Technische Innovation, theoretische Sackgasse? Chancen und Grenzen der automatisierten Inhaltsanalyse in Lehre und Forschung

Boczek, Karin; Hase, Valerie

Erstveröffentlichung / Primary Publication
Konferenzbeitrag / conference paper

Empfohlene Zitierung / Suggested Citation:
Boczek, K., & Hase, V. (2020). Technische Innovation, theoretische Sackgasse? Chancen und Grenzen der automatisierten Inhaltsanalyse in Lehre und Forschung. In J. Schützeneder, K. Meier, & N. Springer (Hrsg.), *Neujustierung der Journalistik/Journalismusforschung in der digitalen Gesellschaft: Proceedings zur Jahrestagung der Fachgruppe Journalistik/Journalismusforschung der Deutschen Gesellschaft für Publizistik- und Kommunikationswissenschaft 2019, Eichstätt* (S. 117-128). Eichstätt: Deutsche Gesellschaft für Publizistik- und Kommunikationswissenschaft e.V. https://doi.org/10.21241/ssoar.70828

Technische Innovation, theoretische Sackgasse? Chancen und Grenzen der automatisierten Inhaltsanalyse in Lehre und Forschung

Karin Boczek, Valerie Hase*

**beide Autorinnen haben gleichermaßen zu diesem Beitrag beigetragen*

Journalistisches Seminar, Institut für Publizistik, Johannes Gutenberg-Universität Mainz

IKMZ, Universität Zürich

Zusammenfassung

Die automatisierte Inhaltsanalyse wird auch in der Journalismusforschung zunehmend genutzt, um Texte (teil)-automatisiert zu analysieren. Sie hat damit zu einer Methodeninnovation im Fach beigetragen, die jedoch selten kritisch diskutiert wird. Der Beitrag gibt einen Überblick über Chancen und Grenzen der automatisierten Inhaltsanalyse in Forschung und Lehre. Er argumentiert für die Vereinbarkeit der Methode mit klassischen Journalismustheorien, weist aber auch auf Grenzen, etwa fehlende Analysetiefe oder mangelnde Validität, hin. Auch für die Vermittlung von Computational Methods werden Probleme offengelegt und Lösungsansätze auf organisatorischer, technischer und didaktischer Ebene diskutiert. Insgesamt bedeutet die automatisierte Inhaltsanalyse für die Journalismusforschung keine Neujustierung, sondern eine Erweiterung – sofern Grenzen der Methode beachtet werden. Anders sieht es bei der Lehre aus, wo eine verstärkte Vermittlung von Computational Methods gefordert wird.

Keywords: Automatisierte Inhaltsanalyse, Computational Methods, quantitative Methoden, Journalismus, Lehre

Summary

The use of automated content analysis has rapidly increased. In journalism studies, (partly) automized text analysis has led to a methodical innovation in the field. However, implications of using this method are seldom discussed. This overview summarizes chances and limitations of the application of automated content analysis in research and teaching. We argue that the method integrates with pertinent theories but that limitations, e.g., a lack of depth in analysis or a lack of validity, also need to be acknowledged. We also shed light on problems of teaching computational methods and offer solutions on an organizational, technical and didactical level. In sum, the use of automated content analysis does not imply a readjustment, but an amplification of research in journalism studies – that is, if its limitations are kept in mind. However, when it comes to teaching, we argue for a readjustment by increasingly including computational methods in current syllabi.

Keywords: Automated content analysis, computational methods, quantitative methods, journalism, teaching

1. Neujustierung der Journalistik durch die automatisierte Inhaltsanalyse?

In einer digitalen Gesellschaft fallen vermehrte Datenspuren an, die maschinenlesbar sind. Gleichzeitig gibt es mehr Rechenkapazitäten und Methoden, um diese automatisiert auszuwerten. Dies trifft nicht nur auf Zahlen wie beispielsweise die Verweildauern bei Onlinemedien zu, sondern auch auf Texte (*Text-as-Data*). Im Journalismus sind für den Datenjournalismus nicht nur Zahlen aus amtlichen Statistiken, sondern auch politische Reden oder Literatur relevant. Bekannte Beispiele für journalistische Produkte, die auf automatisiert ausgewerteten Textdaten basieren, sind die Zeit-Online-Veröffentlichung „Bundeswörter" (Biermann et al., 2019), in der Bundestagsreden für Leser*innen durchsuchbar gemacht wurden, oder die Datenrecherche „Blaue Bücher, rosa Bücher" (Brunner et al., 2019) von der Süddeutschen Zeitung, die untersuchte, welche Geschlechter-Klischees in Kinderbüchern stecken.

Auch in der Journalismusforschung ergeben sich neue Potentiale durch die automatisierte Inhaltsanalyse. Forschende können Textdaten auf innovative Art und Weise, schneller und ressourcenschonender auswerten. Schon jetzt können sie auf eine zunehmende Anzahl von maschinenlesbaren journalistischen Inhalten zugreifen, beispielsweise digitalisierte Medienarchive, die Berichterstattung über lange Zeiträume erfassen. Künftige Potentiale von automatisierter Inhaltsanalyse, die Journalismusforschung voranzutreiben, entstehen auch aus der automatischen Transkription von Audioinhalten wie klassischem Radio oder Podcasts, die in der Journalismusforschung aufgrund des Aufwands bisher weniger stark bei Medieninhaltsanalysen berücksichtigt wurden.

Es gibt bereits zahlreiche Publikationen, die einen Überblick über verschiedene Verfahren der automatisierten Inhaltsanalyse geben (Grimmer & Stewart, 2013; Günther & Quandt, 2016; Maier et al., 2018; Trilling & Jonkman, 2018; Welbers et al., 2017). Nach einer kurzen Einführung in die automatisierte Inhaltsanalyse soll daher in diesem Beitrag gezeigt werden, wie man diese Verfahren mit Theorien der Journalismusforschung verknüpfen kann und wo die Grenzen der Methode liegen. Da die Methode bisher noch nicht Teil der Standard-Curricula der Journalistik-Ausbildung ist, wird in diesem Beitrag außerdem von den Lehrerfahrungen mit der Methode anhand eines beispielhaften didaktischen Konzepts für einen Workshop berichtet[i] und thematisiert, welche Schlüsse andere Lehrende daraus ziehen können. Der Beitrag soll eine methodische Neujustierung der Journalistik in Bezug auf zwei Aspekte diskutieren: einerseits aufzeigen, welche Potentiale in der automatisierten Inhaltsanalyse für die Journalismusforschung in einer digitalen Gesellschaft liegen und wo ihre Grenzen liegen, und andererseits Hinweise geben, was Lehrende bei der Vermittlung beachten sollten, um Studierende auf die verstärkte Nachfrage von Arbeitgeber*innen in Forschung und Journalismuspraxis nach diesen Fähigkeiten vorzubereiten.

2. Eine kurze Einführung in die automatisierte Inhaltsanalyse

Die automatisierte Inhaltsanalyse beschreibt eine Methode der Textanalyse, bei der ein Teil der Analyse von Textdaten automatisiert von Maschinen übernommen wird. Sie ist teilautomatisiert, da in vielen Verfahren der automatisierten Inhaltsanalyse manuelle Codierungen weiterhin eine Rolle für die Vorbereitung als auch die Validierung der Analyse spielen (Song et al., 2020). Zudem müssen Forschende zahlreiche Entscheidungen treffen, wenn es um die Bereinigung von Text zur Vorbereitung der Analyse oder das gewählte Analyseverfahren geht, die Maschinen nicht für sie übernehmen können. Ähnlich wie bei der manuellen Inhaltsanalyse ist das Ziel vieler automatisierter Verfahren, Texte explorativ zu verstehen (qualitative Inhaltsanalyse) oder das Vorkommen von Konstrukten in Texten numerisch zu erheben (quantitative Inhaltsanalyse). Ein zusätzliches Ziel vieler automatisierter Verfahren ist die Prognose und Klassifizierung unbekannter Texte, die auch für die Journalismusforschung von Interesse sein kann (de Grove et al., 2020). Ein klassisches Beispiel aus der journalistischen Praxis ist die automatisierte Klassifizierung von Hate Speech in sozialen Netzwerken.

Ein weiterer wichtiger methodischer Unterschied zwischen der automatisierten und der manuellen Inhaltsanalyse ist dabei ein unterschiedliches Verständnis von Text: Während manuelle Analysen Text auf Basis ganzer Artikel oder Sätze inklusive ihrer Syntax interpretieren, brechen automatisierte Analysen Texte auf sogenannte *Features* unter. Dies sind oft einzelne Wörter[ii] in einem *Bag-of-Words*-Ansatz, bei dem vor allem das numerische

Vorkommen dieser Wörter wichtig ist, aber ihre Reihenfolge oder Position im Text ignoriert wird. Auf Basis des numerischen Vorkommens von Wörtern sollen dahinter liegende, latente Konstrukte von Interesse gemessen werden. Ein einfaches Beispiel dafür ist die Diktionär-Analyse: Diktionäre sind Listen von Wörtern (z. B. „fröhlich", „positiv", „Fortschritt"), deren Vorkommen erhoben wird, um beispielsweise Rückschlüsse auf das Vorkommen des latenten Konstrukts ‚positives Sentiment' zu ziehen. Interpretationsbasis für die automatisierte Analyse wären also numerische Wortfrequenzen, nicht wie bei der manuellen Analyse ganze Texte oder Sätze – weshalb man die Methode auch als *Text-as-Data* bezeichnet. Da von manifesten Indikatoren (z. B. Wortfrequenzen) auf latente Konstrukte (z. B. positives Sentiment) geschlossen wird, ist die automatisierte Inhaltsanalyse anfällig für Unsicherheiten und Fehler bei diesen Schätzungen (Grimmer & Stewart, 2013), ähnlich wie manuelle Analysen (Benoit et al., 2009).

Die automatisierte Inhaltsanalyse umfasst klassischerweise vier Schritte (Wilkerson & Casas, 2017): (1) die Sammlung von Textdaten, (2) die Bereinigung, durch welche Text in numerische Daten transformiert wird, (3) die Analyse und (4) die Validierung der Analyse. Für den dritten Schritt bietet die automatisierte Inhaltsanalyse eine Reihe deduktiver und induktiver Verfahren (Boumans & Trilling, 2016; Grimmer & Stewart, 2013; Günther & Quandt, 2016). Die in der Journalistik bekanntesten Verfahren sind aktuell Diktionäre, *Supervised Machine Learning* und *Unsupervised Machine Learning* in der Form von *Topic Modeling*. Diktionäre beschreiben ein deduktives Verfahren, bei dem Listen von Wörtern erstellt werden, deren Vorkommen erhoben wird, um Rückschlüsse auf bekannte latente Konstrukte wie z. B. Sentiment zu ziehen. Auch *Supervised Machine Learning* ist als deduktives Verfahren zu verstehen: Hier wird das Konstrukt von Interesse, z. B. das Vorkommen von negativem vs. positivem Sentiment in Text, zunächst manuell in einem Trainingsdatensatz codiert. Klassifikationsalgorithmen wie *Naive Bayes* werden dann auf Basis dieses Trainingsdatensatzes trainiert, um unbekannte Textdaten in einem Testdatensatz automatisch in vorgegebene Sentiment-Kategorien zu klassifizieren. Zuletzt wird *Topic Modeling* als eher induktives Verfahren des *Unsupervised Machine Learning* in der Journalistik häufig genutzt, um Themen und ihre Prävalenz in Texten anhand von häufig gemeinsam auftretenden Wörtern explorativ zu identifizieren. Viele der genannten Übersichtsarbeiten (Grimmer & Stewart, 2013; Günther & Quandt, 2016; Maier et al., 2018; Trilling & Jonkman, 2018; Welbers et al., 2017) geben detailliertere Einführungen in diese und weitere Verfahren der automatisierten Inhaltsanalyse. Der Fokus dieses Beitrags liegt auf der journalismustheoretischen Einbettung der Methode und zugehöriger Verfahren, weniger auf ihrer methodischen Umsetzung. Daher werden die im nachfolgenden genannten Verfahren und Metriken der automatisierten Inhaltsanalyse nur überblicksartig in Tabelle 1 erläutert. Für eine detaillierte Erklärung der jeweiligen Methoden wird daher auf die weiterführenden Quellen in der Tabelle verwiesen.

Wichtig ist, dass viele der genannten automatisierten Verfahren mit manuellen Inhaltsanalysen und anderen Methoden, wie beispielsweise Netzwerkanalysen, kombiniert werden können. Studien, die manuelle und automatisierte Inhaltsanalysen in Form von Mixed-Methods-Ansätze verbinden, nutzen automatisierte Verfahren beispielsweise, um Codebücher für qualitative Analysen zu erstellen (Waldherr et al., 2019) oder Stichproben für manuelle Analysen zu ziehen (Hase et al., 2020). In Kombination können diese Methoden nicht nur interdisziplinäre Zusammenarbeit fördern (Ophir et al., 2020), sondern Stärken und Schwächen der jeweiligen Methoden dabei ausbalancieren und Textdaten dadurch umfassender und detaillierter analysieren.

Tabelle 1. **Übersicht über Analyseverfahren und -Metriken**

Verfahren/Metrik	Beschreibung	Weiterführende Quelle(n)	Anwendungsbereich
Ähnlichkeitsmaße	Numerischer Wert für Ähnlichkeit zwischen zwei Objekten, hier Texten, basierend z.B. auf Worthäufigkeiten oder Wortketten. Genannte Beispiele: Kosinus-Ähnlichkeit, Jaccard-Index	Mozer et al. (2020)	Agenda-Setting und Agenda-Building & Nachrichtenwertforschung
Shingling	Vergleich von Texten anhand von Wortketten, d.h. aufeinander folgenden Wörtern	Manning et al. (2008), Nicholls (2019)	
Diktionäre	Wörterliste zu einem bestimmten Konstrukt (z.B. „fröhlich"; „positiv"; „Fortschritt" als Wörterliste für das Konstrukt ‚positives Sentiment'). Das Vorkommen dieser Wörter wird automatisch erhoben, um das Vorkommen des Konstruktes zu analysieren	Boukes et al (2019), Lind et al. (2019), Muddiman et al. (2019)	Nachrichtenwert-forschung
Named Entity Recognition	Automatische Erfassung von Personen, Organisationen oder Orten mittels z.B. Diktionären oder Supervised Machine Learning	Marrero et al. (2013)	
Supervised Machine Learning	Klassifikationsalgorithmen werden auf Basis eines Trainingskorpus trainiert, um zu erkennen, welche Textinformationen (z.B. Wortfrequenzen) die Klassifikation der Texte in bestimmte Kategorien (z.B. Sentiment, Themen) erlauben. Der trainierte Algorithmus wird genutzt, um bisher unbekannte Texte, d.h. den Testkorpus, automatisch zu klassifizieren	Mirończuk & Protasiewicz, (2018), Scharkow (2012)	Nachrichtenwert-forschung & Framing-Theorie
Topic Modeling	Verfahren des Unsupervised Machine Learning, bei dem Themen durch das gemeinsame Auftreten von Wörtern explorativ identifiziert werden	Maier et al. (2018)	Framing-Theorie

3. Forschung: Vereinbarkeit der automatisierten Inhaltsanalyse mit Theorien der Journalismusforschung

Durch die aus der Computerlinguistik, Informatik und Informationswissenschaft stammende Methode der automatisierten Inhaltsanalyse können z. B. Themen, Akteur*innen oder Sentiment als klassische Variablen manueller Inhaltsanalysen inzwischen automatisiert erfasst werden (Hase, im Druck). Wenige Studien diskutieren dabei kritisch, inwiefern die interdisziplinäre Methode empirisch das erfasst, woran die Journalismusforschung theoretisch interessiert ist – und wo ihre Grenzen liegen.

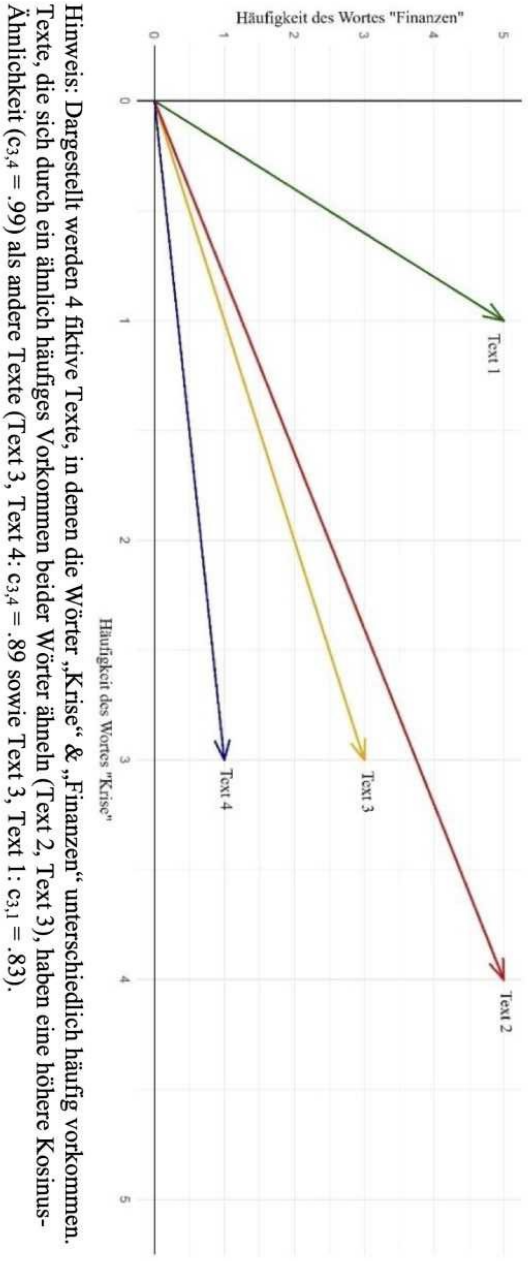
3.1 Chancen der Methode für die Forschung

Mit Blick auf „Agenda-Setting und Agenda-Building", die „Nachrichtenwerttheorie" und die „Framing-Theorie" als ausgewählte Journalismustheorien mittlerer Reichweite (Löffelholz & Rothenberger, 2016), wird im Folgenden diskutiert, inwiefern die automatisierte Inhaltsanalyse mit Fragestellungen und Theorien der Journalismusforschung vereinbar ist. Es bleibt abzuwarten, ob diese Methodeninnovation auch zu neuen Journalismustheorien führen kann.

Agenda-Setting und Agenda-Building

„Agenda-Setting und Agenda-Building"-Studien analysieren, warum „die Medien über bestimmte gesellschaftliche Probleme [berichten] und über andere nicht" (Maurer 2016, S. 427) und welche Effekte dies auf das Publikum hat. Die Journalismusforschung ist dabei vorrangig an der Entstehung medialer Themenagenden interessiert, insbesondere inwiefern andere Akteur*innen dort inhaltliche Schwerpunkte setzen können. Automatisierte Verfahren im Bereich von *Text-Re-Use* (Mozer et al., 2020) können für solche Fragestellungen genutzt werden: Studien untersuchen z. B., inwiefern Pressemitteilungen (Vogler & Schäfer, 2020) oder Agenturtexte (Welbers et al., 2018) in journalistische Artikel übernommen werden. *Text-Re-Use*-Verfahren zählen z. B. Wörter in Texten und verorten Texte dadurch als Vektoren in einem n-dimensionalen Vektorraum. Ähnlichkeiten zwischen Texten können dann über Kosinus-Ähnlichkeiten, d. h. den Kosinus des Winkels zwischen zwei Textvektoren, bestimmt werden (s. Abbildung 1). Ein anderes Verfahren in diesem Bereich ist *Shingling*. Dort werden Wortketten, etwa eine Kette von sieben aufeinander folgenden Wörtern, identifiziert, um zu analysieren, inwiefern diese Wortkette in anderen Texten genauso vorkommt (Manning et al., 2008; Nicholls, 2019), oft in Kombination mit Ähnlichkeitsmaßen wie dem Jaccard-Index als Maß für die Schnittmenge zweier Texte (Vogler & Schäfer, 2020). Unklar ist, ab welchem Grenzwert Texte als Duplikate gelten bzw. andere Akteur*innen „erfolgreich" ihre Inhalte in der medialen Agenda platziert haben. Gerade hier braucht es Journalismusforschende, die solche Fragen theoretisch klären.

Abbildung 1. Identifikation von Text Re-Use durch Kosinus-Ähnlichkeiten



Hinweis: Dargestellt werden 4 fiktive Texte, in denen die Wörter „Krise" & „Finanzen" unterschiedlich häufig vorkommen. Texte, die sich durch ein ähnlich häufiges Vorkommen beider Wörter ähneln (Text 2, Text 3), haben eine höhere Kosinus-Ähnlichkeit ($c_{3,4} = .99$) als andere Texte (Text 3, Text 4: $c_{3,4} = .89$ sowie Text 3, Text 1: $c_{3,1} = .83$).

Nachrichtenwerttheorie

Die „Nachrichtenwerttheorie" konzeptualisiert, welche Faktoren ein Ereignis „nachrichtenwürdig" machen. Automatisierte (wie auch manuelle) Inhaltsanalysen, die Nachrichtenfaktoren anhand der journalistischen Berichterstattung über ein Ereignis identifizieren, zeigen, inwiefern journalistische Beschreibungen mit medialer Aufmerksamkeit zusammenhängen (Eilders, 2016).[iii] Hier kann die automatisierte Analyse an vielen Stellen unterstützen: Das Vorkommen bestimmter Akteur*innen (Nachrichtenfaktor: „Elite-Nationen" oder „Personalisierung") kann automatisiert über das automatisierte Auslesen von Personalpronomen via *Part-of-Speech-Tagging* (Bednarek & Caple, 2014) oder *Named Entity Recognition*, d. h. der automatischen Erkennung von Personen, Organisationen oder Orten, erhoben werden. Diktionäre, d. h. Wörterlisten, oder *Supervised Machine Learning*, vortrainierte Klassifikationsalgorithmen, kommen dafür zum Einsatz (Burggraaff & Trilling, 2020). Wie geografisch nah ein Ereignis dem Publikum ist (Nachrichtenfaktor: „Nähe"), kann durch das automatisierte Auslesen von Orten aus Texten berechnet werden (Watanabe, 2018). Wie die journalistische Darstellung von Ereignissen den Nachrichtenwert eines Ereignisses beeinflusst (Nachrichtenfaktor: „Negativität"), lässt sich durch Sentimentanalysen mittels Diktionären oder *Supervised Machine Learning* erfassen (Stine, 2019). Und schließlich lassen sich bereits publizierte Artikel zu einem Ereignis (Nachrichtenfaktor: „Kontinuität") über genannte Ähnlichkeitsanalysen identifizieren (Burggraaff & Trilling, 2020). Auch bei der Analyse von Nachrichtenfaktoren – soweit diese auf Basis der journalistischen Beschreibung eines Textes erfasst werden sollen –, kann die automatisierte Analyse also einen Mehrwert bieten.

Framing-Theorie

Eine weitere relevante Theorie ist die „Framing"-Theorie, v. a. im Hinblick auf Medien-Frames als inhaltsanalytisch erfassbare Deutungsrahmen (Scheufele & Engelmann, 2016). Es gibt dabei sowohl induktive als auch deduktive Verfahren der automatisierten Erfassung von Frames: Induktive Verfahren nutzen *Topic Modeling* zur Identifizierung von Themen der Berichterstattung, die anschließend durch Netzwerkanalysen (Walter & Ophir, 2019) oder Clusteranalysen (van der Meer et al., 2019) zu Frames verdichtet werden. Dabei sind diese Frames nicht im Vorhinein bekannt, sondern werden explorativ identifiziert. Deduktive Verfahren leiten Frames aus der Theorie ab, lassen das Vorkommen dieser in einem Trainingskorpus manuell codieren und identifizieren Frames in bisher unbekannten Texten durch *Supervised Machine Learning* (Burscher et al., 2014). Problematisch ist hier allerdings gerade bei induktiven Verfahren, dass Studien selten validieren, inwiefern das, was automatisierte Analysen als „Frame" identifizieren, tatsächlich dem journalismustheoretischen Verständnis eines Frames entspricht.

3.2. Grenzen der Methode für die Forschung

Obgleich die automatisierte Inhaltsanalyse klassische Variablen manueller Inhaltsanalysen automatisiert erfassen kann und mit einigen Fragestellungen und Theorien der Journalismusforschung vereinbar ist, schränken vier Aspekte die Aussagekraft und das Potential der Methode ein.

Fehlender theoretischer Kontext: Strittig ist, inwiefern automatisierte Analysen komplexere theoretische Konzepte hinreichend erfassen können. Ein bereits erwähntes Beispiel ist die induktive Analyse von Frames: Ob die durch automatisierte Verfahren gefundenen Frames (van der Meer et al., 2019; Walter & Ophir, 2019) dem journalismustheoretischen Verständnis von Frames entsprechen, wird bisher kaum validiert. Die Gleichsetzung automatisiert erfasster Frames mit dem theoretischen Verständnis von Frames in unserem Fach könnte zu einem weiteren Zerfransen des ohnehin eher uneinheitlichen Framing-Begriffes führen. In vielen Studien werden empirische Ergebnisse also nicht genügend vor dem Hintergrund theoretischer Annahmen diskutiert.

Ohne Syntax keine Semantik: Die meisten automatisierten Verfahren arbeiten mit dem genannten *Bag-of-Words*-Ansatz, bei dem die Reihenfolge und der Kontext von Wörtern ignoriert werden. Für viele Fragestellungen ist aber Syntax eine Voraussetzung für Semantik – d. h. der Satzbau hat einen Einfluss auf die Bedeutung eines Satzes und zu erfassender Inhalte. Studien der Journalismusforschung beschäftigen sich beispielsweise damit, welche Inhalte von welchen Akteur*innen im medialen Diskurs vorangetrieben werden (Agenda-Setting und -Building) oder wie bestimmte Akteur*innen von wem bewertet werden (Framing). In vielen Bereichen können automatisierte Verfahren Fragestellungen, die journalismustheoretisch von Interesse sind, also (noch nicht) mit genü-

gend Tiefe untersuchen, auch wenn syntaktische Beziehungen zunehmend Einzug in Analysen erhalten (Fogel-Dror et al., 2019).

Fehlende Validierung: Der wichtigste Schritt automatisierter Analysen, die Validierung von Ergebnissen etwa durch Vergleich zu manuellen Codierungen, wird noch nicht zu Genüge durchgeführt (Song et al., 2020). Ein klassisches Beispiel für die Problematik mangelnder Validierung betrifft *Topic Modeling*. Hier werden Themen anhand von häufig gemeinsam vorkommenden Wörtern automatisch identifiziert. Ohne Prüfung der Validität – etwa durch manuelle Inspektion von Texten, in denen diese Themen mit hoher Wahrscheinlichkeit vorkommen – können Forschende, die sich nur auf die Ergebnisse der automatisierten Analyse verlassen, schnell zu Fehlinterpretationen darüber verleitet werden, welche Themen diese Wortlisten beschreiben (Quinn et al., 2010).

Fehlende Reliabilität: Zuletzt ist auch die automatisierte Analyse kein vollständig reliables Verfahren, selbst wenn das Verfahren durch die Syntax von Computerprogrammen scheinbar beliebig reproduzierbar scheint. Unabhängig von Entscheidungen bei der Bereinigung von Text, die Ergebnisse beeinflussen (Denny & Spirling, 2018), können Modelle zu lokalen Modi konvergieren, d. h. Ergebnisse können sich z. B. auf Basis von Starteinstellungen zwischen Rechnern unterscheiden (Roberts et al., 2016. Für die Kommunikationswissenschaft wäre es daher wichtig, sich ähnlich wie bei Intercoderwerten auf akzeptierte Reliabilitätsindizes und zugehörige Grenzwerte zu einigen, die zumindest bei der Codierung von Trainingsdatensätzen für *Supervised Machine Learning* auch zunehmend gefordert werden (Song et al., 2020).Ähnlich wie in der Journalismusforschung verschiedene Indizes existieren und sich der Grenzwert für akzeptable Intercoderwerte in einem gewissen Rahmen bewegen kann, wird es auch für automatisierte Analysen weder „den einen" Index noch „den einen" Grenzwert für reliable Analysen geben. Reliabilitätstests und -Ausweise verstärkt einzufordern, wäre jedoch auch eine Aufgabe für unser Fach, das diese Gütekriterien für manuelle Inhaltsanalysen mitentwickelt hat.

Zusammenfassend lässt sich sagen, dass automatisierte Inhaltsanalysen dann den größten Mehrwert bieten, wenn ihre Vorteile – etwa die vereinfachte Bereinigung durch Duplikatserkennung – mit den Vorteilen manueller Inhaltsanalysen – etwa der Detailtiefe der Analyse – kombiniert werden. Dabei sollte sowohl auf die theoretische Rückbindung der Operationalisierung als auch die Validität und Reliabilität automatisierter Analysen verstärkt geachtet werden. Die Journalismusforschung unterliegt also keinesfalls einer radikalen Neujustierung durch die automatisierte Inhaltsanalyse. Vielmehr sind ihre Theorien als Basis für automatisierte Analysen und ihre Gütekriterien als Messkriterien für die Validität und Reliabilität automatisierter Analysen gefragter denn je.

4. Lehre: Vermittlung automatisierter Inhaltsanalyse und ihre Herausforderungen

Nach einem Blick auf die mögliche Neujustierung der Forschung durch die automatisierte Inhaltsanalyse wird sich der folgende Abschnitt mit der Vermittlung der Methode in der Lehre und zugehörigen Problemen beschäftigen.

4.1 Beispielhaftes didaktisches Konzept für einen kompakten Workshop

Unsere Erfahrungen beruhen unter anderem auf einem Workshop zum Thema „Automatisierte Inhaltsanalyse" am Tag vor der Fachgruppentagung. Der Workshop war explizit auch für Teilnehmende ohne Vorkenntnisse geöffnet und umfasste ein Zeitfenster von drei Stunden sowie zwei Stunden für einen optionalen interaktiven Teil. Didaktisch bestand er aus fünf Elementen: (1) theoretische Einführung, (2) praktische Tutorials, (3) gemeinsames Vertiefen mit Expert*innen, (4) individuelles Vertiefen des Erlernten an eigenen Projekten mit Nachfragemöglichkeiten, (5) Syllabus zum Nachlesen.

Im Workshop wurde mit dem Statistikprogramm R anhand von frei verfügbaren Beispieldatensätzen von Wikinews[iv] gearbeitet. Alle Teilnehmenden bekamen vorher die Datensätze sowie R-Skripte gestellt, so dass sie den Code nicht während des Workshops mitschreiben mussten und sich aufs Zuhören und Ausprobieren konzentrieren konnten. Da der Workshop im Zweier-Team geleitet wurde, konnte die nicht vortragende Person bei spontanen Problemen der Teilnehmenden helfen. Nach einer theoretischen Einführung in die Logik der automatisierten Inhaltsanalyse lernten die Teilnehmenden in praktischen Tutorials das Einlesen und Bereinigen von Textdaten, einfache Frequenzanalysen und fortgeschrittene Methoden wie *Topic-Modeling* und

Sentiment-Analysen kennen. Nach den Tutorials konnten sich die Teilnehmenden im optionalen Teil für ein Skype-Gespräch mit Expert*innen entscheiden. In den Videotelefonaten wurden Fragen zur praktischen Umsetzung automatisierter Inhaltsanalysen im Alltag von Forschenden besprochen.[v] Danach arbeiteten die Teilnehmenden in der Diskussion mit den Dozierenden und der Gruppe daran, das im Workshop erlernte auf eigene Forschungsfragen und Datensätze anzuwenden. Anschließend bekamen alle Teilnehmenden einen Syllabus mit einer Zusammenstellung von nützlicher Literatur zu den Methoden, der Anwendung dieser in empirischen Studien und praktischen Tutorials zur Vertiefung.[vi]

4.2 Herausforderungen für die Lehre

Auf Basis unserer Lehrerfahrungen werden wir im Folgenden thematisieren, welche Herausforderungen man bei der Lehre von Computational Methods, hier der automatisierten Inhaltsanalyse, beachten sollte.

Organisatorische Herausforderungen: Solange Kurse in automatisierter Inhaltsanalyse nicht fest in die Curricula für Journalistik-Studierende integriert sind, bleibt wenig Zeit für die (teilweise freiwillige) Methodenausbildung. Das Didaktikkonzept-Beispiel (siehe 4.1) zeigt, wie es möglich ist, in einem kurzen Zeitfenster einen Überblick über die automatisierte Inhaltsanalyse zu geben. Um die Zeitrestriktionen abzufangen, war es neben der Bereitstellung von Materialien vor und nach dem Kurs hilfreich, im Zweier-Team zu lehren, so dass die andere Person während des Vortrags und der Übungen dringende Fragen beantworten konnte.

Technische Herausforderungen: Bei der Software für automatisierte Inhaltsanalyse ist aktuell R in der Journalismusforschung dominant, obgleich andere Disziplinen weitere Programmiersprachen wie beispielsweise Python nutzen. Bei der Entscheidung für einzelne R-Pakete sollten sich Lehrende ganz bewusst mit Lock-In-Effekten beschäftigen, die dazu führen, dass Teilnehmenden und auch Lehrenden selbst dauerhaft bestimmte Analyse-Arbeitsabläufe verschlossen bleiben, die alternative R-Pakete geboten hätten (s. Link zum Workshop-Syllabus in den Fußnoten). Eine weitere Herausforderung ist es, dass der Einstieg in die eigentliche Methoden-Lehre durch den unterschiedlichen Wissensstand der Teilnehmenden in Bezug auf R erschwert wird. Unserer Erfahrung nach dauert es relativ lange bis die Teilnehmenden auf alle benötigen Dateien in R zugreifen können, um den inhaltlichen Aspekten zu folgen. Dies wird zusätzlich erschwert, wenn die Teilnehmenden mit eigenen Laptops und damit unterschiedlichen Systemvoraussetzungen und Betriebssystemen arbeiten. Beispielsweise muss das R-Working-Directory auf Mac und Windows anders angesteuert werden, was auch im Skript durch die Teilnehmenden angepasst werden muss. Diese Herausforderung kann dadurch abgemildert werden, dass es einen vorgeschalteten R-Kurs für alle Teilnehmenden gibt, die wenig Vorerfahrung mit dem Programm haben.

Didaktische Herausforderungen: Der heterogene Wissenstand der Kursteilnehmenden ist auch die größte didaktische Herausforderung bei Methodenkursen, weil die Motivation schnell sinkt, wenn der Fortschritt zu schnell oder zu langsam ist. Um ihr zu begegnen, kann die flexible Vor- und Nacharbeit mit einem Syllabus helfen. Außerdem kann es nützlich sein, Teilnehmende mit unterschiedlichem Wissensstand in eine Gruppe einzuteilen, sodass diese sich gegenseitig helfen können. Zusätzlich ist es wichtig, bei aufeinander aufbauenden Elementen Wiedereinstiegsmöglichkeiten für die Teilnehmenden zu schaffen, beispielsweise indem nach einer Übung zur Datenbereinigung ein bereinigter Datensatz bereitgestellt wird, mit dem alle weiterarbeiten können, die nicht mitgekommen sind.

Bei neuen Verfahren der automatisierten Inhaltsanalyse ist es häufig so, dass auch die Lehrenden selbst nicht alltäglich mit R-Paketen, Techniken und Tools arbeiten. Um trotzdem tiefergehende Erfahrungen zu aktuellen Entwicklungen und Methoden zu vermitteln, können Skype-Gespräche mit Expert*innen helfen. Allerdings braucht es auch dafür Ressourcen – sei es in Form von Bezahlung oder Zeit der wenigen ausgebildeten Kolleg*innen in diesem Bereich, die ihr Fachwissen oft unentgeltlich weitergeben.

5. Zusammenfassung

Insgesamt zeigt der Beitrag, dass die automatisierte Inhaltsanalyse bei Beachtung ihrer Grenzen eine methodische Erweiterung für die Journalismus-forschung bedeuten kann. Dazu bedarf es keiner Neujustierung der Disziplin, weil viele Verfahren der Methode mit bestehenden Journalismustheorien integrierbar sind, sondern nur einer Einigung auf gemeinsame Standards bei ihrer Nutzung.

In der Lehre muss sich die Journalistik auf die Vermittlung dieser Methode neu einstellen. Diese Neujustierung der Lehre sollte beinhalten, dass sich alle Studierenden und Wissenschaftler*innen im Laufe ihrer Ausbildung mit automatisierter Inhaltsanalyse beschäftigen können. Aktuell ist dies an vielen Hochschulen noch nicht ausreichend der Fall, da es zumindest in der Journalismusforschung noch kein etabliertes Angebot an Lehre in diesem Bereich gibt. In einer digitalen Gesellschaft sind Studierende wie Forschende mit diesem Wissen besser für den Arbeitsmarkt in Journalismus und Forschung vorbereitet, der in Stellenausschreibungen verstärkt eine Ausbildung in Programmierkenntnissen und Computational Methods nachfragt.

6. Literatur

Bednarek, M., & Caple, H. (2014). Why do news values matter? Towards a new methodological framework for analysing news discourse in critical discourse analysis and beyond. *Discourse & Society*, *25*(2), 135–158. https://doi.org/10.1177/0957926513516041

Benoit, K., Laver, M., & Mikhaylov, S. (2009). Treating word as data with error: Uncertainty in text statements of policy positions. *American Journal of Political Science, 53*(2), 495–513. https://doi.org/10.1111/j.1540-5907.2009.00383.x

Biermann, K., Blickle, P., Drongowski, R., Ehmann, A., Erdmann, E., Gortana, F., Lindhoff, A., Möller, C., Rauscher, C., Scheying, S., Schlieben, M., Stahnke, J., Tröger, J., & Venohr, S. (2019). 70 Jahre Bundestag: Darüber spricht der Bundestag. *Die Zeit*. https://www.zeit.de/politik/deutschland/2019-09/bundestag-jubilaeum-70-jahre-parlament-reden-woerter-sprache-wandel

Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars. *Digital Journalism*, *4*(1), 8–23. https://doi.org/10.1080/21670811.2015.1096598

Boukes, M., van de Velde, B., Araujo, T., & Vliegenthart, R. (2020). What's the tone? Easy doesn't do it: Analyzing performance and agreement between off-the-shelf sentiment analysis tools. *Communication Methods and Measures*, *14*(2), 83–104. https://doi.org/10.1080/19312458.2019.1671966

Brunner, K., Ebitsch, S., Hildebrand, K., & Schories, M. (2019). Blaue Bücher, rosa Bücher. *Süddeutsche.de*https://projekte.sueddeutsche.de/artikel/kultur/gender-wie-gleichberechtigt-sind-kinderbuecher-e970817/

Burggraaff, C., & Trilling, D. (2020). Through a different gate: An automated content analysis of how online news and print news differ. *Journalism*, *21*(1), 112–129. https://doi.org/10.1177/1464884917716699

Burscher, B., Odijk, D., Vliegenthart, R., de Rijke, M., & de Vreese, C. H. (2014). Teaching the computer to code frames in news: Comparing two supervised machine learning approaches to frame analysis. *Communication Methods and Measures*, *8*(3), 190–206. https://doi.org/10.1080/19312458.2014.937527

Denny, M. J., & Spirling, A. (2018). Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it. *Political Analysis*, *26*(2), 168–189. https://doi.org/10.1017/pan.2017.44

Eilders, C. (2016). Journalismus und Nachrichtenwert. In M. Löffelholz & L. Rothenberger (Hrsg.), *Handbuch Journalismustheorien* (S. 431–442). Springer VS. https://doi.org/10.1007/978-3-531-18966-6

Fogel-Dror, Y., Shenhav, S. R., Sheafer, T., & Van Atteveldt, W. (2019). Role-based association of verbs, actions, and sentiments with entities in political discourse. *Communication Methods and Measures*, *13*(2), 69–82. https://doi.org/10.1080/19312458.2018.1536973

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, *21*(3), 267–297. https://doi.org/10.1093/pan/mps028

De Grove, F., Boghe, K., & de Marez, L. (2020). (What) can journalism studies learn from supervised machine learning? *Journalism Studies, 21*(7), 912–927. https://doi.org/10.1080/1461670X.2020.1743737

Günther, E., & Quandt, T. (2016). Word counts and topic models. *Digital Journalism*, *4*(1), 75–88. https://doi.org/10.1080/21670811.2015.1093270

Hase, V. (im Druck). Automated content analysis. In F. Oehmer, S. H. Kessler, E. Humprecht, K. Sommer, & L. Castro Herrero (Hrgs.), *Handbook of standardized content analysis: Applied designs to research fields of communication science*. Springer VS.

Hase, V., Engelke, K., & Kieslich, K. (2020). The things we fear. Combining automated and manual content analysis to uncover themes, topics and threats in fear-related news. *Journalism Studies*, *21*(10), 1384–1402. https://doi.org/10.1080/1461670X.2020.1753092

Lind, F., Eberl, J.-M., Heidenreich, T., & Boomgaarden, H. G. (2019). When the journey is as important as the goal: A roadmap to multilingual dictionary construction. *International Journal of Communication*, *13*, 4000–4020.

Löffelholz, M., & Rothenberger, L. (2016). *Handbuch Journalismustheorien*. Springer VS. https://doi.org/10.1007/978-3-531-18966-6

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H., & Adam, S. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures*, *12*(2–3), 93–118. https://doi.org/10.1080/19312458.2018.1430754

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge.

Marrero, M., Urbano, J., Sánchez-Cuadrado, S., Morato, J., & Gómez-Berbís, J. M. (2013). Named Entity Recognition: Fallacies, challenges and opportunities. *Computer Standards & Interfaces*, *35*(5), 482–489. https://doi.org/10.1016/j.csi.2012.09.004

Maurer, M. (2016). Journalismus und Agenda-Setting. In M. Löffelholz & L. Rothenberger (Hrsg.), *Handbuch Journalismustheorien* (S. 419–430). Springer VS. https://doi.org/10.1007/978-3-531-18966-6

Mirończuk, M. M., & Protasiewicz, J. (2018). A recent overview of the state-of-the-art elements of text classification. *Expert Systems with Applications*, *106*, 36–54. https://doi.org/10.1016/j.eswa.2018.03.058

Monroe, B. L., Colaresi, M. P., & Quinn, K. M. (2008). Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis, 16*(4), 372–403. https://doi.org/10.1093/pan/mpn018

Mozer, R., Miratrix, L., Kaufman, A. R., & Jason Anastasopoulos, L. (2020). Matching with text data: An experimental evaluation of methods for matching documents and of measuring match quality. *Political Analysis*. Advance online publication. https://doi.org/10.1017/pan.2020.1

Muddiman, A., McGregor, S. C., & Stroud, N. J. (2019). (Re)claiming our expertise: Parsing large text corpora with manually validated and organic dictionaries. *Political Communication*, *36*(2), 214–226. https://doi.org/10.1080/10584609.2018.1517843

Nicholls, T. (2019). Detecting textual reuse in news stories, at scale. *International Journal of Communication*, *13*, 4173–4197.

Ophir, Y., Walter, D., & Marchant, E. R. (2020). A collaborative way of knowing: Bridging computational communication research and grounded theory ethnography. *Journal of Communication*, *70*(3), 447–472. https://doi.org/10.1093/joc/jqaa013

Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., & Radev, D. R. (2010). How to analyze political attention with minimal assumptions and costs. *American Journal of Political Science*, *54*(1), 209–228. https://doi.org/10.1111/j.1540-5907.2009.00427.x

Roberts, M. E., Stewart, B. M., & Tingley, D. (2016). Navigating the local modes of big data: The case of topic models. In R. M. Alvarez (Hrsg.), Computational Social Science (S. 51–97). Cambridge University Press. https://doi.org/10.1017/CBO9781316257340.004

Rosengren, K. E. (1974). International news: Methods, data and theory. *Journal of Peace Research*, *11*(2), 145–156. https://doi.org/10.1177/002234337401100208

Scharkow, M. (2012). *Automatische Inhaltsanalyse und maschinelles Lernen*. http://www.epubli.de/shop/buch/2000000013459

Scheufele, B., & Engelmann, I. (2016). Journalismus und Framing. In M. Löffelholz & L. Rothenberger (Hrsg.), *Handbuch Journalismustheorien* (S. 443–456). Springer VS. https://doi.org/10.1007/978-3-531-18966-6

Song, H., Tolochko, P., Eberl, J.-M., Eisele, O., Greussing, E., Heidenreich, T., Lind, F., Galyga, S., & Boomgaarden, H. G. (2020). In validations we trust? The impact of imperfect human annotations as a gold standard on the quality of validation of automated content analysis. *Political Communication*. Advance online publication. https://doi.org/10.1080/10584609.2020.1723752

Stine, R. A. (2019). Sentiment Analysis. *Annual Review of Statistics and Its Application*, *6*(1), 287–308. https://doi.org/10.1146/annurevstatistics-030718-105242

Trilling, D., & Jonkman, J. G. F. (2018). Scaling up content analysis. *Communication Methods and Measures*, *12*(2–3), 158–174. https://doi.org/10.1080/19312458.2018.1447655

van der Meer, T. G. L. A., Kroon, A. C., Verhoeven, P., & Jonkman, J. (2019). Mediatization and the disproportionate attention to negative news: The case of airplane crashes. *Journalism Studies*, *20*(6), 783–803. https://doi.org/10.1080/1461670X.2018.1423632

Vogler, D., & Schäfer, M. S. (2020). Growing influence of university PR on science news coverage? A longitudinal automated content analysis of university media releases and newspaper coverage in Switzerland, 2003-2017. *International Journal of Communication*, *14*, 3143–3164.

Waldherr, A., Wehden, L.-O., Stoltenberg, D., Miltner, P., Ostner, S., & Pfetsch, B. (2019). Induktive Kategorienbildung in der Inhaltsanalyse: Kombination automatischer und manueller Verfahren. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *20*(1). https://doi.org/10.17169/fqs-20.1.3058

Walter, D., & Ophir, Y. (2019). News frame analysis: An inductive mixed-method computational approach. *Communication Methods and Measures*, *13*(4), 248–266. https://doi.org/10.1080/19312458.2019.1639145

Watanabe, K. (2018). Newsmap: A semi-supervised approach to geographical news classification. *Digital Journalism*, *6*(3), 294–309. https://doi.org/10.1080/21670811.2017.1293487

Welbers, K., Atteveldt, W. V., & Benoit, K. (2017). Text analysis in R. *Communication Methods and Measures*, *11*(4), 245–265. https://doi.org/10.1080/19312458.2017.1387238

Welbers, K., van Atteveldt, W., Kleinnijenhuis, J., & Ruigrok, N. (2018). A gatekeeper among gatekeepers: News agency influence in print and online newspapers in the Netherlands. *Journalism Studies*, *19*(3), 315–333. https://doi.org/10.1080/1461670X.2016.1190663

Wilkerson, J., & Casas, A. (2017). Large-scale computerized text analysis in political science: Opportunities and challenges. *Annual Review of Political Science*, 20, 529–544. https://doi.org/10.1146/annurev-polisci-052615-025542

Fußnoten:

[i] Der Workshop wurde vom Nachwuchsnetzwerk Journalismusforschung (NaJoFo) veranstaltet und fand im Rahmen der DGPuK-Fachgruppentagung Journalistik/Journalismusforschung 2019 statt.

[ii] Neuere Verfahren bieten komplexere Varianten der Feature-Auswahl, indem etwa n-grams (zwei, drei oder mehr aufeinander folgende Wörter) einbezogen werden oder Features auf Basis von Part-of-Speech Tagging (d. h. ihrer Funktion im Satz) ausgewählt werden. Auf diese und weitere Möglichkeiten der Feature-Auswahl wird hier der Einfachheit halber nicht eingegangen (Monroe et al., 2008)

[iii] Ein zweiter traditioneller Ansatz der Nachrichtenwertforschung wäre der Vergleich der Berichterstattung mit Extra-Media-Daten (Rosengren, 1974), z. B. der Abgleich terroristischer Anschläge laut einer externen Datenbank mit medialer Berichterstattung.

[iv] https://de.wikinews.org/wiki/Hauptseite

[v] Ein herzliches Dankeschön für die Bereitschaft zur Teilnahme an diesem Experiment an Cornelius Puschmann (Bremen), Anke Stoll (Düsseldorf) und Martin Wettstein (Zürich)!

[vi] Der Syllabus ist unter diesem Link zu finden: https://github.com/valeriehase/automatisierteIA_DGPuK