

H a u s a r b e i t

„Sie hetzen gegen Israel“ – Berichterstattung über Pro-Palästina-Demos in Deutschland

Eine *semi-supervised* Themenanalyse

Jana Borchers

Matrikelnummer: 8007047

Goethe-Universität Frankfurt

Fachbereich 03 Gesellschaftswissenschaften

Forschungspraktikum:

Computational Social Science

Dozent: Dr. Christian Czymara

Inhalt

1. Einleitung	1
2. Daten.....	2
2.1 Selektion und Aufbereitung	2
2.2 Deskriptive Auswertung	3
3. Modell.....	5
3.1 Methode	5
3.2 Keywords	6
3.3 Anzahl der Topics	7
4. Ergebnisse.....	8
5. Fazit	17
Literaturverzeichnis	18

1. Einleitung

Seit dem Angriff der militanten palästinensischen Organisation Hamas in Israel am 7. Oktober 2023 und den darauffolgenden Bombardierungen des Gazastreifen durch Israel, haben weltweit hunderttausende Menschen, darunter vor allem Studierende, gegen das israelische Vorgehen protestiert (Hebling/Trautmüller, 2024; Amnesty International, 2024). Auch in Deutschland kam es in zahlreichen Städten zu Demonstrationen sowie mehrtägigen Protestcamps. Die Demonstrationen lösten dabei starke mediale Aufmerksamkeit aus. So diskutierten beispielsweise Gäste der Talkshow „hart aber fair“ am 12. August 2024 unter dem Titel „Israel im Krieg: Kritik erlaubt?“, welche Art des Protests angemessen sei (ARD, 2024). Auch zahlreiche Politiker*innen äußerten wiederholt Kritik an den Demonstrationen und warfen den Teilnehmenden antisemitische und gewaltverherrlichende Äußerungen vor (Zielezinski, 2024; Böldt, 2024; Lehmann, 2023). Die NRW-Antisemitismusbeauftragte und ehemalige Bundesjustizministerin Sabine Leutheusser-Schnarrenberger (FDP) forderte gar, das Versammlungsrecht für Ausländer einzuschränken (Beck, 2023). Eine Studie der Universität Mannheim aus dem Jahr 2024, die den Zusammenhang zwischen pro-palästinensischen und antisemitischen Einstellungen untersucht hat, bezeichnet den Antisemitismus-Vorwurf hingegen als „vorschnell“ (Hebling/Trautmüller, 2024: 1). Selten scheint es in der Auseinandersetzung um die inhaltlichen Positionen der Demonstrierenden und deren mediale Rezeption zu gehen. Hier soll in dieser Arbeit angesetzt werden: Welche Themen stehen bei der Berichterstattung über Palästina-Demos im Vordergrund? Welche inhaltlichen Aspekte werden abgesehen von dem Vorwurf des Antisemitismus verhandelt?

Dazu werden Medienberichte, die seit dem 7. Oktober 2023 erschienen sind und die die Pro-Palästina-Demonstrationen in Deutschland thematisieren, mithilfe von Topic Modelling daraufhin überprüft, welche Themen in der Berichterstattung vorliegen. Während *unsupervised* („unüberwachte“) Topic Modelle dazu geeignet sind, große Textmengen auch ohne fundiertes Vorwissen nach zentralen Themen zu untersuchen (Asmussen/Møller, 2019: 14), ist dieser Ansatz nicht ohne Schwächen. So ist unter anderem davon auszugehen, dass Forscher*innen nur in seltenen Fällen gänzlich ohne Vorannahmen in die Analyse gehen. Dies ist auch in der vorliegenden Arbeit der Fall: So wird unter anderem angenommen, dass Medienberichte über Pro-Palästina-

Demonstrationen Bezug auf das Thema Antisemitismus nehmen können. Als weitere Möglichkeit der Berichterstattung wird ein Fokus auf die Inhalte sowie auf den Umgang von Polizei und Behörden mit den Protesten erachtet. Gleichzeitig ist anzunehmen, dass die Berichterstattung sich nicht allein auf diese Themenfelder beschränkt und möglicherweise weitere Topics in den Daten vorliegen können. Entsprechend kommt in dieser Arbeit ein *keyword-assisted* Topic Modell zum Einsatz, das als *semi-supervised* Modell eine Erweiterung unüberwachter Modelle darstellt.

Im Folgenden wird zunächst eine Übersicht über die Daten und erste deskriptive Ergebnisse gegeben. Um thematisch nicht relevante Artikel auszuschließen, wird ein *supervised* („überwachtes“) Verfahren angewendet. Danach wird die Methode des KeyATM-Modells vorgestellt und relevante Keywords für die Analyse identifiziert. Anschließend werden die für das Modell geeignete Anzahl an Topics bestimmt. Es folgt eine Darstellung der Ergebnisse sowie Fazit und Ausblick.

2. Daten

2.1 Selektion und Aufbereitung

Für die Analyse wurden Zeitungsartikel über die kommerzielle Datenbank LexisNexis extrahiert. Der Zeitraum der Publikation erstreckt sich vom 7. Oktober 2023 bis zum 21. Januar 2025. Am 7. Oktober 2023 fand der Überfall der Hamas auf ein israelisches Festival statt, das als Anlass für die israelischen Luftangriffe auf den Gazastreifen diente (El-Shewy et al., 2024). Folglich wird durch diese zeitliche Festlegung die Analyse auf Demonstrationen, die auf die genannten Ereignisse Bezug nehmen, eingegrenzt. Die betreffenden Artikel wurden über die Stichworte „Demo“ und „pro-palästinensisch“ gesucht. Da die Bezeichnung „pro-palästinensisch“ allerdings nicht einheitlich verwendet wird und einige Medien die Proteste auch als „Anti-Israel“-Proteste bezeichnen, wurde die Suche zudem auf das Stichwort „Anti-Israel“ ausgeweitet. Die Art der Artikel ist auf Zeitungsberichte und Online-Publikationen beschränkt. Insgesamt wurden so 2116 Artikel extrahiert, die durch das Entfernen von Duplikaten in einem weiteren Schritt auf 1676 reduziert wurden. Dabei wurden nur solche Duplikate entfernt, die sowohl einen identischen Text als auch identische Metadaten aufweisen. Artikel, die zwar im Wortlaut übereinstehen, aber in verschiedenen Publikationen erscheinen, wurden dagegen für die

Analyse beibehalten, da so ein vollständigeres Bild der abgedeckten Themen über verschiedene Medien hinweg erwartet wird.

Im weiteren Verlauf wurde ein Sample von 20 Prozent der Daten manuell daraufhin codiert, ob diese tatsächlich palästinasolidarische beziehungsweise israelkritische Proteste thematisieren. 18 Beobachtungen wurden hierbei als nicht dem Thema entsprechend klassifiziert. Dabei handelte es sich beispielsweise um Leserbriefe oder um solche Artikel, in denen zwar die für die Suche spezifizierten Stichworte vorkommen, diese inhaltlich aber in keinem Zusammenhang stehen. Die codierten Daten wurden genutzt, um ein randomForest-Modell¹ zu trainieren. Mithilfe des Modells wurden Vorhersagen für die restlichen 80% des Datensatzes getroffen (vgl. ähnliche Verfahren: Czymara, 2023; Kaur et al., 2020; Maruf et al., 2015). Ein weiteres Sample von 100 Beobachtungen aus den trainierten Daten wurde per Hand auf ihre Richtigkeit überprüft. Hierbei zeigte sich, dass das Modell alle im Sample enthaltenen Texte korrekt vorhersagen konnte. Entsprechend wurden die 18 manuell als nicht dem Thema entsprechenden Artikel sowie neun weitere durch das Modell vorhergesagte Artikel aus dem Datensatz entfernt, sodass nun 1649 Artikel für die Analyse verbleiben. Anschließend wurden die Artikel für die Analyse vorbereitet, bereinigt, und tokenisiert. Da Möglichkeiten zum Stemming bzw. zur Lemmatisierung für Texte in deutscher Sprache bislang begrenzt und stark fehleranfällig sind (Lepsky, 2023: 173ff.), wurde hierauf in dieser Arbeit verzichtet. Stattdessen wurden einzelne Begriffe, die sich im Laufe der Analyse als relevant herausstellten (z.B. „palästinensisch“, „israelisch“), auf ihre nicht-deklinierte Form zurückgeführt.

2.2 Deskriptive Auswertung

Eine Betrachtung der häufigsten Begriffe im Datensatz zeigt, recht erwartbar, „Israel“, „Hamas“, „Gaza“ sowie „Polizei“ mit den stärksten Ausprägungen an. Auch „jüdisch“ und „Juden“ sowie „Antisemitismus“ und „antisemitisch“ sind unter den 20 am häufigsten auftretenden Wörtern und stützen damit die Anfangsvermutung, dass

¹ RandomForest-Modelle stellen eine Methode überwachten maschinellen Lernen dar, bei der mithilfe der Trainingsdaten mehrere „Entscheidungsbäume“ generiert werden, anhand derer unbekannte Objekte einer Objektart zugeordnet werden können (Hänsch/Hellwich, 2017: 607ff.). Aufgrund der Größe der verwendeten Daten wurde die randomForest-Implementierung „ranger“ nach Wright/Ziegler (2018) verwendet.

Antisemitismus ein zentrales Thema in der Berichterstattung über Palästina-Demonstrationen darstellen könnte.

Werden auch die Verbindungen der häufigsten Begriffe untereinander aufgezeigt, steht zunächst „Israel“ im Zentrum. Das ist insofern naheliegend, da die Demonstrationen explizit Bezug auf Israel und dessen militärisches Vorgehen in Gaza nehmen. Gleichzeitig könnte die Tatsache, dass „Israel“ und nicht – wie ebenso zu vermuten wäre – „Palästina“ dominiert, ein erster Hinweis darauf sein, dass die Zeitungsartikel insbesondere das Verhältnis der Demonstrierenden zum Staat Israel und mögliche daraus abzuleitende ideologische Positionen thematisieren.

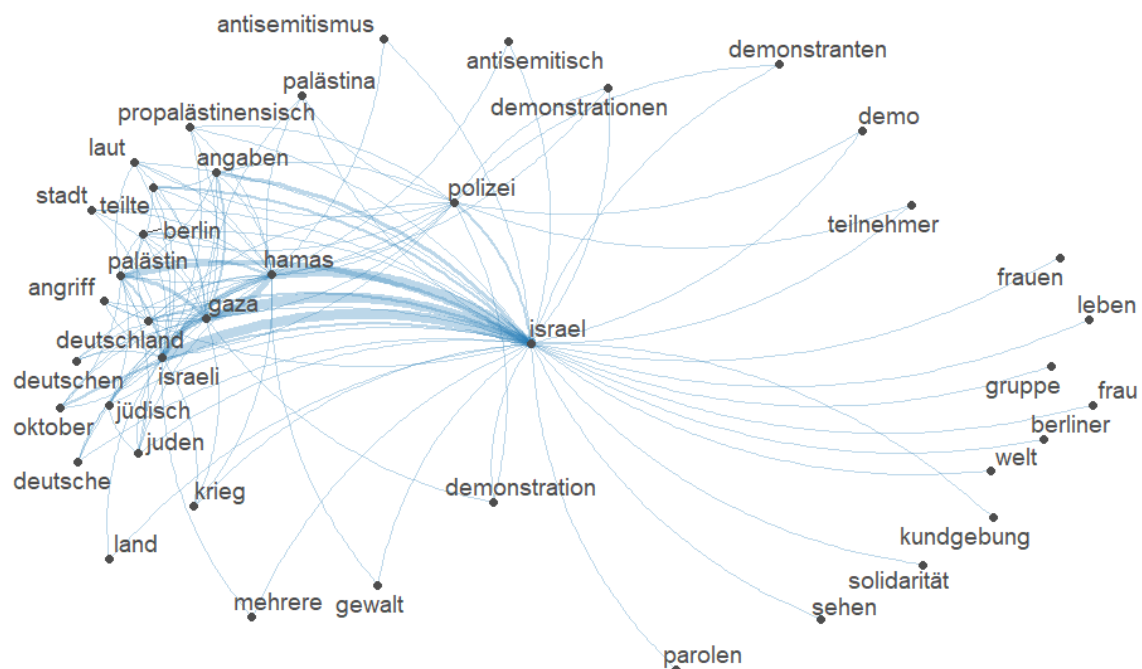


Abbildung 1

Die Verbindungen zu „Gaza“ und „Hamas“ sowie darüber „Angriff“ und „palästin“ („palästinensisch“/„Palästinenser/innen“) deuten auf politische und geopolitische Hintergründe hin, die so möglicherweise ein weiteres Thema in der Berichterstattung darstellen könnten. Interessant ist, dass „jüdisch“/„Juden“ sowie „Oktober“ (vermutlich: 7. Oktober, der Tag des Überfalls der Hamas) in dem Textplot in räumlicher Distanz zu „antisemitisch“ und „Antisemitismus“ stehen und keine direkte Verbindung aufweisen.

Möglicherweise werden diese also in unterschiedlichen Kontexten verwendet und bilden jeweils ein eigenes Sub-Thema ab.

Eine Betrachtung von Kollokationen – also von zwei oder mehr Wörtern, die häufiger und nicht nur zufällig zusammen auftreten (Evert, 2007: 2; McKeown/Radev, 2000: 2) – zeigt weit oben Begriffspaare wie „to the“, „from the“ sowie „river to“. Diese weisen auf die Parole „From the river to the sea“ hin, die im Kontext von Pro-Palästina-Demonstrationen zu hören ist und bereits wiederholt Gegenstand politischer sowie rechtlicher Auseinandersetzungen war (Vehrenkotte, 2024). Werden Kollokationen mit drei verbundenen Begriffen angezeigt, wird dies noch deutlicher:

Collocation	Count	Length
river to the	297	3
from the river	297	3
to the sea	293	3
the river to	291	3
fridays for future	186	3
palestine will be free	128	3

Tabelle 1

Der Slogan „Free Palestine“ tritt zudem als Bigramm 299 Mal auf. Entsprechend werden "from the river to the sea", "fridays for future", "free palestine" und "palestine will be free" im Corpus durch Platzhalter ersetzt, sodass die Modelle sie im weiteren Verlauf als Kollokationen berücksichtigen können

3. Modell

3.1 Methode

Unsupervised Topic Modelle stellen eine automatisierte Möglichkeit dar, große Datenmengen an Text nach darin vorkommenden latenten Themen zu untersuchen (Guo et al., 2016). Eshima et al. (2024) argumentieren allerdings, dass die so identifizierten Themen nicht immer interpretierbar seien. So werden teils verschiedene Themen zusammen gruppiert, in anderen Fällen werden ähnliche Inhalte zu verschiedenen Themen zugeordnet. Zudem kann eine Interpretation erst im Anschluss an das Model

Fitting vorgenommen werden – wenngleich Forscher*innen das Modell meist bereits unter bestimmten Vorannahmen spezifizieren (Eshima et al., 2024; 731).

Einen Ansatz, diese Problematik zu umgehen, bieten *keyword assisted* Topic Modelle wie das von Eshima et al. entwickelte keyATM Modell, das auch in dieser Arbeit Anwendung findet. Dieser *semi-supervised* Ansatz basiert wie die von Blei et al. (2003) entwickelte *Latent Dirichlet allocation* (LDA) auf einer Dirichlet-Verteilung. Das KeyATM-Modell erweitert das *unsupervised* LDA-Modell aber insofern, als dass es neben der explorativen Themenfindung auch ermöglicht, weitere Themen um zuvor spezifizierte Keywords zu „gruppieren“ (Tushev et al., 2022). Gleichzeitig bleibt die Möglichkeit, weitere Topics zu ergründen, die nicht über vorspezifizierte Keywords identifiziert wurden (Eshima et al., 2024: 731).

3.2 Keywords

Für eine inhaltliche Auseinandersetzung mit den Demonstrationen wurde zwischen „Kritik“ und „Forderungen“ differenziert. Um hier relevante Keywords zu identifizieren, wird auf eine Erklärung eines Bündnisses von 50 Organisationen, darunter Amnesty International, zurückgegriffen, das am 15. Februar 2024 in Berlin unter dem Motto „Für einen gerechten Frieden in Palästina und Israel“ demonstrierte (Der Tagesspiegel, 2025; Amnesty International, 2025). Um ein möglichst breites Spektrum der Bewegung abzudecken, wurde zudem der Text einer Petition des mittlerweile aufgelösten Frankfurter Vereins Palästina e.V. sowie der „Jüdischen Stimme für gerechten Frieden in Nahost“ zur Identifikation von Keywords genutzt (Palästina e.V., 2023; Jüdische Stimme, 2025). Entsprechend wurden für das Thema „Kritik“ die Stichwörter „Genozid“, „Völkermord“, „Besatzung“, „Kriegsverbrechen“, „Luftangriff“, „Kolonialismus“ und „Apartheid“ ausgewählt. Für das Thema „Forderungen“ wurden die Keywords „Waffenruhe“, „Waffenstillstand“, „Waffenlieferungen“ sowie „humanitäre Hilfe“ festgelegt. Schließlich wurde das Thema „Antisemitismus“ über die Begriffe „Antisemitismus“, antisemitisch“, „jüdisch“, „Juden“/„Jüdinnen“ sowie „Hass“ und „Hetze“ charakterisiert. Amnesty International hat zudem den Umgang von Polizei und Behörden mit den Protesten kritisiert (Amnesty International, 2024). Gleichzeitig haben seit Oktober 2023 mehrere Prozesse gegen Demonstrierende, denen beispielsweise „volksverhetzendes“ Verhalten vorgeworfen wurde, stattgefunden (Kastner, 2025;

Gehrke, 2024; Hildebrand, 2024; Barkey, 2024) Mithilfe der Keywords „Repressionen“, „verboten“, „Polizei“, „Polizeigewalt“, „Gericht“ und „Prozess“ wird so ein weiteres Topic spezifiziert, das im Folgenden als „Repressionen“ bezeichnet wird.

3.3 Anzahl der Topics

Eine Herausforderung in *unsupervised* und *semi-supervised* Modellen stellt das Festlegen der Anzahl an zu extrahierenden Topics dar (Ding et al., 2023: 6). Um dem zu begegnen, werden Modelle mit fünf bis 30 Topics in Fünferschritten auf die Kennwerte Kohärenz und Exklusivität getestet. Während die Kohärenz misst, wie stark sich Wörter ähneln, die häufig zusammen auftreten, zeigt die Exklusivität an, wie charakteristisch die Wörter für ein einzelnes Thema sind. Eine hohe Kohärenz bedeutet dementsprechend eine bessere Interpretierbarkeit – hohe Exklusivität wiederum eine gute Unterscheidbarkeit der einzelnen Themen. (Tolochko et al., 2024: 350; Ding et al., 2023: 6; Airoidi/Bischof, 2016: 1395).

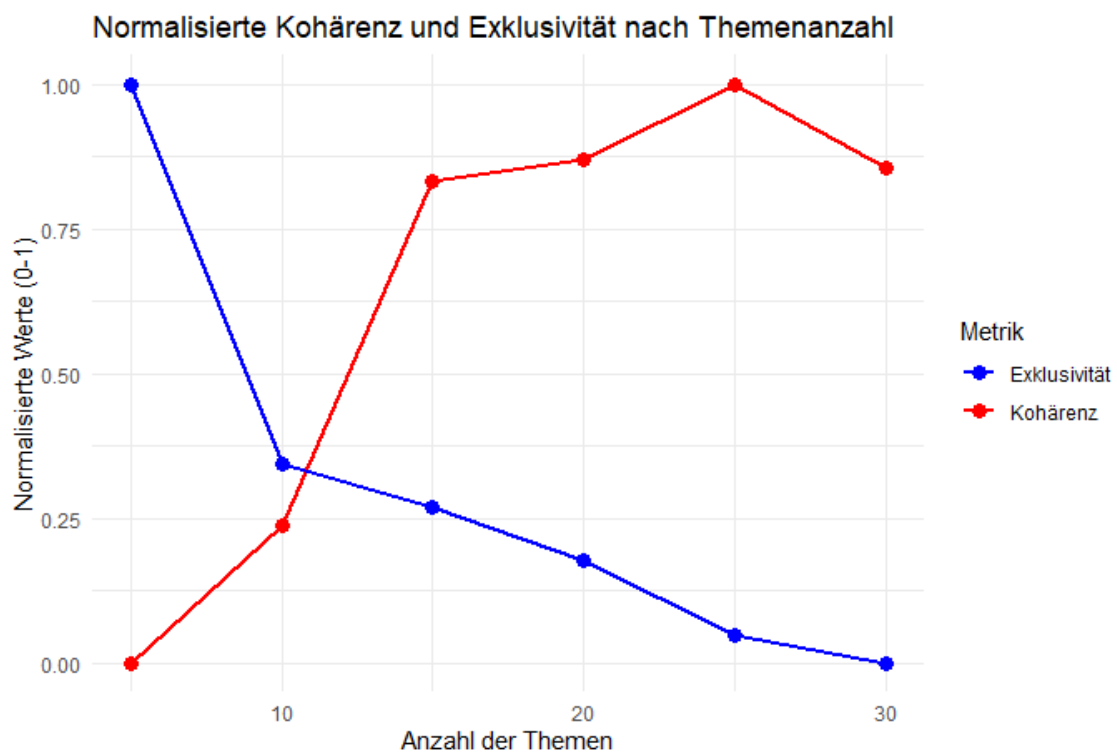


Abbildung 2

In der graphischen Darstellung zeigt sich ein steiler Anstieg der Kohärenz nach zehn Topics, die danach mit zunehmender Anzahl recht gleichmäßig verläuft und bei 25 Topics

ihren Höhepunkt hat. Die Exklusivität hingegen fällt nach fünf Topics stark ab und zeigt danach ein schwaches Absinken. Dies deutet darauf hin, dass die dem Modell zugrundeliegenden Texte sehr viele unterschiedliche Themen umfassen: Bei einer niedrigen Anzahl Topics können diese so allgemein gefasst werden, dass sie sich noch klar voneinander abgrenzen lassen, während die Interpretierbarkeit jedoch gering ist – was auch die niedrige Kohärenz an dieser Stelle bestätigt. Eine höhere Anzahl Topics erleichtert die Interpretierbarkeit der einzelnen Themen, allerdings überschneiden die darin vorkommenden Begriffe sich nun stärker. Die Graphik deutet zudem darauf hin, dass die *keyword-assisted* Topics in dem Modell sehr spezifisch gefasst sind, was auch dadurch nahelegt, dass sehr viele Keywords für die ersten vier Topics gewählt wurden. So lassen sich bei wenigen Topics diese gut um die Keywords herum gruppieren, während es bei hoher Anzahl zwangsläufig stärkere Überschneidungen gibt.

Die Anzahl der Topics wird im Folgenden auf 20 festgelegt: Hier hat sich die Kohärenz stabilisiert und es lässt sich von einer guten Interpretierbarkeit der Themen ausgehen. Die Exklusivität ist nun zwar vergleichsweise niedrig, allerdings ist die Steigung hier nur noch schwach, sodass ab einer Topic-Anzahl von zehn von keinem substanziellen Verlust mehr auszugehen ist. Gleichzeitig wird eine gute Interpretierbarkeit für das Forschungsanliegen als wichtiger erachtet als eine eindeutige Abgrenzung der Themen untereinander, sodass die deutlich höhere Kohärenz an dieser Stelle priorisiert wird.

4. Ergebnisse

Bei der inhaltlichen Interpretation zeigt sich, dass nicht alle Topics im gleichen Maße klar umrissen sind.

Topic	Keywords	Beispieltext
„Kritik“	"völkermord" "genozid" "besatzung" "kriegsverbrechen" „luftangriff“ „kolonialismus“ „apartheid“	„Sie hetzen gegen Israel, verharmlosen den Hamas-Terror: Die internationale "Fridays for Future (FFF) "-Organisation um Greta Thunberg (20) schockt mit Israel-Hass und Antisemitismus. Erst schwieg die Klima-Greta-Truppe tagelang zum Mord an 1400 Juden. Dann

	<p>Weitere:</p> <p>„thunberg“ „greta“ „fridays for future“ „bewegung“ „neubauer“</p>	<p>stellten sie sich knallhart gegen Israel. Jetzt die Eskalation: "Fridays for Future International" wirft Israel einen "Genozid" (Völkermord) vor, obwohl Israels Armee sich gegen den Terror verteidigt. Die Gruppe spricht im Hamas-Sprech von "Märtyrern" - so nennen die Terroristen ihre Getöteten. Sie schwafelt von "imperialistischen Regierungen", die westliche Medien kontrollieren sollen. Es ist die alte Antisemiten-Mär von Juden, die die Welt kontrollieren. Ihre Hass-Forderung im Islamisten-Slang: "Free Falastin" (Befreit Palästina). Kritik am Terror: Fehlanzeige (...)“.</p>
„Forderungen“	<p>"waffenruhe" "waffenstillstand" "waffenlieferungen" "humanitäre hilfe"</p> <p>Weitere:</p> <p>„hamas“ „angaben“ „armee“ „geiseln“ „getötet“ „netanyahu“ „zivilisten“ „libanon“ „iran“</p>	<p>„24-Jähriger bei Pro-Palästina-Demo in Berlin verhaftet. Nach teils gewalttätigen Zusammenstößen mit der Polizei bei pro-palästinensischen Demonstrationen befindet sich ein 24-Jähriger in Untersuchungshaft. (...)“</p> <p>Israel bekräftigte am Dienstag, dass die Waffenruhe auf unbestimmte Zeit verlängert werden könne, solange die Hamas weiterhin mindestens zehn Geiseln pro Tag freilasse. Da die Hamas jedoch nur noch wenige Frauen und Kinder gefangen hält, könnte die Waffenruhe über den Mittwoch hinaus nur verlängert werden, wenn sie erstmals auch einige israelische Männer freilässt. Die am Montag um zwei Tage ausgedehnte Waffenruhe schien auch den</p>

		fünften Tag zu halten (...)“.
„Antisemitismus“	"antisemitismus" "antisemitisch" "hass" "hetze" „juden“ „jüdinnen“ „jüdisch“ Weitere: „hamas“ „oktober“ „deutschland“ „gewalt“ „palästina“	„(...) Der hessische Antisemitismusbeauftragte Uwe Becker hat ein Verbot der für den 7. Oktober geplanten propalästinensischen Demonstration in Frankfurt gefordert. „Wenn am ersten Jahrestag der barbarischen Hamas-Massaker im Süden Israels Sympathisanten des Terrors zur Demonstration in Frankfurt aufrufen, dann ist dies eine zutiefst antisemitische Unmenschlichkeit, eine absolute Provokation, die so nicht stattfinden darf“, sagte Becker. Wer den größten Massenmord an Jüdinnen und Juden seit der Schoah zum Anlass nehme, um Hass und Hetze gegen Israel zu verbreiten, der verhöhne die Opfer und ihre Familien (...)“.
„Repressionen“	"repressionen" "polizei" "verboten" "polizeigewalt" Weitere: „demonstration“ „propalästinensisch“ „teilnehmer“ „versammlung“ „parolen“ „angemeldet“	„Bei einer propalästinensischen Demo ist es in Berlin zu Zusammenstößen zwischen der Polizei und Teilnehmern der Versammlung gekommen. Dabei wurden nach Angaben der Polizei 7 Demonstranten und 17 Polizisten verletzt. «Insgesamt kam es zu Freiheitsbeschränkungen beziehungsweise Freiheitsentziehungen gegen 26 Versammlungsteilnehmende», teilte die Polizei heute mit. Demnach wurden bei der Demo gestern 28 Strafanzeigen aufgenommen – unter anderem wegen Beleidigung, Körperverletzung, versuchter

		Gefangenbefreiung, widerstands gegen vollstreckungsbeamte und volksverhetzung (...)“
„Other 4“ / Regionalspezifische Texte: Raum Essen/NRW	„essen“ „samidoun“ „duisburg“ „reul“ „cdu“ „islamisten“ „düsseldorf“	„Nach der umstrittenen Pro-Palästina-Demonstration in Essen vom Freitag ermittelt die Polizei gegen einen der Redner wegen Volksverhetzung. Die SPD-Landtagsfraktion beantragte eine Debatte zu dem Thema im Innenausschuss. Essens Oberbürgermeister Thomas Kufen zeigte sich empört: "Nur schwer erträglich. Islamisten, Antidemokraten und Judenhasser ziehen geschützt durch das deutsche Grundgesetz durch Essen (...)“.
Other_5 / „Queers for Palestine“ / LGBTQ* und Haltung zu Palästina/Israel	„queers“ „pride“ „juden“ „januar“ „csd“	„(...)Die Kölner CSD-Szene sollte aufgemischt werden, in anderen Städten wird dies auch der Fall sein; in Berlin treffen sich die linken Queers ohnehin immer separat, viele Jahre beim Alternativen CSD in Kreuzberg, gern die Verbrennung einer israelischen Fahne inklusive, aber auch jetzt: Der große Berliner CSD am 27. Juli soll durch eine Kultur gestört werden, die in der Formel gebündelt ist, die man inzwischen landläufig auf Transparenten zu sehen bekommt: Queers for Palestine . Jene, die sich hinter diesem Schild sammeln, sind, man muss es so deutlich sagen, Feinde schwuler, lesbischer und trans Emanzipation. Sie dämonisieren Israel, sie behaupten, das LGBTI*-Leben in diesem jüdischen Land, besonders in Tel Aviv, sei eine Marketingaktion des zionistischen Staates, um davon abzulenken, dass es sich in Wahrheit um eine

		Strategie gegen Palästinenser handelt.(...).
„Other 14“ / Geopolitische Hintergründe/ Krieg im Nahen Osten	„osten“ „feuerpause“ „soldaten“ „deutsche“ „gaza-streifen“ „krieg“ „eskalation“	(...) Die israelische Armee hat nach eigenen Angaben "die Zerschlagung der militärischen Struktur der Hamas im nördlichen Gaza-Streifen abgeschlossen". Die militärische Struktur der Hamas im Norden des Gaza-Streifens sei "demontiert". Sprecher Daniel Hagari sagte am Samstag, die Hamas habe vor Kriegsbeginn vor drei Monaten im Norden des Küstenstreifens über zwei Brigaden mit zwölf Regimentern verfügt. "Insgesamt waren es etwa 14.000 Terroristen", sagte Hagari. Es seien seitdem zahlreiche Kommandeure getötet sowie Waffen und Munition zerstört worden (...).

Tabelle 2

Besonders gut interpretierbar ist das Thema „Antisemitismus“. So sind alle zuvor definierten Keywords prominent in den dem Thema zugeordneten Texten vertreten. Werden exemplarisch einige Texte betrachtet, dann sind auch diese inhaltlich klar zuzuordnen; behandeln also den antisemitischen Gehalt der auf Demonstrationen getätigten Aussagen oder die vermeintlich antisemitische Gesinnung der Demonstrierenden.

Texte, die dem Thema „Kritik“ zugeordnet sind, nehmen zwar inhaltlich schlüssig Bezug auf einige der spezifizierten Keywords – darunter insbesondere „Genozid“ und „Völkermord“ sowie „Apartheid –, wider Erwarten findet hier aber in der Regel keine tiefere inhaltliche Auseinandersetzung mit den geäußerten Positionen statt. Stattdessen werden die genannten Begriffe eher als Indiz herangezogen, um hier ebenfalls Antisemitismus nachzuweisen: „Jetzt die Eskalation: "Fridays for Future International" wirft Israel einen "Genozid" (Völkermord) vor, obwohl Israels Armee sich gegen den Terror verteidigt“. Auffällig ist auch, dass ein Großteil der Texte, die unter dieses Topic fallen, Bezug auf die Bewegung „Fridays for Future“ und deren Gründerin Greta Thunberg nimmt: „Auf ein Statement zum Irrlichter-Kurs von Gretas "Fridays for Future"-Bewegung warteten die Zuhörer vergeblich. Auf X (früher Twitter) verweist Neubauer lediglich auf einen Beitrag des deutschen "Fridays for Future"-Ablegers. In diesem

heißt es: "Nein, der internationale Account spricht - wie zuvor betont - nicht für uns". Auch werden die Begriffe „greta“, „thunberg“, „bewegung“ und „neubauer“ sowie die Kollokation „fridays for future“ durch das Modell neben den vordefinierten Keywords als weitere dominante Begriffe für das Topic identifiziert. Der Vorwurf des Genozids wird also insbesondere hinsichtlich der ideologischen Positionierung der Klimaaktivistin Greta Thunberg diskutiert.

In Bezug auf das Topic „Forderungen“ zeigt sich, dass dies nicht das zu messen scheint, was über die Spezifizierung der Keywords intendiert wurde: Stattdessen informieren die zugeordneten Texte eher über die Hintergründe in der Region, auf die die Demonstrationen lediglich Bezug nehmen, was auch über die Keywords „hamas“, „armee“, „soldaten“, „getötet“ und „netanyahu“ deutlich wird. Die Begriffe „Waffenruhe“ und „Waffenstillstand“ werden im Zusammenhang mit der politischen Lage vor Ort und nicht als Inhalt der Positionen der Demonstrierenden thematisiert: „Israel bekräftigte am Dienstag, dass die Waffenruhe auf unbestimmte Zeit verlängert werden könne“.

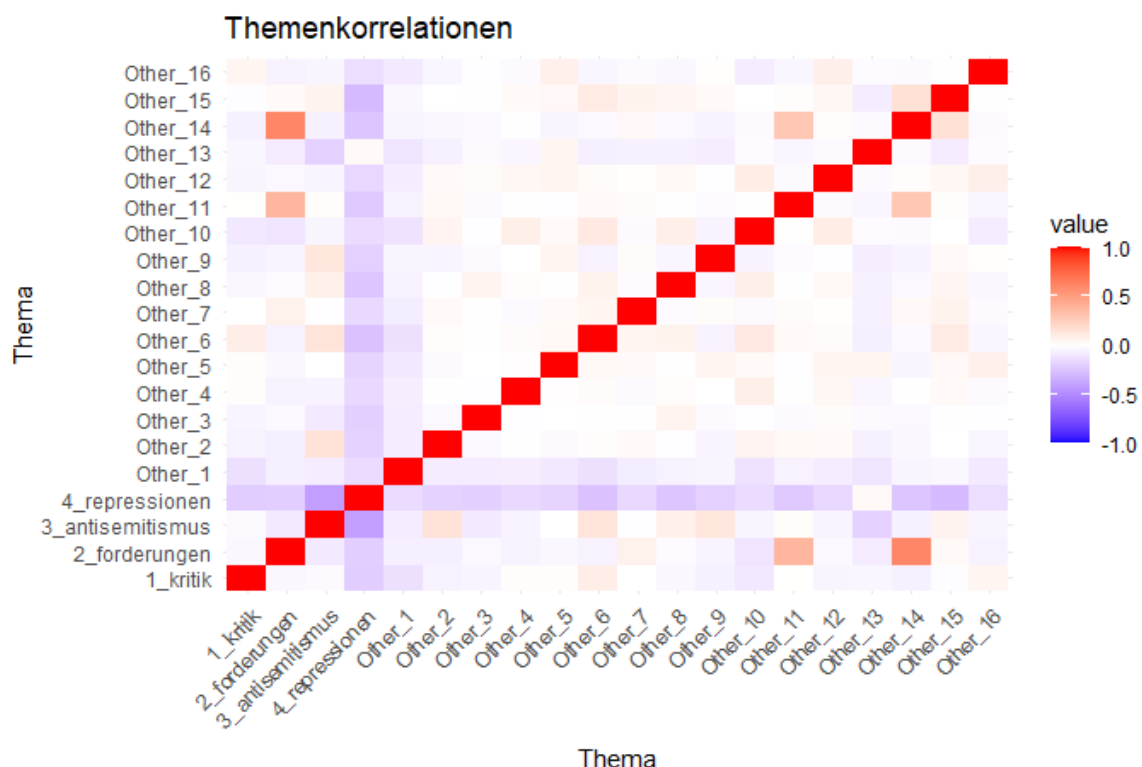


Abbildung 3

Werden die Korrelationen der verschiedenen Topics untereinander betrachtet, zeigt sich kein Zusammenhang zwischen „Kritik“ und „Forderungen“ – auch das entspricht nicht der ursprünglichen Erwartungen und verdeutlicht zusätzlich, dass über das hier modellierte Topic ein anderes latentes Thema gemessen wird.

Eine hohe Korrelation ist hingegen zwischen „Forderungen“ und dem Topic „Other_14“ zu beobachten. Noch deutlicher wird hier, dass dieses das latente Thema der geopolitischen Hintergründe im Nahen Osten abbildet, insbesondere im Kontext kriegerischer Auseinandersetzungen, was auch Keywords wie „feuerpause“, „soldaten“, „gaza-streifen“, „krieg“ und „eskalation“ verdeutlichen. Ein weiteres kohärentes Thema stellt „Other_5“ dar – die Keywords „queers“, „pride“, „juden“, „csd“ verweisen auf LGBTQ*-Rechte, gleichzeitig scheint es aber auch um einen Zusammenhang mit jüdischem Leben zu gehen: „Sie dämonisieren Israel, sie behaupten, das LGBTI*-Leben in diesem jüdischen Land, besonders in Tel Aviv, sei eine Marketingaktion des zionistischen Staates.“ Entsprechend lässt sich dieses Topic auch als die Frage nach der Vereinbarkeit von Pro-Palästina- und LGBTQ*-Aktivismus interpretieren.

Einige Topics scheinen eher sehr spezifische Unterthemen zu behandeln, so beispielsweise regionalspezifische Nachrichten. Das zeigt sich auch anhand des ersten weiteren Topics (other_1), das regionale Nachrichten aus dem Raum Essen/NRW umfasst. Texte, die Demonstrationen in dieser Region thematisieren, werden zusammen gruppiert, wenngleich eine solche Zuordnung für die übergeordnete Forschungsfrage inhaltlich nur wenig sinnvoll ist. Ein möglicher Grund hierfür ist die recht niedrige Fallzahl: Da insgesamt nur eine vergleichsweise kleine Anzahl an Zeitungsartikeln vorliegt, werden teilweise Dokumente, die dasselbe Ereignis behandeln oder in derselben Region erscheinen, vom Modell einem Topic zugeordnet. Mit einer größeren Fallzahl ließe sich diese Einschränkung möglicherweise beheben². Einige der darüber hinaus modellierten Themen scheinen eher „Garbage“-Topics darzustellen, deren zugeordneten

² Über LexisNexis wurden zwar alle im genannten Zeitraum aufgelisteten Print- und Online-Artikel extrahiert – Berichte aus dem Online-Auftritt von TV- oder Hörfunkformaten (zb. Tagesschau oder Deutschlandfunk) konnten dabei allerdings nicht berücksichtigt werden. Zudem sind nicht alle in Deutschland erscheinenden Publikationen in der Datenbank vertreten. Möglicherweise könnte über weitere Webscraping-Methoden die Fallzahl entsprechend erhöht werden.

Begriffe kein inhaltlich schlüssig interpretierbares Thema bilden (z.B. Topic Other_12: „ukraine“, „frauen“, „jungs“, „märz“, „sonnenallee“, „euro“).

Auch in einer Visualisierung zeigt sich deutlich, dass die einzelnen Topics sehr unterschiedlich gut performen. Dabei bilden insbesondere „Kritik“ und „Forderungen“ klar erkennbare Cluster, die deutlich von den anderen Topics abzugrenzen sind.

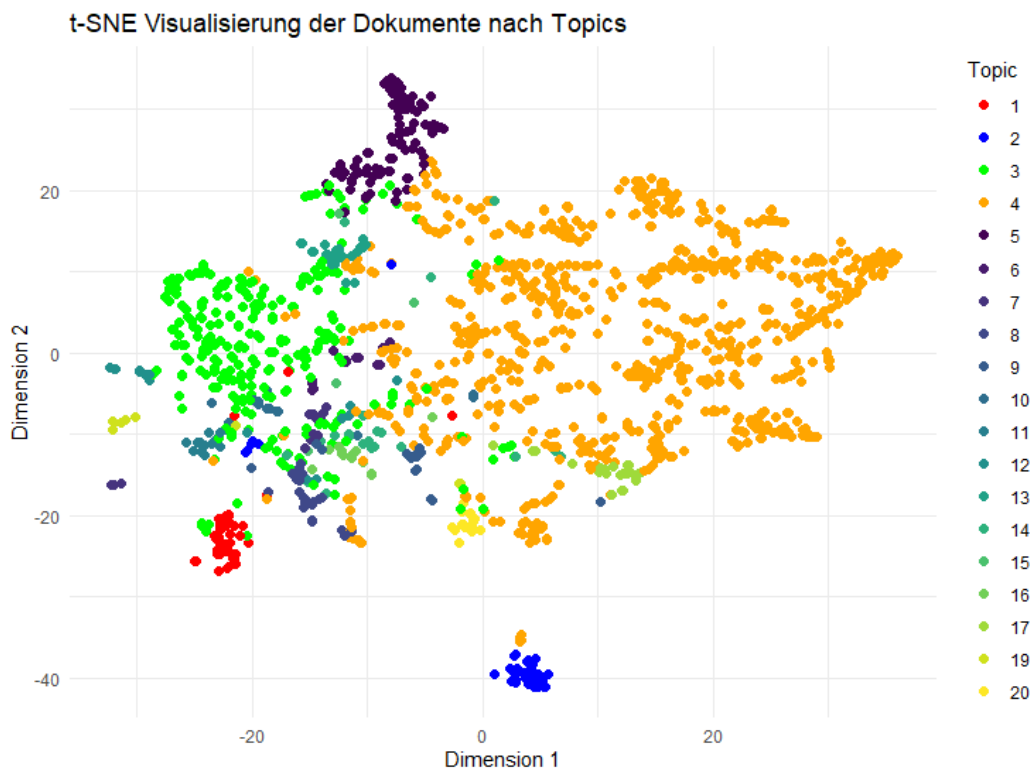


Abbildung 4

Besonders weit gestreut ist das Topic „Repressionen“: Hier liegt nahe, dass die Keywords sich stark über fast alle Dokumente verteilen – zum Beispiel deswegen, da fast alle Berichte über Demonstrationen Angaben der Polizei enthalten. Möglicherweise trägt also auch die Festlegung der Keywords, so wie sie in dieser Arbeit erfolgt ist, zu einer Schwächung der Modellgüte bei. Etwas überraschender ist die Streuung des Topics Antisemitismus – wenngleich das Topic insofern klar umrissen ist, als dass alle spezifizierten Keywords in den zugeordneten Dokumenten dominieren, und eine gute Interpretierbarkeit des Themas möglich ist, ist in der graphischen Darstellung kein klares

Cluster erkennbar. Auch hier liegt der Schluss nahe: Der vermeintliche antisemitische Gehalt der Pro-Palästina-Demonstrationen wird in einem Großteil der Zeitungsberichte aufgegriffen, sodass es in vielen Fällen zu Überschneidungen mit anderen Themen kommt.

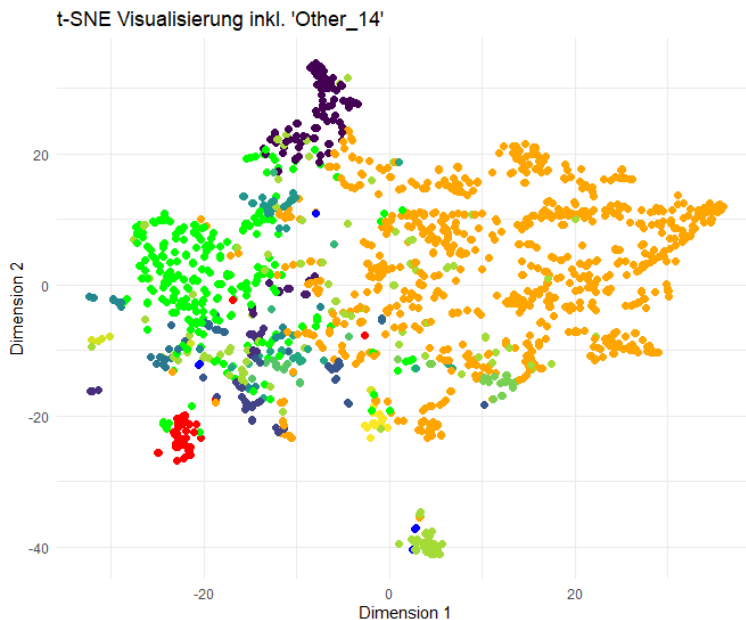


Abbildung 5

Auch die hohe Korrelation zwischen „Forderungen“ und „Krieg im Nahen Osten“ (Other_14) wird hier deutlich. Letzteres taucht in der Graphik gar nicht mehr auf – kein Dokument weist für dieses Thema die höchste theta-Wahrscheinlichkeit auf, sodass diesem Topic überhaupt kein Dokument mehr zugewiesen wurde. Wird stattdessen ein

Schwellenwert (hier: 0,5) definiert, der die Dokumente dann einem Thema zuweist, wenn sie statt eines maximalen Wertes lediglich diesen Schwellenwert überschreiten, zeigt sich in der graphischen Darstellungen, dass beide Themen sich nun vollständig überschneiden.

Betrachtet man die Kohärenz der einzelnen Topics, zeigt sich, dass diese insgesamt in einem niedrigen Bereich zwischen 0,03 und 0,2 schwanken, was auf eine eher schlechte Interpretierbarkeit hinweist. In Einklang mit der obigen Graphik performt hier das Topic „Repressionen“ besonders schlecht (0,039). „Antisemitismus“ und „Kritik“ liegen mit 0,147 bzw 0,142 eher im oberen Bereich, während „Forderungen“ dazwischen rangiert. Den höchsten Wert weist mit 0,2 das Topic „Krieg im Nahen Osten“ auf. Hier lässt sich vermuten, dass ein Verzicht auf die Spezifizierung der als ‚Forderungen‘ begriffenen Keywords, möglicherweise ein Zusammenführen der beiden Themen und damit insgesamt eine höhere Kohärenz und Modellgüte ermöglicht hätte.

Für eine zusätzliche quantitative Validierung wurde das keyATM-Modell mit verschiedenen Seeds berechnet³. Hier zeigt sich, dass die *keyword-assisted* Topics relativ stabil sind und das Modell überwiegend die gleichen Begriffe um die vordefinierten Keywords herum gruppiert. Bei den weiteren durch die Modelle identifizierten Topics zeigen sich hingegen teils deutliche Unterschiede. So tauchen in einigen Topics andere *Top words* auf, teils wurden Themen generiert, die in anderen Modellen gar nicht vertreten sind. Dies bestätigt den Eindruck, dass neben einigen wenigen prominenten Themen in den Dokumenten so spezifische Sub-Themen vorliegen, dass es dem Modell schwerfällt, inhaltlich sinnvoll interpretierbare Cluster zu identifizieren.

5. Fazit

In der Berichterstattung über Demonstrationen, die gegen die israelischen Angriffe im Gazastreifen protestieren, nimmt das Thema Antisemitismus einen wichtigen Stellenwert ein. Eine quantitative Auswertung mithilfe des *semi-supervised* KeyATM-Modells nach Eshima et al. (2024) konnte zeigen, dass wichtige Keywords prominent in den Daten vertreten sind und die Cluster, die um diese Keywords identifiziert wurden, schlüssig als „Antisemitismus“ interpretierbar sind. Dabei ist wichtig festzuhalten, dass hierbei nicht ein antisemitischer Gehalt der Berichterstattung identifiziert werden sollte, sondern die Berichte mögliche antisemitische Inhalte der thematisierten Demonstrationen zum Gegenstand haben. Anders als eventuell anzunehmen wäre, findet – wie exemplarische Texte zeigen – keine tiefergehende theoretische Reflexion zu diesem Vorwurf statt; so beispielsweise eine Auseinandersetzung mit dem verwendeten Slogan „*From the river to the sea, Palestine will be free*“, dessen antisemitischer Gehalt auch in der rechtlichen Beurteilung umstritten ist. Darüber hinaus greifen die Texte auch geopolitische Hintergründe auf oder thematisieren regionalspezifische innenpolitische Ereignisse in Zusammenhang mit den Demonstrationen. Letzteres ist allerdings auch mit der geringen Fallzahl des Samples zu erklären: Das Modell findet so zwar Cluster in den Daten – für Aussagen über die bundesweite Berichterstattung und darin dominierende latente

³ Alternativ ließe sich der Datensatz in Trainings- und Testdaten aufteilen, wobei die Trainingsdaten genutzt werden, um die Topics für die Testdaten vorherzusagen. Da der Datensatz hier aber ohnehin eine nur geringe Fallzahl aufweist, wird befürchtet, dass eine erneute Aufteilung in Test- und Trainingsdaten so zu starken Schwankungen und Overfitting führen könnte (vgl. Jiang et al., 2020: 682).

Themen sind diese Muster aber weniger aussagekräftig. Zugleich zeigt die Auswertung, dass das Topic „Repressionen“ tendenziell eher kein sinnvolles Thema abbildet. Das liegt unter anderem daran, dass das Keyword „Polizei“ in fast allen Dokumenten vorkommt. Auf einen Ausschluss häufig auftretender Begriffe wurde in der Datenvorbereitung verzichtet – auch dies trägt vermutlich zur Schwächung des Modells bei. In einer Überarbeitung wäre womöglich auch eine stärkere Orientierung an den Daten (s. Kapitel 2.2) gegenüber einer theoriegeleiteten Spezifizierung der *keyword-assisted* Topics sinnvoll. Die Hypothese, dass in der Berichterstattung auch inhaltliche Positionen der Demonstrierenden tiefergehend analysiert werden, konnte durch diese Arbeit nicht bestätigt werden. Insgesamt aber bietet das Modell nur mäßige Erklärungskraft. Für ein tieferes Verständnis aktueller Debatten um den Nahostkonflikt, seine Rezeption hierzulande und politische Reaktionen, sind möglicherweise auch weitergehende qualitative Analysen der hier untersuchten Daten hilfreich. Auch ein Vergleich der Berichterstattung über Pro-Palästina-Demonstrationen und der über rechtsextreme Proteste sowie dort verbreiteten Antisemitismus könnte sich als aufschlussreich erweisen.

Literaturverzeichnis

Airoldi, E. M. & Bischof, J. M. (2016). Improving and Evaluating Topic Models and Other Models of Text. *Journal of the American Statistical Association*, 111(516), 1381–1403.

Amnesty International (2024). Recht auf Protest für alle: Zu aktuellen Einschränkungen von Palästina-solidarischen Protesten in Deutschland. Verfügbar unter: <https://www.amnesty.de/aktuell/deutschland-einschraenkung-pro-paleastinensischer-proteste> (Zugriff am 05.03.2025).

Amnesty International. (2025). Kundgebungen für einen gerechten Frieden in Palästina und Israel. Berlin. Verfügbar unter: <https://www.amnesty.de/termine/kundgebung-fuer-einen-gerechten-frieden-palaestina-israel> (Zugriff am 19.03.2025).

ARD (2024): "hart aber fair": Israel im Krieg - Kritik erlaubt?, ARD. Verfügbar unter: <https://www.ardmediathek.de/video/hart-aber-fair/israel-im-krieg-kritik-erlaubt/das-erste/Y3JpZDovL3dkci5kZS9CZWl0cmFnLXNvcGhvcnEtZjkyYWE4NGMtYzkyZi00MzYwLWI1N2YtODIyOTU4NWZkZmY0> (Zugriff am 05.03.2025).

- Asmussen, C. B. & Møller, C. (2019). Smart literature review: a practical topic modelling approach to exploratory literature review. *Journal of Big Data*, 6(1), 1–18.
- Barkey, S. (2024). „From the River to the Sea“ gerufen: Berliner Demonstrantin wegen Volksverhetzung verurteilt. *Berliner Zeitung*. Verfügbar unter: <https://www.berliner-zeitung.de/news/from-the-river-to-the-sea-gerufen-erster-prozess-in-berlin-um-umstrittene-parole-li.2242176> (Zugriff am 18.03.2025).
- Beck, K. (2023). „Grundrecht, das nur Deutschen zusteht“: FDP will Versammlungsrecht für Ausländer einschränken. *Frankfurter Rundschau*. Verfügbar unter: <https://www.fr.de/politik/antisemitismus-demonstration-palaestina-fdp-grundrecht-leutheusser-schnarrenberger-zr-92671042.html> (Zugriff am 05.03.2025).
- Blei, D. M., Ng, A. Y. & Jordan, M. I. (2002). Latent Dirichlet Allocation. In T. G. Dietterich, S. Becker & Z. Ghahramani (Eds.), *Advances in Neural Information Processing Systems 14* (pp. 601–608). The MIT Press.
- Böldt, D. (2024). „Friedlicher Protest diskreditiert“: Berliner Linken-Spitze kritisiert Pro-Palästina-Kundgebung. *Der Tagesspiegel*. Verfügbar unter: <https://www.tagesspiegel.de/berlin/friedlicher-protest-diskreditiert-linken-spitze-kritisiert-pro-palastina-kundgebung-12251219.html> (Zugriff am 06.03.2025).
- Czymara, C. S. (2024). Real-World Developments Predict Immigration News in Right-Wing Media: Evidence from Germany. *Mass Communication and Society*, 27(1), 50–74.
- Dale, R., Moisl, H. & Somers, H. (Hrsg.). (2001). *Handbook of Natural Language Processing* (Bd. 27).
- Dietterich, T. G., Becker, S. & Ghahramani, Z. (Hrsg.). (2002). *Advances in Neural Information Processing Systems 14*: The MIT Press.
- Ding, K., Niu, Y. & Choo, W. C. (2023). The evolution of Airbnb research: A systematic literature review using structural topic modeling. *Heliyon*, 9(6), e17090.
- Dwyer, M. B., Damian, D. & Zeller, A. (Hrsg.). (05212022). *Proceedings of the 44th International Conference on Software Engineering*. New York, NY, USA: ACM.
- El-Shewy, M., Griffiths, M. & Jones, C. (2025). Israel's War on Gaza in a Global Frame. *Antipode*, 57(1), 75–95.

Eshima, S., Imai, K. & Sasaki, T. (2024). Keyword-Assisted Topic Models. *American Journal of Political Science*, 68(2), 730–750.

Gehrke, K. (2024). Nach Besetzung der FU Berlin: Hörsaal-Besetzerinnen bleiben straffrei. *Der Tagesspiegel*. Verfügbar unter: <https://www.tagesspiegel.de/berlin/nach-besetzung-der-fu-berlin-horsaal-besetzerinnen-bleiben-straaffrei-12659843.html> (Zugriff am 18.03.2025).

Guo, L., Vargo, C. J., Pan, Z., Ding, W. & Ishwar, P. (2016). Big Social Data Analytics in Journalism and Mass Communication. *Journalism & Mass Communication Quarterly*, 93(2), 332–359.

Hänsch, R. & Hellwich, O. (2017). Random Forests. In C. Heipke (Hrsg.), *Photogrammetrie und Fernerkundung* (S. 603–643). Berlin, Heidelberg: Springer Berlin Heidelberg.

Heipke, C. (Hrsg.). (2017). *Photogrammetrie und Fernerkundung*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Hildebrand, L. (2024). Freispruch für Pro-Palästina-Besetzerin: Humboldt-Universität verliert im Streit über Institutsbesetzung gegen Studentin. *DER SPIEGEL*. Verfügbar unter: <https://www.spiegel.de/panorama/justiz/berlin-humboldt-universitaet-erleidet-niederlage-im-streit-gegen-studenten-a-663d14a7-6441-4b1b-8fea-b3aa0e57eff7> (Zugriff am 18.03.2025).

Jiang, T., Gradus, J. L. & Rosellini, A. J. (2020). Supervised Machine Learning: A Brief Primer. *Behavior Therapy*, 51(5), 675–687.

Jüdische Stimme für einen gerechten Frieden in Nahost. (2025). Nicht in unserem Namen! Über uns. Verfügbar unter: <https://www.juedische-stimme.de/#about-info> (Zugriff am 19.03.2025).

Kastner, B. (2025). Amtsgericht: Palästina-Aktivistin verurteilt, weil sie Hamas-Kennzeichen verwendet habe. *Süddeutsche Zeitung*. Verfügbar unter: <https://www.sueddeutsche.de/muenchen/muenchen-urteil-palaestina-aktivistin-verbot-li.3219616> (Zugriff am 18.03.2025).

Kaur, R. P., Kumar, M. & Jindal, M. K. (2020). Newspaper text recognition of Gurumukhi script using random forest classifier. *Multimedia Tools and Applications*, 79(11-12), 7435–7448.

Kuhlen, R., Lewandowski, D., Semar, W. & Womser-Hacker, C. (Hrsg.). (2022). *Grundlagen der Informationswissenschaft*: De Gruyter.

Lehmann, R. (2023). Armin Laschet kritisiert Pro-Palästina-Demos in Deutschland. *noz.de*. Verfügbar unter: <https://www.noz.de/deutschland-welt/politik/artikel/armin-laschet-kritisiert-pro-palaestina-demos-in-deutschland-45738442> (Zugriff am 06.03.2025).

Lepsky, K. (2022). B 3 Automatisches Indexieren. In R. Kuhlen, D. Lewandowski, W. Semar & C. Womser-Hacker (Hrsg.), *Grundlagen der Informationswissenschaft* (S. 171–182). De Gruyter.

McKeown, K. R. & Radev, D. R. (2001). Collocations. In R. Dale, H. Moisl & H. Somers (Hrsg.), *Handbook of Natural Language Processing* (Bd. 27, S. 507–523). Verfügbar unter: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=479d9c0308f824eb27fd586f1071e4322598ba45>

Palästina e.V. (2023). Freiheit für Palästina! Demokratie für Deutschland! (Zugriff am 19.03.2025. Verfügbar unter: <https://verein-palaestina.org/>

Pinto Gurdíel, L., Morales Mediano, J. & Cifuentes Quintero, J. A. (2021). A comparison study between coherence and perplexity for determining the number of topics in practitioners interviews analysis. Retrieved from <https://repositorio.comillas.edu/xmlui/handle/11531/67714>

Shah, K., Patel, H., Sanghvi, D. & Shah, M. (2020). A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification. *Augmented Human Research*, 5(1), 1–16.

Tagesspiegel (2025). Demonstrationen: 650 Menschen bei Demo für Frieden in Palästina und Israel. Verfügbar unter: <https://www.tagesspiegel.de/berlin/demonstrationen-650-menschen-bei-demo-fur-frieden-in-palastina-und-israel-13210894.html> (Zugriff am 18.03.2025).

Tolochko, P., Balluff, P., Bernhard, J., Galyga, S., Lebernegg, N. S. & Boomgaarden, H. G. (2024). What's in a name? The effect of named entities on topic modelling interpretability. *Communication Methods and Measures*, 18(4), 349–370.

Tushev, M., Ebrahimi, F. & Mahmoud, A. (05212022). Domain-specific analysis of mobile app reviews using keyword-assisted topic models. In M. B. Dwyer, D. Damian & A. Zeller (Hrsg.), *Proceedings of the 44th International Conference on Software Engineering* (S. 762–773). New York, NY, USA: ACM.

Verenkotte, C. (2024). "From the river to the sea – Palestine will be free": Woher kommt der umstrittene Slogan. BR24. Verfügbar unter: <https://www.br.de/nachrichten/bayern/from-the-river-to-the-sea-palestine-will-be-free-woher-kommt-der-umstrittene-slogan,UI2Pxht> (Zugriff am 22.03.2025).

Waal, A. de & Barnard, E. (Hrsg.). (2008). Evaluating topic models with stability. *Evaluating topic models with stability*. Verfügbar unter: <http://hdl.handle.net/10204/3016>

Wright, M. N. & Ziegler, A. (2017). ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, 77(1), 1–17.

Wu, X., Nguyen, T. & Luu, A. T. (2024). A survey on neural topic models: methods, applications, and challenges. *Artificial Intelligence Review*, 57(2), 1–30.

Zhao, W., Chen, J. J., Perkins, R., Liu, Z., Ge, W., Ding, Y. et al. (2015). A heuristic approach to determine an appropriate number of topics in topic modeling. *BMC Bioinformatics*, 16 Suppl 13(Suppl 13), S8.

Zielezinski, J. (2024). Palästina-Demos: Bürgermeister verurteilt Demonstranten als Verfassungsfeinde. t-online. Verfügbar unter: https://www.t-online.de/nachrichten/ausland/internationale-politik/id_100501760/palaestina-demos-buergermeister-verurteilt-demonstranten-als-verfassungsfeinde.html (Zugriff am 06.03.2025).