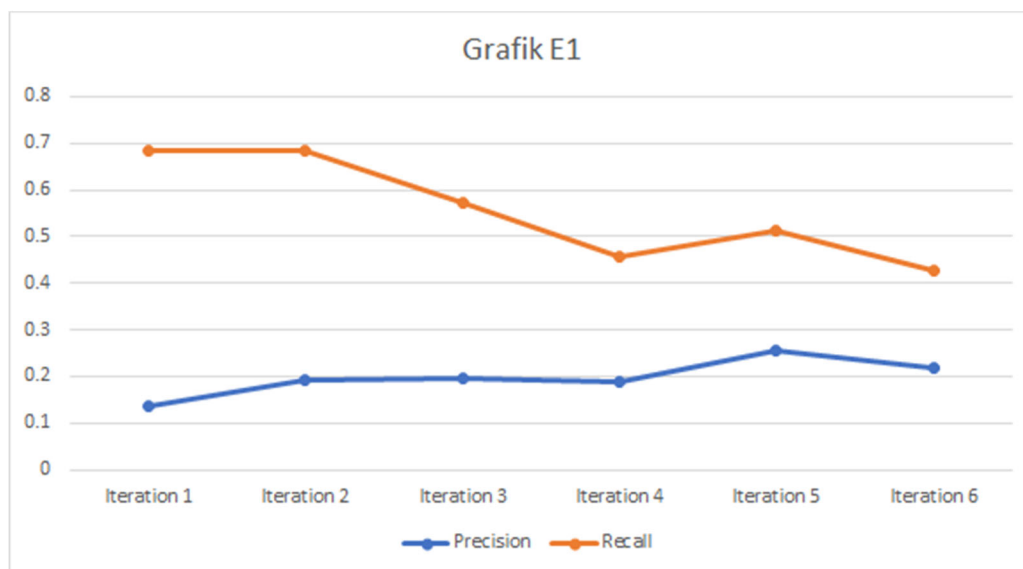


Abbildung 15: Entwicklung von Precision und Recall über sechs Iterationen

Um die zweite Forschungsfrage: “Welche Faktoren beeinflussen die Qualität eines SML Trainingssets?” zu beantworten wurden verschiedene Experimente durchgeführt.

Im bereits beschriebenen iterativen Prozess, wurde das Trainingsset stetig erweitert dabei überwiegend False Positives und True Positives hinzugefügt, da nur Absätze mit einer hohen Wahrscheinlichkeit berücksichtigt wurden. Wie Grafik E1 zeigt sank mit fast jeder Iteration der Recall des Learners von anfänglich .69 auf .43 während die Precision von .13 auf .22 im Vergleich deutlich weniger anstieg.

Um die Gründe für den Abfall des Recalls über die Iterationen hinweg zu untersuchen wurde das Trainingsset auf verschieden Arten manipuliert und die Ergebnisse mit der letzten Iteration verglichen. Hierzu wurden

- in Experiment 1 dem Trainingsset zusätzliche zufällige Negatives hinzugefügt,
- in Experiment 2 alle False Positives, sowie True Negatives aus den Iterationen aus dem Trainingsset entfernt, wobei die Anzahl an Positives und Negatives in etwa gleich hoch gehalten wurde,
- in Experiment 3 alle False Positives aus dem Trainingsset entfernt, jedoch alle True Negatives behalten,
- in Experiment 4 wurde das Set aus Experiment 3 zusätzlich um zufällige Negatives erweitert,
- in Experiment 5 wurde das Set aus Experiment 3 nur um so viele zufällige Negatives erweitert, dass genau so viele Positives wie Negatives im Set enthalten waren.

Die Ergebnisse dieser fünf Experimente, wenn Absätze ab einer Wahrscheinlichkeit $>.50$ als populistisch gewertet werden, sind in Abbildung 16 zusammengefasst.

	<i>Experiment 1</i>	<i>Experiment 2</i>	<i>Experiment 3</i>	<i>Experiment 4</i>	<i>Experiment 5</i>
True Positives	20	24	26	23	29
True Negatives	775	717	668	742	673
False Positives	90	148	197	123	192
False Negatives	15	11	9	12	6
Precision	.18	.14	.12	.16	.13
Recall	.57	.69	.74	.66	.83

F-Score	.28	.23	.20	.25	.23
Selectivity	.98	.98	.99	.98	.99
Accuracy	.88	.82	.77	.85	.78
Youden's Index	.55	.67	.73	.64	.82
N im Trainingsset	591	280	248	528	275
P im Trainingsset	275	275	275	275	275

Abbildung 16: Verschiedene Experimente in den Iterationen

Der finale Learner aus Experiment 5 erreichte bei einem Eta von 0.3 und einer maximalen Tiefe von 3, einen Recall Wert von .83, bei einer Precision von .13, eine Accuracy von .78 und eine Selectivity von .99.

Die wichtigsten Wörter und Bi-Gramms des Modells waren:

bevolker, polit, mensch, burg, osterreich, mehr, lass, wer, europa, geht, sei, parlament, dass_die, partei, fur_die, herr, sieht, in_der, volk, nicht_mehr, gar, arbeit, in_ein, an_der, steuerzahl, leut, stimm, mehrheit, imm, muss, wahlerin, problem, die_osterreich, gross, eigent, neu, kind, bei_den, vertret, sagt, wahl, jahr, regier, dabei, wien, wort, demokrati, ist_das, meist, die_mensch, von_ein, ist_nicht, die_spo, reich, geg_die, auf_die, macht, rot, an_die, nach_ein, uber_die, los, deshalb, halt, der_spo, koalition, schon, end, vor_all, herrn, vorschlag, die_fpo, drei, steht, und_die, bei_der, welt, einzig, fur_das, leb, richtig, kommend, nach, in_osterreich, offent, sag, stell, das_ist, ganz, nach_dem, der_polit, der_mensch, viel, ja, geh, wahlkampf, gut, thema, weiss,...

Abbildung 17 stellt die Gewichtung der wichtigsten Features, welche am häufigsten in Entscheidungsbäumen vorkommen, grafisch dar. Es zeigt sich wie vermutet, dass Wörter die mit People-Centrism konnotiert sind, an der Spitze stehen. Mögliche Gründe dafür werden weiter unten diskutiert.

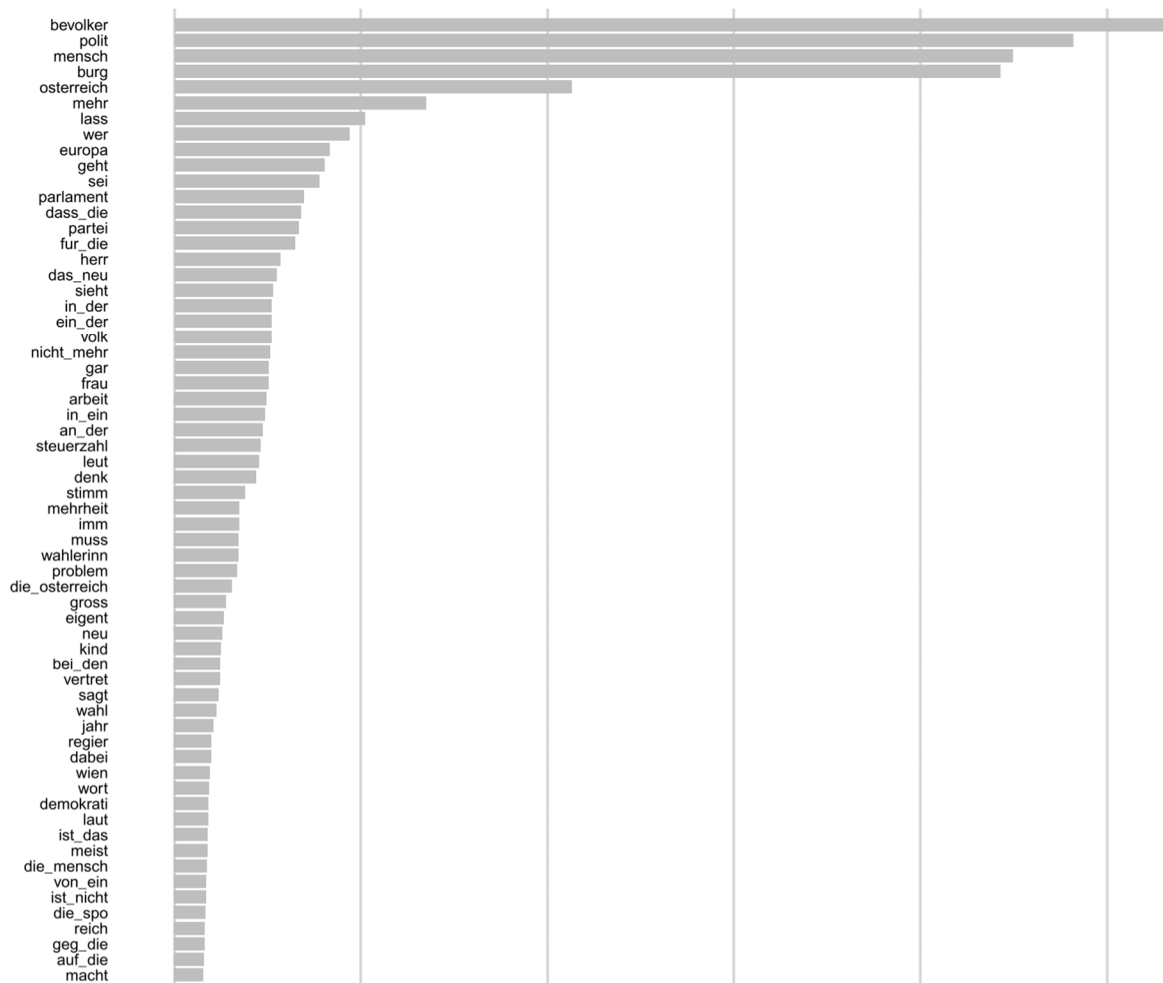


Abbildung 17: Darstellung der wichtigsten "Features" in den Entscheidungsbäumen

Aus diesen Wörtern bildete das Modell 1000 Decision-Trees, die wichtigsten zwei sind exemplarisch in den Abbildungen 18 und 19 dargestellt.

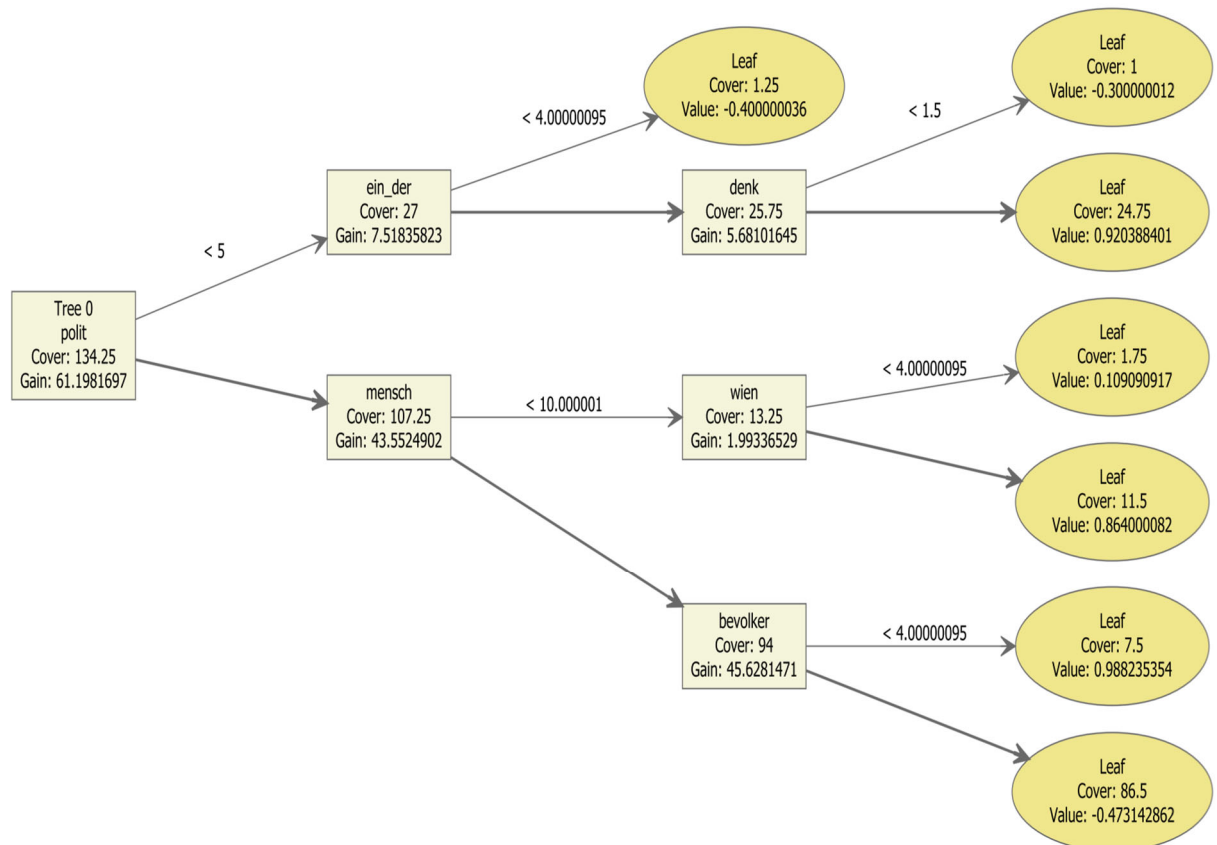


Abbildung 18: Exemplarischer Descision Tree I

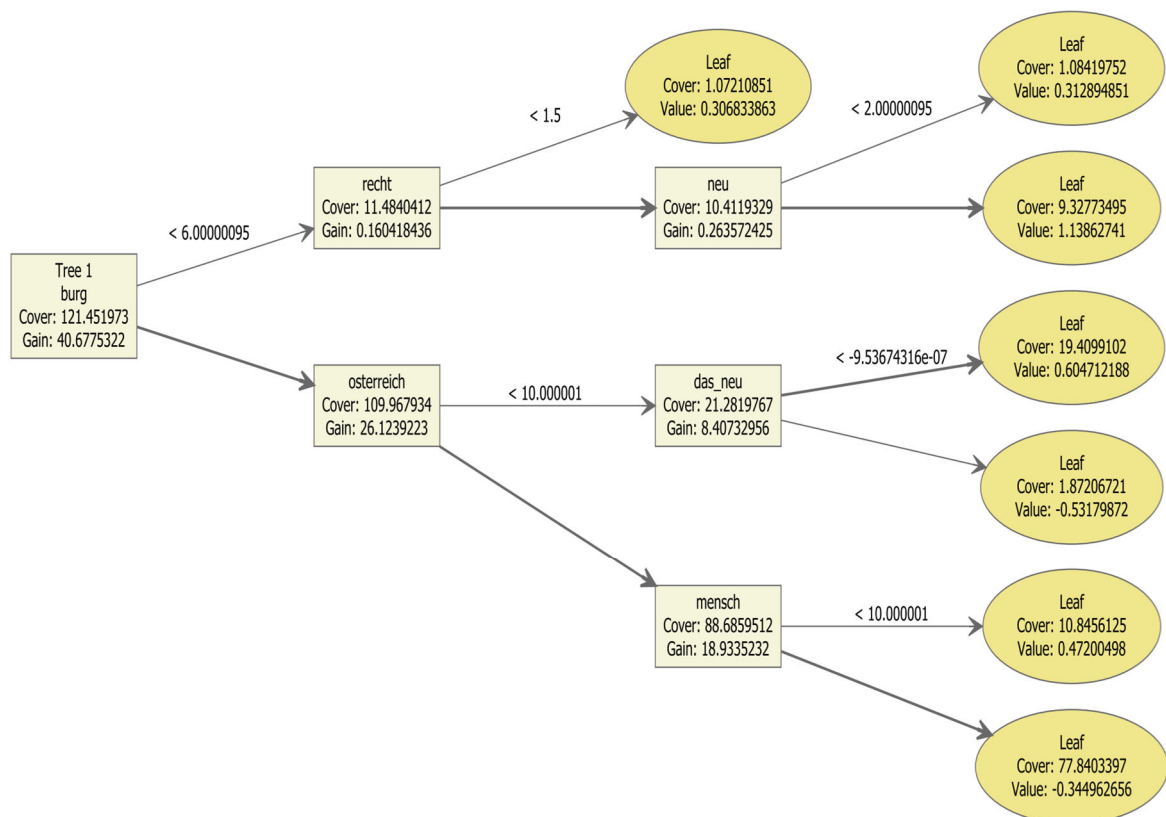


Abbildung 19: Exemplarischer Descision Tree II

6. Ergebnis Diskussion

6.1. Faktoren für die Qualität eines Trainingssets

Autor: Gregor Rettenegger

Wie sind die Ergebnisse aus den Iteration sowie des finalen Learners nun zu deuten? Grundsätzlich lässt sich feststellen, dass sich der hier vorgestellte iterative Prozess durchaus als hilfreich erwiesen hat in relativ kurzer Zeit zusätzliche Positives für das Trainingsset zu generieren. Während für die ersten 200 positiven Absätze, über 600 Artikel kodiert werden mussten, die bereits mit einer Wörterbuch-basierten Methode vorausgewählt waren, wurden die nächsten 75 positiven Absätze deutlich schneller gefunden. Hierzu wurden aus 6 Iteration im Durchschnitt 50 Absätze kodiert, insgesamt also in etwa 300 Absätze. Geht man davon aus, dass ein Artikel aus ca. 6 Absätzen besteht, konnte die Anzahl an Absätzen die manuell kodiert werden müssen um einen zusätzlich positiven zu finden, von 18 auf 4 reduziert werden. Diese stellt eine erhebliche Erleichterung und Kostenreduktion für die manuelle Kodierung dar.

Will man die in den Iterationen gewonnen Dokumente in das Trainingsset einfließen lassen gilt es jedoch einige Punkte zu beachten. Wie Abbildung 15 zeigt, brach der Recall Wert von anfänglich .69 auf .43 massiv ein. Unsere Vermutung war, dass dieser Rückgang eine Folge von False Positives welche ins Set aufgenommen worden waren, wurde in Experiment 2 - 4 bestätigt. Sobald die False Positives aus den Iteration aus dem Trainingsset entfernt wurden erhöhte sich der Recall von .43 (Iteration 6) auf .69 (Experiment 2) um 26%. Das heißt aus den Iterationen gewonnen False Positives - und mit der hier vorgestellten Methode sind das bei einer geringen Precision ein erheblicher Anteil der Dokumente, da nur die wahrscheinlichsten überprüft werden - sollten nicht in das Trainingsset aufgenommen werden. Der Einbruch des Recall-Wertes wenn das Trainingsset um False Positives erweitert wird, ist durchaus nachvollziehbar. Der Learner erstellt in seinem Modell Regeln, mit denen Ausprägungen von populistischer Kommunikation als solche klassifiziert werden; in jeder Iteration werden diese Regeln nun überprüft. Ist nur jeder fünfte Fall der positiv klassifiziert wurde tatsächlich ein Positive und werden alle anderen Dokumente dem Trainingsset nun als Negatives hinzugefügt, muss der Learner zwangsweise viele der bis dahin gewonnenen Regeln revidieren, wodurch es schwieriger wird tatsächliche Negatives korrekt zu klassifizieren. Experiment 2 zeigt das

die Aufnahme von True Negatives ins Trainingsset keine negativen Auswirkungen auf den Recall hat, was die hier angestellten Überlegungen bestätigt.

Grundsätzlich scheint, wie besonders Experiment 1 und 4 belegen, eine hohe Anzahl an Negatives die Precision zu steigern und den Recall zu reduzieren. Vergleicht man jedoch die Recall Werte von Experiment 1 und 4 zeigt sich relativ eindeutig die negative Auswirkung von False Positives im Trainingsset. Anders sind die beinahe 10% Unterschied des Recall Wertes zwischen .57 und .66 nicht zu erklären, da das relativ überschaubare Mehr an Negatives im Trainingsset von Experiment 1 keinen so starken Unterschied machen sollten.

Ein weiterer Faktor der bei der iterativen Erstellung eines Trainingssets zu beachten ist scheint ein ausgewogenes Verhältnis zwischen Positives und Negatives zu sein. Die mit Abstand höchste Recall Wert wurde mit .83 in Experiment 5 erreicht, bei dem die Anzahl von Negatives und Positives im Set exakt gleich hoch waren, gefolgt von Experiment 3 mit .74 . Experiment 5 fand mit 29 von insgesamt 35 tatsächlichen Fällen von populistischer Kommunikation, die meisten True Positives in den Validierungs-Dokumenten und erreichte dadurch die höchsten Recall, Selectivity und Youden's Index Werte. Jedoch wurden mit 192 auf vergleichsweise viele Dokumente fälschlicherweise als Positives klassifiziert, wodurch sich der im Vergleich mit den anderen Experimenten eher niedrige Precision, Accuracy und F-Score Wert ergibt.

Das Beste Modell ist unter bei diesen Werten nicht eindeutig bestimmbar und hängt vom jeweiligen Einsatzgebiet ab. Sollen möglichst alle Fälle von Populismus gefunden werden gilt es auf einen hohen Recall Wert zu achten, sollen möglichst alle Dokument die keinen Populismus beinhalten ausgeschlossen werden wird die Precision umso wichtiger. Ein Vorteil der hier vorgestellten Methode ist die prozentuale Skala mit der populistische Kommunikation klassifiziert wird. damit haben es die Forscher in der Hand Dokumente ab 1% bis 100% Wahrscheinlichkeit als Populismus zu werten. Wertet man Dokumente ab .25 als Populismus, erreicht der Learner aus Experiment 5 einen Recall Wert von .94 bei einer immer noch relative soliden .11. Werden nur Dokumente mit einer Wahrscheinlichkeit von 100% angenommen liegt die Precision jedoch weiterhin nur bei .26, bei einem Recall von .33. Forscher haben hier aber eine hilfreiche Stellschraube zur Hand, auch wenn der maximale Precision Wert weiterhin nicht ausreichend ist. Wie Abbildung 17 zeigt, stieg die Precision über die Iterationen und mit der Größe des Trainingssets relativ konstant an, was vermuten lässt, dass mit einer größeren Anzahl an Dokumenten hier bessere Werte

möglich wären. Weitere Vorschläge wie die hier vorgestellte Methode verbessert werden könnte folgen im nächsten Kapitel 7: Conclusio.

6.2 Beispiele für False Positives und False Negatives

Autor: Maximilian Haselsteiner

Es konnte also gezeigt werden, dass vor allem die Precision durch die Iterationen konstant gestiegen ist, bei relativ konstant hohen Recall. Dennoch ist eine Precision von $\sim 0,2$ auf den ersten Blick nicht beeindruckend. Zwar relativieren sich diese Werte im Vergleich zu ähnlichen Ansätzen, mehr dazu findet sich allerdings in Kapitel 7: Conclusio. Der Grund für die geringe Precision ist die hohe Anzahl an False Positives. Es gibt mehrere Gründe wieso Artikel als Positiv angezeigt werden, ohne tatsächlich Populistisch zu sein. Ein Problem etwa herrscht bei sehr kurzen Absätzen mit mehreren wichtigen Signalwörtern oder sehr langen Artikeln mit vielen Wörterketten (im Sinne der Entscheidungsbäume) die negative Ladungen haben. Dieses Problem würde wohl durch weitere Iterationen immer marginaler werden, einige Probleme werden jedoch wahrscheinlich nie gänzlich vermeidbar sein. Ein evidentes Beispiel dafür sind selbstkritische Aussagen:

[...] Soll man es mit fünf Hausärzten bespielen? Ich kann mir nicht vorstellen, wie das funktionieren soll", so der Volkspartei-Ortschef Bernhard, der sich übergangen fühlt. „Es ist schon arg, wie man über den größten Bezirk Österreichs drüberfährt. Und die Bevölkerung glaubt uns als Kommunalpolitikern sowieso nichts mehr, egal was wir sagen."

Hier erkennt der Algorithmus richtig das Politikern die Nähe zum Volk versagt wird, was ursprünglich ein valider Indikator für Populismus ist, da jedoch der Kommunikator über sich selbst spricht, kann man in diesem Fall nicht von einer Kritik an irgendeiner Elite außer ihm selbst gelten. Der Algorithmus ist also nicht in der Lage, herauszufinden ob der Kommunikator sich selbst oder jemanden anderen mein, wahrscheinlich ist dies ein Vorgang, der nur von menschlichen Kodierern korrekt erfasst werden kann. Weitere False Positives werden produziert, wenn jemand im Namen einer österreichischen Gesellschaft oder einem Verein spricht.

Peter Lammerhuber, CEO der Medienplaners GroupM, sieht den österreichischen Online-Werbemarkt durch Programmatic Advertising gefährdet - und Österreichs private Medien durch die E-Privacy-Richtlinie. ORF-Generaldirektor Alexander

Wrabetz verwies auf sein jüngst präsentiertes Positionspapier und trat neuerlich dafür ein, bei den ORF-Gebühren die "Streaminglücke" zu schließen. Ernst Swoboda (Verband Österreichischer Privatsender) wünscht sich eine Politik für einen "starken Medienmarkt und Medienstandort insgesamt".

Bei diesem False Positive schlägt der Algorithmus aus zwei Gründen an, erstens, weil der **österreichische** Werbemarkt **gefährdet** ist. Zweitens glaubt der Algorithmus das der Kommunikator im Namen "der Österreicher" (= embodiment of the people) spricht. Wegen des Stemming wird der Begriff "Österreichischer" als das Feature "österreich*" gesehen, während er das Verhältnis mit "Verband" und "Privatsender" ohne Word Embedding (mehr dazu im Kapitel 7) nicht nachvollziehen kann. In dieselbe Kerbe schlägt die Sportberichterstattung; hier werden Medaillen "für Österreich" gewonnen - oder aber Absätze im Zusammenhang mit der Zeitung "ÖSTERREICH" von Wolfgang Fellner. Allgemein werden Entscheidungsbäumen mit Signalwörtern zum Thema People-Centrism teilweise ein etwas zu hohes η^9 zugeordnet. Dies liegt daran, dass öfter Anti-Elitismus vorkommt ohne dass der Absatz holistisch betrachtet populistisch ist, zweitens ist die Vielfalt der Features, welche auf People-Centrism oder gar „Sovereignty to the people“ scoren, signifikant geringer (siehe Kapitel 6.1.), veranschaulicht kann dies an folgendem Beispiel werden:

„Auch an der Stärkung der direkten Demokratie wird man, nachdem man sie angekündigt hat, nicht mehr vorbeigehen können“, erklärte Holzinger. Er habe hier auch nicht diese Befürchtungen, die andere hegen. Zuletzt hatte Verwaltungsgerichtshof-Präsident Clemens Jabloner im „Presse“-Interview schwere Bedenken gegen mehr direkte Demokratie geäußert.“

Es kommt also durchaus vor, dass ein Artikel der nur ein populistisches Element enthält, von dem Instrument einen hohen „Populismus“ Wert hat. In diesem Fall orientiert sich der Algorithmus an der „Stärkung + direkten_Demokratie“ und „gegen + direkte_Demokratie“. Der Algorithmus errechnet hier einen Wert von 0.28 (nach Version 5, siehe Experiment 5, Kapitel 6.1.). Verwendet man die für Voruntersuchungen empfohlene Schwelle von 0.25 wird dieser Artikel als Populistisch gewertet. Wäre der Kommunikator der gleiche, könnte man hier auf jeden Fall von Populismus sprechen, weil die „Bedenken gegen direkte

⁹ Lässt sich am besten wohl mit Trainingsstärke oder Klassifikator-Stärke beschreiben.

Demokratie“ welche einen Verwaltungsgerichtshof-Präsidenten zugeschrieben wird, auch als Anti-Elitismus aufgefasst werden kann. So sind es jedoch zwei getrennte Aussagen bzw. Statements zur direkten Demokratie. Der Grad ist jedoch eher schmal und auch für manuelle Kodierer wäre dieser Absatz nicht klar zu deuten. Außerdem würde dieses Problem durch einen nicht-holistischen Ansatz wohl beseitigt werden, mehr dazu findet sich im Conclusio (Kapitel 7). Weitere Faktoren, welche False-Positives produzieren, sind Artikel allgemein über Populismus und wenn sehr pauschal über “die Österreicher” gesprochen wird:

“Die Österreicher sind freundlich und gastfreundlich, auch uns Piefkes gegenüber [...] Für uns ist Österreich irgendetwas, das an Deutschland dranhängt - historisch, politisch, wirtschaftlich und kulturell [...]”

Teils finden sich außerdem kurze Politikerbiografien oder historische Abrisse in dem False Positives. Viele ähnliche Fälle wurden durch die Einführung verschiedener Stoppwords zwischen den Iterationen bereits unterbunden. Die eher hohe False Positives Rate verbessert sich außerdem vor allem durch die Vergrößerung des Trainingssets, wie weiter oben bereits gezeigt wurde. Ein gewisser Rest wird beim Erkennen eines solche komplexen Konstrukt wie Populismus höchstwahrscheinlich immer bleiben. Es ist also auf jeden Fall zu empfehlen, die mit dem Instrument gefundenen Artikel manuell nachzukodieren, über die praktische Anwendungsgebiete wird im Conclusio (Kapitel 7) diskutiert. Was jedoch sehr viel problematischer als False Positives sind, sind die False Negatives, denn diese Artikel können eben nicht problemlos nachkodiert werden. Hierzu müsste man alle Artikel, auch die negatives, manuell nachkontrollieren. Deshalb wurde in dieser Arbeit auch schon zum wiederholten Mal die Wichtigkeit eines hohen Recall- Wertes betont. Denn bei einem Phänomen, welches so selten vorkommt kann man es sich nicht leisten einen bedeutenden Teil der tatsächlich populistischen Absätze zu verlieren. Man könnte mit einem geringen Recall Wert auch niemals statistische Aussagen tätigen. Es gibt mehrere erklärbare Faktoren welche False Negatives produzieren, allen voran die Verwendungen von Analogien oder Metaphern wie hier:

“In Richtung FPÖ richtete Kern aus: „Bevor die FPÖ wirklich die Interessen des kleinen Mannes vertritt, wird Mick Jagger in Spielberg den Erzherzog-Johann-Jodler anstimmen.”

Ähnlich verhält es sich mit sprachlichen Stilmitteln wie Satire oder Parodie. Leider kann hier nicht von einer realistischen Lösung ausgegangen werden. Schwierig wird es auch bei solchen Artikeln:

“Die FPÖ hat 17 Prozent selbstständige Kandidaten, bei den Grünen sind es 13, Schlusslicht ist die SPÖ mit sieben. Der Wiener Notar Harald Stefan, der seit zwölf Jahren die Hälfte seiner Arbeitszeit der FPÖ-Landes-, später der Bundespolitik widmete, hält eine repräsentative Vertretung für wichtig: "Man versteht die Sorgen besser, man weiß, was Lohnnebenkosten sind.”

Diese sind selbst für menschliche Kodierer sehr schwer zu fassende Konstrukte. Man kann “hält eine repräsentative Vertretung für wichtig” etwa sowohl verstehen als, „dass alle Menschen vertreten sein sollten“ (=“Sovereignty to the People”) oder aber es wird im Vergleich zu direkter Demokratie gestellt, wobei es sich eher um ein gegenteiliges Sentiment handeln würde. Auch “wenn man weiß, was Lohnnebenkosten sind” kann verstanden werden als jemand, der “die Sorgen der Menschen kennt”, aber auch als jemand, der als selbständiger Unternehmer Teil der Elite ist. Jene Absätze sind also sowohl für menschliche als auch für automatische Verfahren also schwer zu fassen. In dieser Arbeit wurde der Artikel in der manuellen Kodierung unterschiedlich und nach längerer Diskussion positiv gewertet, der automatisierte gibt dem Artikel einen sehr geringen Wert. In diesem Fall stellt sich sogar die Frage ob der Algorithmus den manuellen Kodierern nicht sogar überlegen war. Es darf hier jedoch auch nicht verschwiegen werden, dass sich viele False Positives nicht erklären lassen.

7. Conclusio

Autor: Maximilian Haselsteiner

Das in dieser Arbeit erstellte Instrument ist bei weitem noch kein perfektes Werkzeug zur automatisierten Messung von Populismus in Medientexten, die Frage “kann Populismus in der politischen Berichterstattung mit Hilfe von (Supervised) Machine Learning quantifiziert werden?” kann jedoch mit einem vorsichtigen Ja beantwortet werden. Es gibt in seiner jetzigen Form schon mehrere praktische Anwendungen für das Instrument.

Erstens, wie bereits erwähnt, kann man den Algorithmen der Programmbibliothek XGBoost lehren, einen Unterschied zwischen populistischen und nicht- populistischen Textkörpern zu erkennen. Zweitens, es gibt die Möglichkeit, durch das Ausführen des R-Scripts (siehe Anhang) mit sehr hoher Wahrscheinlichkeit einen Populistischen Artikel zu finden. Sucht man also für qualitative Arbeiten beliebige populistische Artikel, kann man sich aus einem Set mit hoher Wahrscheinlichkeit solche auswerfen lassen (= Precision-Maximierung). Setzt man den Prediction-Classifier sehr hoch ($x = 0.99$) fand sich aus unseren Validation Set mit 900 Absätzen 26 Absätze von welchen 10 tatsächlich populistisch sind. Zwar kann man hier immer noch nicht von einer wirklich hohen Precision sprechen, man muss jedoch bedenken, dass es sich um ein Phänomen handelt, das nur etwa in 2-4% aller Absätze vorkommt¹⁰. Allerdings kann dies nicht für quantitative Arbeiten empfohlen werden, da der Recall dadurch sehr gering wird (~ 0.3). Drittens, und das ist die wohl signifikanteste Anwendungsmöglichkeit für dieses Instrument, kann es als Werkzeug zur Voruntersuchung von Artikeln benutzen werden. Auch ohne spezifisches informationstechnisches Verständnis ist es möglich Unmengen von Zeit für künftige Forschung zu sparen, dazu vorher noch ein kurzer Exkurs;

Die manuelle Kodierung von Populismus hat zwei große Probleme, erstens ist das Konzept sehr komplex, was häufig zu einer geringen Interkoderreliabilität bzw. hohen Kosten bei dem Training der Kodierer führt, zweitens ist die Zahl der True Positives, also der tatsächlich Populistischen Artikel teils extrem gering. Zwar stellte Jager und Walgraves (2007) Untersuchung einen Durchbruch in der Messung von Populismus dar, doch Reliabilitäts- und Validitäts- Probleme wurden kaum beachtet (vgl. Pauwels, 2017: 125).

¹⁰ Bei dem völlig randomisierten Validierungsset von 500 Absätzen finden sich 3,2% als manuell populistisch kodierte Absätze.

Die Untersuchungen von Hawkins (2009; Hawkins & Silva, 2016) unter Verwendung einer holistischen Methodik haben zwar zufriedenstellende Werte, wurden jedoch hauptsächlich auf Partei-Manifestos angewendet. Das Untersuchen von Partei Manifesten ist aber vor allem im deutschsprachigen Raum mit nur geringem Erkenntnisgewinn verbunden. Bei der Untersuchung von Pauwels (2017) etwa, haben sowohl DIE LINKE als auch die ALTERNATIVE FÜR DEUTSCHLAND, welche von Experten und Laien als populistische Parteien klassifiziert werden, sehr geringe Werte. Das liegt wohl daran, dass die Manifeste eher vom bürgerlich-liberalen Flügel der Partei verfasst worden sind, man wolle „auch im Westen ankommen“ (AFD, 2015; zitiert durch Franzmann 2016: 32), außerdem gibt es eine Reihe von Wörtern und Bezeichnungen, welchen man aus historischen Gründen um jeden Preis ausweicht. Ruth Wodak bezeichnet das Phänomen, dass sich rechtspopulistische Parteien bürgerlich und staatstragend geben, als „Haiderisierung“ der Rechten (vgl. Wodak, 2015). In dieser Arbeit wurde Populismus in ‚durch Medien vermittelte Botschaften‘ untersucht, neben Manifesten und Parteipresse gibt es ein neues Untersuchungsfeld in dem ungefiltert und intentionell Populismus als Teil der politischen Kommunikationsstrategie verbreitet wird. Als künftig vielleicht sogar wichtigstes Untersuchungsmaterial empfiehlt sich also die Social Media Kommunikation von Parteien. Aufgrund der anderen, legeren Sprache mag es zwar zu intervenierenden Variablen kommen welche bereits in Kapitel 4.2. und 6.2. Erwähnung fanden, dennoch sollte das Script vor allem aufgrund der vielen direkten Reden vermutlich besser funktionieren als bei medienvermittelter Kommunikation. Ob dafür ein eigenes Trainingsset von Nöten wäre, oder das hier erstellte generalisiert werden kann, wäre ebenfalls eine Frage, dessen Beantwortung eine sinnvolle Erweiterung des Forschungsstandes darstellen würde.

Um diese daraus resultierenden riesigen Textmengen untersuchen zu können, bietet das in dieser Arbeit erstellte Instrument eine Möglichkeit dar die zu untersuchenden Texte stark einzuschränken. Dazu ist es nötig einen hohen Recall Wert zu erreichen, damit möglichst alle populistischen Aussagen in dem neu erstellten Datenmaterial enthalten sind. Durch das erstellte Instrument war es uns möglich, aus einem händisch vorher kodierten Set von 900 Absätzen, von welchen 35 Absätze als Populismus kodiert wurden, 648 Absätze als True Negatives auszuschließen, bei nur vier False Negatives. Damit kann man mit einem geringen Verschleiß beinahe drei Viertel aller Absätze im Vorhinein ausschließen. Diese dritte Anwendungsmöglichkeit kann eventuell schon bald in weiterführenden Forschungsarbeiten angewendet werden. Es muss hier jedoch angemerkt werden, dass die

Größe des Validierungssets etwas zu klein ist um endgültige Aussagen zu treffen. Laut Stryker (2006) müssten etwa 20.000 Absätze für ein Validierungsset verwendet werden um einen “Gold Standard” zu erreichen, wenn ein Merkmal dermaßen selten vorkommt wie Populismus. Dies ist jedoch aus forschungsökonomischen Gründen unrealistisch.

Diese Arbeit muss als explorative Studie verstanden werden. Eine weitere Beschäftigung, Verbesserung und Adaptierung wird für zukünftige Arbeiten auf jeden Fall empfohlen. Einige Unzulänglichkeiten des Instruments wurden bereits am Ende des vorigen Kapitels gezeigt. Im Folgenden sollen hier deshalb noch auf weitere, sowohl technische als auch theoretische, Schwierigkeiten und Fehler angesprochen werden.

In Retrospektive ist es ein theoretisches Versäumnis, dass Populismus in dieser Arbeit eng definiert wurde. Zwar ist Muddes Definition klar State of the Art, allerdings wäre statt eines Einschränkens der Definition auf die drei Grundelemente, eine Erweiterung, vor Allem mit Rechtspopulismus sehr fruchtbar gewesen. Die Generalisierbarkeit ist wohl das beste Argument für Muddes Definition. Natürlich macht das für diese Thematik auch Sinn, so kann man kann diese Arbeit als Anleitung lesen, die Algorithmen mit Artikeln aus anderen Sprachen zu lehren. Bei einigen Sprachen, wie Englisch, würde das wohl besser funktionieren, zumindest geht aus den meisten komparativen Studien hervor, dass die Übersetzungen in die deutsche Sprache der Wörterbuch Studien sogar am schlechtesten funktioniert (vgl. etwa Rooduijn und Pauwels, 2011¹¹), ebenfalls sind Word-Embedding (dazu siehe weiter unten) Programme für Englisch und Spanisch weiter fortgeschritten. Eine komparative Studie mit Supervised Machine Learning Algorithmen und verschiedensprachigen Trainingssets wäre ohnehin eine wichtige Ergänzung für dieses noch eher junge Forschungsgebiet. Es gibt eine Fülle von sprachlichen Elementen mit denen Rechtspopulismus operationalisiert werden könnte. Einen guten Überblick, was dem rechtspopulistischen Diskurs ausmacht, findet man bei Ruth Wodaks “The politics of fear: What right-wing populist discourses mean” (2015). Man könnte hier Anti-Elitismus mit allgemeiner „manichäischen“ Dichotomie erweitern. So definieren Rechtspopulisten nicht nur Eliten als “die Anderen”, welche im Gegensatz zum reinen und guten Volk stehen, sondern ziehen diese Abgrenzung vor allem durch Religion, Ethnizität und Nationalität (vgl. Pelinka 2012: 15), Pelinka nennt dieses den Populisten fundamentale Element “defining others”. Weitere Elemente die SML Algorithmen - unserer Einschätzung nach -

¹¹ Allerdings handelt es sich dabei um die Analyse von Manifesten der Parteien.

auf jeden Fall gelehrt werden können sind außerdem der Charismatische Führer. Der Begriff “Charismatisch” geht auf Max Weber zurück und meint eine “außeralltägliche [...] Qualität einer Persönlichkeit um derentwillen sie als mit übernatürlichen oder übermenschlichen [...] nicht jedem zugänglichen Kräften oder Eigenschaften, oder als gottgesandt [...] gewertet werden” (2002: 141¹²). Rechtspopulisten versuchen mit enormen Aufwand und durch konsistente und höchst professionelle strategische Medieninszenierung einen Führer zu zeichnen, “who sheds their elitist aura and become ‘one of us’, [...] while is still seen as untouchable and als living in a different world” (Pels, 2003: 59). Mudde und Kaltwasser (vgl. 2014, 386f) sehen dies in Hinblick auf dezentrale populistische Bewegungen wie die amerikanische Tea-Party, aber auch “Occupy Wall Street” oder die “weißen Märsche” in Belgien nicht als generalisierbares und daher definierendes Element von Populismus. Die von Wodak definierte “Arroganz der Ignoranz” (vgl. ebd., 2013: 27) meint grob gesagt eine Kommodifizierung der Politik, Anti-Intellektualismus und das Glorifizieren des “gesunden Hausverstandes” („common sense“). Zuletzt können noch emotionale und affektive Elemente, wie die Konstruktion und Instrumentalisierung von Angst (vgl. Altheide, 2002; Wodak, 2015; Schemer, Wirth & Matthes, 2010) verwendet werden um einen “Rechtspopulismus-Predictor” zu erstellen. Man könnte diesen als exklusiven, ergänzenden Predictor errechnen lassen. Diese Erweiterung würde vor allem für Untersuchungen im westeuropäischen und nordamerikanischen Raum einen starken Erkenntnisgewinn bei moderaten Aufwand gewähren.

Mit mehr Zeit und Ressourcen wäre es auch sinnvoll das Trainingsset größer zu machen. Burscher et al. zeigen zwar, dass der höchste Anstieg bei der Genauigkeit und Präzision eines Machine Learning Algorithmus im Bereich bis ~500 Artikel liegt (vgl. Abbildung 7; Burscher et.al. 2014: 201) vor allem bei komplexen Konstrukten - in jenem Fall “Morality”, bei uns Populismus - steigt die Classification Accuracy jedoch bis 4000 Artikel im Trainingsset konstant an. Auch in dieser Arbeit konnte beobachtet werden, dass sich das Instrument durch die verschiedenen Iterationen verbessert hat, je größer das Trainingsset wurde (vgl. Kapitel 5). Vor allem die Precision steigt immer noch weiter, der Recall wird dadurch jedoch etwas geringer, während kombinierte Kennzahlen wie Informedness noch steigen. Es ist also anzunehmen, dass mit mehr Ressourcen ein noch besseres Instrument für die in diesem Kapitel erwähnten praktischen Anwendungsmöglichkeiten erschaffen werden kann. Wie in Kapitel 6.1. diskutiert, führten die Iterationen nicht immer zu einem

¹² In fünfter Auflage, erste Auflage 1921, Hrsg. Marianne Weber.

ausschließlich positiven Effekt. Das mag daran liegen, dass der Algorithmus selbst aufgestellte Regeln revidieren muss, welche vor allem auch in Bezug auf die holistische Erhebung des Konzepts teilweise nicht ganz transparent sind. Wenn man ihn nun mit seinem eigenen Ergebnissen füttert, können sich Misskonzeptionen verhärten, während er falsche Regeln oft nicht separat anpassen kann und daher weniger sinnvolle Verbindungen knüpft. Auch angemerkt muss werden, dass schon das in Kapitel 4.3. erstellte, grundlegende Trainingsset durch eine Voruntersuchung bearbeitet wurde. Hier wurden zwar nur etwa knapp 30.000 Artikel der insgesamt 59.000 auf AmCat verfügbaren Artikel ausgeschlossen, ein intervenierender Effekt kann trotzdem nicht zur Gänze ausgeschlossen werden.

Eine wichtige Frage die sich aufdrängt ist es, ob Populismus (dem Algorithmus) tatsächlich wie in dieser Arbeit, und von Hawkins (2009, 2010) propagierten holistischen Weise gelehrt werden sollte. Es bedarf einer empirischen Untersuchung, ob es nicht sinnvoller wäre, einzelne Elemente des Populismus, also Anti-Elitismus, People-Zentrismus und eventuell auch rechtspopulistische Anti-Migrations-Sentiments oder Exklusivismus allgemein einzeln zu trainieren und für die Analyse an einem späteren Zeitpunkt zusammenzufügen. Dies würde nicht nur einen allgemeinen Erkenntnisgewinn generieren, sondern den Algorithmus erlauben einfachere und widerspruchsfreie Regeln zu erstellen. Bei der Validierung wurden oft Absätze als False Positive klassifiziert, die entweder nur People-Centrism oder Anti-Elitismus Elemente enthielten. Dieses Problem könnte durch ein aufteilen dieser Konzepte in unterschiedliche Learner umschifft werden. Zwar haben Burscher et al. (2016) in ihrer Arbeit einen holistischen mit einem sogenannten indikator-basierenden Ansatz verglichen, die von ihnen untersuchten Frames und die dazugehörigen Indikator Fragen der manuellen Kodierung lagen sprachlich jedoch nah beieinander. So macht es vermutlich keinen großen Unterschied ob die Fragen “Is there a reference to the financial costs/degree of expense involved, or to financial losses or gains?”, und “Is there a reference to economic consequences of pursuing or not pursuing a course of action?” einzeln (indikator-basierend) oder gemeinsam (holistisch) trainiert werden. Texte die in Bezug auf diese Fragen positiv klassifiziert werden verwenden mit hoher Wahrscheinlichkeit sehr ähnliche Wörter, wodurch es von Vorteil sein kann beide Fragen zu einem Konzept zu verbinden. Gänzlich anders stellt sich die Situation bei People-Centrism, Anti-Elitismus und potentiellen anderen Ausprägungen von Populismus in Texten dar. Diese sind mitunter sprachlich sehr weit voneinander entfernt. Damit entstehen für den SML Learner teils erhebliche Schwierigkeiten. Erstellt er korrekte Regeln für Anti-

Elitismus, ist aber in Folge mit einem Text konfrontiert der laut der hier verwendeten Definition keine populistische Kommunikation ist, weil keine People-Centrism Elemente enthalten sind, muss er diese korrekten Regeln revidieren und nach anderen Regeln suchen. Dies ist unserer Meinung nach der Grund für die fatalen Folgen die die Aufnahme von False Positives in das Trainingsset, im Zuge unseres iterativen Prozesses, auf den Recall-Wert hatte. Es scheint deshalb durchaus vielversprechend bei Populismus, bzw. allgemeiner gesagt, bei Konzepten die sprachlich klar voneinander abtrennbare Elemente haben, diese Elemente einzeln zu kodieren und in Folge damit den Learner zu trainieren. Eine sinnvolle technische Erweiterung wäre es die einzelnen Wörter in den Dokumenten durch Wort-Vektoren zu ersetzen, eine Methode die als „Word Embeddings“ bekannt ist. Hierbei werden Wörtern ein gewisser Abstand zu anderen Wörtern zugeschrieben. So liegen die Wörter „Mann“ und „Frau“ oder „Bürger“ und „Menschen“ relativ nah beieinander. Diese Methode wurde von Mikolov et. al. (2013) entwickelt und fand seither viel Beachtung. Wie genau sie funktioniert ist selbst für Experten auf dem Gebiet der Linguistik und Statistik nicht einfach nachvollziehbar. So schlussfolgern Goldberg und Levy (2014:) beispielsweise:

“Why does this produce good word representations? Good question. We don’t really know. The distributional hypothesis states that words in similar contexts have similar meanings. [...] Intuitively, this means that words that share many contexts will be similar to each other”.

Diese Methode wurde für die vorliegende Studie jedoch nicht verwendet da sie sehr rechenaufwendig ist und dadurch Iteration deutlich schwerfälliger macht. Word embeddings empfehlen sich deshalb eher um das finale Trainingsset weiter zu verbessern. Als R-Package dafür empfiehlt sich „word2vec”.

Als Fazit lässt sich also feststellen, dass eine Quantifizierung von Populismus sehr wohl möglich ist. Mit einer größeren Menge an Ressourcen, den nötigen Adaptionen und der Klärung einiger geringeren Unklarheiten ist es nach Ansicht der Autoren möglich, dass ein Instrument erstellt werden kann, welches die Forschung für kleine Forscherteams revolutionieren kann, da diese „Big Data“ effizienter verarbeiten können. Schließlich sollte ein Vergleich mit mehreren automatisierten, eventuell auch manuellen Methoden in den Ausblick gestellt werden. Es mag zwar sein, dass noch einige Punkte offen sind und eine Adaptierung des Instruments und des Erstellungsprozesses noch einiges an geistigen

Ressourcen zukünftiger Forscher in Anspruch nehmen wird, doch konnte hier gezeigt werden, dass es den Aufwand wert ist.

8. Literaturverzeichnis

Akkerman, A., Mudde, C., & Zaslove, A. (2013): How populist are the people? Measuring populist attitudes in voters. *Comparative Political Studies*, 0010414013512600.

Aslanidis, P. (2018). Measuring populist discourse with semantic text analysis: an application on grassroots populist mobilization. *Quality & Quantity*, 52(3), 1241-1263.

Bauer, M. W. (2000). Classical content analysis: A review. *Qualitative researching with text, image and sound*, 131-151.

Bekkerman, R. (2015): The Present and the Future of the KDD Cup Competition.
URL: <https://www.kdnuggets.com/2015/08/kdd-cup-present-future.html> (16.1.2019)

Bélangier, É., Meguid, B. M. (2008). Issue salience, issue ownership, and issue-based vote choice. *Electoral Studies*, 27(3), 477-491.

Bernhard, L., Kriesi, H., & Weber, E. (2015). The populist discourse of the Swiss People's Party.

Bernhard, L. (2016). Left or Right? Populist communication of political parties in recent Western European elections (No. 92). NCCR Democracy Working Paper.

Blumler, J.G., 2003. Foreword: Broadening and deepening comparative research. In: Mazzoleni, G., Stewart, J., Horsfield, B. (Eds.), *The Media and Neo-Populism: A Contemporary Comparative Analysis*. Praeger, Westport, CT.

Boomgaarden, H. G., & Vliegenthart, R. (2007). Explaining the rise of anti-immigrant parties: The role of news media content. *Electoral studies*, 26(2), 404-417.

Bos, L., van der Brug, W., & de Vreese, C. (2011). How the Media Shape Perceptions of Right-Wing Populist Leaders, *Political Communication*, 28:2, 182-206, DOI: 10.1080/10584609.2011.564605

- Breiman, L. (2001). Random forests. *Mach. Learn.*, 45(1):5–32.
- Burscher, B., Odijk, D., Vliegenthart, R., De Rijke, M., & De Vreese, C. H. (2014). Teaching the computer to code frames in news: Comparing two supervised machine learning approaches to frame analysis. *Communication Methods and Measures*, 8(3), 190-206.
- Chang, Y. W., Hsieh, C. J., Chang, K. W., Ringgaard, M., & Lin, C. J. (2010). Training and testing low-degree polynomial data mappings via linear SVM. *Journal of Machine Learning Research*, 11(Apr), 1471-1490.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). ACM.
- Chomsky, N. (1989): *Necessary Illusions: Thought Control in Democratic Societies*. London: Pluto Press.
- Cohen, B. (1963). *The press and foreign policy*. Princeton, NJ: Princeton University Press.
- Ernst, N., Engesser, S., & Esser, F. (2017). Bipolar Populism? The Use of Anti-Elitism and People-Centrism by Swiss Parties on Social Media. *Swiss Political Science Review*, 23(3), 253-261.
- Goldberg, Y., & Levy, O. (2014). word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method. *arXiv preprint arXiv:1402.3722*.
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis*, 21(3), 267-297.
- Hawkins, K. A. (2009). Is Chávez populist? Measuring populist discourse in comparative perspective. *Comparative Political Studies*, 42(8), 1040-1067.

Hawkins, K. A. (2010). *Venezuela's chavismo and populism in comparative perspective*. Cambridge University Press.

Hawkins, K. A., Riding, S., & Mudde, C. (2012). *Measuring populist attitudes*. Committee on Concepts and Methods.

Heinisch, R., & Mazzoleni, O. (2017, October). ANALYSING AND EXPLAINING POPULISM: BRINGING FRAME, ACTOR AND CONTEXT BACK IN. In *Political Populism* (pp. 105-122). Nomos Verlagsgesellschaft mbH & Co. KG.

Hillard, D., Purpura, S., & Wilkerson, J. (2007, April). An active learning framework for classifying political text. In *Annual Meeting of the Midwest Political Science Association*, Chicago.

Hillard, D., Purpura, S., & Wilkerson, J. (2008). Computer-assisted topic classification for mixed-methods social science research. *Journal of Information Technology & Politics*, 4(4), 31-46.

Ionescu, G., & Gellner, E. (Eds.). (1969). *Populism: its meaning and national characteristics*. Macmillan.

Jagers, J., & Walgrave, S. (2007). Populism as political communication style: An empirical study of political parties' discourse in Belgium. *European Journal of Political Research*, 46(3), 319-345.

Joachims, T. (1998). *Text categorization with support vector machines: Learning with many relevant features*. Berlin, Germany: Springer

Kaltwasser, C. R., Taggart, P. A., Espejo, P. O., & Ostiguy, P. (Eds.). (2017). *The Oxford Handbook of Populism*. Oxford University Press.

Kneuer, M. (2017). *Politische Kommunikation und digitale Medien in der Demokratie*. Harald Gapski, Monika Oberle und Walter Staufer (Hg.): Medienkompetenz.

Herausforderung für Politik, politische Bildung und Medienbildung (S. 43-52). Bonn: Bundeszentrale für politische Bildung (bpb).

Krafft, T. (2017). Qualitätsmaße binärer Klassifikatoren im Bereich kriminalprognostischer Instrumente der vierten Generation. Masterarbeit; Technische Universität Kaiserslautern. URL:<https://arxiv.org/pdf/1804.01557.pdf> (25.1.2019)

Kumar, A. (2014). Looking back at Obama's campaign in 2008: "True Blue Populist" and social production of empty signifiers in political reporting. *Journal of Communication Inquiry*, 38(1), 5-24.

Laclau, E. (1977). *Politics and ideology in Marxist theory: Capitalism, fascism, populism* (P. 173). London: Nlb.

Laver, M., & Garry, J. (2000). Estimating policy positions from political texts. *American Journal of Political Science*, 619-634.

Lippmann, R. P. (1987). An introduction to computing with neural nets. *ASSP Magazine*, 4(2), 4-22.

Lopez-Escobar, E., Llamas, J. P., McCombs, M., & Lennon, F. R. (1998). Two levels of agenda setting among advertising and news in the 1995 Spanish elections. *Political Communication*, 15, 225-238.

Manucci, L. (2017). Populism and the Media. In Kaltwasser, C. R., Taggart, P. A., Espejo, P. O., & Ostiguy, P. (Eds.). (2017). *The Oxford Handbook of Populism*. Oxford University Press. 467.

Maratea, A., Petrosino, A., & Manzo, M. (2014). Adjusted F-measure and kernel scaling for imbalanced data learning. *Information Sciences*, 257, 331-341.

Mazzoleni, G., & Schulz, W. (1999). "Mediatization" of politics: A challenge for democracy? *Political Communication*, 16, 247-261.

Mazzoleni, G. (2008). Populism and the Media. In *Twenty-first century populism* (pp. 49-64). Palgrave Macmillan, London.

Meyer, T. M., & Müller, W. C. (2013). The issue agenda, party competence and popularity: An empirical analysis of Austria 1989–2004. *Journal of Elections, Public Opinion & Parties*, 23(4), 484-500.

McCombs, M. E., Donald L. Shaw, D. L. (1972). The agenda-setting function of mass media." *Public opinion quarterly* 36.2 (1972): 176-187.

McCombs, M. E. (2004). The evolution of agenda-setting theory. Paper presented at the annual meeting of National Communication Association, Chicago.

McNair, B. (1995). *An introduction to political communication*. London, England: Routledge

Mikolov, T., Yih, W. T., & Zweig, G. (2013). Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 746-751).

Mudde, C. (2004). The populist zeitgeist. *Government and opposition*, 39(4), 541-563.

Mudde, C. (2007). *Populist radical right parties in Europe*. Cambridge, England: Cambridge University Press.

Mudde, C. (2017). Populism: An Ideational Approach. In Kaltwasser, C. R., Taggart, P. A., Espejo, P. O., & Ostiguy, P. (Eds.). (2017). *The Oxford Handbook of Populism*. Oxford University Press.

New York Times; Confessore & Yourish (2016): The 2016 Race: \$2 Billion Worth of Free Media for Donald Trump. URL: <https://www.nytimes.com/2016/03/16/upshot/measuring-donald-trumps-mammoth-advantage-in-free-media.html> (3.1.2019)

Nielsen, D. (2016). Tree Boosting With XGBoost-Why Does XGBoost Win" Every" Machine Learning Competition? (Master's thesis, NTNU).

Petrocik, J.R., (1996). Issue ownership in presidential elections, with a 1980 case study. *American Journal of Political Science* 40 (3), 825-850.

Pauwels, T. (2014). *Populism in Western Europe: Comparing Belgium, Germany and the Netherlands*. Routledge.

Pauwels, T. (2017). MEASURING POPULISM: A REVIEW OF CURRENT APPROACHES. In *Political Populism* (pp. 123-136). Nomos Verlagsgesellschaft mbH & Co. KG.

Papcke, S. (2001). *Gesellschaft der Eliten. Zur Reproduktion und Problematik sozialer Distanz*. Münster: Westfälisches Dampfboot.

Plasser, F., & Ulram, P. A. (2003). Striking a responsive chord: Mass media and right-wing populism in Austria. In G. Mazzoleni, J. Stewart, & B. Horsfield (Eds.), *The media and neo-populism: A contemporary comparative analysis* (pp. 21–43). Westport, CT: Praeger.

Plasser, F., Sommer, F. (2018). *Wahlen im Schatten der Flüchtlingskrise: Parteien, Wähler und Koalitionen im Umbruch*. Facultas

Powers, D. M. W. (2011). ["Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation"](#) (PDF). *Journal of Machine Learning Technologies*. 2 (1): 37–63.

Pozzolo, A., Caelen, O., Johnson R. and Bontempi, G. (2015). Calibrating Probability with Undersampling for Unbalanced Classification. In *Symposium on Computational Intelligence and Data Mining (CIDM)*, IEEE, 2015; Also URL: <https://www.kaggle.com/lct14558/imbalanced-data-why-you-should-not-use-roc-curve> (25.1.2019)

Roberts, C. W. (1997). Semantic text analysis: On the structure of linguistic ambiguity in ordinary discourse. *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*, 55-77.

Rooduijn, M., & Pauwels, T. (2011). Measuring populism: Comparing two methods of content analysis. *West European Politics*, 34(6), 1272-1283.

Saloutos, T. (1970). Reviewed Work: *Populism: Its Meaning and National Characteristics* by Ghita Ionescu, Ernest Gellner. In *Technology and Culture* Vol. 11, No. 2 (Apr., 1970), pp. 327-329

Scharkow, M. (2013). Thematic content analysis using supervised machine learning: An empirical evaluation using German online news. *Quality & Quantity*, 47(2), 761–773.

Schulz, W. (2009). *Politische Kommunikation: theoretische Ansätze und ergebnisse empirischer Forschung*. Springer-Verlag. Spier, T. (2010). Springer Fachmedien.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM computing surveys (CSUR)*, 34(1), 1-47.

Skovsgaard, M., & Van Dalen, A. (2013). Dodging the gatekeepers? Social media in the campaign mix during the 2011 Danish elections. *Information, communication & society*, 16(5), 737-756.

Stanley, B. (2008). The thin ideology of populism. *Journal of Political Ideologies*, 13(1), 95-110.

Stewart, J., Mazzoleni, G., Horsfield, B., (2003). Conclusion: power to the media managers. In: Mazzoleni, G., Stewart, J., Horsfield, B. (Eds.), *The Media and Neo-Populism: A Contemporary Comparative Analysis*. Praeger, Westport, CT.

Stryker, J. E., Wray, R. J., Hornik, R. C., & Yanovitzky, I. (2006). Validation of Database Search Terms for Content Analysis: The Case of Cancer News Coverage. *Journalism & Mass Communication Quarterly*, 83(2), 413–430

Taggart, P. (2000). *Populism*. Buckingham.

Učeň, P. (2007). Parties, populism, and anti-establishment politics in East Central Europe. *SAIS Review of International Affairs*, 27(1), 49-62.

Van der Meer, T. G., & Verhoeven, P. (2013). Public framing organizational crisis situations: Social media versus news media. *Public Relations Review*, 39(3), 229-231.

Van Kessel, S., (2011). Explaining the electoral performance of populist parties: the Netherlands as a case study. *Perspectives on European Politics and Society*, 12 (1), 68 - 88.

Vargo, C. J., Guo, L. (2017). Networks, Big Data, and Intermedia Agenda Setting: An Analysis of Traditional, Partisan, and Emerging Online US News. *Journalism & Mass Communication Quarterly*, 94(4), 1031-1055.

Vonbun, R., Königslöw, K. K. V., & Schoenbach, K. (2016). Intermedia agenda-setting in a multimedia news environment. *Journalism*, 17(8), 1054-1073.

Walgrave, S., de Swert, K., (2004). The making of the (issues of the) Vlaams Blok: the media and the success of the Belgian extreme-right party. *Political Communication* 21 (4), 479e500.

Walgrave, S., & Van Aelst, P. (2006). The contingency of the mass media's political agenda setting power: Toward a preliminary theory. *Journal of communication*, 56(1), 88-109.

Wirth, W., et al. (2016). The appeal of populist ideas, strategies and styles: A theoretical model and research design for analyzing populist political communication. NCCR Democracy, University of Zurich.

Wodak, R. (2013). *Anything Goes. The Haiderization of Europe. Right-Wing Populism in Europe: Politics and Discourse*, 23-37.

Wodak, R. (2015). *The politics of fear: What right-wing populist discourses mean*. Sage.

9. Appendix

9.1. Abstract, deutsche Version

Populismus ist in der Mitte der Gesellschaft angekommen und gilt immer mehr als akzeptabler Kommunikationsstil von Politikern sowohl in Wahlkampfzeiten, in der Opposition aber auch in der Regierung. Forscher die das Phänomen untersuchen stehen nicht nur vor dem Problem, dass das Konzept abstrakt und schwer definierbar ist, sie müssen – auch in Hinblick auf neue Kommunikationskanäle – immer größere Textmengen untersuchen. Diese Arbeit behandelt die Frage, ob es möglich ist, Populismus mit automatisierten, also computergestützten Methoden quantifizierbar zu machen. Mithilfe des „Supervised Machine Learning“ Algorithmus XGBoost wird ein Instrument erstellt, der kleinen Forscherteams die Möglichkeit geben soll, größere Mengen an Datenmaterial als bisher zu untersuchen.

Zuerst wird dabei versucht, die byzantische Begriffsverwirrung um Populismus etwas aufzulösen. Es wird dazu auf die einzelnen Elemente die fundamental für das Konzept sind eingegangen. Danach werden verschiedene Wege vorgestellt, wie Populismus bisher, sowohl manuell als auch computergestützt, gemessen wurde. Nach einer oberflächlichen Einführung in das „Machine Learning“ und der Begründung für die Wahl der SML Programmbibliothek XGBoost wird mit einer manuellen Inhaltsanalyse ein Trainingsset erstellt. Mit diesem Trainingsset wird ein Instrument erstellt, welches in verschiedenen iterativen Verfahren weiter verbessert wird und durch verschiedene Experimente angepasst wird.

Es zeigt sich, dass es sehr wohl möglich ist, einem Algorithmus den Unterschied zwischen Populismus und nicht-Populismus zu lehren. Es ergeben sich praktische Anwendungsmöglichkeiten, die es etwa ermöglichen, durch Voruntersuchungen beinahe drei Viertel der nicht-populistischen Absätzen aus einem Untersuchungsmaterial mit einen sehr zufriedenstellenden Recall Wert auszuschließen. Im Anschluss werden die Schwierigkeiten und Verfehlungen des Projekts diskutiert und teils allgemeine, teils sehr spezifische Verbesserungen vorgeschlagen.

9.2. Abstract, english version

Populism is everywhere. More than ever it is seen as an acceptable communication style in politics, not only in opposition and campaigning, but also in government. Researchers do not only have problems finding a common definition for the abstract construct, but also – with regards to new forms of communication – they have to deal with ever-larger quantities of text. This study tries to answer the question, if it is possible to quantify populism with automatic machine learning methods. With the use of the algorithm- library XGBoost we try to create an instrument which should give small teams of researchers the possibility to study big data.

First, we try to make sense of the byzantine confusion of definitions by exploring the state of the art of Populisms' basic elements. Then we show different approaches and methods of measuring populism in manual as well as automated ways. After a brief introduction in Supervised Machine Learning, with a special emphasis on the program library XGBoost, we create a trainingset using a manual content analysis. With this trainingset we create a tool, which we improve using an iterative design and several experiments.

It shows that it is indeed possible for the XGBoost algorithms to learn the difference between populist and non-populist paragraphs. In this state there are already practical uses for the created tool, like using it for a preliminary investigation of the research sample. It is possible to eliminate nearly three quarters of the non-populist paragraphs from the research sample with a very satisfying recall value. In the last chapter we investigate the problems and failures of the research process. We also offer and discuss several general and some very specific advancements for future research.

9.3. R-Script

```
library(quanteda)
library(amcatr)
library(xgboost)
library(dplyr)
library(DiagrammeR)
library(DiagrammeRsvg)
library(rsvg)

conn <- amcat.connect("http://autnes.amcat.nl")

mystopwords <- c("heinz-christian", "hc", "heinz", "christian",
"strache", "hc_strache", "christian_strache", "heinz-christian_strache",
"strache_heinz", "strache_heinz-christian", "strache_hc", "c_strache",
"strache_h",
                "peter", "pilz", "peter_pilz", "pilz_peter",
                "kern", "christian_kern", "kern_christian",
                "sebastian", "kurz", "sebastian_kurz",
"kurz_sebastian",
                "ulrike", "lunacek", "ulrike_lunacek",
"lunacek_ulrike",
                "matthias", "strolz", "matthias_strolz",
"strolz matthias",
                "corinna", "milborn", "corinna_milborn",
"milborn_corinna",
                "armin", "wolf", "armin_wolf", "wolf_armin", "dass",
                "kickl",
                "spö", "spo", "spoe", "sozialdemokrat*",
                "fpö", "fpo", "fpoe", "freiheitlich*",
                "övp", "oevp", "ovp", "liste kurz", "liste_kurz", "neue
volkspartei", "neue volkspartei",
                "gruene", "die gruenen", "die_gruenen", "gruen",
"gruenen", "grüne", "die grünen", "die_grünen", "grün", "grünen",
                "neos",
                "liste pilz", "liste_pilz",
                "behinderung_menschen", "behinderte_menschen",
"beeinträchtigte_menschen", "beeinträchtigung_menschen",
                "blinde_menschen",
                #"bei_der", "in_der", "von_der", "dass_die", "ist_das",
"bei_den", "in_die", "in_den",
                #"beim", "auf_der", "an_die", "mit_der", "und_auch",
"das_ist", "es_ist",
                "der_övp", #"und_das",
                "burg"
)

mystopwords <- c(mystopwords, stopwords("german"))

##### read in and prepare texts#####
setwd('C:/Users/rette/Desktop/Populismus/Validation')

coding <- read.csv("coding_use1.csv", na.strings=c("",
"NA"), stringsAsFactors=F) # Read in coding

coding$populism <- ifelse(coding$populism=="Yes", 1, 0) # Yes/ No -> 1/0
```

```

table(coding$populism)

corpuscoding <- corpus(coding, text_field = "text")

codingdfm <- dfm(corpuscoding,
  tolower = TRUE,
  remove = mystopwords,
  remove_url=TRUE,
  remove_numbers = TRUE,
  remove_punct = TRUE,
  remove_symbols = TRUE,
  remove_twitter = TRUE,
  remove_hyphens = FALSE,
  ngrams=1:2, verbose=TRUE)

#word stemming
codingdfm <- dfm_wordstem(codingdfm, language = "german")

#MIA - WOrd embedding

#delete all features that occur only once
codingdfm <- dfm_trim(codingdfm, min_termfreq = 2, verbose = TRUE)

# split in training and test sets
set.seed(123)
training <- sample(1:nrow(coding), floor(.80 * nrow(coding))) # 80% goes
to Training
test <- (1:nrow(coding))[1:nrow(coding) %in% training == FALSE] # rest
to Test

#train model

# converting matrix object
X <- as(codingdfm, "dgCMatrix")
# parameters to explore
tryEta <- c(0.3,0.6)
tryDepths <- c(2,3,4)
# placeholders for now
bestEta=NA
bestDepth=NA
bestAcc=0

for(eta in tryEta){
  for(dp in tryDepths){
    bst <- xgb.cv(data = X[training,],
      label = coding$populism[training],
      max.depth = dp,
      eta = eta,
      nthread = 4,
      nround = 500,
      nfold=5,
      print_every_n = 100L,
      objective = "binary:logistic")
    # cross-validated accuracy
    acc <- 1-mean(tail(bst$evaluation_log$test_error_mean))
    cat("Results for eta=",eta," and depth=", dp, " : ",
      acc," accuracy.\n",sep="")
    if(acc>bestAcc){
      bestEta=eta
      bestAcc=acc
      bestDepth=dp
    }
  }
}

```

```

    }
  }
}

cat("Best model has eta=",bestEta," and depth=", bestDepth, " : ",
    bestAcc," accuracy.\n",sep="")

# running best model
rf <- xgboost(data = X[training,],
              label = coding$populism[training],
              max.depth = bestDepth,
              eta = bestEta,
              nthread = 4,
              nround = 1000,
              print_every_n=100L,
              objective = "binary:logistic")

# out-of-sample accuracy
preds <- predict(rf, X[test,])

## function to compute accuracy
accuracy <- function(ypred, y){
  tab <- table(ypred, y)
  return(sum(diag(tab))/sum(tab))
}

# function to compute precision
precision <- function(ypred, y){
  tab <- table(ypred, y)
  return((tab[2,2])/(tab[2,1]+tab[2,2]))
}

# function to compute recall
recall <- function(ypred, y){
  tab <- table(ypred, y)
  return(tab[2,2]/(tab[1,2]+tab[2,2]))
}

cat("\nAccuracy on test set=", round(accuracy(preds>.60,
coding$populism[test]),3))
cat("\nPrecision(1) on test set=", round(precision(preds>.60,
coding$populism[test]),3))
cat("\nRecall(1) on test set=", round(recall(preds>.60,
coding$populism[test]),3))
cat("\nPrecision(0) on test set=", round(precision(preds<.60,
coding$populism[test]==0),3))
cat("\nRecall(0) on test set=", round(recall(preds<.60,
coding$populism[test]==0),3))

# feature importance
labels <- dimnames(X)[[2]]
importance <- xgb.importance(labels, model = rf, data=X,
label=coding$populism)
importance <- importance[order(importance$Gain, decreasing=TRUE),]
head(importance, n=100)

# adding sign - Preparing indicator visualization
sums <- list()
for (v in 0:1){
  sums[[v+1]] <- colSums(X[coding[, "populism"]==v,])
}
sums <- do.call(cbind, sums)
sign <- apply(sums, 1, which.max)

```

```

df <- data.frame(
  Feature = labels,
  sign = sign-1,
  stringsAsFactors=F)
importance <- merge(importance, df, by="Feature")

## best predictors
for (v in 0:1){
  cat("\n\n")
  cat("value==", v)
  importance <- importance[order(importance$Gain, decreasing=TRUE),]
  print(head(importance[importance$sign==v,], n=100))
  cat("\n")
  cat(paste(unique(head(importance$Feature[importance$sign==v], n=100)),
collapse=", "))
}

#Visualizatio

treebuild1 <- xgb.plot.tree(model = rf)

plottedimpo <- xgb.plot.importance(importance)

gr <- xgb.plot.tree(model=rf, trees=0:100, render=FALSE)

treenumber = 101
pdfheight = 1800 * treenumber

export_graph(gr, 'tree0-100_finalo.pdf', width=3000, height=pdfheight)

##### running on other corpus#####

testsetid = 3753 #amcat article set id

#create corpus
amcatcorp <- quantda.corpus(conn, 3, testsetid,
                           textcolumns = c("title", "text"),
                           metacolumns = c("date", "medium", "byline",
"section", "pagenr"))

# change to paragraphs and remove paragraphs shorter than 4 words
#amcatcorp <- corpus_reshape(corpus_trim(amcatcorp, what ="paragraphs",
min_ntoken = 5), "paragraphs")

mytestset <- read.csv("test_set.csv", na.strings=c("",
"NA"),stringsAsFactors=F) # Read in coding

#coding$populism <- ifelse(coding$populism=="Yes", 1, 0) # Yes/ No ->
1/0
table(coding$populism)

amcatcorp <- corpus(mytestset, text_field = "texts")

#amcatcorp <- corpus_reshape(amcatcorp, "paragraphs")

#get sample of 400 paragraphs
set.seed(123)
amcatcorp <- corpus_sample(amcatcorp, 500, replace = FALSE)

#create dfm
amcatdfm <- dfm(amcatcorp, tolower = TRUE,

```

```

        remove = mystopwords,
        remove_url=TRUE,
        remove_numbers = TRUE,
        remove_punct = TRUE,
        remove_symbols = TRUE,
        remove_twitter = TRUE,
        remove_hyphens = FALSE,
        ngrams=1:2, verbose=TRUE)

#word stemming
amcatdfm <- dfm_wordstem(amcatdfm, language = "german")

#only words min 2 times
amcatdfm <- dfm_trim(amcatdfm, min_docfreq = 2, verbose=TRUE)

#match the training set and current dfm
amcatdfm_match <- amcatdfm[, intersect(colnames(amcatdfm),
colnames(codingdfm))]

for(col in setdiff(colnames(codingdfm), colnames(amcatdfm))){
  amcatdfm_match <- cbind(amcatdfm_match, 0)
  colnames(amcatdfm_match)[ncol(amcatdfm_match)] <- col
}
amcatdfm_match <- amcatdfm_match[, colnames(codingdfm)] # sort features
as in codingdfm
amcatX <- as(amcatdfm_match, Class = "dgCMatrix")

##### Predict #####
amcat_predict <- predict(rf, amcatX)

#Prepare result data
amcat_results <- amcatcorp$documents
amcat_results$pred <- round(amcat_predict, 2)

write.csv(amcat_results, file = "results_1.csv")

#Save Indicatorwords
indicatorwords <- data.frame()
indicatorwords <- c(indicatorwords, importance)
write.csv(indicatorwords, file = "indicatorwords_1.csv")
#View(indicatorwords)

```

9.4. Codebuch

Introduction

Introducing the concept of statement coding and some basic rules and examples of when and how to code statements.

Identifying Statements

For this study **attributions** found in **statements** are the units that need to be coded. This is an overview of those two concepts:

Statements are what a **speaker** (politician, member of elite, common person, journalist,...) says about **target actors** such as other politicians, supranational organizations, elites, or the people.

Example of a **statement**: Politician A (**speaker**) says about Politician B (**target actor**) that she/he is incompetent (**attribution**).

Identifying these statement is the first and a crucial step in identifying populism in media coverage. The utterance of a speaker in any text may contain one or several **statements** on one or different target actors.

A target actor can take many forms (foreign countries, politicians,...) but relevant for this study are only the following (which will be explained in more detail):

- Elites (Political Elites, The Press, Scientific Elites, Financial Elites...)
- The People (Society as a whole, The common sense, “The will of the people”,...)
- self (i.e., the “populist” or conveyor of a populist message)

If the speaker makes a statement about his own person, this also may be coded.

Statements about historic, mythic, or hypothetical actors (e.g., founding fathers) are not to be coded. **In addition, if the statement is only about the feelings or emotions of a passive target, this also may not be coded!**

Any statement coded in this content analysis has a strictly defined shape as it consists of a **speaker** (subject of the statement), a **target actor** (object of the statement) and the content of the statement (**attribution**). This content may be a description of personal traits and attributes, a description of actions and statements, or an allegation of purposes.

Each of these different properties of the target may be **ascribed or denied** by the speaker.

Statements containing such an ascription or denial of properties are subsumed under the term '**attribution**'. Attributions are thereby the relationship between subject and object in any statement (Figure 7). A statement may contain several **attributions**. Utterances not containing any attributions are mere mentions of an actor. They are not counted and not coded as statements.

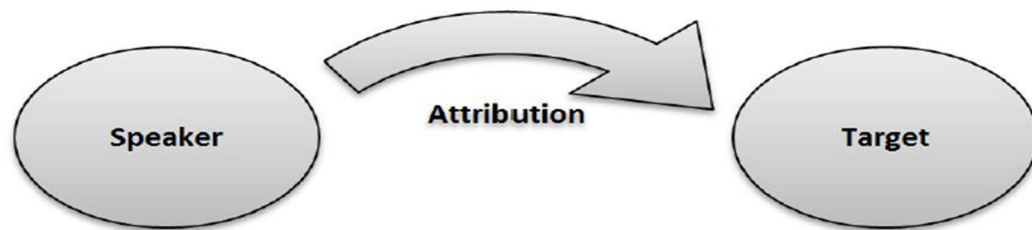


Figure 7: Prototypical form of a target evaluation

Each statement thus contains **at least** a target actor and one of these elements:

- Attributions
- Positive Traits
- Negative Traits
- Impact/Responsibility
- Actions
- Power

All elements must be coded for each statement of any target. A statement might contain several attributions (e.g. “politicians are corrupt and out of touch with the people on the street”) in which case both attributions **should be coded separately**.

- 1.) Speaker A says politicians (target actor) are corrupt (attribution)
- 2.) Speaker A says politicians (target actor) are out of touch with the people (attribution)

The same attribution is not coded twice in the same article. If a speaker uses different synonyms to describe the same target actor (e.g. die Mitte-Links-Parteien and SP und die Grünen) this also may be coded as one single statement.

Target actors not to be coded as statements:

There are some statements on other persons or institutions which must not be coded as statements.

This includes:

- Mere mentioning of how persons/organizations are affected by something
- Mere mentioning of the feelings of persons/institutions
- Mere connection of a person/institution to an issue without indicating a position or elaboration.

If a statement consists of merely mentioning this target and her/his/their opinion, this is not considered an a statement because it lacks an evaluation of a target actor. This limitation is especially important when considering the statements of the author on the text toward actors within the text.

Examples:

The people will suffer from this policy (Mere affected group and not a target actor! Therefore not to be coded as statement.)

Cameron also talked about immigration in his speech (No target actor, merely mentioning of events taking place. Therefore not to be coded as statement.)

Merkel was downcast after the results of the poll were presented to her (mere description of feelings/affection. No Target actor and therefore not to be coded as statement.)

"Recent polls suggest that the Brits are suffering from the increasing numbers of immigrants in the cities" (Mere affected group and not a target actor! Therefore not to be coded as statement.)

Speaker

One text may contain numerous speakers, each with numerous statements toward other actors (here: target actors). By aggregating the single speakers or statements to the level of articles, the content of each article may be described fine-grained.

Definition:

A speaker is an actor which is quoted in the text either directly or indirectly to utter any statement on a target actor covered in the article. Each text has at least one speaker (**the journalist or author**) which may talk about the issue at hand or any actor connected to this issue.

Only cases where a statement is directed at either the people or some form of elites are relevant for this study!

In addition to the journalist, political actors are considered speakers if they are given the opportunity to express their opinion or position towards an actor by means of one of the following style elements:

- Soundbite/Direct Speech: A statement is literally transcribed or presented in the text.
- Indirect Speech: A statement is paraphrased and explicitly assigned to a specific actor.
- Explanation of position: The position or attitude of an actor is transported without using direct or indirect speech (e.g. "Steinbrück lehnt eine grosse Koalition jedoch entschieden ab").

All persons, organizations, or collective actors being cited in such a way are coded as a speaker.

A **statement on a target actor** is a statement on another person, collective, or abstract person or institution.

Such a statement must at least contain the following:

- One or more explicitly mentioned attributions or denial of some personality trait, action, impact, power, or relation to the people.
- OR an explicitly stated agreement or disagreement with the speaker.

Coding

A list of the variables that must be coded for this study. Definition gives an overview of the concept. Values are the actual codes that have to be entered.

V01_Target Specification

Is the target actor of a statement either people, elite or the speaker himself?

Definition:

The target of a statement may be any actor, collective or abstract which is described, evaluated, or addressed by the speaker. This includes all specifically named actors (Persons, Organizations, States...), proportions of the population (e.g.: socioeconomic layers, ethnic groups, the people as a whole...) and the speaker himself.

For this study only People, Elites and Own Person are relevant categories all others are merely mentioned to distinguish those from other categories.

Values: 1 The People, 2 Elites, 3 Own Person

The following are definitions and examples for the single values:

Definition “The People”

General

References to the people are directed toward the whole of the society, population, nation, or people or to a specific part of this collective which is considered the true people. In some cases, the people may be addressed by means of an inclusive 'we' such as in 'we have to solve that problem together'. If this inclusive 'we' is explicitly addressed to the people (e.g. in the context of the sentence). If, however, the reference to the people is implicit and the speaker may as well refer to the own party or any other collective he is part of, the object of the statement has to be coded as the 'Own Person'.

The four categories 'Nation', 'Ethnos', 'Societal Layer', and 'Function' are definitions of the people as a whole. However, they are strongly dependent on the definition of people and especially of the definition of persons that do not belong to the people. In some languages there are words for these different definitions of the people. In most cases, however, the definition is only identifiable by the explicit exclusion of persons not belonging to the people. See the definitions of the single categories for examples and language-specific cues.

Positive Examples:

People as a nation vs. foreign actors

People as an ethnos vs. other origins

People as a societal layer

Hypothetical prototype of the people (common man, man in the street)

People as a function (e.g: voters, citizens, inhabitants)

Not to be coded as “The people”:

Majority of the people

Socioeconomic group

Socioeconomic group: Rich people

Socioeconomic group: Poor people / Welfare recipients

Origin group

Origin group: Immigrants / Foreign

Origin group: Asylum Seekers / Refugees

Religious group

Religious group: Muslim

Age group

Education group

Ideology group

Gender Group (male/female)

Unemployed persons

Readers

other specification

In addition any reference to 'the people' where the function, extent or meaning of this group of persons is not made explicit is also coded as the people. This includes the single notion of 'Schweizer' in Switzerland or 'Britons' in England without any contrast to other people, groups or nations. It also includes mentions of 'Some people' or 'Lots of people' (Ger: "Alle", "Die Leute"). Also if the reference is regarding a survey of the people which refers to their opinions about an issue, this may also be coded.

“Die Schweizer sind gute Wintersportler.”

“Gewisse Leute sind nur auf ihren eigenen Vorteil bedacht.”

People will not suffer this decision and hit the streets.

People as a nation

Any reference to the whole people in contrast to persons and organizations outside the own country. This definition of the people is always recognizable by references to international or foreign actors which are put in contrast to the people.

The people must not tolerate this interference by foreign politicians. We should stand our ground.

“Die Schweizer werden von Brüssel immer stärker fremdbestimmt. Sie sollten mehr Macht haben, in ihrem eigenen Land zu bestimmen.“

„Recent polls suggest that the Brits are strongly opposed to the increasing number of Immigrants“

People as an ethnos

Any reference to the whole people in contrast to persons within the own country but of different origin. This also includes notions of the 'True Englishman', or the 'Urschweizer'/'Eidgenossen'.

“Die Schweizer haben im Gegensatz zu Ausländern eine gute Arbeitsmoral.“

"Recent polls suggest that the Brits are strongly opposed to the increasing numbers of immigrants in the cities"

People as a societal layer

Any reference to the people in contrast to persons within the own country but of different social standing (e.g: rich or poor). This includes notions of the 'working population' and the 'Proletariat'.

This is not to be confused with a socioeconomic group within the population. The people as a societal layer always refers to the whole of the population, except the rich.

Socioeconomic groups may be the poor, the rich, or people with an income in a certain range. They are always considered just a part of the population.

Alle ausser die Superreichen sind gegen diese Vorlage.

“In contrast to other Brits, the working population would suffer from this decision.”

Hypothetical prototype of the people

Any reference to a hypothetical person representing the whole population as a prototype. Please see the extended definition in style: Reference to the ‘common man’.

“Der Otto Normalverbraucher vergleicht beim Kauf seiner Winterausrüstung die Preise der verschiedenen Anbieter.”

Explain this to Johnny Normal!

People as a function

Any reference to the people in a specific function within the society (e.g: voters, electorate, citizens, inhabitants). This definition may either be identified by a contrast to the political elite or by the usage of a term describing a specific function (e.g: tax-payer, Wähler, consumer). The simple notion of participants or readers of a survey may not be coded as the people.

“Das Schweizerstimmvolk hat entschieden, die ...“

The tax payer

The voters

Definition “Elites”

Any kind of elite is coded with this category. The elite may thereby be the leading class in different sectors of the society.

The **political elite** comprise all politicians (as a whole), the executive, legislative, judicative, and administration.

The **financial elite** comprise all very rich persons and banks, the wealthy 1% of the population ("Die oberen 10'000"), or any other notion of influential rich persons.

The **economic elite** comprises heads of companies, large enterprises, or employers. The media elite comprises all producers of mass media.

The **media elite** are journalists (as a whole), media companies (news papers, tv-stations), the press in general or other well known opinion leaders (e.g. social media influencer). The **security elite** consists of institutions and leaders of the military, police and secret service.

Also **supranational institutions** like the UN, Nato, EU are to be coded as elites when they are the target of an attribution.

Finally, there is an **unspecific elite** which is just referred to as the 'mighty ones' the 'ones behind the curtain' or 'the ones at the core of power'.

The speaker might also refer to a **combination** of any of those elites.

Definition “Own person”

Any reference to the own person or family is coded in this category. This explicitly excludes any notion of the own party (>specific party) and the own people (>people).

This does not include statements as 'I think' or 'In my opinion'. These phrases are only an indicator for the personal opinion as a Sourcing of the statement.

Coding Attributions

When coding attributions it is important to be aware of the fact that they can be either attributed (the target actor DOES have those features) or denied (the target actor DOES NOT have those features) by the speaker. Denying a positive attributions is coded as a negative attribution (e.g. “*In an interview AfD leader Lucke accused Merkel of not being able to solve the immigration problems Germany are confronted with*”), while denying a negative attribution is counted as a positive attribution (e.g. “*Obama claims illegal immigrants don’t burden taxpayers.*”

Each attribution comes with a list of categories (examples) that cover all cases that should be coded in this variable. If an attribution can not clearly be located in one (or more) of those categories the attribution is not to be coded!

V03_Impact: Attribution of suspected impact and responsibility

Values: -1 Negative Allegation, **9 not mentioned**, +1 positive Allegations

Definition:

The speaker may attribute impacts and effects to an actor or his actions. By doing this it is possible to shift the blame or give credit to an actor in public.

Questions:

- **Does the speaker attribute NEGATIVE allegations or supposed impacts or effects on society to the target?**
 - **Attribution of: Posing a threat, Posing a burden, Accountable for negative development/situation**
 - **Denial of: Posing an enrichment, Accountable for positive development/situation, Ability to alleviate/solve a problem**
- + **Does the speaker attribute POSITIVE allegations or supposed impacts or effects on society to the target?**
 - + **Attribution of: Posing an enrichment, Accountable for positive development/situation, Ability to alleviate/solve a problem**
 - + **Denial of: Posing a threat, Posing a burden, Accountable for negative development/situation**

The following are categories within V03_Impact that cover and illustrate all cases that are to be coded within this variable:

Posing a Threat to the people, identity, nation, or society in general

In a political context, fear or threat appeals can be used by political actors to promote their political view or to discredit the view of their opponents respectively, by stressing negative consequences of a policy or a development in society. Therefore also negative consequences of (planned) actions or intentions by the target may be coded as a negative allegation.

"ConDem policy kills hope, jobs, communities and social cohesion. This is the battleground on which we must fight for our future." (Attribution of being a threat to society)

Posing a burden

Especially when mentioning certain groups within the population, political actors may describe them as a burden to society, the state or the population. This is especially the case with groups receiving benefits from the state (e.g: unemployed persons, farmers, single parents) or immigrants accepting lower wages.

Negative allegations will be coded wherever a speaker talks about other actors as posing either:

- A financial burden to the state
- An ongoing problem for society
- A limiting factor in positive development of the country

„Der Anteil von Migranten bei Hartz-IV-Empfängern ist überdurchschnittlich hoch. Die jahrelang ungesteuerte Einwanderung von bildungsfernen Migranten und Geringqualifizierten ist eine Belastung unserer Sozialsysteme. Hier besteht Handlungsbedarf“, sagte der Obmann der CDU-Fraktion. (Attribution of posing a burden)

Denial of posing a burden (= positive allegation) may be coded in cases where such an allegation is dismissed.

“Obama claims illegal immigrants don’t burden taxpayers.”

Posing an Enrichment

Opposed to the attribution of posing a burden to the society, speakers may choose to label someone as enrichment to society and thereby presenting the actor as entirely positive in his/her/their impact on society.

The enrichment may thereby be:

- Financial (e.g.: bringing money to the country)
- Cultural (e.g.: providing new customs and worldviews)
- Social (e.g.: helping other people / caring for neighbors)
- Professional (e.g.: doing jobs that would not be done otherwise)

All notions of an actor or a group of actors having a positive impact on society will be coded as positive allegation.

“Der Nachrichtenagentur dpa sagte er (Innenminister) nach seinem Besuch, der Alltag dort zeige, dass Asylbewerber eine Bereicherung für das gemeinsame Leben seien.”

„Kanzlerin Angela Merkel und Ministerpräsident Volker Bouffier sind die Garanten für die Fortführung der positiven Entwicklung der vergangenen Jahre. Deshalb haben wir allen Grund, optimistisch das Jahr 2014 anzugehen.”

“Anyone who reads this letter - and I've read the letter - will see that what he is saying is that British Muslims make a great contribution to our country..., he (Cameron) said.”

Accountability/Responsibility for negative development/situation

Accountability or responsibility may be pointed out in various ways. Due to national (resp. linguistic) differences between texts any differentiation is omitted here.

Accountability or responsibility for a negative situation may be coded as negative allegation if:

- An actor is explicitly stated as having caused a situation.
- An actor is explicitly stated as not having (been able to) prevented a situation.
- An actor was in a position in which he has to take responsibility for the situation (e.g.: A minister is responsible for any problems within his or her department)
- An actor is explicitly stated as having to acknowledge responsibility.

“Together, we’ve agreed to cut the deficit by more than \$2 trillion. And we’ve put in place new rules to hold Wall Street accountable, so a crisis like this never happens again.”

Accountability/Responsibility for positive development/situation

Accountability or responsibility may be pointed out in various ways. Due to national (resp. linguistic) differences between texts any differentiation is omitted here.

Accountability or responsibility as a positive allegation may be coded if:

- An actor is explicitly stated as having caused a situation.
- An actor is explicitly stated as not having (been able to) prevented a situation.
- An actor was in a position in which he has to take responsibility for the situation (e.g.: A minister is responsible for any problems within his or her department)
- An actor is explicitly stated as having to acknowledge responsibility.

Attention: If only a future prognosis is made about how someone will be responsible for a probable future development, do not code it here!

Ability to solve a problem

This covers cases of a speaker attributing or denying an actor with the ability of solving or at least alleviating a problem. This category only applies if the speaker explicitly states this attribution.

“In an interview AfD leader Lucke accused Merkel of not being able to solve the immigration problems Germany are confronted with”.

V04_ People: Attribution of Relation to the People

Values: -1 Deceiving people/closeness denied , 9 not mentioned, +1 close to the people/deceiving denied

Definition:

Especially when talking about political actors or members of the Elite of the country, the connection to the people they represent, govern or influence is frequently debated. In this content analysis the ascribed connection of political actors to the people is of utmost importance since the link to the people is important for populist actors. For the media, the connection to the people or their readership is of high importance as well.

The relation to the people may be attributed on different levels either by another actor or by the speaker himself. In both cases the attribution will be coded.

Question:

How is the relation of this target to the people described?

Belonging to the people

The actor is described as a member of the people he or she represents.

“Ich (Filippo Leutenegger) wohne seit vielen Jahren in Zürich, kaufe hier ein, benutze den ÖV, bringe die Kinder hier zur Schule. Das macht mich ohne Zweifel zu einem normalen Bürger der Stadt.“

“He (Obama) grew up in this community and he therefore is definitely one of us.”

Being close to the people

If a certain actor aims to communicate with the people in a more direct, i.e. unmediated, way – as it is usually the case in representative democracies – the statement falls into this category. The respective actor may for example offer criticism with regards to the fact that ‘the people’, for the most part, tends to be separated from the political-administrative sphere.

In order to demonstrate closeness to the people the leader may convey his/her rejection of the classical isolation of the political-administrative sphere, which is increasingly alienated from the people, whom the politicians, belonging to this sphere, claim to represent. This may also be achieved by means of a certain casual dress code etc.

However, he/she may as well explicitly state it, thereby intensifying the positive connection he/she maintains with the people.

"Mein wahrer Koalitionspartner ist die Bevölkerung", sagt Seehofer. (Self-Attribution of Closeness to the people. Not belonging since the people is called a partner)

"(Blick) Herr Maurer war es wichtig, keine abgeschottete Exkursion zu machen, sondern sich

direkt unters Volk zu mischen“, weiss Nehmer." (Attribution of being close to the people)

"Ich zeige damit, dass ich im Alltag nahe bei den Leuten bin', begründet Leutenegger den

Werbespruch."

"Honest UKIP can speak directly to the British people, says Leo McKinstry."

Knowing the needs of the people

An actor claims to know what 'the people' wants or how 'the people' feel. They assert to know what problems (in the name of 'the people') shall be solved and how the current state of mind of the population looks like.

Since the populist leader embodies the peoples' they will naturally know what the people want or feel. He or she connects with the 'common man' by talking about the peoples' needs or stressing the preservation of their overall well-being.

"(Grant Shapps) ,People have sent a message, we get it, we hear what people are saying, people are concerned that we get on with big issues facing hard-working people in this country, like fixing the economy, sorting out the welfare system, helping hard-working people to get on.'" (Self-Attribution of knowing the needs of the people)

„I understand why some people who've supported us before didn't support us again. They

want us to do even more, to work for hard working people, to sort out the issues they care about. More to help with the cost of living. More to turn the economy around...They will be

our focus, they are our focus, but we've got to do more. "" (Self-Attribution of knowing the needs of the people)

Cameron has no sense for the needs of the people. (Denial)

Caring for the people

An actor expresses that he cares about what the people want or claim. He shows empathy or tries to shed light upon the unheard objectives of "the common man". The populist actor expresses a similar emotional state as the people. For example he would express personal suffering if a major crisis is claimed to exist by the respective actor.

A populist actor embodies the peoples' will and therefore knows, without asking, what the people really want or what their concerns are. But because they themselves are part of the people the peoples' concerns are automatically the leaders concerns, too, which is why they are in the same emotional state as the 'common man'.

"(NYT) 'I am worried about those who are unemployed,' he told reporters at his ranch near

Crawford, Tex. 'I think about the overall economy and how best to help those folks who are

looking for work.'"" (Self-Attribution of caring for the people)

Denial of caring for the people may be coded when an actor is explicitly stated as not caring for the needs of the people.

"Warum sollten wir wählen gehen, die sind doch eh alle korrupt und kümmern sich nicht um

uns", sind sich die drei Gäste am Stand des "China-Thai-Bistros" auf dem Parkplatz vor Norma und Lidl einig. (Denial of caring for the people)

Speaking on behalf of the people

The populist actor portrays himself as the legitimate spokesperson for the people and his aim is to push through what the people really want. In order to preserve the legitimacy of his claims (since he represents the true will of the people) he has to defend the peoples' will if a certain party refuses to implement the true objectives of the people.

"Ich weiss, was das Volk wirklich will und werde daher den Willen des Volkes auch im Parlament würdig vertreten."

"I'm the only one who knows what the people of Britain want and what their needs are."

Empowering the people

Actors may be described as giving power to the people or taking the power away from them. This may also be coded in this variable. This only applies to cases where the empowerment of the people by an actor is explicitly mentioned.

"Die SVP setzt sich dafür ein, dass das Volk in dieser Sache mitbestimmen kann"
(Attribution of Empowerment)

"Her mission is to give the people their power back."

Denial of empowering the people may be coded when an actor is explicitly stated as not empower the people.

"Wenn es nach dem Parlament geht, hat das Volk in dieser Sache nichts zu sagen"
(Denial of empowerment)

Embodiment/Representation of people's will

Populist actors have a twofold self-concept which allows them to be a populist leader in spite of the strong aversion against leaders and elites in populist ideology. On the one hand, this self-concept holds that the leader is still a member of the people in spite of the current position within the political system.

On the other hand, populist self-concept holds that the populist actor is not an actor in his own right longing for power (such as all other elites) but the necessary embodiment of the people's will within the political system. Therefore the populist actor represents the will of the people and does not pursue any own ambitions, aims or schemes. In order for this self-concept to work the populist leader has to claim that "the people" has one unified will or agreement which may be represented by him.

Stating that the will of the people is fragmented and the actor represents a proportion of the people (even the majority) may not be counted as populist strategy.

Wir haben vom Schweizer Volk den Auftrag bekommen.

We have to follow the will of the people here. The British have repeatedly challenged these plans of the government.

“I have to follow the will of the people here. The British have repeatedly challenged these plans of the government.”

Agreement with the people

In addition to evaluating target actors with respect to personality attributes, impact, and actions, a speaker may also state agreement or disagreement with an actor, either generally or specifically. General agreement would mean that the speaker points out that he agrees to a large part to the positions and opinions of the target actor (in this case - “The People”). This should be coded as “+1 close to the people”.

Important: Agreement or disagreement is only to be coded if it is made explicit by the speaker. Just because a speaker has the same opinion as the target does not mean that the agreement is part of the statement of the speaker toward the target.

Performing Everyday actions

Everyday actions are considered actions that normally need to be done on every day by everybody – not only political active people. A speaker can describe a political actor doing grocery shopping, eating lunch or exercising. If the speaker does not explicitly draw a connection between this action and the actor's political functions then this action will be considered only an everyday action (even if a symbolic component is clearly present).

“(Washington Post) For a look outside presidential bubble, Obama reads 10 personal letters each day.” (Attribution of everyday action: Reading letters. The action may be symbolic in reality but is not explicitly presented as such.)

„Regelmäßig ist sie bei Proben und Auftritten der Jazz- und Gospelgruppe ‚Kirchenwecker‘ dabei und Freizeitschauspielerin bei der Theatergruppe der DJK Rohr.“

Deceiving the people

This negative allegation applies if an actor is allegedly deceiving the people, disseminating false information or just fooling the population. This includes treason to the people and acting against previously made arrangements.

"Wie lange wollen Sie die Bevölkerung noch für dumm verkaufen, Frau Sommaruga?" (Attribution of Deceit)

"Wir können dann genüsslich zusehen, wie die Frau ohne Charakter Widmer-Schlump noch und noch weitere Bücklinge macht und den Deutschen in den Hintern kriecht, unter Begleitung des Gesamtbundesrats, dieses Verräterrats." (Attribution of Treason to the people)

"Ein Politiker, der solche Vorschläge macht, ist unredlich und täuscht vorsätzlich die Öffentlichkeit." (Attribution of Deceit)

“With this action, Cameron intentionally mislead the British society.”

Counter-Example:

"Und der Sonntagsblick gibt sich noch zur Publikation eines solchen Quatsches hin. Wo bleibt die journalistische Hinterfragung?" (No attribution of deceiving the people since there is no people in this statement. The word 'publishing' does not explicitly state a link to the people.)

V05_Mono: Monolithic people

Is the People referred to as a monolithic actor?

(This is only to be coded if there is an “Attributions of Relation to the People” present)

Values: -1 People diverse, 1 People common feeling, desires, opinions, 9 not defined

Definition:

When referring to the people a speaker may differentiate between two concepts of the people. On one hand, there is a "monolithic people" who shares a common understanding of the world and has common feelings and desires. In this concept of the people, it is possible to speak of 'the will of the people' or 'the people's fears'. On the other hand, there is the concept of a fractured population of individuals with their own feelings, opinions, and desires which do disagree in some points.

Question:

Please indicate whether the people (as specified above) shares a common will, opinion, feeling, or desire.

Filter

Only to be coded for reference to the people (Target = People (not part of the people)) AND for attributions of relation to the people.

Common feelings, desires, opinion

When treating the people as a monolithic actor, the people is described as exhibiting a common understanding of the world. This may be expressed by common feelings (such as fear or joy), common opinions (approvement / disapprovement), or a common will (demands, plans, orders to the government).

Important to this category is the absence of any relativization to the extent of this common feeling. There is no mention of certain shares of the people that think otherwise or would disagree.

Wir haben vom Schweizer Volk den Auftrag bekommen.

People fear that there may be more cuts.

Diverse feelings, desires, opinion

This category applies as soon as there are any relativizing formulations which point toward a fractured opinion within the population. This may be either a qualitative information (e.g. some people / most of the people / many citizens) or quantitative information (e.g.: 57% of the people). This information to the extent of a common sharing of any feeling, desire or opinion implies that there are some individuals that might disagree.

The majority of the Swiss population voted for the ban of minarets.

Studies show that 30% of the population supports Chancellor Osborne.

V06_Power: Attribution/Description of power

Values: (Rating: 1=Attribution, -1=Denial, 2=Should, -2=Should not, 9=Not mentioned)

Definition:

Next to the impact of actors, their power may also be discussed or evaluated by the speaker. Target actors may be described to gain, lose, or have power. At the same time

the speaker may evaluate the right of certain target actors by stating that they should have more or less power.

Question: Is the power of the target discussed or described by the actor?

Gain/ Increase Power

Gaining or increasing power may be coded if a speaker explicitly argues that an actor (mostly politicians or parties) has to gain or increase his power. This may be coded if the actor either has no power at the moment but should gain this in the future or if the actor already has power but should reinforce this.

“I am (Emmanuel Atefuah Danso) still holding on to that to seek for renewal of mandate to work harder for the party...all we want is to see NPP gain political power again in 2016 to rescue Ghanaians from the slavery imposed on them by NDC”, he disclosed.”

Lose Power

In contrast to gaining power this attribution may be coded if a speaker explicitly states that the target actor should lose his power. For this attribution the actor needs to be in power at the moment.

Die von der Islamischen Republik finanzierte und hochgerüstete libanesische Hisbollah hat im syrischen Bürgerkrieg an der Seite Assads bereits mehr Kämpfer verloren als 2006 im Krieg gegen Israel. Dieser Verschleiß droht die iranische Machtposition in der Region langfristig zu schwächen.

“Yesterday a Panelbase poll showed that on current voting intentions Scottish Labour would lose half its seats with the SNP adding an impressive 29 seats to the half dozen they currently hold.”

Have Power

The attribution of having power may be coded if a speaker explicitly argues that the actor has or should have power. In the Swiss political context speaker may attribute that the people should have power by stressing that the people should decide on a specific issue in the context of a “Volksinitiative”. The mere mention of the word “Volksinitiatvie” however is not enough to be coded as having power.

Doch gibt es vorläufig wenig Anzeichen dafür, dass Janukowitschs Machtposition irgendwie angeschlagen wäre.

„Jetzt muss aber sofort eine neue Volksinitiative gegen das Armeebudget her.“

“Citizens should have the power to call constitutional conventions”

Counter-Example:

„Ständerat Roberto Zanetti schätzt die Kosten einer Annahme der Volksinitiative auf 0,8 Prozent der jährlichen Lohnsumme.“

V07_Action: Attribution/Description of committed actions

Values: -1 Wrong stupid action/ Denial of successfull action, 1 important successfull action/denial of stupid action, 2 Everyday action

Definition:

Next to personality traits and supposed impact/accountability, certain actions may be attributed to an actor. An action is seen as the fact or process of doing something.

Typically this action is done to achieve a certain aim. In print articles an actor can be described doing certain things and only these descriptions will be coded.

Question:

Which of these actions or statements are attributed to the target?

Wrong, stupid action or mistake

An actor can be said to have made a singular mistake or several mistakes in a row. This has to be mentioned explicitly by a speaker.

"Der viel gescholtenen schwarz-gelben Regierung muss man ein Kompliment machen, dass sie wie in der Atompolitik nicht den Weg des geringsten Widerstands gegangen ist, sondern mit einer denkbar maßvollen Erhöhung der Regelsätze um wohl zehn Euro den Transferempfängern signalisiert, dass sie mit staatlich garantierten Lohnzuwächsen für das

Nichtstun nicht rechnen dürfen." (Attribution of committing correct action)

„Der IMF hat zugegeben, dass in den Rettungsprogrammen für Griechenland Fehler begangen wurden.“

"Sowing that kind of doubt so soon in his term was a mistake that even free streetcar rides for tots should not obscure."

Immoral/Indecent Behavior

Immoral or indecent actions are considered to contravene against prevailing norms and values yet not against current law. They are morally questionable and thus mostly evaluated negatively. An actor involved in intrigues would for example act immorally. An intrigue is dealt with when an actor makes secret plans to do something detrimental to someone. This is considered as acting morally wrong but not necessarily illicit (in contravention of the law). In order for this category to apply, the speaker has to condemn the action explicitly as immoral, of low moral standards, or morally wrong. The words used for this explication may differ but there has to be a condemnation of the action.

„Union und FDP sollten sich, historisch betrachtet, lieber in mehr Demut üben. Vor allem diese Parteien duldeten und deckten über Jahrzehnte Prominente in ihren Reihen, die Mittäter in der Nazizeit gewesen waren.“

"Die SPD verkauft sich als Moralapostel-Partei und faellt der eigenen Landesregierung in den Rücken, was schon schäbig genug ist."

"Ein Politiker, der solche Vorschläge macht, ist unredlich und täuscht vorsätzlich die Öffentlichkeit."

Um von dort ganz unchristlich gegen böse Ausländer zu pöbeln, die Eure schönen Straßen einfach so befahren, ohne dafür zu zahlen. Dafür hat der Führer damals schließlich nicht die Autobahnen gebaut! Kurz gesagt: Wenn Ihr bis in alle Ewigkeit ungestört weiterregieren wollt, schert Euch einfach einen Dreck um Eure eigenen „Werte“. Lasst die Leute wählen, schimpfen, saufen, fressen und vögeln, wie sie wollen. Dann habt Ihr's bald so schön wie in Bayern.

"The once in charge should better practice humility instead of promising things they can never achieve."

Counterexample:

"Wir können dann genüsslich zusehen, wie die Frau ohne Charakter Widmer-Schlump noch und noch weitere Bücklinge macht und den Deutschen in den Hintern kriecht, unter Begleitung des Gesamtbundesrats, dieses Verräterrats." (No attribution of immoral action since there is no reference to moral, values, or the violation of any norms. Although the action is clearly immoral, there is no explicit passage within the text).

Unser Premier ist ein Lügner und Heuchler, und die Leute haben genug von seiner Art zu politisieren. (attribution of negative traits)

„Putin ist ein Verbrecher“ (attribution of negative traits)

Criminal or unlawful behavior

Criminal or unlawful actions include any forms of behavior on behalf of an actor which is in conflict with current law. This behavior may range from misdemeanor to major crimes

committed by any actor and brought into the discussion. By labeling the actor as a criminal his arguments are rejected.

"Blocher hätte diese Unterlagen nie besitzen dürfen. Das war ein klarer Rechtsverstoss"

"I can not believe that he is still in office – after all the criminal activities he has been charged with."

Important, successful, right action

An actor may be said to have done something good, important or beneficial. This also includes having been successful at an important task.

On The Andrew Marr Show on BBC1, Gove said: "I absolutely agree with [Cameron] and I do think it's practical politics. The prime minister has struck exactly the right note on migration which is to celebrate the achievements of people who have come here, to recognise that migration has to work for those who are already here."

Jetzt bleibt Europa keine andere Wahl, als Strukturreformen durchzuführen und die Hemmnisse für ein Wachstum der Produktivität zu entfernen. Spanien, Portugal und Griechenland haben hier große Schritte unternommen und ernten bereits die ersten Früchte.

V08_Person: Attribution of personality traits

Values: 9 none, -1 negative personality traits, 1 positive personality traits

Definition:

A speaker may attribute or deny positive or negative personality traits to another actor. By explicitly linking personality traits to the description of an actor, the speaker is able to shape the image of this actor in the media. Attributions of personality traits are necessary in the measurement of populist strategies such as "Stressing virtues" and "Discrediting".

Positive Personality traits:

Good Character, Virtuousness, Morality, Benevolence

This category of positive personality traits encompasses all traits which describe the character of the actor. This includes both internal character traits such as character, manners, or morality and external character traits such as friendliness or benevolence.

Likewise, expressions of a good character may be coded in this category. These expressions may be also an abstinent living (e.g. no drugs / no liaisons) or dutiful living (e.g. being a good citizen).

It's a beautiful place full of friendly, hard-working people and that's what makes its financial situation so incredibly sad." (Attribution of good character)

Denial of these personality traits is not equivalent with the attribution of negative personality traits. Denial of good character may be coded when the good character of a person is seriously doubted by the speaker.

"Wir können dann genüsslich zusehen, wie die Frau ohne Charakter Widmer-Schlump nach und nach weitere Bücklinge macht und den Deutschen in den Hintern kriecht, unter Begleitung des Gesamtbundesrats, dieses Verräterrats." (Denial of good character)

Charisma, Positive Appearance, Sympathy

This category of personality traits encompasses the appearance of actors and the effect they have on other people. These traits range from purely physical positive appearance (attractiveness, sexual appeal) over general good impressions (sympathy, likability) to charisma (i.e.: Compelling attractiveness or charm that can inspire devotion in others).

"If you are looking for the most powerful EU leader to be a deal maker, then she (Merkel) is the perfect leader. The more you talk to people about her the more you like her. I warmed to her enormously." (Attribution of a positive effect on people / Likeability)

Anstößig an diesem zuletzt von Peer Steinbrück verliehenen Titel ist allenfalls, dass er das

Phänomen Silvio Berlusconi völlig verfehlt. Denn in Wirklichkeit ist der "Clown" ein Großbandit – ein veritables Genie der Korruption und der Verführung, ein skrupelloser und schillernder Machtmensch von jener Sorte, wie sie Shakespeare in seinen Königsdramen beschrieben hat. (Attribution of charisma, strong effects on others)

Counterexample:

Steinbrücks Schwäche sei dagegen, "dass er zu schnell redet. Wenn er langsamer reden würde, käme er noch besser an." (No explicit attribution of positive effect on others. Although there may be an implicit hint in the fact that he would do 'better'. So he is already good.)

"I really like that she (Merkel) doesn't care about her image," Mr Marr said. (No explicit mention of her effect on others. This would be an 'other positive attribute')

Denial of these personality traits is coded when the absence of charisma, sympathy or positive appearance is attributed to an actor.

Common Sense, Intelligence, Understanding

This category of personality traits encompasses all traits which are connected with the knowledge, understanding or information of an actor. This may range from a general and unspecific intelligence (e.g. smartness, intelligence, common sense, education) to specific intelligence on an issue (e.g.: understanding, being informed)

Die AfD bezeichnet sich als Vertreterin eines „gesunden Menschenverstands“.
(Attribution of
common sense)

Denial of these personality traits may be coded when the lack of knowledge or understanding of an actor is explicitly named.

"Und der Sonntagsblick gibt sich noch zur Publikation eines solchen Quatsches hin. Wo bleibt die journalistische Hinterfragung?" (Denial of Knowledge of the facts)

*"Mainstream Britain still thinks that the UKIP leader Nigel Farage is not intelligent."
(Denial)*

Credibility, Trustworthiness, Loyalty

This category of personality traits encompasses all traits related to the question how much a person may be trusted or believed. This may be either in an unspecific context (e.g. general trustworthiness, loyalty, credibility) or linked to a specific issue on which a person may be trusted or believed in. Denial of these personality traits may be coded when the absence of credibility or trustworthiness is explicitly mentioned.

„Albig hält Stegner für einen der fähigsten Sozialdemokraten überhaupt. Er kenne auch kaum jemanden, mit dem er so erfolgreich und loyal zusammenarbeiten könne - ein dickes Lob für jemanden, der ihn im Kampf um die Spitzenkandidatur einst anging.“

*Die kulturpolitische Sprecherin der Berliner Grünen, Sabine Bangert, sagte dem rbb, Schmitz' Verhalten sei "für einen Staatsdiener absolut inakzeptabel". Er habe jede politische Glaubwürdigkeit verloren. "Er muss die Konsequenzen ziehen und jetzt zurücktreten.«
(Denial)*

Counter-Example:

"Und der Sonntagsblick gibt sich noch zur Publikation eines solchen Quatsches hin. Wo bleibt die journalistische Hinterfragung?"(No denial of credibility. This would merely implicit but has not been pointed out.)

"The researchers used their findings to assess images of the three main party leaders, and concluded that Gordon Brown had the least trustworthy appearance due to his "thick eyebrows, wide nose and the size of his mouth." (Denial)

Skills, Competence (e.g. Leadership skills)

This category of personality traits refers to the skills and abilities an actor possesses. These skills may lie in different fields of expertise. No field is excluded from the coding. Most frequently, however, leadership skills or rhetoric skills of a political actor will be named. Leadership skills refer to skills that are important for actions of leading a group of people or an organization. Rhetorical skills are skills that refer to the art of effective or persuasive speaking or writing, especially the exploitation of figures of speech and other compositional techniques.

„An Christian Lindner kommt niemand vorbei. Er ist jemand, der die Partei aus der Lethargie herausführen kann“, sagte Kubicki. (Attribution of Leadership Skills)

“If you are looking for the most powerful EU leader to be a deal maker, then she (Merkel) is the perfect leader. The more you talk to people about her the more you like her. I warmed to her enormously.” (Attribution of Negotiation and Leadership Skills)

Anstößig an diesem zuletzt von Peer Steinbrück verliehenen Titel ist allenfalls, dass er das Phänomen Silvio Berlusconi völlig verfehlt. Denn in Wirklichkeit ist der "Clown" ein Großbandit

ein veritables Genie der Korruption und der Verführung, ein skrupelloser und schillernder Machtmensch von jener Sorte, wie sie Shakespeare in seinen Königsdramen beschrieben hat. (Attribution of skills to corrupt and entice. Not really positive but it is explicitly described as a skill.)

Denial of Skills or competence is coded whenever the skills are seriously doubted or criticized by the speaker.

Steinbrücks Schwäche sei dagegen, "dass er zu schnell redet. Wenn er langsamer reden würde, käme er noch besser an." (Denial of Rhetoric Skills)

*„Her public speaking style is as inspiring as the eurozone quarterly growth figures.”
(Denial of rhetoric skills)*

"Rechnen ist wohl nicht ihre Stärke, wie?" (Denial of calculus skills)

"Und der Sonntagsblick gibt sich noch zur Publikation eines solchen Quatsches hin. Wo bleibt die journalistische Hinterfragung?" (Denial of journalist skills)

Consistency, Straightness, Strong Conviction

Target actors may be said to be straightforward and strong of conviction if they hardly ever change their mind and possess a certain (healthy) stubbornness. On the other hand, they may be denied this positive personality trait if they often change their minds, argue against their own position or take unpredictable positions.

"Er sei erstaunt, dass der Vorschlag ausgerechnet von Neuhaus komme. Im Herbst habe er sich in der Debatte zum Raumplanungsgesetz noch 'für das exakte Gegenteil' stark gemacht." (Denial of Consistency)

"Levine always stays true to his values. We may rest assured to have him on the side of the welfare reform" (Attribution of Consistency)

Negative Traits

Malevolence, viciousness, bad character

Malevolence is the state or condition of being malevolent (=having or showing a wish to do evil to others). Malevolence or hostility describes a mindset that is consciously bad. It can be attributed to all kind of actors – from single persons to groups and institutions. Other than corruption or criminality, malevolence is less concrete and not linked to specific actions. It rather describes a state of mind, a bad character the actor may possess.

"Ein Politiker, der solche Vorschläge macht, ist unredlich und täuscht vorsätzlich die

Öffentlichkeit." (Attribution of Viciousness/Immorality)

Anstößig an diesem zuletzt von Peer Steinbrück verliehenen Titel ist allenfalls, dass er das

Phänomen Silvio Berlusconi völlig verfehlt. Denn in Wirklichkeit ist der "Clown" ein Großbandit

ein veritables Genie der Korruption und der Verführung, ein skrupelloser und schillernder

Machtmensch von jener Sorte, wie sie Shakespeare in seinen Königsdramen beschrieben hat. (Attribution of Viciousness)

"Luis Guterrez, a Democrat, accused Republicans of "viciousness" for trying to make it easier to deport immigrants brought to the U.S. as children."

Denial of these personality traits may be used in defense against attributions by other actors.

Criminal, corrupt, unlawful

Criminal or corrupt personality does not refer to criminal actions the actor has committed but

explicitly to his character. Especially in the public debate on politics, politicians may be called 'corrupt' without explicitly accusing them of bribery. Corruptness and criminality is just seen as an especially bad mindset an actor may possess.

"Warum sollten wir wählen gehen, die sind doch eh alle korrupt und kümmern sich nicht um

uns", sind sich die drei Gäste am Stand des "China-Thai-Bistros" auf dem Parkplatz vor Norma und Lidl einig. (Attribution of corruption as a personality trait)

Anstößig an diesem zuletzt von Peer Steinbrück verliehenen Titel ist allenfalls, dass er das

Phänomen Silvio Berlusconi völlig verfehlt. Denn in Wirklichkeit ist der "Clown" ein Großbandit

ein veritables Genie der Korruption und der Verführung, ein skrupelloser und schillernder Machtmensch von jener Sorte, wie sie Shakespeare in seinen Königsdramen beschrieben hat. (Attribution of corruption)

“The perks Mr McDonnell enjoyed might not raise eyebrows in states like Illinois, whose governors have a well-known record of corruption.”

Laziness, inactivity, inertia

Laziness or inactivity is the quality of being unwilling to work or use energy. It thus refers to a

behavior of willingly not coming up to one's possibilities. It will be attributed more likely to persons than to institutions.

„Die Südländer gingen zu früh in Rente und leisteten sich zu viel Urlaub, rügte Merkel. Berlin

werde nur dann helfen, wenn "die anderen sich auch anstrengen". Vielen ihrer CDU-Anhänger sprach Merkel mit diesem recht unverhohlenen Vorwurf der Faulheit aus der Seele,,

"Our politicians are, of course, infamous for their greed, laziness, unfitness and being tightfisted."

Denial of these personality traits may be used in defense against attributions by other actors:

"Leider denken immer noch viele, dass alle Langzeitarbeitslosen dumm und faul sind und es sich in der Hängematte des Wohlfahrtsstaates gemütlich machen", sagt Dörre. (Denial of laziness and stupidity)

Counterexample:

„Auch Wochen nach den Enthüllungen über die Überwachung des Internets durch Geheimdienste verhält sich die Bundesregierung weitgehend passiv“ (Denial of other action)

Stupidity, lack of understanding, being wrong

This set of negative attributes refers to the knowledge or understanding of an actor.

Stupidity is a behavior that shows a lack of good sense or judgment. Stupidity refers to the mental capacity of an actor, thus next to “stupid” also expressions such as “incapable”, “ignorant” or “foolish” fall into this category, as well as several forms of insults. It is important though that an actor, and not an issue (e.g. an idea/proposal of an actor) is addressed. As stupidity is linked to cognition, this category is most likely to be coded for people or groups of people and less likely for institutions.

„Jetzt sind die in Bern ganz durchgeknallt“

Antal Rogan: “The European Union is incapable of protecting its member states, and for this reason these states, including Hungary, must protect themselves.”

Counterexample:

"Rechnen ist wohl nicht ihre Stärke, wie?" (No attribution of stupidity, just denial of skills)

Denial of a lack of understanding or intelligence may be coded in cases where such an allegation is dismissed.

"Leider denken immer noch viele, dass alle Langzeitarbeitslosen dumm und faul sind und es sich in der Hängematte des Wohlfahrtsstaates gemütlich machen", sagt Dörre. (Denial of laziness and stupidity)

Extremism

Extremism may be attributed to actors either in combination with being right- or left-wing or as an isolated property. Any kind of political or social extremism may be coded with this category. If it appears in combination with right/left-wing, both extremism and the respective political orientation is to be coded.

“CDU-Generalsekretär Hermann Gröhe nannte die Anschlagsserie ‘empörend und beschämend’. ‘Ich hoffe, dass die Täter, die offenkundig aus dem linksradikalen Lager stammen, schnell gefasst und zur Verantwortung gezogen werden.’ (Attribution of Extremism and Being left-wing)

“In my opinion the political elite is nothing else than purely extremist.”

Racism

Racism or Xenophobia may be attributed to actors. In order to code this category, an aversion of the target toward foreigners has to be made explicit by the speaker. This explication may contain words as 'racist', 'antisemite' but may also be stated soberly (e.g. 'He has a very negative attitude toward strangers').

„Edathy forderte ein NPD-Verbot, warnte vor No-Go-Areas. Und als sein Parteikollege Thilo

Sarrazin Thesen zur Einwanderung aufstellte, attestierte er ihm ‚Rassismus pur‘“

“The Conservative Party are 'essentially racist' according to former Tory aide.”

Undemocratic

Being undemocratic may be attributed to actors. In order to code this category, the undemocratic behavior of the target must be made explicit by the speaker.

„Der Linke-Parteivorstand hat am Wochenende eine Passage im Leitantrag zum Europa-Programm gestrichen. Darin hieß es, die EU sei zu einer ‚neoliberalen, militaristischen und

weithin undemokratischen Macht‘ geworden.“

"The Senate is undemocratic and it matters."

V09_Name: Name Calling

Values: -1 Yes, negative, +1 yes positive, 9 none/no clear valenz

Definition:

Sometimes speakers use nicknames to refer to certain actors, parties or organizations.

These names are recognizable by the audience and are most often judgmental. Nicknames may be any name an actor is called except from the official name of this actor. This includes diminutives and slight changes to the spelling of the name (e.g: "Charly Osborne"; "Barry O." for Obama; "Angie" for Merkel) as well as the attachment of pre- and suffixes (e.g: "Problem-Peer", "Mama Merkel", "Schmidt-Schnauze für Helmut Schmidt "Mövenpick-Partei" für die FDP) and the creation of names not related to the actual name (e.g: "Twerp from Downing Street" for Cameron).

Only name calling will be coded in this category. Any other insults are explicitly excluded from this definition.

Il Cavalliere

Mama Angie

Güllner erwirkte eine Unterlassungsverfügung dagegen und kündigte an, den Sprecher künftig „Lügen-Lucke“ zu nennen.