Argumentanalyse in digitalen Textkorpora

Butt, Miriam
Universität Konstanz, Deutschland
miriam.butt@uni-konstanz.de

Heyer, Gerhard
Universität Leipzig, Deutschland
heyerasv@informatik.uni-leipzig.de

Holzinger, Katharina
Universität Konstanz, Deutschland
katharina.holzinger@uni-konstanz.de

Kantner, Cathleen
Universität Stuttgard, Deutschland
cathleen.kantner@sowi.uni-stuttgart.de

Keim, Daniel A.
Universität Konstanz, Deutschland
daniel.keim@uni-konstanz.de

Kuhn, Jonas
Universität Stuttgard, Deutschland
jonas.kuhn@ims.uni-stuttgart.de

Schaal, Gary
Helmut-Schmidt-Universität, Universität der Bundeswehr, Hamburg
gschaal@hsu-hh.de

Blessing, André
Universität Stuttgard, Deutschland
andre.blessing@ims.uni-stuttgart.de

Dumm, Sebastian
Helmut-Schmidt-Universität, Universität der Bundeswehr, Hamburg
sebastian.dumm@hsu-hh.de

El-Assady, Mennatallah
Universität Konstanz, Deutschland
mennatallah.el-assady@uni-konstanz.de

Gold, Valentin
Universität Konstanz, Deutschland
valentin.gold@uni-konstanz.de

Hautli-Janisz, Annette
Universität Konstanz, Deutschland
annette.hautli@uni-konstanz.de

Lemke, Matthias
Helmut-Schmidt-Universität, Universität der Bundeswehr, Hamburg
lemkem@hsu-hh.de

Müller, Maike
Universität Konstantz, Deutschland
maike.mueller@uni-konstanz.de

Niekler, Andreas
Universität Leipzig, Deutschland
aniekler@informatik.uni-leipzig.de

Overbeck, Maximilian
Universität Stuttgard, Deutschland
maximilian.overbeck@sowi.uni-stuttgart.de

Wiedemann, Gregor
Universität Leipzig, Deutschland
gregor.wiedemann@uni-leipzig.de

Inhalt

1. Zusammenfassung der Sektion

Valentin Gold, Annette Hautli-Janisz, Andreas Niekler, Maximilian Overbeck und Gregor Wiedemann

Die Extrahierung und Annotation von Argumentationsstrukturen hat im Bereich der automatischen Diskursanalyse in den letzten Jahren an Bedeutung gewonnen, sei es in juristischen Dokumenten (Mochales / Moens 2011; Bach et al. 2013), wissenschaftlichen Texten (Kirschner et al. 2015), Zeitungsartikeln (Feng / Hirst 2011) oder Online-Diskussionen (Bex et al. 2013, 2014; Oraby et al. 2015). Vor diesem Hintergrund haben sich in den vergangenen Jahren die drei interdisziplinären Projekte *e-Identity*, *ePol* und *VisArgue* im Rahmen der eHumanities-Förderlinie des BMBF mit der semi-automatischen Identifikation und Analyse von Argumenten auseinandergesetzt.

Die Herausforderung, die allen Projekten gemein ist, ist die, dass die jeweilige Fragestellung über den eigentlichen Prozess der Argumentationsanalyse hinausgeht: Im Falle von *VisArgue* soll die Deliberativität des Diskurses approximiert werden, bei *ePol* geht um Ökonomisierungstechniken neoliberalen Sprechens, Begründens und Argumentierens in der politischen Öffentlichkeit und bei *e-Identity* um die Mobilisierung unterschiedlicher kollektiver Identitäten in politischen Debatten zu bewaffneten Konflikten und humanitären militärischen Interventionen. Daher sind diese Projekte beispielhaft für die Anforderung der eHumanities: Trotz des gemeinsamen Zieles der Argumentationsextraktion wird der Begriff des Arguments und dessen Rolle in den einzelnen Projekten konzeptionell sehr verschieden gefasst und muss daher im Hinblick auf die jeweilige inhaltliche Fragestellung und die zu untersuchende Datenbasis unterschiedlich operationalisiert werden.

Den Kern im Projekt *VisArgue* bildet die Extrahierung von kausalen und adversativen Argumentstrukturen (Bögel et al. 2014), um Instanzen von Begründungen, Schlussfolgerungen und Gegenargumenten im Diskurs herausfinden zu können. Dies geschieht mithilfe eines linguistisch motivierten, regelbasierten Systems, das explizite Diskurskonnektoren automatisch disambiguiert und die Teile des Arguments im Diskurs verlässlich annotiert. Diese Annotationen dienen als Basis für die Visualisierung von deliberativen Mustern über den Diskurs hinweg und die damit einhergehende Interpretation desselben. Im Gegensatz zur regelbasierten Extrahierung werden im Projekt *ePol* maschinelle Lernverfahren angewandt, die jene Abschnitte in Zeitungstexten für eine inhaltsanalytische Auswertung identifizieren, die sprachliche Muster ökonomisierter Begründungen für Politik enthalten. Allerdings finden sich in Zeitungstexten nur sehr wenige explizite Argumentstrukturen, die einer formalen Anforderung expliziter Formulierung von beispielsweise Prämisse, Schlussregel und Schlussfolgerung genügen. Muster der hier eher implizit enthaltenen Begründungsstrukturen können anhand einer Menge von annotierten Beispielargumenten gelernt und zur Identifikation ähnlicher Textabschnitte angewendet werden, ohne dass eine bestimmte Form der Argumente explizit vorgegeben wird. Im *e-Identity* Projekt wurden die Potentiale für computer- und korpuslinguistische Methoden erschlossen, die eine interaktive und flexible Tiefenanalyse der Mobilisierung unterschiedlicher kollektiver Identitäten in öffentlichen Debatten über Krieg und Frieden ermöglichen. Maschinelle Lernverfahren kamen dabei sowohl bei der inhaltlichen Bereinigung der mehrsprachigen Textkorpora sowie bei der halb-automatischen Identifikation der unterschiedlichsten kollektiven Identitäten zum Einsatz.

In dieser Sektion wird daher der Frage nachgegangen, wie unterschiedliche theoretische und methodische Ansätze für die (semi-)automatische Identifikation und Analyse von Argumenten eingesetzt werden. In den Vorträgen werden die heterogenen Ansätze vor dem Hintergrund der jeweiligen Fragestellungen und daraus resultierender Anforderungen einzelnen eHumanities-Projekte im Detail vorgestellt. Dabei liegt der Fokus der Vorträge auf den Anwendungen, Ergebnissen und auf Perspektiven für die Evaluation. Insbesondere der Gütekontrolle räumen die Vorträge mehr Raum ein, um die Leistungsfähigkeit unterschiedlicher Ansätze und die Auswirkung auf Ergebnisse transparent darzustellen. Als prototypische Anwendungen von Argumentanalysen in den Humanities zeigen die Vorträge methodische Perspektiven und Ideen für Verwendungsmöglichkeiten jenseits der vorgestellten Projekte.

2. Vortrag 1: Deliberation in politischen Verhandlungen: Eine linguistisch-motivierte visuelle Analyse

Mennatallah El-Assady, Valentin Gold, Annette Hautli-Janisz, Maike Müller, Miriam Butt, Katharina Holzinger, Daniel A. Keim

2.1. Einleitung

Das *VisArgue* Projekt hat zum Ziel, automatisch zu erfassen, ob Verhandlungsteilnehmer deliberativ agieren, d. h. ob sie ihre Positionen u. a. respektvoll und rational begründen und sich schlussendlich dem besten Argument fügen. Die Datenbasis sind dabei transkribierte reale Verhandlungen, wie zum Beispiel die Schlichtungsgespräche zu Stuttgart 21. Zusätzlich zur Erfassung von Argumentationsmustern spielen bei der Argumentanalyse auch noch andere Faktoren eine Rolle, insbesondere die Beziehung des Sprechers zum Gesagten, die Beziehungen der Sprecher untereinander und die Struktur der Diskussion insgesamt. Mithilfe eines innovativen Visualisierungssystems werden diese vielschichtigen Muster aufgearbeitet, damit die einzelnen Faktoren von Argumentation, aber auch die Beziehungen der einzelnen Faktoren untereinander, interpretierbar gemacht werden können.

In diesem Beitrag wird am Beispiel der automatischen Erfassung von Argumentationsmustern aufgezeigt, wie das Projekt mit den generellen Herausforderungen der eHumanities umgeht: Das Konzept der Deliberation ist (computer)linguistisch gesehen eher abstrakt

und bedarf einer konkreten Operationalisierung, damit das Konzept in den Daten fassbar gemacht wird. Die Zusammenhänge zwischen den verschiedenen Faktoren, die den Diskurs bestimmen, werden dann mithilfe eines Visualisierungssystems interpretierbar gemacht.

Im Folgenden werden die verschiedenen Dimensionen von Deliberation vorgestellt, gefolgt von einer Beschreibung der automatischen Argumentationsextraktion und der Annotation anderer deliberationsrelevanter Merkmale. Abschließend wird anhand eines konkreten Beispiels gezeigt, wie das Visualisierungssystem die Interpretation von Argumentationsmustern im Diskurs erlaubt.

2.2. Die Operationalisierung des Konzeptes der Deliberation

Das Konzept der Deliberation wird, wie in der folgenden Abbildung gezeigt, operationalisiert durch vier Dimensionen, die für die automatische Extraktion deliberativer Muster im Text relevant sind: Teilnahme (Participation), Atmosphäre und Respekt (Atmosphere & Respect), Argumentation und Rechtfertigung (Argumentation & Justification) und Entgegenkommen (Accomodation) (Gold / Holzinger 2015). In der Dimension 'Argumentation & Justification' werden unter anderem kausale Argumentationsketten annotiert, die darauf hindeuten, dass die Teilnehmer im Prozess der Entscheidungsfindung sind und Argumente austauschen ('Reason-giving'). In der Subdimension 'Information Certainty' wird auf der Basis von Ausdrücken epistemischer Modalität wie 'mit Sicherheit', 'wahrscheinlich' etc. annotiert, wie sicher sich die Sprecher des Gesagten sind. In der Dimension 'Accomodation' werden solche Einheiten im Diskurs annotiert, die entweder auf eine Einigung in der Verhandlung abzielen oder eine Uneinigkeit bekräftigen. Informationen, ob Sprecherbeiträge emotional oder sachlich sind, ob Sprecher andere Redner unterbrechen oder ob sie sich höflich verhalten, werden in der Dimension 'Atmosphere & Respect' gebündelt.

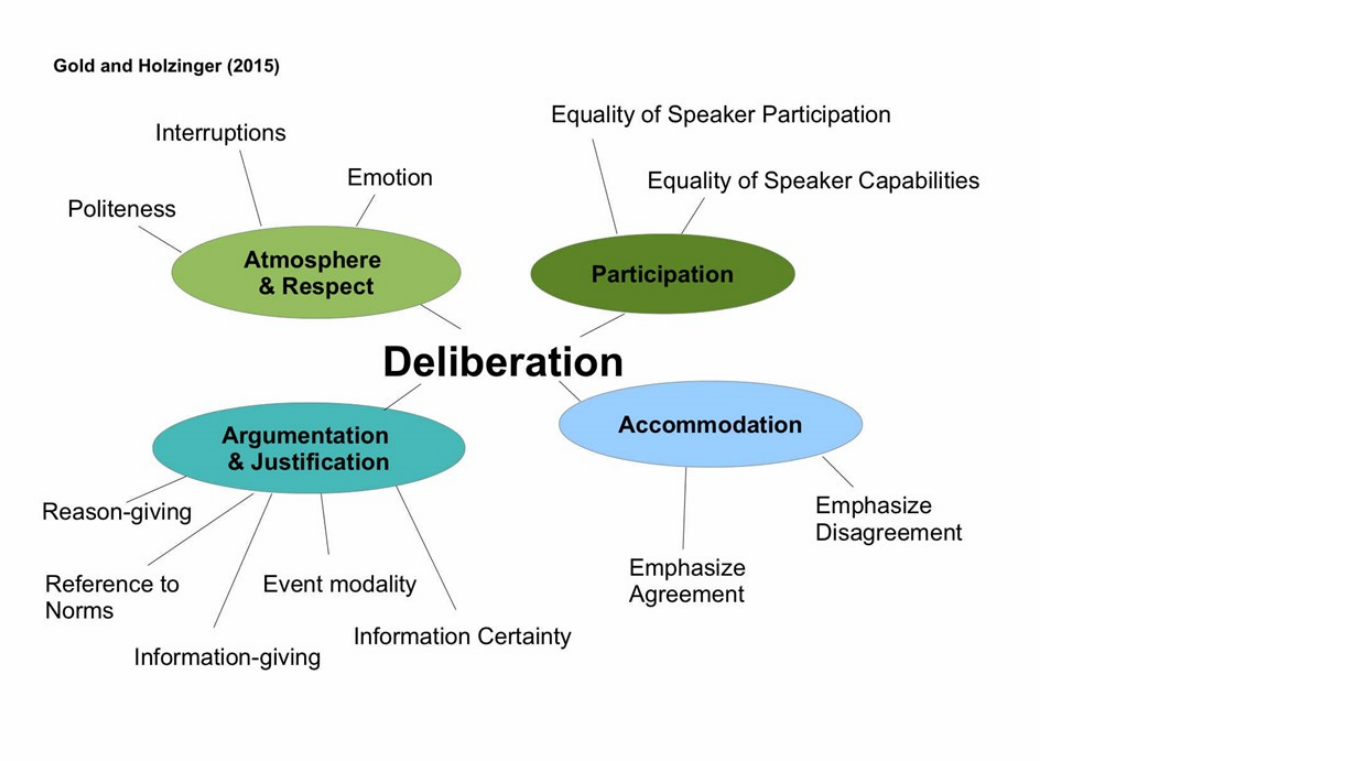

Figure 1. ***Abb. 1****: Dimensionen der Deliberation*

Auf Basis dieser Konkretisierung des Begriffs der Deliberation wird im Folgenden anhand der Dimensionen 'Argumentation & Justification' und 'Accomodation' gezeigt, wie die verschiedenen Ebenen innerhalb des Diskurses konkret annotiert werden. Zusammengenommen dienen diese Annotationen als Basis für die Visualisierung, um die Muster im Diskurs im Sinne der Deliberation interpretieren zu können.

2.3. Argumenterfassung

Als Datenbasis dienen transkribierte Verhandlungen, die entweder in projektinternen Verhandlungssimulationen gewonnen wurden oder von realen politischen Verhandlungen stammen, wie z. B. dem Schlichtungsverfahren von Stuttgart 21. Diese Daten werden in ein XML Schema übertragen, auf dessen Basis der Diskurs annotiert wird. Dazu werden die Äußerungen der Teilnehmer in Sätze aufgeteilt, die wiederum in kleinere Einheiten, sogenannte "elementary discourse units (EDUs)" eingeteilt werden, unter der Annahme, dass jede dieser Diskurseinheiten ein Event darstellt (Polanyi et al. 2004).

Ein Modul in der Annotation ist die Extrahierung von kausalen Argumentstrukturen (Bögel et al. 2014), was mithilfe eines linguistisch motivierten, regelbasierten Systems geschieht, das explizite Diskurskonnektoren automatisch disambiguiert und die einzelnen Teile eines Arguments im Diskurs verlässlich annotiert. Kausale Diskurskonnektoren wie 'weil', 'da' und 'denn' etc. leiten die Begründung einer Schlussfolgerung ein und geben so Hinweise auf argumentative Phasen in der Diskussion. Diese Informationen sind Teil der Ebene 'Reason-giving' in der Dimension 'Argumentation & Justification'. Im Gegensatz dazu stehen adversative Konnektoren wie 'aber', 'allerdings', 'jedoch' etc., die eine gegensätzliche Aussage zum Hauptsatz zum Ausdruck bringen und eine Ablehnung des Sprechers indizieren. Diese Äußerungen sind Teil der Subdimesion 'Emphasize Disagreement' in der Dimension 'Accomodation'.

Für die automatische Annotation derjenigen EDUs, die Teil der kausalen oder adversativen Einheiten bilden, werden den EDUs verschiedene Werte des XML Attributes 'discrel' zugeordnet, zum Beispiel discrel="reason" und discrel="conclusion" für kausale

Argumentationsketten und discrel="opposition" für adversative Strukturen.

Zusätzlich zu der Information, dass die Teilnehmer Argumente austauschen oder sich zustimmend oder ablehnend in einer Diskussion verhalten, wird in der Unterdimension 'Information Certainty' in 'Argumentation & Justificiation' herausgearbeitet, wie sicher sich der Sprecher mit dem Inhalt seines Beitrages ist, d. h. welchen Kenntnisstand er vorgibt zu haben. Dies wird sichtbar durch Ausdrücke epistemischer Modalität, wie zum Beispiel 'wahrscheinlich', 'vielleicht' oder 'mit Sicherheit'. Um deren Bedeutung messbar zu machen, wird die Skala von Lassiter (2010), der die sogenannten "modes of knowing" von 0 (unmöglich – impossible) bis 1 (mit Sicherheit – certain) quantifiziert, herangezogen, und entsprechend annotiert: Der epistemische Ausdruck wird auf der Lexem-Ebene identifiziert und seine Bedeutung auf der Ebene der EDU mit dem XML-Attribut 'epistemic_value' versehen.

Ein weiterer Faktor, der für Deliberation relevant ist, ist die Haltung des Sprechers zum Gesagten. Dabei bleibt der Wahrheitsgehalt der Aussage unberührt, aber der Sprecher zeigt, wie er sich im Diskurs positioniert. Diese pragmatisch-relevante Ebene, die aus theoretisch-linguistischer Sicht schon vielseitig analysiert wurde, wird insbesondere von Partikeln wie 'ja', 'halt' und 'doch' ausgelöst (u. a. Kratzer 1999; Karagjosova 2004; Zimmermann 2011) und ist linguistisch gesehen eine konventionelle Implikatur ('conventional implicature') (Potts 2012). Eine Herausforderung ist die Ambiguität der Partikel in der gesprochenen Sprache. Beispielsweise wird 'ja' häufig dazu verwendet, das gemeinsame Wissen der Diskussionteilnehmer zu betonen, auch verstanden als 'common ground' ("Sie wissen ja, dass …."). Allerdings kann 'ja' auch noch Zustimmung oder Ungeduld ("ja ja…") signalisieren, oder aber Hinhaltetechnik sein ("ja [Pause] ja"). Mithilfe eines regelbasierten Systems, das den Kontext vor und nach den Partikeln untersucht, werden die unterschiedlichen Bedeutungen herausgefiltert und als konventionelle Implikatur (CI) annotiert.

Diese Ebenen, die die klassische Argumentationsstruktur komplettieren, sind hochrelevant für die Analyse im Sinne der Deliberation: Neben der Frage, ob und wann argumentiert wird, ist auch noch relevant, WIE argumentiert wird: Argumentiert der Sprecher auf der Basis gemeinsamen Wissens (common ground), oder ist er sich seiner Schlussfolgerung sicher? Die Visualisierung muss daher die verschiedenen Bedeutungsebenen, die für die Herausarbeitung deliberativer Muster relevant sind, einzeln, aber auch im Zusammenspiel darstellen. Dazu wird im Folgenden das VisArgue Visualisierungssystem vorgestellt und gezeigt, wie Muster von Argumentationsstrukturen und Sprecherhaltung visuell über den Diskurs hinweg dargestellt werden können.

2.4. Visualisierung

Neben der Visualisierung von thematischen Blöcken in politischen Verhandlungen (Gold / Rohrdantz et al. 2015; Gold / El-Assady et al. 2015), ist ein Ziel der Visualisierung, Muster von Deliberation über den Diskurs hinweg, aber auch aggregiert für einzelne Sprecher so darzustellen, dass die zugrundeliegenden Daten, aber auch das große Ganze sichtbar wird. Eine Herausforderung ist hierbei die Mehrdimensionalität der Information, da zum einen die Ebene OB argumentiert wird, zum anderen aber auch die Information WIE argumentiert wird, visuell dargestellt werden soll. Dazu wird beispielsweise die Argumentationsdichte mit den Partikeln gemeinsam visualisiert: Jede Äußerung eines Sprechers wird als Glyph (Abbildung 2) dargestellt, wobei die Größe des Glyphen bestimmt wird durch die Länge der Aussage. Innerhalb des Glyphen sind die verschiedenen Werte der konventionellen Implikaturen abgetragen. Die zwei äußeren Ringe um den Glyphen zeigen Argumentationsmuster von 'reason' und 'conclusion' in einer Äußerung an; je größer die Teilringe, desto mehr EDUs sind Teil einer kausalen Argumentation. Die zugrundeliegende Äußerung kann mit einem Doppelklick auf den Glyph eingesehen werden (Abbildung 3).

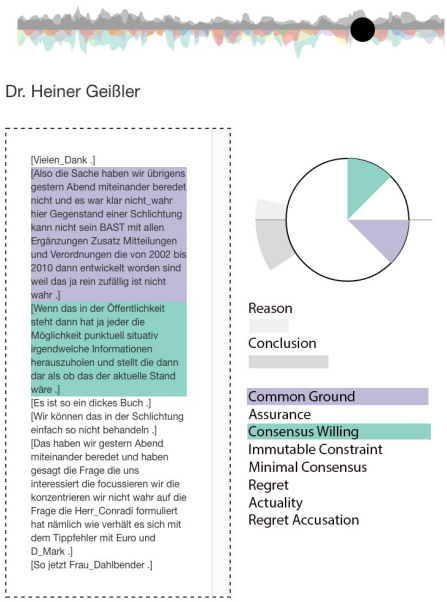

Figure 2. ***Abb. 2****: Detailansicht Glyph*

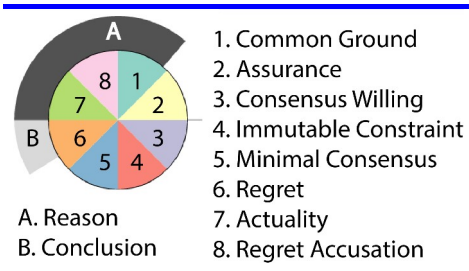

*Figure 3. **Abb. 3**: Glyphdarstellung von Argumenten*

Diese Glyphen werden für jede Äußerung über den Diskurs hinweg erstellt, wie in Abbildung 4 gezeigt. Durch die Interaktivität lässt sich ein Überblick über die Äußerungen der einzelnen Sprecher herstellen, und dabei Rückschlüsse ziehen, welche Rolle ein Sprecher in der Verhandlung eingenommen hat. Abbildung 4 zeigt einen Verhandlungstag der Schlichtungen zu Stuttgart 21: In der obersten Zeile findet sich der Mediator der Schlichtung, Dr. Heiner Geißler, wieder, dessen Beiträge von einem hohen Maß an 'consensus willing' und 'common ground' geprägt sind, wobei relativ wenig Argumente angeführt werden. Im Gegensatz dazu zeigt einer der Befürworter des Projektes, Dr. Volker Kefer, ein anderes Muster auf, nämlich einen deutlich höheren Anteil an argumentativen Redebeiträgen, die geprägt sind von Zusagen ('assurance') und unabänderlichen Vorgaben ('immutable constraint'). Der hohe Grad an kausaler Argumentation findet sich auch bei einem Gegner des Projektes, Boris Palmer, der sich in seinen Beiträgen mehrheitlich auf den 'common ground', d. h. das gemeinsame Wissen der Verhandlungsteilnehmer, beruft.

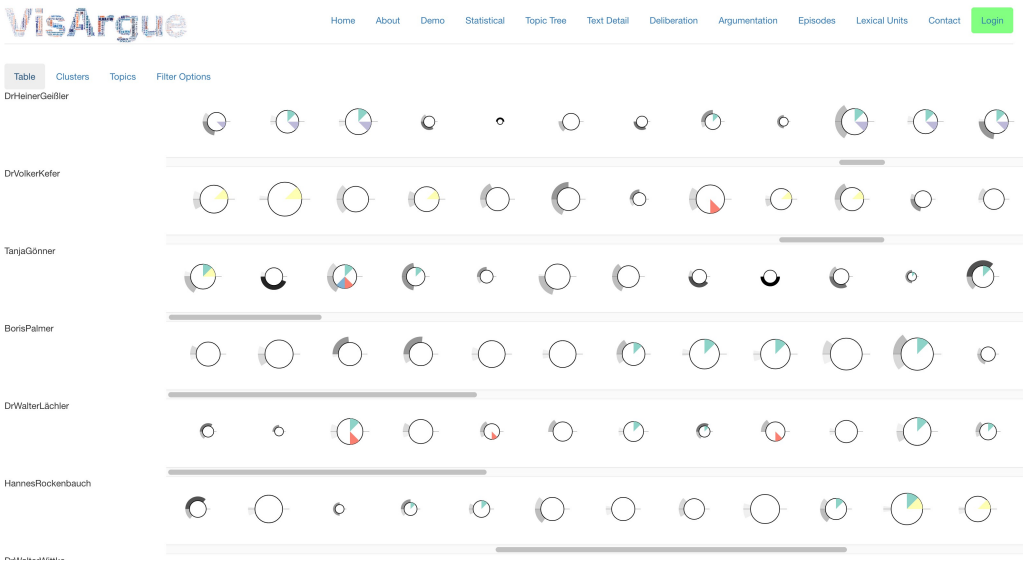

*Figure 4. **Abb. 4**: Glyphen pro Sprecher*

Im Sinne der Deliberation sind diese Muster relevant, weil sie zeigen, dass Verhandlungsteilnehmer verschieden argumentieren und sich damit unterschiedlich in der Verhandlung positionieren. Diese Muster tragen wesentlich dazu bei, den Verlauf und den Ausgang der Verhandlung zu erklären und Segmente von intensiven deliberativen Debatten zu identifizieren.

Zukünftige Arbeiten werden sich insbesondere mit dem Thema befassen, wie die weiteren Dimensionen der Deliberation in die Glyphenstruktur eingearbeitet werden können und inwiefern die linguistische Analyse weitere Anhaltspunkte von Argumentation und ihre Ausprägung aus dem Text extrahieren kann.

2.5. Zusammenfassung

Das *VisArgue*-Projekt zeigt am Beispiel der Argumentationserfassung, wie ein Ziel der Digital Humanities erreicht werden kann, nämlich der interdisziplinäre Austausch von Konzepten und Methoden: Durch die Kooperation von Politikwissenschaft, Linguistik und Informatik werden regelbasierte Analyse und visuelle Darstellung kombiniert und dadurch eine valide Basis für die Interpretation von Deliberation in politischen Verhandlungen möglich.

3. Vortrag 2: (Semi)-automatische Klassifikation für die Analyse neo-liberaler Begründungen und Argumentationen in großen Nachrichtenkorpora

Sebastian Dumm, Matthias Lemke, Andreas Niekler, Gregor Wiedemann, Gerhard Heyer, Gary S. Schaal,

Für die Analyse großer Mengen qualitativer Textdaten stehen den Sozialwissenschaften unterschiedliche konventionelle und innovative Methoden der Inhalts- und Diskursanalyse zur Verfügung. Die klassische sozialwissenschaftliche Inhaltsanalyse kann methodisch mit Verfahren des überwachten maschinellen Lernens verbunden werden (Scharkow 2012). Zur effizienten Generierung von Trainingsbeispielen kann eine solche (semi-)automatische Textklassifikation zu einem Active Learning Prozess erweitert werden (Dumm / Niekler 2015). Dabei werden schrittweise vom Computer vorgeschlagene Textbeispiele als Kandidaten für eine inhaltsanalytische Kategorie manuell evaluiert, und der Klassifikationsprozess mit den manuell bewerteten Beispielen erneut ausgeführt. Auf diese Weise können schnell mehrere hundert repräsentative Beispiele für eine Kategorie in großen Textkollektionen identifiziert werden. Ein solches Untersuchungsdesign ist im Rahmen des Projekts „ePol - Postdemokratie und Neoliberalismus“ methodologisch entworfen und technisch umgesetzt worden (Wiedemann et al. 2013). Der Vortrag beschreibt Ergebnisse und Lessons Learned aus diesem Projekt.

Das Projekt *ePol* greift die politiktheoretische Diskussion um die Erscheinungsformen gegenwärtiger westlicher Demokratien auf, welchen mit dem Konzeptbegriff Postdemokratie unter anderem eine Ökonomisierung des Poltischen unterstellt wird. 1 Die Ökonomisierung in den Begründungen politischer Entscheidungen untersuchen wir anhand von Sprachgebrauchsmustern in der politischen Öffentlichkeit, speziell in einem Korpus aus 3,5 Millionen Artikeln deutscher Tages- und Wochenzeitungen im Zeitraum von 1949 bis 2011. Unter neoliberalem Plausibilisieren verstehen wir dabei „Ökonomisierungstechniken“, die Argumente, Behauptungen und Metaphern zur Legitimation von politischem Output einsetzen und somit zum öffentlichen Sprachspiel der Politik gerechnet werden können. Den Gebrauch solcher qualitativer Begründungsmuster quantitativ im Zeitverlauf zu verfolgen und dessen Zu- oder Abnahme in Bezug auf bestimmte Randbedingungen zu testen (z. B. Zeitung oder Politikfeld) ist Ziel des Projekts. Dazu wurde ein modulares

Forschungsdesign in drei Schritten umgesetzt:

- Selektion relevanter Artikel aus dem Korpus von 3,5 Millionen Artikeln, welche eine hohe dichte an neoliberalen Begründungsmustern erwarten lassen,
- Manuelle Annotation von Textstellen, welche neoliberale Begründungsmuster enthalten. Unterschieden werden zwei Kategorien von Ökonomisierungstechniken, die des Argumentierens und die des Behauptens.
- Automatische Klassifikation der beiden Kategorien auf dem Gesamtdatenbestand zur Identifikation von Trends im Sprachgebrauch ökonomisierter Begründungen.

3.1. Selektion relevanter Artikel

In einem ersten Schritt wird eine Dokument-Retrieval-Strategie auf das gesamte Korpus angewendet, um Artikel mit (potenziell) möglichst hoher Dichte an neoliberalem Sprachgebrauch und Begründungsmustern zu identifizieren. Die Dokumente werden mit Hilfe eines einfachen Wörterbuches von 127 Argumentmarkern (Dumm / Lemke 2013) und eines kontextualisierten Wörterbuches (Wiedemann / Niekler 2014) nach Relevanz bewertet. Das kontextualisierte Wörterbuch enthält typischen Sprachgebrauch, der aus 36 in deutscher Sprache verfügbaren Schriften der Mitglieder des neoliberalen Think Tanks „Mont Pélerin Society“ extrahiert wurde. Dies umfasst eine Liste mit 500 Schlüsselbegriffen (z. B. Markt, Freiheit, Preis) sowie Statistiken über deren typische Kontexte (z. B. persönliche Freiheit, unternehmerische Freiheit). Die Berechnung eines Ähnlichkeitsmaßes des Sprachgebrauchs in diesem Vergleichskorpus mit den Artikeln aus unserem Zeitungskorpus hinsichtlich neoliberaler Sprachgebrauchsmuster und Argumentmarker führen zu einer sortierten Liste von Artikeln, welche als Ausgangspunkt für den Prozess der (semi-)automatischen Kodierung dient. Die 10.000 höchst bewerteten Dokumente werden für die Folgeschritte selektiert.

3.2. (Semi-)autiomatische Kodierung als Active Learning

Nachrichtenartikel enthalten für gewöhnlich nur wenige detaillierte und elaborierte argumentative Strukturen, welche den formalen Anforderungen einer vollständigen Argumentation folgen. Aus diesem Grund betrachten wir zwei Kategorien von Begründungsmustern: Argumente und Plausibilisierungen in neoliberalen Begründungszusammenhängen. Diese Kategorien werden in einem theoretisch begründeten Codebuch formal definiert. Im Gegensatz zu Argumenten, welche die Vollständigkeit von Argumentationsmustern durch Vorhandensein von Prämisse, Kausalmarker und Schlussfolgerung voraussetzen, sind Plausibilisierungen durch Behauptungen und idiomatische Referenzen auf vermeintlich akzeptiertes Wissen gekennzeichnet (z. B. „Tatsache ist ...“, „selbstverständlich“). Anschließend werden in den 100 relevantesten Artikeln aus Schritt 1 Textstellen annotiert, die den Codebuch-Definitionen entsprechen. Zur Überprüfung der Qualität der Codebuch-Definitionen und der Arbeit der Kodierer kann die Intercoder-Reliabilität bestimmt werden – ein Maß, welches die (zufallsbereinigte) Übereinstimmung zweier Kodierer auf demselben Text angibt. Insofern es sich bei den in unserem Projekt verwendeten Kategorien um zwei recht abstrakte Konzepte handelt, sind die Übereinstimmungsmaße eher am unteren Ende der akzeptablen Werte für eine verlässliche Kodierung angesiedelt. Im Gegensatz zu typischen Kodes wie Thema oder Affektposition wird hier die Schwierigkeit bei der Operationalisierung komplexer politiktheoretischer Konzepte deutlich. Insbesondere die Kategorie des Behauptens zeichnet sich durch eine große sprachliche Varianz aus, welche sowohl manuelle als auch automatische Kodiermethoden vor große Probleme stellt. Insofern es uns aber eher um die Bestimmung von Kategorieproportionen und Trends in sehr großen Datenmengen geht, als um die exakte Bestimmung von Einzelereignissen in den Daten, sind diese Ungenauigkeiten hinnehmbar. In diesem initialen Annotationsprozess wurden 218 Absätze mit Argumentationszusammenhang und 135 Absätze mit Plausibilisierungszusammenhang in den 100 relevantesten Artikeln annotiert.

Diese initiale Trainingsmenge muss für eine valide Trendbestimmung mit Hilfe automatischer Textklassifikation noch deutlich erweitert werden. Um effizient mehr gute, das heißt die Kategorien gut beschreibende, Textbeispiele zu finden, wird ein Active-Learning-Ansatz angewendet. Dazu wird ein maschineller Lernalgorithmus auf Basis der aktuell annotierten Textbeispiele trainiert und auf die noch nicht annotierten Dokumente aus den 10.000 zuvor selektierten, potenziell relevanten Dokumenten angewendet. Auf der technologischen Ebene nutzen wir eine Support Vector Machine (SVM) mit einem linearen Kernel. Wir extrahieren eine große Vielfalt von Texteigenschaften (Features) aus den Trainingsbeispielen, um den Klassifikationsprozess auch generisch für andere Probleme nutzen zu können. Die extrahierten Feature-Strukturen beinhalten Wort-N-Gramme, Part-of-Speech-N-Gramme und binäre Features über das Vorhandensein von Begriffen in unseren zwei initial erstellten Diktionären (neoliberaler Sprachgebrauch und Argumentmarker). Wir wenden eine Chi-Square Feature-Selektion an, um für die eigentliche Klassifikation nur Kategorie-relevante Features zu verwenden und übergeben die so vorverarbeitete Trainingsmenge an den Klassifikator. Der Klassifikator liefert eine Menge an Absätze aus den bislang ungesehenen Zeitungsartikeln zurück, welche eine hinreichende Ähnlichkeit in Bezug auf die Merkmalsstrukturen der bereits annotierten Artikel aufweisen. Die Kodierer sind nun gefragt, eine Auswahl dieser Textbeispiele manuell zu evaluieren und so der Trainingsmenge neue Positiv- bzw. Negativ-Beispiele für die zwei Kategorien hinzuzufügen. In je zehn Iterationen dieses Prozesses, bei denen jeweils 200 gefundene Textstellen evaluiert wurden, wurde die initiale Trainingsmenge um 515 Absätze mit Argumentationszusammenhang und 540 Absätze mit Plausibilisierung erweitert.

3.3. Evaluation und automatische Kodierung

Analog zu den Gütekriterien der Sozialforschung werden für Ansätze des Text Mining bzw. des maschinellen Lernens Methoden zur Qualitätssicherung eingesetzt. Die Güte einer Textklassifikation wird in der Regel mit Hilfe der k-fachen Kreuzvalidierung bewertet, für die k mal auf k-1 Teilen der Trainingsdaten ein Klassifikationsmodell trainiert und auf dem verbliebenen ein Teil der Trainingsdaten getestet wird (Dumm / Niekler 2015). Dazu werden Qualitätskennzahlen wie Precision, Recall und ihr gewichtetes Mittel, der F1-Wert, zur Beurteilung der Güte des Verfahrens berechnet. Diese Maße sind verwandt mit den Reliabilitätsmaßen aus den klassischen Methoden der Sozialwissenschaften wie beispielsweise Cohens Kappa. Idealerweise werden F1-Werte um 0,7 analog zu reliablen menschlichen Kodierern angestrebt. Für den oben beschriebenen Active-Learning Prozess lässt sich feststellen, dass die F1-Werte ausgehend von sehr niedrigen Werten um 0,25 im Zuge weiterer Iterationen zunächst schrittweise auf höhere Werte ansteigen, nach ca. sieben Iterationen jedoch kaum noch eine Verbesserung stattfindet. Die Sammlung von Trainingsbeispielen für die Kategorie kann in diesem Fall als weitgehend gesättigt betrachtet werden, insofern das Hinzufügen von neuen Beispielen die Performance nicht mehr allzustark verändert. Gleichzeitig sind nach ca. 7 bis 10 Iterationen genug Trainingsbeispiele vorhanden, um eine valide Klassifikation des Gesamtkorpus aller 3,5 Mio. Dokumente vorzunehmen.

Für die finalen Trainingsmengen werden die folgenden F1-Werte erreicht: $F1_{Argument} = 0{,}608$ und $F1_{Plausibilisierung} = 0{,}491$. Für eine individuelle Klassifikation, welche darauf bedacht ist möglichst genau Einzelereignisse in einer Datenmenge korrekt zu bestimmen,

können diese Qualitätswerte nur bedingt zufrieden stellen. Unser Klassifikator liefert bei relativ hohem Recall auch viele Textstellen zurück, die bei manueller Evaluation nicht in die entsprechende Kategorie einsortiert werden können. Unser Analyseziel liegt jedoch, wie häufig in den Sozialwissenschaften, nicht in der Vorhersage von Einzelereignissen, sondern in der validen Bestimmung von Proportionen und Trends (Hopkins / King 2010). Für diesen Fall kann die Performance des Klassifikators als ausreichend betrachtet werden, da durch die systematische Überschätzung des wahren Anteils an Textbeispielen für eine Kategorie die Änderungen im Verhältnis der Kategorieproportionen zueinander an unterschiedlichen Zeitpunkten des diachronen Korpus nicht verfälscht werden. Auch wenn die Anteile insgesamt durch den Klassifikator als zu hoch eingeschätzt werden mögen, reflektieren die Messungen der Kategorieanteile im Korpus in unterschiedlichen Zeitabschnitten die Zu- bzw. Abnahme der Häufigkeit des Gebrauchs von neoliberalen Argumentations- bzw. Plausibilisierungsmustern korrekt. Im Gesamtkorpus des *ePol*-Projekts werden im Zuge der finalen Klassifikation 105.740 Argumentansätze und 753.653 Plausibilisierungsabsätze identifiziert.

Für die Bestimmung von Trends werden die Dokumente gezählt, in denen eine der beiden Kategorien vorkommt. Daraus lassen sich wiederum Dokumentfrequenzen für bestimmte Zeiträume aggregieren und mit dem Gesamtdatenbestand in diesen Zeiträumen normalisieren. Damit können Zeitverläufe der Kategorien sichtbar gemacht werden, die wiederum politikwissenschaftlich interpretiert werden können (siehe Abbildung 5).

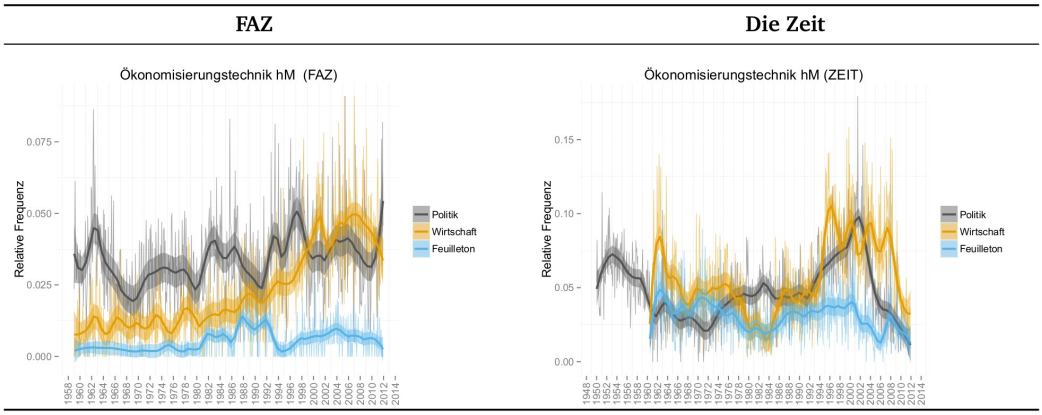

Figure 5. ***Abb. 5****: Relative Frequenzen von Dokumenten in Die Zeit und FAZ welche neoliberale Argumentzusammenhänge enthalten, getrennt nach drei Zeitungsressorts (Politik, Wirtschaft, Kultur)*

3.4. Verallgemeinerung der Ergebnisse

Der Ansatz des Active Learning im *ePol*-Projekt zur Messung abstrakter Kategorien, welche bislang lediglich qualitativ beschrieben worden sind, kann zu einem Ansatz von semi-automatischer Inhaltsanalyse verallgemeinert werden, bei dem die Schritte 1. Dokumentidentifikation, 2. manuelle Kodierung und 3. automatische Kodierung in der beschriebenen Weise miteinander kombiniert werden. Für die Auswertung sehr großer Datenmengen erlaubt der Ansatz nicht nur die Beobachtung von komplexen Kategorien im Zeitverlauf, sondern auch, in Erweiterung des *ePol*-Ansatzes mit einem größeren Kategorienschema, die Beobachtung des gemeinsamen Auftretens von Kategorien für inhaltliche Schlussfolgerungen auf sich gegenseitig bedingende Inhalte. Zusätzlich bietet die beliebige Facettierung der automatischen Analyse einer Vollerhebung Vorteile gegenüber manuellen Analysen, die auf vorab festgelegte Sampling-Strategien beschränkt sind.

4. Vortrag 3: Die Anwendung computer- und korpuslinguistischer Methoden für eine interaktive und flexible Tiefenanalyse der Mobilisierung kollektiver Identitäten in öffentlichen Debatten über Krieg und Frieden – e-Identity

Cathleen Kantner, Jonas Kuhn, André Blessing und Maximilian Overbeck

Internationale Krisenereignisse wie Kriege und humanitäre militärische Interventionen lösen heftige öffentliche Kontroversen aus. Die Menschen machen sich Sorgen und fragen: Welche Effekte hat der Konflikt für unser eigenes Land, für Europa und für die Welt? Wer sind die Opfer, wer die Täter im Krisenland? Soll unser Land Truppen entsenden, verstärken oder ihren Einsatz zum wiederholten Male verlängern? Und falls ja, mit welchem Mandat sollen „unsere“ Truppen agieren – verteilen sie Lebensmittel oder setzten sie Waffen ein? Wie gehen „wir“ (in unserem Land, in Europa, im Westen, …) damit um, wenn Zivilisten oder „unsere“ Soldaten dabei das Leben verlieren?

In öffentlichen Debatten zu kontroversen politischen Themen werden unterschiedlichste politische Positionen oftmals mit Rekurs auf das kollektive Selbstverständnis einer Wir-Gemeinschaft begründet. Die Mobilisierung unterschiedlichster kollektiver – europäischer, nationaler, religiöser usw. – Identitäten stellt somit eine zentrale Argumentationsfigur in der politischen Öffentlichkeit dar. Politische Sprecher begründen ihre Beteiligung an einem militärischen Einsatz oder ihre Enthaltung mit Rekurs auf das kollektive Selbstverständnis einer Wir-Gemeinschaft.

Ein Beispiel: In der europäischen, öffentlichen Debatte über die militärische Intervention in Libyen 2011 wurde auch über das kollektive Selbstverständnis der Europäer verhandelt. Die europäische Identität wurde teils als Problemlösungsgemeinschaft, teils aber auch als Gemeinschaft mit einem normativen Selbstverständnis diskutiert, die sich der Verteidigung der Menschenrechte verpflichtet habe.

Im *e-Identity* Projekt wurden die Potentiale für computer- und korpuslinguistische Methoden erschlossen, die eine interaktive und flexible Tiefenanalyse der Mobilisierung dieser unterschiedlichsten Formen kollektiver Identitäten in öffentlichen Debatten über Krieg und Frieden ermöglichen. 2 Zur methodischen Umsetzung der Forschungsfragen und Überprüfung der Hypothesen untersuchten wir internationale Diskussionen über Kriege und humanitäre militärische Interventionen seit dem Ende des Kalten Krieges 1990. Dabei ging

es uns vor allem darum, das komplexe Geflecht von Identitätsdiskursen in diesen Kontroversen genauer zu analysieren. Im Prozess der Anwendung und Analyse wurden zwei computer- und korpuslinguistische Tools entwickelt, der *Complex Concept Builder* und eine *Explorationswerkbank*.

Eine Explorationswerkbank zur Korpuserstellung, -erschließung und -bearbeitung wurde entwickelt, um Sozialwissenschaftlern auch über das Projektende hinaus als flexibles Bindeglied zu vorhandenen Infrastrukturen (z. B. CLARIN) zu dienen. Sie lässt sich unterschiedlichsten individuellen Forschungsfragen und Textmaterialien anpassen und bildet insbesondere auch die technische Basis für den Complex Concept Builder (Kliche et al. 2014; Mahlow et al. 2014). Im *e-Identity* Projekt wurde somit ein bereinigtes, mehrsprachiges Korpus von 460.917 3

Zeitungsartikeln aus sechs Ländern (Deutschland, Österreich, Frankreich, UK, Irland, USA) generiert, das den Zeitraum von Januar 1990 bis Dezember 2012 abdeckt.

Um in Korpora Textbelege zu finden, in denen Sprecher sich auf eine kollektive Identität beziehen, sind gängige stichwortbasierte Suchtechnologien nicht ausreichend, weil solch ein komplexer Begriff sehr unterschiedlich lexikalisiert und in seiner Interpretation in hohem Maße kontextuell bestimmt sein kann. Gesucht waren daher neue Methoden zur interaktiven inhaltlichen Korpuserschließung.

Um der Vielschichtigkeit der im Korpusmaterial zu untersuchenden Indikatoren ebenso Rechnung zu tragen wie dem erheblichen Korpusumfang und dem Nebeneinander von deutsch-, englisch und französischsprachigen Texten, wurde ein transparenter, vom jeweiligen Forschungsteam individuell nutzbarer *Complex Concept Builder* entwickelt, der sprachtechnologische Werkzeuge und Methoden anbietet, die in den Sozialwissenschaften bislang nur in Ausnahmefällen Anwendung fanden (Blessing et al. 2013). Maschinelle Lernverfahren kamen dabei sowohl bei der inhaltlichen Bereinigung der mehrsprachigen Textkorpora sowie bei der halbautomatischen Identifikation der verschiedenen Identitätstypen zum Einsatz. Komplexe fachwissenschaftliche Begriffe (wie der Identitätsbegriff inklusive der feinen Unterschiede und Nuancen zwischen verschiedenen kollektiven Identitäten) können innerhalb des *Complex Concept Builder* für die Anwendung an alltagssprachlichem Textmaterial operationalisiert werden.

Explorationswerkbank und Complex Concept Builder (CCB) werden im Verlauf dieses Jahres über einen CLARIN Server zugänglich gemacht. Beide Tools erlauben den Export ihrer aggregierten Ergebnisse (z. B. Artikelanzahl, Anzahl der identifizierten Textstellen) zur anschließenden statistischen Analyse. Die für die Fachwissenschaftler transparente Reflexion der Ergebnisse bleibt dabei weiterhin gewährleistet, indem beispielsweise ein Aufsplitten der quantitativen Analysen in die einzelnen qualitativen Analysen möglich ist.

Im Folgenden wird in Kürze angerissen, in welchen Bereichen computer- und korpuslinguistische Methoden sowie Ansätze des maschinellen Lernens Anwendung fanden, um die Analyse kollektiver Identitäten innerhalb der umfangreichen Zeitungstextkorpora durchzuführen. Die Verbindung quantitativer und qualitativer Analyseschritte ermöglichte es, eine komplexe sozialwissenschaftliche Fragestellung auf einer großen Textmenge zu untersuchen und zugleich die Einhaltung der sozialwissenschaftlichen Forschungsstandards der Validität und Reliabilität zu gewährleisten. Im Rahmen unseres Vortrags auf der DHd-Jahrestagung 2016 sollen die folgenden Verfahren genauer präsentiert werden.

- *Inhaltliche Bereinigung der Zeitungstextkorpora von Sampling-Fehlern*: Der Complex-Concept-Builder (CCB) wurde entwickelt, um große mehrsprachige Textmassen nach sozialwissenschaftlich relevanten Aspekten „vorzusortieren“ und er erwies sich bereits bei der Samplebereinigung unter inhaltlichen Gesichtspunkten als äußerst produktiv (Blessing et al. 2015). Mithilfe einer Topic Modellierung wurde eine optimale Vorauswahl einer Trainingsmenge von Texten zur inhaltlichen Dokumentenbereinigung möglich. 4 Die manuelle Annotation erlaubte beispielsweise den sofortigen Ausschluss eines ‚Topics‘ wie Buchrezensionen, die für unsere politikwissenschaftliche Fragestellung nicht relevant sind. Andererseits konnte das schwierige 'Thema' Sport, das sowohl reine Sportberichterstattung mit militärischen Metaphern als auch politische Berichte über Militäreinsätze mit sportlichen Metaphern und Referenzen enthält, detailliert annotiert werden. Maschinelles Lernen setzten wir dann bei der Klassifikation der Dokumente des gesamten Korpus in gute versus off-topic Texte ein.

Die folgende Abbildung (Abb. 6) zeigt im oberen Teil eine herkömmliche Klassifizierung per Zufallsauswahl der manuell-annotierten Trainingsdaten. Unten ist unser Verfahren abgebildet: Topics helfen, die optimale Trainingsmenge zu bestimmen, wobei mindestens aus jedem Topic ein Dokument manuell annotiert wird und damit eine breite Abdeckung gewährt ist. Dadurch wird das Ergebnis des neuen Klassifikators besser (er findet nun z. B. auch Artikel zum Ruanda-Konflikt).

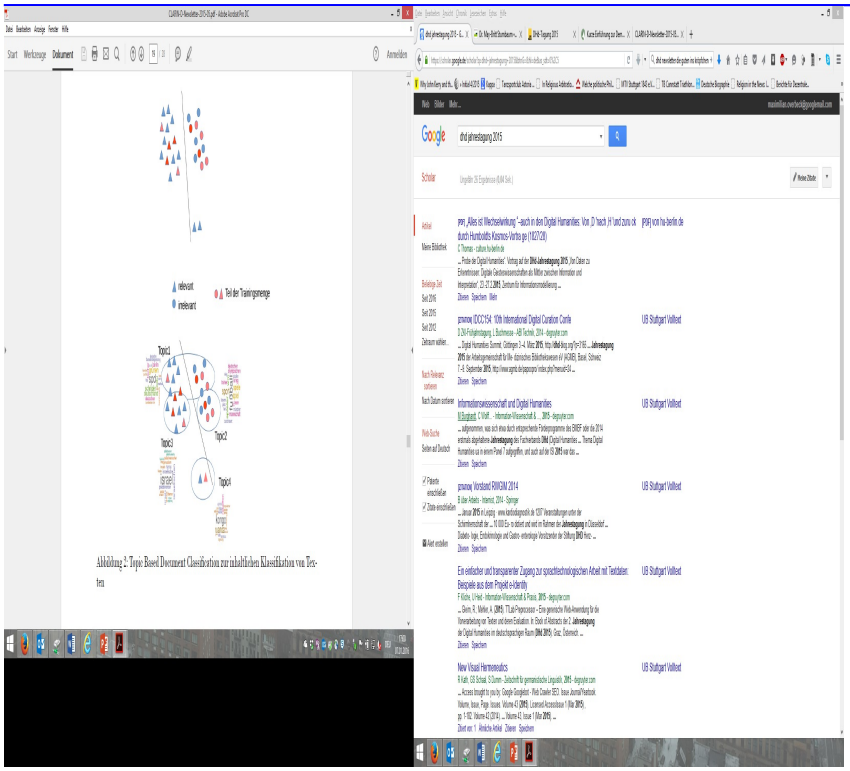

*Figure 6. **Abb. 6**: Topic Based Document Classification zur inhaltlichen Bereinigung von Texten*

Im Anschluss an die Bewertung einer bestimmten Anzahl an Zeitungsartikel wird über maschinelles Lernen die Bewertung auf die Gesamtmenge der Zeitungsartikel angewendet. Wir folgen der Idee von *Dualist*, einem interaktiven Klassifikationsmechanismus (Settles 2011; Settles / Zhu 2012). Die Architektur von Dualist basiert auf MALLET (McCallum 2002) und konnte leicht in unsere Architektur integriert werden. Die Zeitungsartikel, die durch den Computer automatisch aussortiert werden, können in weiteren iterativen Schritten erneut manuell bewertet werden, um den Klassifikator weiter zu optimieren. Eine weitere Abbildung (Abb. 7) zeigt den inhaltlichen Vorgang der Bereinigung im Complex Concept Builder. Die rot markierten Artikel wurden nach einer qualitativen Kodierung automatisch als Sampling Errors identifiziert, während die grün markierten Artikel automatisch dem Issue "Kriege und Humanitäre Militärische Interventionen" zugeordnet wurden.

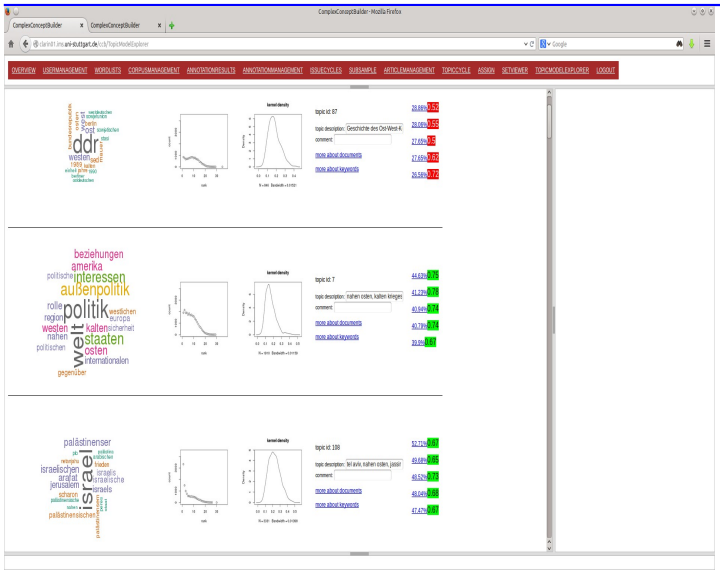

*Figure 7. **Abb. 7**: Halbautomatisierte inhaltliche Bereinigung von Sampling Fehlern im Complex Concept Builder*

Im Fall des *e-Identity* Korpus blieben von insgesamt 766.452 Zeitungsartikeln, die ursprünglich Teil des unbereinigten Korpus waren, lediglich 460.917 Zeitungsartikel übrig (siehe Abbildung 8).

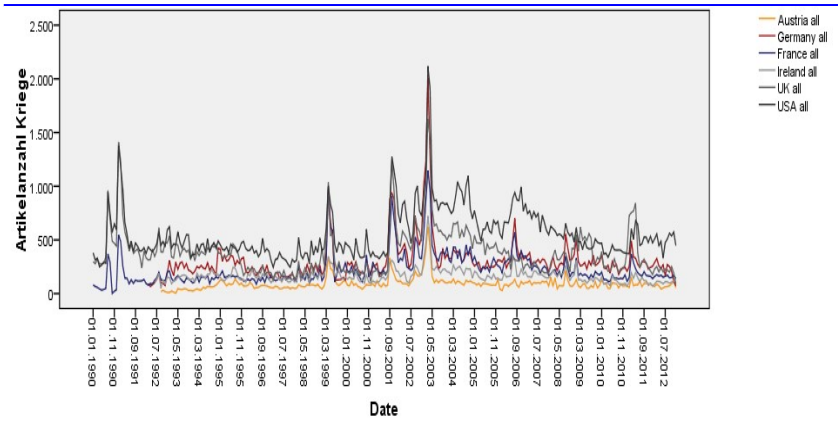

Figure 8. ***Abb. 8****: Das bereinigte Issue-Cycle für das Thema Kriege und Humanitäre Militärische Interventionen (nach Monaten aggregiert, N=460.917)*

Dieses Verfahren eignet sich darüber hinaus zum Aufspüren und Erstellen inhaltlicher Subkollektionen von Texten für spezifische Fragestellungen. Es bildet somit einen der methodischen Ausgangspunkte für das von Prof. Dr. Andreas Blätte an der Universität Duisburg-Essen geleitete Kurationsprojekt der FAG-8, in dem es um die thematische Strukturierung deutscher Parlamentsprotokolle geht.

- *Korpuslinguistische Verfahren für die semi-automatische Identifikation kollektiver Identitäten in den Zeitungstextkorpora:* Es wurden semantische Felder für die unterschiedlichen Identitätsebenen (z. B. nationale, europäische oder transatlantische Identitäten) mitsamt unterschiedlicher Begründungsfiguren (z. B. kulturelle Identität vs. Interessengeleitete Zweckgemeinschaft) generiert. Relevante Terme wurden über komplexe Diktionäre operationalisiert, die sowohl Lemmatisierungen als auch variable Äußerungen der interessierenden Terme innerhalb eines Satzes berücksichtigen können. Die finalen Diktionäre wurden auf die mehrsprachigen Zeitungskorpora angewendet und in Form von Zeitreihen-Plots visualisiert und ausgewertet.

- *Manuelle Kodierung und die halbautomatische Identifikation von der Äußerung kollektiver Identitäten in bewaffneten Konflikten, unterstützt durch maschinelles Lernen*: Aus dem 460.917 Zeitungsartikel umfassenden Gesamtkorpus wurde ein Teilsample gezogen, das die wissenschaftlich üblichen Kriterien der Repräsentativität erfüllt. Auf diesem Teilsample wurde die manuelle Kodierung von insgesamt 5.000 Zeitungsartikeln durchgeführt. Die Unterstützung durch die *Complex Concept Builder*-Oberfläche ermöglichte die gleichzeitige und kontinuierliche Supervision und Datenauswertung der Kodierungen. Die manuell kodierten Textpassagen dienten im Anschluss als Datengrundlage für das Machine Learning Verfahren. Es wurde ein Klassifikator für die halbautomatische Identifikation der Äußerung kollektiver Identitäten trainiert und anschließend auf den Gesamtkorpus von 460.917 Zeitungsartikeln angewendet.

Zusammenfassung:

Die aus sozialwissenschaftlicher Perspektive interessanten und forschungsleitenden Konzepte sind nicht standardisierbar. Im *e-Identity* Projekt vertreten wir daher den Ansatz, dass computer- und korpuslinguistische Ansätze den Forscher dabei unterstützen sollten, ihre auf individuelle Fragestellungen gemünzten Korpora effizient zu managen und zu bereinigen. Sie sollten dem einzelnen Forscherteam Raum für seine eigene Operationalisierung lassen und dabei z. B. im Wechselspiel von manueller und automatischer Annotation in 'lernenden' Anwendungen die Vorteile beider Zugänge intelligent kombinieren. Dies schließt natürlich nicht aus, dass bewährte Operationalisierungen für die im Umfeld dieser komplexen fachlichen Konzepte ausgedrückten Sachverhalte, Bewertungen und Beziehungen usw. wie üblich analysiert werden können. Transparente und flexible CLARIN-Tools, die sich zu Workflows zusammenbinden lassen, die für eine spezifische fachwissenschaftliche Forschungsfrage sensibel bleiben, werden Sozialwissenschaftlern viele kreative Möglichkeiten bieten, interdisziplinären Austausch stimulieren und Spaß bei der Arbeit machen!

Anmerkungen

[1]Ausführliche Informationen zum Projekthintergrund auf http://www.epol-projekt.de. Die hier vorgestellten Analysen wurden mit dem Leipzig Corpus Miner, einer webbasierten Analyseinfrastruktur, durchgeführt (Niekler et al. 2014; Wiedemann / Niekler 2015).

[2]Für weitere Details zum e-Identity Projekt siehe http://www.uni-stuttgart.de/soz/ib/forschung/Forschungsprojekte/eIdentity.html. Für sozialwissenschaftliche Studien, in denen die Tools und Korpora des *e-Identity* Projekts bereits angewendet wurden, siehe Kantner 2015, Kantner et al. (erscheint in Kürze), Overbeck 2015 (im Druck).

[3]Die unbereinigte Textmenge betrug 902.029 Zeitungsartikel. Der umfangreiche und innovative Bereinigungsprozess des Datensatzes von Dubletten und Samplingfehlern war ein zentraler Bestandteil des *e-Identity* Projekts.

[4]Der *Complex Concept Builder* bietet ein Verfahren, um auf der Grundlage von insgesamt 50, 100 oder 200 automatisch erstellten Topics, die auf Grundlage der "Latent Dirichlet Allocation" (LDA) – Methode generiert werden (Blei et al. 2003; Niekler / Jähnichen 2012), inhaltliche Samplingfehler zu identifizieren. Die Visualisierung von Wortwolken einer automatischen Topicanalyse erleichtert die Identifikation von inhaltlichen Samplingfehlern.

Bibliographie

1. **Bach, Ngo Xuan / Nguyen Le Minh / Tran Thi Oanh / Akira Shimazu** (2013): "A Two-Phase Framework for Learning Logical Structures of Paragraphs in Legal Articles", in: *ACM Transactions on Asian Language Information Processing (TALIP)* 12, 1: Nr. 3.
2. **Bex, Floris / Lawrence, John / Snaith, Mark / Reed, Chris** (2013): "Implementing the Argument Web", in: *Communications of the ACM* 56, 10: 66–73.
3. **Bex, Floris / Snaith, Mark / Lawrence, John / Reed, Chris** (2014): "ArguBlogging: An Application for the Argument Web", in: *Journal of Web Semantics* 25:

9–15.

4. **Blei, David M. / Ng, Andrew Y. / Jordan, Michael I.** (2003): "Latent dirichlet allocation", in: *Journal of machine Learning research* 3: 993-1022.
5. **Blessing, Andre / Sonntag, Jonathan / Kliche, Fritz / Heid, Ulrich / Kuhn, Jonas / Stede, Manfred** (2013): "Towards a tool for interactive concept building for large scale analysis in the humanities", in: *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, Sofia.
6. **Blessing, Andre / Kliche, Fritz / Heid, Ulrich / Kantner, Cathleen / Kuhn, Jonas** (2015): "Die Exploration großer Textsammlungen in den Sozialwissenschaften", in: *CLARIN Newsletter* 8: 17-20.
7. **Bögel, Tina / Hautli-Janisz, Annette / Sulger, Sebastian / Butt, Miriam** (2014): "Automatic Detection of Causal Relations in German Multilogs", in: *Proceedings of the EACL 2014 Workshop on Computational Approaches to Causality in Language (CAtoCL)* 20–27.
8. **Dumm, Sebastian / Lemke, Matthias** (2013): "Argumentmarker. Definition, Generierung und Anwendung im Rahmen eines semi-automatischen Dokument-Retrieval-Verfahrens", in: *Schriftenreihe des Verbundprojekts „ePol – Postdemokratie und Neoliberalismus“*, Discussion-Paper 3 .
9. **Dumm, Sebastian / Niekler, Andreas** (2015): "Methoden, Qualitätssicherung und Forschungs design. Diskurs- und Inhaltsanalyse zwischen Sozialwissenschaften und automatischer Sprachverarbeitung", in: Lemke, Matthias / Wiedemann, Gregor (eds.): *Text Mining in den Sozialwissenschaften*. Grundlagen und Anwendungen zwischen qualitativer und quantitativer Diskursanalyse. Wiesbaden: Springer VS 89-116.
10. **Feng, Vanessa Wei / Hirst, Graeme** (2011): "Classifying Arguments by Scheme", in: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics* 987–996.
11. **Gold, Valentin / Holzinger, Katharina** (2015): *An Automated Text-Analysis Approach to Measuring the Quality of Deliberative Communication*. Paper presented at the 73th Annual Meeting of the Midwest Political Science Association (MPSA), San Francisco.
12. **Gold, Valentin / Rohrdantz, Christian / El-Assady, Mennatallah** (2015): "Exploratory Text Analysis using Lexical Episode Plots", in: The Eurographics Association (ed.): *EuroVisShort2015* 85-89 http://dx.doi.org/10.2312/eurovisshort.20151130.
13. **Gold, Valentin / El-Assady, Mennatallah / Bögel, Tina / Rohrdantz, Christian / Butt, Miriam / Holzinger, Katharina / Keim, Daniel** (2015): "Visual Linguistic Analysis of Political Discussions: Measuring Deliberative Quality", in: *Digital Scholarship in the Humanities* http://dx.doi.org/10.1093/llc/fqv033.
14. **Hopkins, Daniel / King, Gary** (2010): "A Method of Automated Nonparametric Content Analysis for Social Science", in: *American Journal of Political Science* 54, 229–247.
15. **Kantner, Cathleen** (2015): *War and Intervention in the Transnational Public Sphere: Problem-solving and European identity-formation*. London: Routledge.
16. **Kantner, Cathleen / Overbeck, Maximilian / Sangar, Eric** (erscheint in Kürze): "Die Analyse ‚weicher' Konzepte mit ‚harten' korpuslinguistischen Methoden: Multiple kollektive Identitäten", in: Behnke, Joachim / Blaette, Andreas / Schnapp, Kai-Uwe / Wagemann, Claudius (eds.): *Big Data: Große Möglichkeiten oder große Probleme?* Baden-Baden: Nomos Verlag.
17. **Karagjosova, Elena** (2004): *The Meaning and Function of German Modal Particles* (= Saarabrücken Dissertations in Computational Linguistics and Language Technology 18). Saarbrücken: Computational Linguistics & Phonetics, Universität des Saarlandes.
18. **Kirschner, Christian / Eckle-Kohler, Judith / Gurevych, Iryna** (2015): "Linking the Thoughts: Analysis of Argumentation Structures in Scientific Publications", in: *Proceedings of the 2nd Workshop on Argumentation Mining (ARG-MINING 2015)* 1-11.
19. **Kliche, Fritz / Blessing, Andre / Sonntag, Jonathan / Heid, Ulrich** (2014): *"The e-identity exploration workbench", in: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik.
20. **Kratzer, Angelika** (1999): *Beyond "oops" and "ouch". How descriptive and expressive meaning interact.* Paper presented at the Cornell Conference on Theories of Context Dependency.
21. **Lassiter, Daniel** (2010): "Gradable epistemic modals, probability, and scale structure", in: *Proceedings of the 20th conference on Semantics and Linguistic Theory (SALT 20)* 197-215.
22. **Mahlow, Cerstin / Eckart, Kerstin / Stegmann, Jens / Blessing, Andre / Thiele, Gregor / Gärtner, Markus / Kuhn, Jonas** (2014): "Resources, Tools, and Applications at the CLARIN Center Stuttgart", in: *Proceedings of the 12th Konferenz zur Verarbeitung natürlicher Sprache (KONVENS 2014)* 11-21.
23. **McCallum, Andrew K.** (2002): "MALLET: MAchine Learning for LanguagE Toolkit" http://mallet.cs.umass.edu/about.php.
24. **Mochales Palau, Raquel / Moens, Marie-Francine**(2011): "Argument Mining", in: *Artificial Intelligence and Law* 19, 1: 1-22.
25. **Niekler, Andreas / Jähnichen, Patrick** (2012): "Matching results of latent dirichlet allocation for text", in: *Proceedings of ICCM 2012, 11th International Conference on Cognitive Modeling* 317-322.
26. **Niekler, Andreas / Wiedemann, Gregor / Heyer, Gerhard** (2014): "Leipzig Corpus Miner - A Text Mining Infrastructure for Qualitative Data Analysis", in: *Proceedings of the Conference on Terminology and Knowledge Engineering 2014*, Berlin.
27. **Oraby, Shereen / Reed, Lena / Compton, Ryan / Riloff, Ellen / Walker, Marilyn / Whittaker, Steve** (2015): "And That's A Fact: Distinguishing Factual and Emotional Argumentation in Online Dialogue", in: *Proceedings of the 2nd Workshop on Argumentation Mining (ARG-MINING 2015*) 116-126.
28. **Overbeck, Maximilian** (im Druck): "Observers turning into Participants: Shifting perspectives on Religion and Armed Conflicts in Western News Coverage", in: *La revue Tocqueville* 36, 2.
29. **Polanyi, Livia / Culy, Chris / van den Berg, Martin / Thione, Gian Lorenzo / Ahn, David** (2004): "Sentential structure and discourse parsing", in: *Proceedings of the 2004 ACL Workshop on Discourse Annotation* 80–87.
30. **Potts, Christopher**(2012): "Conventional implicature and expressive content", in: Maienborn, Claudia / von Heusinger, Klaus / Portner, Paul (eds.): *Semantics* 3 (= Handbücher zur Sprach- und Kommuninikationswissenschaft 33, 3). Berlin: de Gruyter Mouton 2516–2536.
31. **Scharkow, Michael** (2012): *Automatische Inhaltsanalyse und maschinelles Lernen*. Berlin: epubli.
32. **Settles, Burr** (2011): "Closing the loop: Fast, interactive semi-supervised annotation with queries on features and instances", in: *Proceedings of the Conference on Empirical Methods in Natural Language Processing* 1467-1478.
33. **Settles, Burr / Zhu, Xiaojin** (2012): "Behavioral factors in interactive training of text classifiers", in: *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* 563-567.
34. **Wiedemann, Gregor / Lemke, Matthias / Niekler, Andreas** (2013): "Postdemokratie und Neoliberalismus. Zur Nutzung neoliberaler Argumentationen in der Bundesrepublik Deutschland 1949-2011", in: *Zeitschrift für Politische Theorie* 4, 1: 99-115.
35. **Wiedemann, Gregor / Niekler, Andreas** (2014): "Document Retrieval for Large Scale Content Analysis using Contextualized Dictionaries", in: *Proceedings of the Conference on Terminology and Knowledge Engineering 2014*, Berlin.
36. **Wiedemann, Gregor / Niekler, Andreas** (2015): "Analyse qualitativer Daten mit dem 'Leipzig Corpus Miner'", in: Lemke, Matthias / Wiedemann, Gregor (eds.):*Text Mining in den Sozialwissenschaften*. Grundlagen und Anwendungen zwischen qualitativer und quantitativer Diskursanalyse. Wiesbaden: Springer VS 63-88.
37. **Zimmermann, Malte** (2011): "Discourse particles", in: von Heusinger, Klaus / Maienborn, Claudia / Portner, Paul (eds.): *Semantics 2*(= Handbücher zur Sprach- und Kommunikationswissenschaft 33, 2). Berlin: Mouton de Gruyter 2011-2038.