

Lässt sich das Sicherheitsgefühl der Bevölkerung automatisiert erfassen?

Eine Fallstudie zur (künftigen) Rolle sozialwissenschaftlicher Methodologie im Zeitalter von Big Data

Wolfgang Aschauer · Alexander Seymer · Martin Weichbold ·
Thomas Herdin · Andreas Röser

© Der/die Autor(en) 2019

Zusammenfassung Seit Oktober 2015 sind wir als wissenschaftlicher Partner an zwei von der FFG geförderten Projekten im Sicherheitsforschungsprogramm KIRAS beteiligt. Im Zuge des ersten Projekts (Laufzeit bis Oktober 2016) wurde in Kooperation mit Unternehmensberatern und einer IT-Firma die Online-Plattform Foresight-Cockpit entwickelt. Diese hat das Ziel, die österreichischen Ministerien in die Lage zu versetzen, kollaborativ sowie ressortübergreifend frühzeitig auf unerwartete Trendbewegungen und Zukunftsszenarien im Bereich Sicherheit aufmerksam zu werden, sodass dauerhaft die Qualität des Risiko- und Krisenmanagements gesteigert werden kann. Im aktuellen Projekt (Laufzeit bis Oktober 2017) wurde eine softwarebasierte Lösung (genannt „Weblyzard“) zur Analyse von Nachrichtenquellen und Social-Media-Daten in das bestehende Tool integriert, um den künftigen NutzerInnen der Plattform die Analyse sicherheitspolitischer Lageperzeptionen und Stimmungslagen in der Bevölkerung zu erleichtern.

Die vorliegende Fallstudie soll primär die Möglichkeiten und Grenzen von Big-Data-Analysen auf Basis von Medienquellen im Vergleich zu klassischen quantitativen Inhaltsanalysen im Kontext einer Evaluationsstudie aufzeigen. Insofern entschieden wir, mit einer konventionellen Inhaltsanalyse auf Basis von zwei reichweiten-

W. Aschauer (✉) · A. Seymer · M. Weichbold · T. Herdin · A. Röser
Universität Salzburg, Salzburg, Österreich
E-Mail: wolfgang.aschauer@sbg.ac.at

A. Seymer
E-Mail: Alexander.Seymer@sbg.ac.at

M. Weichbold
E-Mail: martin.weichbold@sbg.ac.at

T. Herdin
E-Mail: thomas.herdin@sbg.ac.at

A. Röser
E-Mail: andreas.roeser@sbg.ac.at

starken Online-Medien (Der Standard und die Kronen Zeitung) parallel zum Weblyzard Häufigkeits- und Sentimentanalysen mit zwei unabhängigen BeobachterInnen durchzuführen, indem die Anzahl relevanter Artikel und UserInnen-Kommentare von ForscherInnen ausgezählt und nach Stimmung bewertet wurden. Die Ergebnisse weisen auf deutliche Schwächen und Fehlinterpretationen durch die Software hin, besonders wenn Stimmungsaussagen sowie zynische und irrelevante Statements in die automatisierte Auswertung einfließen.

Die Ergebnisse zeigen deutlich, dass empirische Forschung stets durch theoriegeleitete Interpretationen ergänzt werden sollte und nicht nur auf Potenziale, sondern auch gezielt auf Risiken des aktuellen Big-Data-Hypes hinweisen muss, um der Anziehungskraft automatisierter textbasierter Analysen entgegenzuwirken.

Schlüsselwörter Sicherheitswahrnehmungen · Soziale Medien · Big Data · Evaluationsstudie

Can a population's security perception be automatically measured?

A case study on the (future) role of social science methodology in the age of Big Data

Abstract Since October 2015, we have been involved as scientific partners in two FFG-funded projects (KIRAS security research program). The first project has developed a tool called Foresight-Strategy Cockpit. This software should enable the Austrian ministries to improve crisis management strategies by becoming aware of unexpected trends and future scenarios in security issues. The second project has integrated WebLyzard, a software tool for an automated analysis of online news and social media sources (comments on articles, Facebook postings and Twitter statements), to analyze the media representation of pressing societal issues and citizen's security perceptions.

The WebLyzard software was used in a case study where two independent observers performed a frequency and sentiment analysis in parallel to the automated Weblyzard results. Specific articles and user comments on selected key-topics in two major online newspapers in Austria ("Der Standard" and "Die Krone") were counted and evaluated according to different sentiment categories. The results indicate various weaknesses of the software leading to misinterpretations, and the automated analysis yielded substantially different results compared to the sentiment analysis by the two raters, especially for cynical or irrelevant statements.

Our case study highlights the potentials and limits of big-data analyses of media sources compared to those of conventional, quantitative content analysis. The results clearly show that empirical research should always be accompanied by theory-based interpretation and should highlight not only the potentials but also, specifically, the risks and weaknesses of the big data hype to counteract the allure of automated text-based analyses.

Keywords Security perceptions · Social media · Big data · Evaluation study

1 Einleitung

In einem von der FFG im Rahmen des Sicherheitsforschungsprogramm KIRAS geförderten Forschungsprojekt wurde 2015/16 von SoziologInnen, IT-SpezialistInnen sowie Unternehmens- und RechtsberaterInnen eine Online-Plattform namens Foresight-Cockpit entwickelt. Deren Ziel war es, den ebenfalls in die Entwicklung eingebundenen österreichischen Behörden ein Tool an die Hand zu geben, mit dem kollaborativ sowie ressortübergreifend frühzeitig aktuelle Trends im Bereich Migration und Sicherheit erkannt und Zukunftsszenarien entwickelt werden können. Dabei entstand die Idee, die für den gesellschaftlichen Diskurs in diesem Bereich relevanten Themen und Einschätzungen automatisiert aus verschiedenen Onlinequellen zu erfassen und zu integrieren; dieses Vorhaben war schließlich Teil eines zweiten Projektes („ForStrat-Cockpit“, Laufzeit Oktober 2016 bis September 2017), auf das in der Folge näher eingegangen wird.

Im Internet, insbesondere auf Nachrichtenportalen und Social-Media-Kanälen, können – so die Überlegung – die medialen und öffentlichen Diskurse zum Thema Sicherheit tagesaktuell erfasst werden. Worüber berichtet und diskutiert wird und welche Inhalte geteilt werden, sollte als Indikator für die gesellschaftliche Meinung in die politisch-administrative Arbeit der Ministerien eingehen und entsprechende Berücksichtigung für die Entwicklung von Szenarien finden. Um dies zu leisten, wurde die an der Modul-Universität Wien entwickelte Software „Weblyzard“ integriert, die in der Lage ist, Nachrichtenquellen und Social-Media-Daten automatisiert zu analysieren. Ziel war es, mithilfe des Programmes ein tagesaktuelles Lagebild relevanter Diskurse (basierend auf Nachrichtenquellen) und Sicherheitsempfindungen (basierend auf Social-Media-Quellen) zu erstellen. Das Tool bietet dazu ein breit gefächertes Portfolio zur Analyse und Visualisierung elektronischer Inhalte, unter anderem eine Frequenzanalyse zur Erhebung der Bedeutung von Themen; es kann aber auch die Valenz von Medienberichten im Sinne eines Stimmungsbarometers erhoben werden: Mithilfe komplexer Algorithmen wird gemessen, ob Themen in einem positiven oder negativen Licht dargestellt werden und ob es hinsichtlich der Meinungen zu diesen Themen starke Schwankungen gibt.

Im Folgenden werden der Projektverlauf und die Ergebnisse des Projektes skizziert, um von diesen Erfahrungen ausgehend Potenziale und Risiken von Big Data in den Sozialwissenschaften auszuloten und insbesondere die Konsequenzen für die bisherige quantitative Methodologie zu diskutieren. Wir wollen uns also einerseits *methodenkritisch* mit einer rein explorativen Nutzung von Big Data auseinandersetzen, wobei wir gleichzeitig auch – im Zuge unserer zweijährigen Projekterfahrung – über die schwierige Rolle der Soziologie in einem transdisziplinären Projekt reflektieren möchten.

2 Die Projektstruktur im Forstrat-Cockpit

Beim Projekt ForStrat-Cockpit handelte es sich um Antragsforschung im Rahmen des KIRAS-Sicherheitsforschungsprogramms, das die FFG in enger Abstimmung mit den beteiligten Sicherheitsabteilungen der Ministerien ausschreibt. Die Minis-

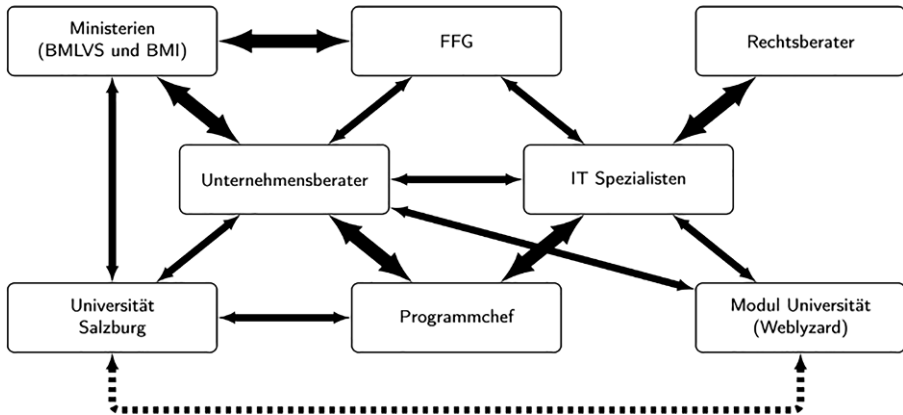


Abb. 1 Übersicht über die Projektstruktur und die beteiligten Akteure. (Quelle: eigene Darstellung)

terien sind frühzeitig in anstehende Projektvorhaben eingebunden, weil ihr Votum zur Bewilligung von Forschungsanträgen mitentscheidend ist. Ein weiteres Kriterium der Antragstellung ist die Beteiligung wissenschaftlicher Partner, die qualitativ hochwertige Forschung im Rahmen der Projekte gewährleisten sollen.

Das hier diskutierte Forschungsprojekt erläutert die Entwicklung einer webbasierten Foresight-Strategie-Plattform, die von einem Zukunftsforscher aus Deutschland initiiert wurde. Der Schlüsselakteur der beiden Projekte wird in der Abb. 1, die das komplexe Geflecht der involvierten Akteure wiedergibt, als „Programmchef“ bezeichnet, wobei die Software mittlerweile im Zuge eines Joint-Ventures vermarktet wird. Der Forschungsantrag wurde in enger Abstimmung mit einer Unternehmensberaterkanzlei und einem IT-Unternehmen formuliert, beide Projektpartner waren mit der Leitung des Projekts betraut. Zudem wurde auf drei wissenschaftliche Projektpartner zurückgegriffen: Ein Rechtsberater, der in enger Verbindung mit den IT-Spezialisten steht, zeichnete sich für die rechtliche Prüfung der Einbettung von Quellen in die Software verantwortlich. Im Rahmen des Projekts wurde entschieden, das Programm „Weblyzard“ der Modul-Universität Wien zur Analyse der Sicherheitsempfindungen zu verwenden. Die Software sollte im Kontext des Projekts themenspezifisch adaptiert und präzise auf die Bedürfnisse der BedarfsträgerInnen zugeschnitten werden. Die Abteilung Soziologie der Universität Salzburg und der Fachbereich Kommunikationswissenschaft wurden in das Projekt integriert, um eine konzeptuelle Analyse gegenwärtiger sicherheitsrelevanter Themen vorzulegen, einen Schlagwortkatalog zum automatisierten Abrufen von Stimmungslagen zu entwerfen und eine repräsentative Auswahl von Medienquellen zu gewährleisten. Im Anschluss an die Adaptierung des Weblyzards bestand unsere Rolle darin, die Funktionsfähigkeit des Weblyzards zu prüfen und die Qualität der automatisierten Extraktion von Medienquellen und Stimmungslagen einzuschätzen. Insofern führten auch wir unsere Tätigkeiten in enger Abstimmung mit dem „Programmchef“, den ProjektleiterInnen und den Ministerien durch. Insbesondere die Kooperation mit der Modul-Universität galt als herausfordernd, weil ein Teil unserer Aufgabe darin bestand, die automatisierte Analyse zu evaluieren und auf Schwachstellen hinzuweisen.

In der Konzeptionsphase des Projekts lag unsere Aufgabe darin, ein umfassendes Modell zu Sicherheitsempfindungen und sicherheitspolitischen Lageperzeptionen zu entwerfen (Abschn. 2.1). Um den Forschungsgegenstand einzugrenzen und eine automatisierte Analyse zu ermöglichen, wurde das Sicherheitsmodell anhand einer Menge von Schlagwörtern operationalisiert (Umsetzungsphase, Abschn. 2.2). Anschließend wurde ein umfangreiches Quellenverzeichnis erarbeitet (Selektionsphase, Abschn. 2.3), auf dessen Basis die automatisierte Analyse durchgeführt werden konnte. Die Ergebnisse der automatisierten Abfassung der Diskurse wurden schlussendlich mittels einer Evaluationsstudie geprüft (Abschn. 3). Auf Basis dieser Erkenntnisse wurden abschließend Qualitätskriterien entwickelt, die auf die Potentiale und Risiken in der Interpretation einzelner Medienquellen hinweisen (Abschn. 4). Abschnitt 5 des Artikels bettet die Ergebnisse der Evaluationsstudie in einen breiteren Kontext ein und weist auf potentielle methodische Fallstricke hin, die im Zuge einer unreflektierten Interpretation von automatisierten Analysen auftreten können.

2.1 Versuch einer ganzheitlichen Konzeption von „Sicherheit“

Die Konzeption des grundlegenden Analysemodells und die Auswahl der Themenfelder orientierte sich an verschiedenen AutorInnen (z. B. Daase 2010), die für einen erweiterten Sicherheitsbegriff plädieren und mehrere Analysedimensionen vorschlagen. Das Modell folgt deshalb einem mehrdimensionalen Sicherheitsbegriff, der über Konzepte wie existierende Kriminalitätsängste (z. B. Lüdemann 2006) oder soziale Unsicherheitslagen (z. B. Hirtenlehner 2009) hinausgeht. Auf persönlicher (und lokaler) Ebene scheinen in Österreich Kriminalitätsängste oder Arbeitsmarktrisiken gering auszufallen und es wird generell von einer hohen (individuellen) Lebenszufriedenheit berichtet. Dennoch scheinen zunehmend *gesellschaftliche Abstiegsängste* aufzubrechen, die sich in einem steigenden Zukunftspessimismus manifestieren. Umfragen (z. B. Heitmeyer 2010) deuten darauf hin, dass aktuell häufig von einer negativen Gesellschaftsentwicklung ausgegangen wird, obwohl die eigene Lebenssituation weiterhin in einem positiven Licht erscheint. Diese ambivalente Stimmungslage ist laut Bigo (2008, S. 34f.) Teil eines zunehmenden Sicherheitsdispositives, das für breite Bevölkerungsgruppen relevant ist und gezielte Gefährdergruppen (Terror, Migration) problematisiert. Damit geht automatisch einher, dass Sicherheit zunehmend transnational verhandelt wird, was zusätzlich zu einem Vertrauensverlust in die Handlungsfähigkeit der nationalen politischen Institutionen führt. Die Einbindung von Medienquellen sollte in dieser Hinsicht aufzeigen, inwieweit objektive, sicherheitsrelevante Indikatoren und subjektive Unsicherheitswahrnehmungen auseinanderdriften und auf welchen Ebenen politischer Handlungsbedarf besteht, Sicherheit adäquat zu kommunizieren und zu vermitteln. In einer *Gesellschaft der Angst* (Bude 2014) ist es notwendig, ganzheitlich Themenfelder zu identifizieren, die von der Bevölkerung als aktuelle Sicherheitsbedrohungen wahrgenommen werden. In umfangreichen Diskussionen im Projektteam wurden sechs zentrale Themenfelder identifiziert, die aus unserer Sicht die Kerndimensionen sicherheitspolitischer Lageperzeptionen verdeutlichen:

- Das Themenfeld „*kulturelle Diversität und Zusammenleben*“ fokussiert auf den Bereich Migration und Integration und thematisiert Herausforderungen, die sich für den gesellschaftlichen Zusammenhalt ergeben. Diese Dimension scheint derzeit die größten Verunsicherungen zu bewirken und ist daher auch im Zentrum des öffentlichen Diskurses.
- Damit in Zusammenhang steht die *Kriminalitätsfurcht*, die sich – im Kontext von Fluchtmigration – von internen Sicherheitsrisiken (z. B. sexuelle Belästigung von Frauen) bis hin zu externen Bedrohungen (z. B. Angst vor Terrorismus) erstreckt.
- Eine stärker zukunftsgerichtete Dimension ist jene des *technologischen Wandels*. In diesem Themenfeld sollen neue technologische Entwicklungen, die künftig die Gesellschaft prägen könnten, berücksichtigt werden, etwa auch die Umwälzungen der Arbeitswelt durch die Digitalisierung, obwohl diese derzeit in der Öffentlichkeit noch keine übermäßige Aufmerksamkeit erfahren.
- Damit in Verbindung steht die *ökonomische Dimension*. Themenbereiche wie soziale Ungleichheit in Europa (z. B. Fredriksen 2012) sind dabei genauso relevant wie Dynamiken in der gesellschaftlichen Mitte (z. B. Burzan und Berger 2010) oder Prekarisierungstendenzen (z. B. Standing 2011), die weitreichende Verunsicherungen bewirken.
- Der *strukturelle Wandel der Gesellschaft* (durch Globalisierung, Überalterung, Zuwanderung) stellt die Politik vor große Herausforderungen und begünstigt Gefühlslagen der Anomie (z. B. Bohle et al. 1997). Viele Menschen erleben sich als unbeteiligte ZuschauerInnen und haben das Gefühl, den rasanten gesellschaftlichen Veränderungen nichts entgegensetzen zu können.
- Folglich geraten auch die *gesellschaftlichen Institutionen* in die Krise. Es steigt die Politikverdrossenheit (z. B. Huth 2004), weil PolitikerInnen keine Lösungskompetenz in Bezug auf die drängenden Fragen der Gegenwart zugetraut wird. Vielfach sind die Menschen von der Komplexität der gesellschaftlichen Entwicklung überfordert und suchen nach simplen Erklärungsansätzen und einfachen Lösungsstrategien. Aus dieser Dynamik heraus könnten sich postdemokratische Entwicklungen auch in westlichen Gesellschaften weiter fortsetzen (z. B. Crouch 2008; Blühdorn 2013).

Um der Gefahr zu entgehen, bei der Analyse von Sicherheitsbedenken den Ängsten der Bevölkerung zu breiten Raum zu geben und eine einseitige Problemfokussierung vorzunehmen, braucht es einen adäquaten Mix von sicherheitspolitischen Herausforderungen und Maßnahmen (Lösungsstrategien). Deshalb ist im Analysemodell (siehe Abb. 2) neben der Sachdimension (Themenfelder) auch eine Handlungsdimension (Maßnahmen) integriert. Dies zeigt sich am augenscheinlichsten im Feld *kulturelle Diversität und Zusammenhalt*. Neben Problemwahrnehmungen wie *beidseitige Intoleranz* bestehen auch Chancen wie *kulturelle Verständigung und Integration*. Maßnahmen beziehen sich folglich auf Aspekte der gesellschaftlichen Eingliederung, wie beispielsweise *Sprachförderungen* oder *Arbeitsmarkteinbindung*.

Abschließend ist zu konstatieren, dass die österreichspezifischen Diskurse, die im Fokus des Projektes stehen, naturgemäß in einen nationenübergreifenden Kontext eingebunden sind (Raumdimension). Exemplarisch wurden für jedes Themenfeld externe Bedrohungslagen definiert, die den nationalen Diskurs zur gesellschaftlichen

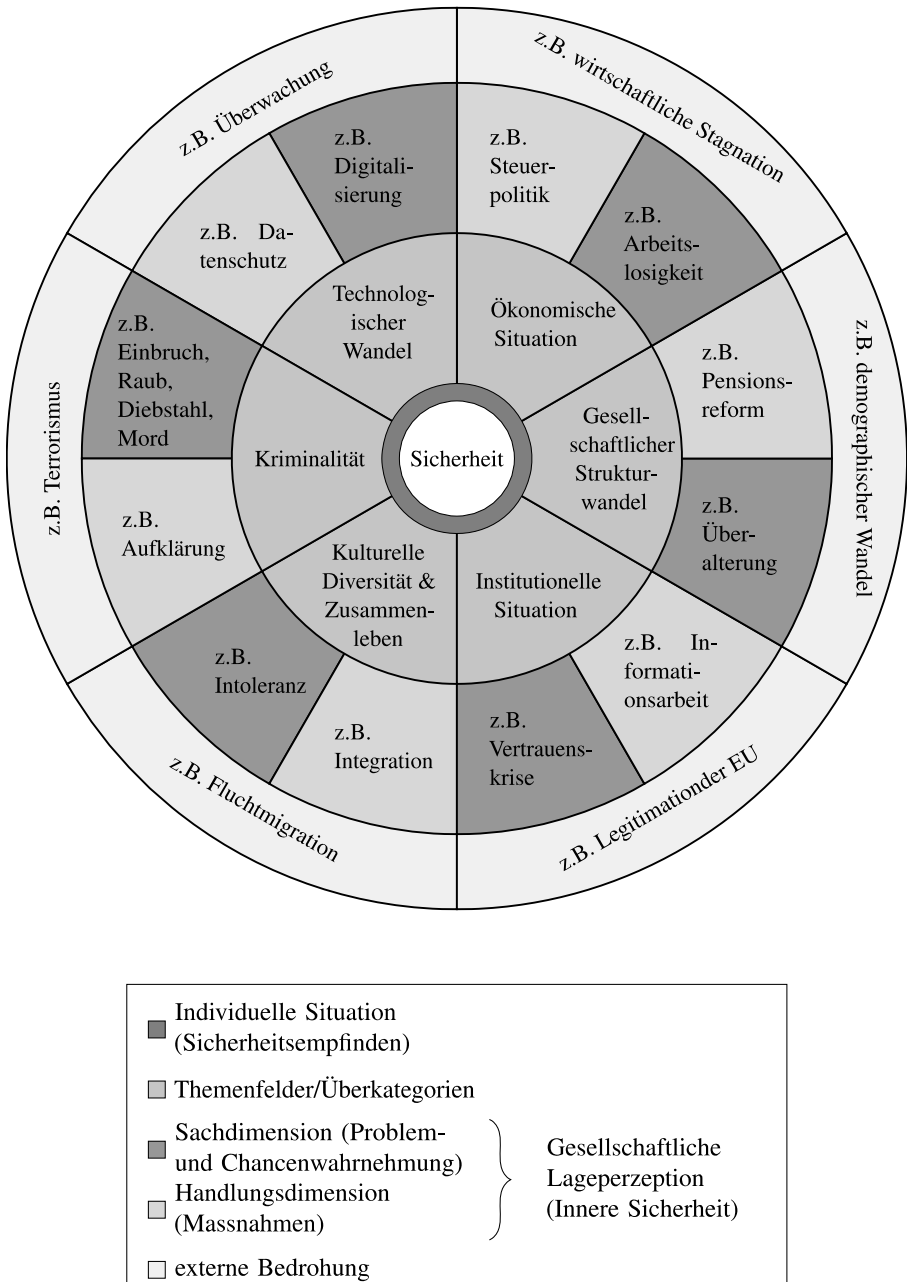


Abb. 2 Analysemodell zu den Sicherheitsempfindungen der ÖsterreicherInnen. (Quelle: eigene Darstellung)

Lage maßgeblich beeinflussen. Bezüglich der wirtschaftlichen Entwicklung besteht beispielsweise ein weitreichender Zukunftspessimismus, dass Europa im Banne aufstrebender Mächte an Einfluss und Wirtschaftskraft verliert (ökonomische Stagnation). Zusätzlich stehen etliche Länder Europas vor demografischen Problemen (z. B. Überalterung), wobei manche Maßnahmen (z. B. Kompensation des Bevölkerungsrückgangs durch Zuwanderung) auf Widerstand stoßen. Die EU sieht sich derzeit generell mit einer verstärkten Legitimationskrise konfrontiert, die sich vor allem durch das wenig erfolgreiche europäische Krisenmanagement im Zuge der Flüchtlingskrise weiter verschärft hat. Vielfach besteht der Eindruck, dass die Nationalstaaten der Sogwirkung globaler Entwicklungen nahezu ungeschützt ausgesetzt sind und an Handlungsmächtigkeit verlieren. Aus diesem Kontext heraus können auch die weit verbreitete EU-Skepsis, das gesellschaftliche Unbehagen (z. B. Aschauer 2017) und die stärkere Hinwendung zum Nationalstaat in vielen Ländern erklärt werden. Dennoch ist offensichtlich, dass globale Herausforderungen des Klimawandels, der Fluchtmigration oder der terroristischen Bedrohung nur über transnationale Risikogemeinschaften (z. B. Beck 2003) adressiert werden können.

2.2 Operationalisierung des Modells mittels Schlagwörtern

Anschließend wurden die genannten Themenfelder mittels einer umfangreichen Schlagwortliste für den Weblyzard operationalisiert, um eine automatisierte textbasierte Erfassung von sicherheitsrelevanten Diskursen zu ermöglichen. In der Auswahl der Schlagwörter wurde darauf geachtet, eine Balance zwischen positiven und negativen Termini zu erreichen und die Themenfelder zum Sicherheitsmodell vollständig und ganzheitlich abzudecken. Auf Basis eines umfangreichen Brainstormings (u. a. mittels eines Kreativworkshops mit Studierenden) und nach Durchführung einschlägiger Recherchen in Social-Media-Quellen wurde eine Liste von mehr als 1000 Schlagwörtern erarbeitet. Diese wurde gemeinsam mit den anderen Projektpartnern auf die Hälfte der Schlagwörter reduziert ($n = 517$), wobei in der Folge eine weitere „Glättung“ durch die Modul-Universität vorgenommen wurde. Besonders drastisch zeigte sich diese Reduktion der Schlagwörter im Themenfeld des strukturellen Wandels, siehe Tab. 1.

Es wurden für die Analyse ausschließlich jene Begriffe berücksichtigt, die ökologische Herausforderungen thematisieren. Sicherheitsbedenken, die auf gesundheitliche Entwicklungen, demografische Prozesse und soziale Problemlagen der Zukunft fokussieren, wurden weitgehend ausgeklammert und können mit der automatisierten Software folglich nur begrenzt abgegriffen werden. Diese Themenfelder finden zudem nur selten einen Weg in die massenmediale Berichterstattung, da sie aufgrund ihrer „Struktur“ (Schulz 1976, S. 69) einen geringen Nachrichtenwert (Schulz 1976; Galtung und Ruge 1965) besitzen. Je diffuser sich Themenfelder darstellen, desto seltener wird über sie berichtet, während über Themen, die thematisch eindeutig sind (Nachrichtenwert „unambiguity“, Galtung und Ruge 1965, S. 65), häufiger berichtet wird. Insofern resultierte zugunsten einer arbeitsökonomischen Anwendung der Software eine deutliche Verengung des Konzepts, wodurch Einschränkungen der Inhaltsvalidität der Themenfelder zu erwarten sind (z. B. Vijver und Tanzer 2004).

Tab. 1 Termliste „Struktureller Wandel“. (Quelle: eigene Darstellung)

Negativ	Positiv/neutral	Maßnahmen
Altersarmut	Aktivismus	Arbeitszeitverkürzung
<i>Atomenergie</i>	Alleinerziehend	<i>Artenschutz</i>
<i>Atomkraft</i>	Bildungschancen	Bildung
Ausgrenzung	Gewerkschaft	Karenzgeld
Burn-out	Ehrenamtlich	Kindergeld
Depression	Feminismus	<i>Klimaschutz</i>
Depressiv	Frauenbild	Medizinischer Fortschritt
Epidemie	<i>Erneuerbare Energie</i>	Papamonat
Erkrankung	Geburtenrate	Pensionsreform
Existenzangst	Genforschung	Pflegegeld
<i>Feinstaub</i>	<i>Gentechnik</i>	<i>Umweltschutz</i>
Generationskonflikt	Gleichberechtigung	Vorsorgeuntersuchung
Genderwahn	Gleichgeschlechtlich	
Parallelgesellschaft	Globalisierung	
<i>Hochwasser</i>	Lebenserwartung	
<i>Hungersnot</i>	Lebensgemeinschaft	
Jugendarbeitslosigkeit	Lebensstandard	
<i>Klimaerwärmung</i>	<i>Nachhaltigkeit</i>	
<i>Klimawandel</i>	Notschlafstellen	
Krebsrisiko	Patchwork-Familie	
Nahrungsmittelknappheit	Pflegebedürftig	
Zukunftssorgen	Pflegepersonal	
<i>Luftverschmutzung</i>	Privatvorsorge	
<i>Massentierhaltung</i>	Scheidungsrate	
Obdachlos	Sozialstaat	
<i>Ölkatastrophe</i>	Suppenküche	
Polarisierung	<i>Umweltaktivisten</i>	
Psychische Krankheiten	<i>Umweltbewusstsein</i>	
Schwarzarbeit	Wertewandel	
Sexismus	Wohlfahrtsstaat	
Suchtgefahr	Wohlstand	
Tierquälerei	Neoliberalismus	
Überalterung	Leistungsgesellschaft	
Zensur		
<i>Umweltverschmutzung</i>		
<i>Verseuchung</i>		
<i>Wegwerfgesellschaft</i>		
Zukunftsangst		
Zukunfts pessimismus		

Kursiv markiert sind jene Terme, die letztlich aufgenommen wurden

2.3 Der Korpus an Medienquellen im Weblyzard

Neben der Operationalisierung der Inhalte durch eine Schlagwortliste bestimmt eine umfassende Auswahl an Medienquellen die Güte einer automatisierten Analyse. Es ist sicherzustellen, dass Online-Nachrichtenmedien und Social-Media-Quellen ein breites Meinungsspektrum abdecken und möglichst nah an das Ideal einer umfassenden und objektiven Darstellung des öffentlichen Diskurses in Österreich herankommen, um valide Schlussfolgerungen über die Einschätzungen der sicherheitspolitischen Lage zu ermöglichen. Die bereits existierende Nachrichtendatenbank des Weblyzards stellte dabei eine gute Basis dar: Neben den wichtigsten Online-Tages- und Wochenzeitungen sind die relevantesten Magazine und Regionalmedien Österreichs angeführt und auch Nischenmedien finden eine adäquate Berücksichtigung. Zur Darstellung des öffentlichen Diskurses in Österreich wurde auf Kommentare in drei führenden Online-Medien (Standard, Kronenzeitung und ORF)¹ sowie auf die relevantesten Facebook- und Twitter-Accounts von Persönlichkeiten und Institutionen aus Politik, Wirtschaft, Medien und Zivilgesellschaft zugegriffen.² Zusätzlich ermöglicht das Tool auch eine Schlagwortsuche über die Datenbank aller Twitter-UserInnen, um die öffentliche Meinung zu sicherheitsrelevanten Themen abzubilden.

3 Evaluationsstudie zur Nutzbarkeit des Weblyzard im Rahmen des Projekts

Mit der Evaluation sollte untersucht werden, wie gut das Programm die tatsächliche Stimmungslage in Österreich im projektrelevanten Kontext abbilden kann. Als Benchmark wurde eine manuelle Medienanalyse durchgeführt. Da Inhaltsanalysen mit manuellem retrieval sehr aufwendig sind, musste eine Eingrenzung vorgenommen werden. Exemplarisch wurden die Onlineinhalte der Kronenzeitung (www.krone.at) und des Standard (www.derstandard.at) ausgewählt, da diese laut Österreichischer Web-Analyse (2018) die größte Online-Leserschaft und Reichweite besitzen. Zudem sind beide Nachrichtenportale aufgrund ihrer ausgeprägten Kommentarsektionen starke deliberative und partizipatorische Plattformen (vgl. Seethaler 2015, S. 76). Des Weiteren repräsentieren der Standard als überregionaler Vertreter von

¹ Im Zuge der Vorbereitungsarbeiten für die Evaluationsstudie ergab sich eine intensive Kooperation mit der Modul-Universität. Von Beginn an wurde kommuniziert, dass auch Diskussionsforen und Postings zu einschlägigen Artikeln zur Analyse von Stimmungslagen zentral sind. Die Modul-Universität hat dankenswerterweise diesen Vorschlag aufgegriffen und zusätzlich zu Facebook und Twitter auch noch Kommentare in den Weblyzard eingepflegt. Man orientierte sich dabei bewusst an zwei zentralen Onlinemedien (Kronenzeitung und Standard), weil diese beiden Online-Medien auch die Basis unserer Evaluation waren. Defizite in der Extraktion von Kommentaren können deshalb nur beschränkt als Schwäche des Tools gewertet werden, weil die Einbindung von Postings in die automatisierte Analyse unter großem Zeitdruck erfolgen musste.

² Die Relevanz der Medien orientiert sich stärker an Meinungsführern als an einem repräsentativen Querschnitt der Medienlandschaft. Dies ist ein Resultat aus dem potenziellen Anwendungsgebiet für die BedarfsträgerInnen im Projekt. Die Liste an Medienquellen wurde in einer Diskussionsrunde mit allen Projektpartnern konsensual finalisiert.

Qualitätszeitungen und die Kronenzeitung als Boulevardmedium verschiedene journalistische Klassifizierungen (vgl. Magin und Stark 2011, S. 107). Dies zeigt sich unter anderem durch die Verwendung unterschiedlicher sprachlicher Stilmittel, welche sich auf die automatisierte Codierung auswirken könnten. Die fünf Suchterme bilden insgesamt drei Themenbereiche aus den Sach- und Handlungsdimensionen der gesellschaftlichen Lageperzeption ab (siehe Abb. 2). Der Begriff Technologie erfasst den Themenbereich Digitalisierung, der Suchterm Arbeitsmarkt ist identisch zum Themenbereich Arbeitsmarkt, und der Themenbereich Radikalisierung wird über die Suchterme Radikalisierung, Rassismus und Muslime umgesetzt. Die Entscheidung für drei Suchterme für das Themenfeld Radikalisierung ergab sich aus der geringen Frequenz und der Eindimensionalität einzelner Suchterme, welche die vielfältigen Formen der Radikalisierung nicht erfassen können.

3.1 Datenbasis der manuellen Analyse

Die manuelle Medienanalyse erfolgte durch zwei wissenschaftliche MitarbeiterInnen der Abteilung für Soziologie und des Fachbereichs Kommunikationswissenschaft der Universität Salzburg, welche dieselben Artikel und Kommentare unabhängig voneinander einer Sentimentanalyse unterzogen. In einem zweiwöchigen Zeitraum vom 29.05.2017 bis 11.06.2017 wurden alle themenspezifischen Online-Artikel und -Kommentare aus den beiden Medien identifiziert und analysiert. Als Kommentare wurden alle Postings³ verwendet, welche innerhalb von 24h nach der Veröffentlichung eines identifizierten Artikels zu diesem auf der Webseite angezeigt wurden. Es wurden alle Kommentare gezählt, allerdings sind nur jeweils die ersten 200 in die Sentimentanalyse eingegangen, weil eine vollständige Sentimentanalyse die zeitlichen Kapazitäten der beiden RaterInnen überschritten hätte.⁴

Im Weblyzard erfolgt das Einlesen der Artikel und Kommentare im Referenzzeitraum mindestens täglich computergestützt und die Sentimentanalyse basiert auf einem Algorithmus (Weichselbraun et al. 2017, 2014, 2013). Die Daten für den Vergleich wurden am 11.07.2017 aus dem Weblyzard extrahiert.

Für die manuelle Analyse wurden alle Artikel, die über eine manuelle Suche nach den Suchtermen gefunden wurden, in einer Datenbank mit Titel, URL und Datumstempel gespeichert. Es erfolgte jeweils ein Abgleich der Google-Alerts-Suche (<https://www.google.at/alerts>) und der Suchergebnisse der Suchmaschine von Google News (<https://news.google.com>). Bei Google News wurde der Suchterm unter Einschränkung der Webseite der Zeitung und des Zeitraums definiert. Ein Suchergebnis bei Google Alerts oder Google News war hinreichend, um in die Datenbank aufgenommen zu werden. Wurde ein Artikel bei mehreren Suchabfragen identifiziert, so erfolgte die Zuordnung zu einem Themenbereich mit Verweis auf

³ Unterhalb der Artikel gibt es auf den Webseiten der beiden Medienangebote die Möglichkeit, Kommentare zu dem jeweiligen Artikel zu verfassen. Diese könnten auch als Postings bezeichnet werden, sollen hier allerdings zur leichteren Abgrenzung von Postings in den sozialen Medien als Kommentare referenziert werden.

⁴ Wir danken an dieser Stelle den beiden MitarbeiterInnen Lena Stöllinger und Patric Messner für die versierte und kompetente und vor allem auch zeitintensive Arbeit im Projekt.

Tab. 2 Identifizierte Artikel der RaterInnen pro Suchbegriff getrennt nach Quelle und Erhebungswoche. (Quelle: eigene Auswertungen)

	Der Standard			Kronenzeitung			Beide Quellen
	Woche 1	Woche 2	Gesamt	Woche 1	Woche 2	Gesamt	
Technologie	33	15	48	9	4	13	61
Arbeitsmarkt	16	11	27	2	1	3	30
Radikalisierung	2	3	5	3	1	4	9
Rassismus	13	8	21	2	1	3	24
Muslime	20	17	37	10	11	21	58
Gesamt	84	54	138	26	18	44	182

Woche 1 entspricht dem Zeitraum 29. Mai–04. Juni 2017 und die Woche 2 dem Zeitraum 05. Juni–11. Juni 2017

die Doppelnennung in den anderen Themenbereichen. Außerdem wurden Artikel, welche als inhaltlich irrelevant beurteilt wurden, einer entsprechenden Kategorie zugeordnet. Insgesamt resultierten aus den Suchanfragen 180 Artikel im Standard und 55 Artikel in der Kronenzeitung, siehe Tab. 2. Abzüglich der irrelevanten und doppelten Artikel blieben 138 bzw. 44 Artikel für die Evaluation übrig. Diese Artikel wurden in positives, negatives oder neutrales Sentiment eingestuft.

Im definierten Zeitraum von 24h ab Veröffentlichung eines Artikels wurden zu den 182 Artikeln insgesamt 26.440 Kommentare identifiziert. Pro Artikel wurden nur die ersten 200 bewertet, sodass die BeurteilerInnen jeweils 12.195 Kommentare einschätzten. Bei den Kommentaren wurde beim Sentiment zwischen positiv, negativ, neutral, zynisch und irrelevant unterschieden. Die Kategorie irrelevant wurde eingeführt, um Kommentare ohne klaren inhaltlichen Bezug zum Artikel zu klassifizieren. Außerdem wurde noch „zynisch“ als Kategorie aufgenommen, da manche Kommentare nicht in das Bewertungsschema eingeordnet werden konnten, aber dennoch eine klar destruktive Meinung artikulierten.

Artikel wie Kommentare wurden von den BeurteilerInnen unabhängig voneinander gelesen und bewertet. Die CodiererInnen erhielten kein spezifisches Training für die Aufgabe, verfügen aber über Codiererfahrung und haben das Projekt inhaltlich begleitet. Nach Abschluss der Codierung wurde das gesamte Untersuchungsmaterial einem Interrater-Reliabilitäts-Test unterzogen. Die statistisch bestimmte Interrater-Reliabilität ist als sehr hoch einzustufen, sodass von einer verlässlichen Messung ausgegangen werden kann.⁵

⁵ Für die Zeitungsartikel konnte ein Cohens Kappa von 0,871 (Kronenzeitung) und 0,92 (Standard) nachgewiesen werden. Die Korrelation zwischen den Beurteilungen der Kommentare war mit Werten von mindestens 0,7 ebenfalls sehr gut (zur Interrater-Reliabilität vgl. Greve und Wentura 1997, S. 111). Auch im Vergleich zu Ergebnissen einer Metaanalyse von 200 Studien (Lombard et al. 2002) ist die Interrater-Reliabilität als sehr gut einzustufen, denn die Autoren kommen zu der Schlussfolgerung, dass „[c]oefficients of 0.90 or greater are nearly always acceptable, 0.80 or greater is acceptable in most situations, and 0.70 may be appropriate in some exploratory studies for some indices“ (Lombard et al. 2002, S. 600).

3.2 Datenbasis der automatisierten Analyse

Analog zu den Suchkriterien der manuellen Analyse wurden Suchabfragen in den fünf Datenquellen des Weblyzard durchgeführt: Nachrichten, Facebook, Twitter-Accounts, Twitter-UserInnen und Kommentare, wobei der Vergleich zur manuellen Analyse sich auf die Nachrichten und Kommentare beschränkte. Die anderen Quellen wurden ausgewiesen, um einen Vergleich innerhalb des Weblyzards über die Gewichtung der Quellen aufzuzeigen. Um mögliche Abweichungen zwischen der manuellen Analyse und den Ergebnissen von Weblyzard erklären zu können, musste der Datenerfassungs- und Datenverarbeitungsprozess des Tools berücksichtigt werden. Die Datenerfassung erfolgte für die Datenquellen über unterschiedliche Prozeduren. Die Inhalte für die Kategorie Twitter-UserInnen wurden direkt über eine Programmierschnittstelle (API) über die vordefinierten Suchbegriffe abgegriffen. Die Daten entsprechen also einer direkten Suchabfrage in der Datenbank von Twitter. Über die Twitter-Accounts waren alle Tweets von 259 Twitter-Konten von Nachrichtenmedien, Institutionen und Personen öffentlichen Interesses verfügbar. Bei Facebook sind die Postings von 185 Facebookseiten berücksichtigt. Dabei wurden alle Postings auf der Facebookseite abgegriffen, solange der/die InhaberIn dies über die Einstellungen zur Privatsphäre erlaubte. Im Vergleich zu Twitter ist die Datenextraktion aus Facebook somit deutlich eingeschränkt. Nachrichten wurden über zwei parallellaufende Prozeduren abgegriffen: Einerseits erfolgte regelmäßig eine Vollspiegelung der Webseite der Nachrichtenquelle und andererseits wurde mehrmals täglich der RSS-Feed ausgelesen.⁶ Basierend auf den Artikeln, welche über die Nachrichten extrahiert wurden, werden die Kommentare berücksichtigt. Da die Artikel in der Vollspiegelung inklusive Kommentaren abgegriffen wurden, können die Kommentare im Anschluss extrahiert werden (vgl. Pollach et al. 2009; Scharl und Wechselbraun 2008).

Neben der computergestützten Sentimentanalyse wurden alle Dokumente, mit Ausnahme der Facebook-Postings und der Kommentare, auf die inhaltliche Relevanz geprüft. Die Prüfung erfolgte mithilfe der unter Abschn. 2.2 vorgestellten Schlagwortliste mit über 500 Wörtern, die aus einem theoretischen Modell abgeleitet wurden. Es wurden nur relevante Inhalte in die Datenbank aufgenommen, solche mit einer zu geringen Relevanz hingegen verworfen; dieser Punkt ist von besonderer Bedeutung, da in der Evaluation nicht nachweisbar ist, ob eine fehlende Übereinstimmung zwischen den Ergebnissen der manuellen Analyse und des Weblyzard in einer Diskrepanz der Suchergebnisse oder der Beurteilung der Relevanz begründet liegt. Wie der Weblyzard die Informationen aus den Quellen in die Daten in der finalen Datenbank transformiert, bleibt eine Blackbox, welche die Abweichung zwischen den Ergebnissen der manuellen und der automatisierten Analyse wesentlich erklärt. Die Datenbank bildet jedenfalls die Datengrundlage für den Weblyzard.

⁶ Ein Phänomen, welches bei der Prüfung der Daten aufgefallen ist, sind Dubletten. Da die Nachrichten mehrfach abgegriffen wurden, gelangten Artikel auch doppelt in die Suchergebnisse. Die identischen Nachrichten sind nicht nur mehrfach in der Datenbank enthalten, sondern auch die Bewertung des Sentiments und der Relevanz divergieren. Diese Dubletten sind nur bei der Kronenzeitung aufgetreten und würden die Anzahl der identifizierten Artikel fast halbieren.

Tab. 3 Übersicht über die extrahierten Ergebnisse aus dem Weblyzard^a. (Quelle: eigene Auswertungen)

	Twitter-Accounts	Facebook	Twitter-UserInnen	Artikel	Kommentare
Arbeitsmarkt	17	18	1812	35	11
Muslime	17	80	16.267	42	38
Radikalisierung	4	8	274	15	6
Rassismus	3	23	4778	22	8
Technologie	4	14	251	61	4
Gesamt (Suche)	45	143	23.382	175	67
Gesamt (Zeitraum)	14.143	12.121	966.748	37.423 ^b	7185 ^c
Suche/Zeitraum	0,003	0,012	0,024	0,005	0,009

^aTwitter-Accounts sind alle Tweets von 259 Twitter-Accounts, Facebook berücksichtigt die extrahierbaren Feeds von 185 Facebookseiten, Twitter-UserInnen erfassen die Ergebnisse der stichwortbasierten Abfrage der kompletten Twitter-Datenbank. Artikel und Kommentare entsprechen den Onlineinhalten von journalistischen Medien, welche im Vergleich zur manuellen Analyse herangezogen wurden

^bdavon 2021 aus dem Standard und 1931 aus der Kronenzeitung

^cdavon 4624 aus dem Standard und 986 aus der Kronenzeitung

Tab. 3 stellt die Frequenz der Dokumente nach den in der Evaluation berücksichtigten Suchbegriffen und der Datenquelle für den Untersuchungszeitraum dar.

Insgesamt standen mehr als eine Million Dokumente zur Verfügung, davon sind allerdings ca. 95 % Tweets von NutzerInnen (vgl. Tab. 3). Besonders auffallend ist die große Diskrepanz zwischen Twitter-Accounts und Facebook auf der einen Seite und Twitter auf der anderen Seite. Da für die Twitter-Accounts und Facebook nur bestimmte einzelne Seiten ausgelesen wurden, ist die damit abgebildete Stimmungslage wohl sehr eingeschränkt.

Mit 37.423 Dokumenten, davon ca. 4000 aus dem Standard und der Kronenzeitung, sind Nachrichtenmedien deutlich besser abgebildet. Insgesamt zeigt sich jedoch, dass Kommentare aus diesen Medien kaum abgegriffen wurden, da die Extraktion der Kommentare aus den Webseiten ungenügend umgesetzt wurde. Insofern ist der Vergleich der extrahierten Kommentare aus dem Weblyzard mit der manuellen Analyse kaum durchführbar, weil die Ergebnisse aus dem Weblyzard nur bedingt als Vergleichskriterium herangezogen werden können. Insgesamt zeigt die Verteilungen der Häufigkeiten über die fünf Suchbegriffe sehr deutlich, dass für sehr spezifische Abfragen nur wenige Aussagen zur Bewertung vorliegen. Es ließ sich zudem nicht vollständig rekonstruieren, wie die Daten im Weblyzard in die Datenbank aufgenommen werden.

3.3 Frequenzanalyse

In einem ersten Schritt sollen nun eine vergleichende *Frequenzanalyse* der extrahierten Artikel und Kommentare im Vergleich zur manuellen Analyse erfolgen. Im Standard zeigt sich für die Suchterme Muslime und Technologie eine geringere Häufigkeit beim Weblyzard; auch zum Suchterm Rassismus wurden nur halb so viele Artikel gefunden. Bei der Kronenzeitung findet sich eine erstaunlich hohe Übereinstimmung in der Frequenz der Artikel, allerdings sind in den Rohdaten des

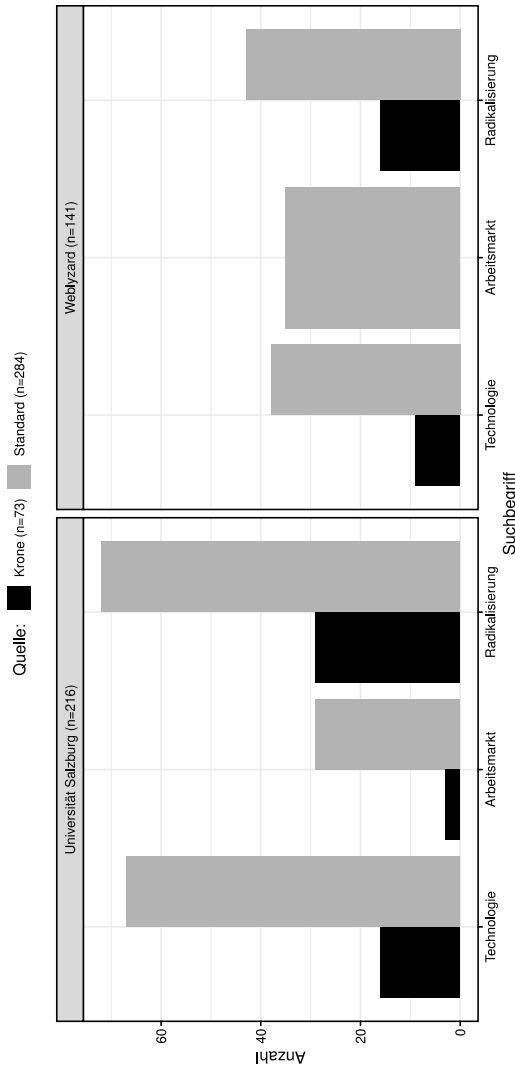


Abb. 3 Frequenz der Artikel nach Themenbereich, Quelle und Analyse (exkl. Dubletten). (Quelle: eigene Auswertungen)

Weblyzards knapp die Hälfte der 50 Artikel aus der Kronenzeitung Dubletten. Nach Bereinigung der Dubletten zeigt sich, dass im Weblyzard durchgehend weniger Artikel gefunden wurden.

Abb. 3 verdeutlicht die Problematik im Detail. Während nach den Ergebnissen des Weblyzard im Standard über alle drei Themenbereich gleich oft berichtet wurde, zeigt sich in der manuellen Analyse, dass der Arbeitsmarkt nur halb so oft Gegenstand von der Berichterstattung war wie etwa Radikalisierung oder Digitalisierung. Auch in der Kronenzeitung ist eine Verzerrung in der Frequenz erkennbar.

Im Themenbereich Radikalisierung werden die Differenzen zwischen den beiden Analysen erneut deutlich, da hier zwei Schlüsselereignisse in den Untersuchungszeitraum fallen, nämlich der Terroranschlag in London (3. Juni 2017) und die Prä-

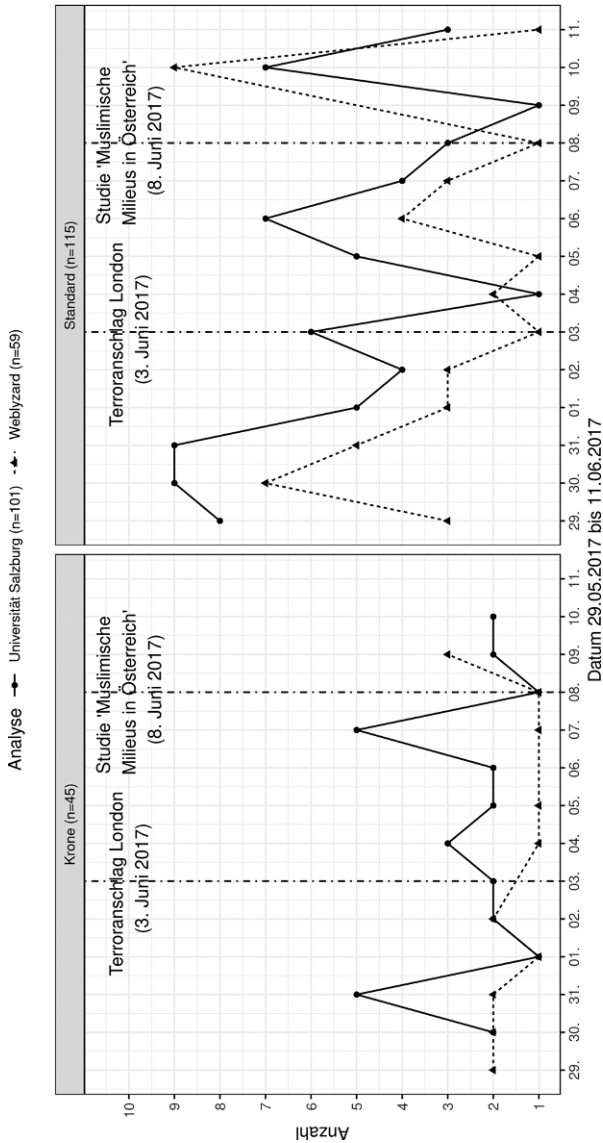


Abb. 4 Frequenz der Artikel für den Themenbereich Radikalisierung über den zeitlichen Verlauf. (Quelle: eigene Auswertungen)

sensation einer Studie zum Thema „Muslimische Milieus in Österreich“. Besonders auffallend ist die fehlende Berücksichtigung der Berichterstattung zum Terroranschlag in London. Die Präsentation der Studie erzeugte eine starke Diskussion im Kommentarbereich des Standards mit 1180 Kommentaren innerhalb der ersten 24h, siehe Abb. 4. Im Vergleich der Frequenz der Artikel über die Zeit ist die Kurve für den Standard ähnlich mit leichten Differenzen am Beginn und Ende des Untersuchungszeitraums. Bei der Kronenzeitung gibt es in der manuellen Analyse zwei deutliche Spitzen, die so nicht im Weblyzard abgebildet sind.

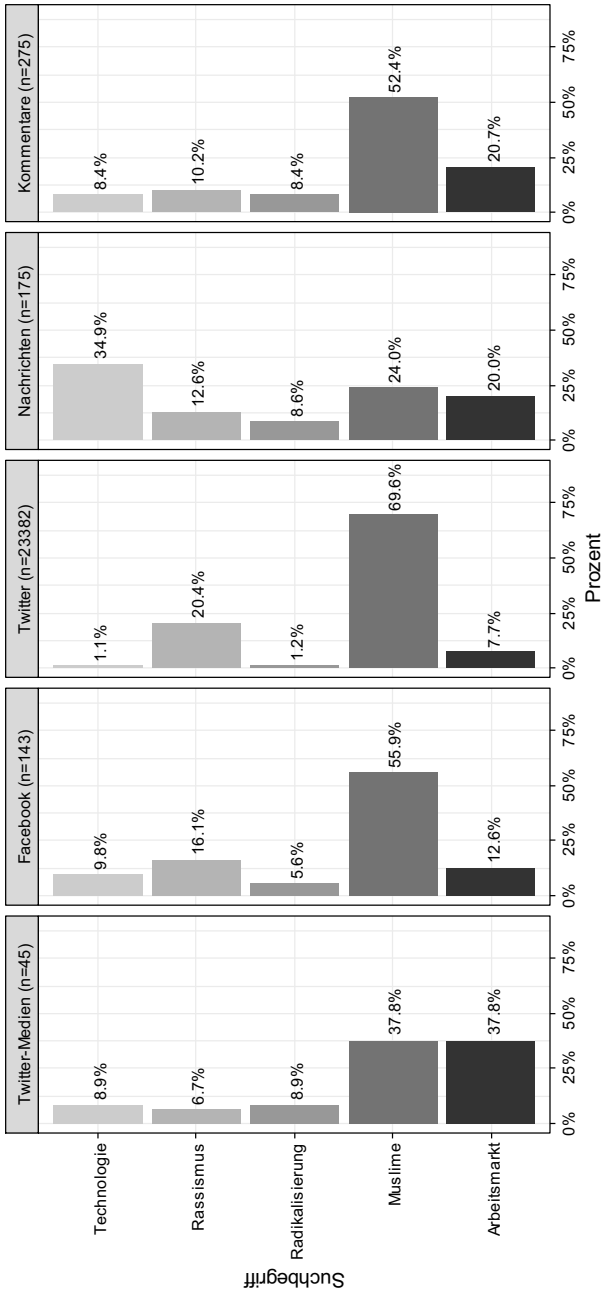


Abb. 5 Vergleich der relativen Häufigkeiten der Suchbegriffe nach Datenquelle für den Weblyzard (Twitter-Accounts sind alle Tweets von 259 Twitter-Accounts, Facebook berücksichtigt die extrahierbaren Feeds von 185 Facebookseiten, Twitter-UserInnen erfassen die Ergebnisse der stichwortbasierten Abfrage der kompletten Twitter-Datenbank. Artikel und Kommentare entsprechen den Onlineinhalten von journalistischen Medien, welche im Vergleich zur manuellen Analyse herangezogen wurden). (Quelle: eigene Auswertungen)

Insgesamt kann festgehalten werden, dass der Weblyzard bei der Frequenz der Artikel noch verhältnismäßig gut abschneidet, aber auch hier schon Verzerrungen hinsichtlich der Relationen der Themenbereiche erkennbar sind. Der Unterschied bei den Kommentaren ist jedoch gravierender. Während in der manuellen Analyse 26.440 Kommentare identifiziert wurden, umfasst die Datenbank des Weblyzards insgesamt lediglich 67 Kommentare (Standard: 61 und Kronenzeitung: 6) für den gleichen Zeitraum, was einen zeitlichen Vergleich der Frequenz der Kommentare hinfällig macht.

Der Weblyzard bietet neben den Nachrichten und Kommentaren noch drei zusätzliche Datenquellen an, die nicht Gegenstand der manuellen Analyse waren. Abb. 5 stellt die Verteilung der Suchbegriffe nach den Datenquellen dar.

Es ist zu beachten, dass Twitter-Accounts und Facebook-Daten von festen Twitter- und Facebook-Seiten auslesen (PolitikerInnen, Parteien, Nachrichtenmedien u. ä.), während die Extraktion der Daten von Twitter-UserInnen auf einer vollständigen Datenbankabfrage beruht. Folglich beinhalten Twitter-Accounts und Facebook eher einen Mix aus Nachrichten und Kommentaren und die Twitter-Userdaten sind eher mit den Kommentaren vergleichbar, denn hier kann jeder Internetnutzer die eigene Meinung artikulieren. Die Datenbank der Twitter-UserInnen sollte somit das Stimmungsbild tendenziell besser abbilden als die Kommentare, welche in moderierten Foren stattfinden.

Im Vergleich der Medienquellen ist augenscheinlich, dass die Nachrichten deutlich mehr über Technologie berichten und dass Themen wie Muslime und Arbeitsmarkt deutlich weniger häufig auftreten. Hier zeigt sich, dass eine Frequenzanalyse durchaus dazu geeignet ist, die Salienz von Themen in den Nachrichtenmedien und in der Öffentlichkeit zu erfassen. Es ist durchaus plausibel, dass arbeitsmarktrelevante Themen und Integrationsherausforderungen (bei Muslimen) mehr öffentliche Resonanz erzeugen. Dies bestätigen auch die relativ gleichmäßigen Frequenzschätzungen bei Social Media (Facebook, Twitter-UserInnen und Kommentare) im Vergleich zu den Nachrichtenmedien. Der Vergleich von Twitter-Accounts und Twitter-UserInnen macht deutlich, dass die für das Stimmungsbild relevanten Diskussionen nicht immer entlang von Informationen klassischer Opinion-Leader (aus Politik und Medien) verlaufen müssen. Innerhalb der Twitter-UserInnen-Gemeinde wird mehrheitlich über Muslime und Rassismus und in geringerem Umfang über arbeitsmarktrelevante Problemstellungen diskutiert, die Technologie bleibt völlig unbedeutend. Insofern ist bei einseitigen Frequenzschätzungen aus spezifischen Quellen (die ein spezifisches Klientel der Bevölkerung abbilden) von starken Verzerrungen auszugehen (vgl. hierzu auch näher Abschn. 4).

3.4 Sentimentanalyse

Neben dem Vergleich der Frequenz ist das zweite Vergleichskriterium die Übereinstimmung in der Abbildung positiver und negativer Stimmungsbilder. Tab. 4 stellt die Verteilungen der Sentimenteinstufungen für die Artikel getrennt nach Themenbereich und Medienquelle dar. Dabei wird bei der manuellen Analyse ersichtlich, dass ein nicht unerheblicher Teil der Artikel als irrelevant eingestuft wurde bzw. unterschiedliche Bewertungen der BeurteilerInnen erkennbar werden. Diese Kategorien

Tab. 4 Vergleich der Artikel Sentimentbewertung der beiden Analysen nach Quelle und Suchbegriff. (Quelle: eigene Auswertungen)

	Sentiment	Der Standard		Kronenzeitung	
		Universität Salzburg	Weblyzard	Universität Salzburg	Weblyzard
Technologie	Positiv	7,5% (5)	68,4% (26)	12,5% (2)	33,3% (3)
	Neutral	34,3% (23)	5,3% (2)	43,8% (7)	22,2% (2)
	Negativ	25,4% (17)	26,3% (10)	12,5% (2)	44,4% (4)
	Uneindeutig	4,5% (3)	0,0% (0)	12,5% (2)	0,0% (0)
	Irrelevant	28,4% (19)	0,0% (0)	18,8% (3)	0,0% (0)
Arbeitsmarkt	Positiv	10,3% (3)	65,7% (23)	33,3% (1)	0,0% (0)
	Neutral	55,2% (16)	20,0% (7)	33,3% (1)	0,0% (0)
	Negativ	27,6% (8)	14,3% (5)	33,3% (1)	0,0% (0)
	Uneindeutig	0,0% (0)	0,0% (0)	0,0% (0)	0,0% (0)
	Irrelevant	6,9% (2)	0,0% (0)	0,0% (0)	0,0% (0)
Radikalisierung	Positiv	1,4% (1)	18,6% (8)	3,4% (1)	18,8% (3)
	Neutral	27,8% (20)	25,6% (11)	24,1% (7)	18,8% (3)
	Negativ	52,8% (38)	55,8% (24)	62,1% (18)	62,5% (10)
	Uneindeutig	5,6% (4)	0,0% (0)	6,9% (2)	0,0% (0)
	Irrelevant	12,5% (9)	0,0% (0)	3,4% (1)	0,0% (0)

sind im Weblyzard nicht vorgesehen, da die Auswahl und eindeutige Zuordnung vorab erfolgt.⁷ Im Vergleich der drei Einstufungen (positiv, neutral, negativ)⁸ ist augenscheinlich, dass die Beurteilung des Weblyzard die Artikel des Standards deutlich positiver einschätzt. Mit Blick auf die Kronenzeitung ist der Unterschied weniger deutlich.

Abb. 6 weist die Sentimentbewertungen differenziert nach Suchbegriff und Datenquelle aus. Die Sentimentverteilung für das am stärksten diskutierte Thema Muslime scheint am ähnlichsten; hier wird erstaunlicherweise in Nachrichtenmedien am kritischsten berichtet.⁹ Die großen Differenzen bei Rassismus zwischen den Kommentaren und Nachrichten auf der einen Seite und den Social-Media-Quellen auf der anderen Seite könnten dadurch erklärt werden, dass Nachrichten und die damit einhergehende Kommentarfunktionen sehr sensibel auf Rassismus reagieren und entsprechende Kommentare beispielsweise schnell entfernt werden. Gleichzeitig sind die sozialen Medien ein Ort, wo gesellschaftliche Diskussionen über das Thema Rassismus gespiegelt werden. Auch die Ergebnisse für den Arbeitsmarkt erscheinen plausibel, da Nachrichten Ängste weniger stark reflektieren als die sozialen Medien. Dass die Kommentare sehr positiv sind, passt hingegen nicht ins Bild. Radikalisie-

⁷ Dies ist ein deutliches Indiz dafür, dass in die automatischen Verfahren zum Extrahieren der Daten aus dem Internet entsprechende Kontrollmechanismen implementiert werden müssen.

⁸ Die metrische Skale des Weblyzards wurde umcodiert in $<-0,33$: negativ, $-0,33$ bis $0,33$: neutral und $>0,33$: positiv.

⁹ Dies könnte auch auf die fehlenden (bzw. falschen) automatisierten Einschätzungen sarkastischer Aussagen zurückzuführen sein, die oftmals wohl fälschlicherweise als „positiv“ bewertet werden.

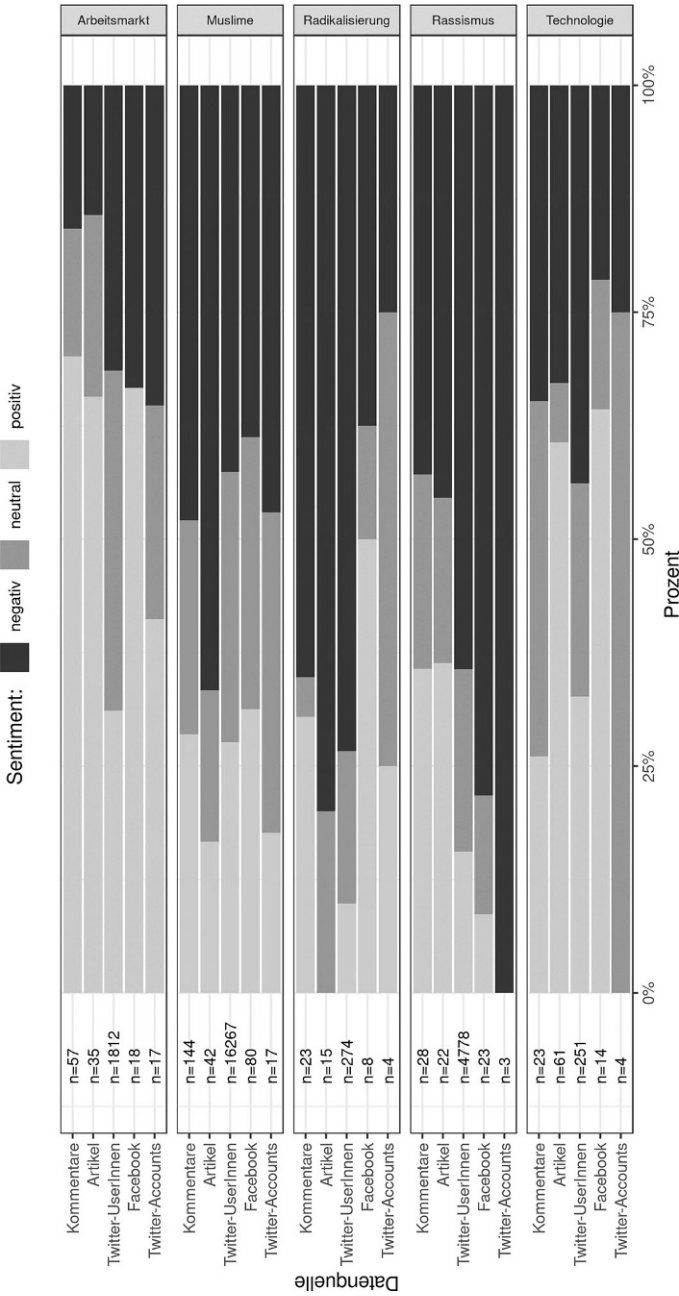


Abb. 6 Vergleich Sentimentbewertung nach Datenquelle und Suchbegriff für den Weblyzard (Twitter-Accounts sind alle Tweets von 259 Twitter-Accounts, Facebook berücksichtigt die extrahierbaren Feeds von 185 Facebookseiten, Twitter-UserInnen erfassen die Ergebnisse der stichwortbasierten Abfrage der kompletten Twitter-Datenbank. Artikel und Kommentare entsprechen den Onlineinhalten von journalistischen Medien, welche im Vergleich zur manuellen Analyse herangezogen wurden). (Quelle: eigene Auswertungen)

rung und Technologie zu betrachten, ist aufgrund der geringen Bedeutung in den sozialen Medien nicht möglich.

Insgesamt zeigt der Vergleich zwischen manueller Analyse und Weblyzard, dass eine automatisierte Erhebung des gesellschaftlichen Diskurses und der Stimmung in der Bevölkerung (noch) auf erhebliche Schwierigkeiten stößt. Die Frage, auf welcher Ebene diese Schwierigkeiten angesiedelt sind – in der Operationalisierung des konkreten Projektes, in dem verwendeten Programm Weblyzard (welches ja auch laufend weiterentwickelt wird) oder in grundsätzlichen Schwierigkeiten einer automatisierten Inhaltsanalyse –, lässt sich nicht eindeutig klären. Doch auch die Potenziale einer solchen Vorgangsweise sind unübersehbar.

Die Sentimentanalyse zeigt sich grundsätzlich als brauchbar bei der Identifikation negativer Sentiments. Die fehlende Transparenz bezüglich irrelevanter und zynischer Kommentare, Postings oder Tweets erschwert die Validität allerdings deutlich. Außerdem analysiert der Algorithmus die Textpassagen auf der Basis zweier Wörterbücher, eines mit positiven und eines mit negativen Wörtern. Dementsprechend kann die Sentimentanalyse nichts über die Beurteilung des Inhalts sagen, sondern nur etwas über die Art und Weise der Darstellung.

Die Nachrichtendatenbank scheint gut geeignet, ein vollständiges Stimmungsbild der Medien abzubilden, das Problem der Dubletten sollte technisch lösbar sein. Die Kommentare waren im Projekt nicht brauchbar, denn die sehr geringe Anzahl, deren Auswahl zudem nicht nachvollzogen werden konnte, kann nicht als valide Datenbasis betrachtet werden. Die sozialen Medien liefern plausible Ergebnisse, allerdings erfordern auch im Kontext von Twitter-Accounts und Facebook die geringen Fallzahlen eine sorgfältige Interpretation der Daten. Es ist aktuell nicht möglich, diese beiden Quellen für ein hinreichend valides Abbild der Stimmung in der Bevölkerung heranzuziehen, zudem ist eine Überschneidung mit anderen deutschsprachigen Räumen leicht möglich. Twitter ist zumindest hinsichtlich der Quantität ein deutlicher Zugewinn und eignet sich zur Identifikation von „hot topics“. Das über Twitter-Themen gewonnene Stimmungsbild muss aber sinnvoll mit anderen Quellen ergänzt werden, weil hier keine Repräsentativität der öffentlichen Meinung vermutet werden kann. Dies liegt nicht nur an einer geringen Durchdringung der Plattform im deutschsprachigen Raum. Durch unterschiedliches Nutzungsverhalten auf Twitter, bei dem nicht jede Handlung einer konkreten Meinungsäußerung entspricht (bspw. Likes aufgrund von Gruppenzugehörigkeit), kommt es zu Verzerrungen, die sich in den erhobenen Daten niederschlagen. Hinzu kommt, dass durch Algorithmen Themen künstlich gestärkt werden und die NutzerInnen nicht „unbeeinflusst und frei“ (Lischka und Stöcker 2017, S. 25) darüber entscheiden, welche Beiträge sie bevorzugen. Die Social-Media-Plattformen sowie auch diverse Suchmaschinen stellen dabei Informationsintermediäre dar. Diesen kann eine hohe Relevanz für die öffentliche Kommunikation zugeschrieben werden. Sie haben eine starke wirtschaftliche Stellung und sind dabei „keineswegs nur neutrale Tools zu Verarbeitung und Darstellung von Information“ (Schulz und Dankert 2017, S. 352). Durch das Sammeln und Verarbeiten großer persönlicher sowie sozialer Datenmengen versuchen diese Intermediäre, den Geschmack („Taste“) ihrer NutzerInnen zu überwachen, um daraufhin gezielt das Informationsangebot zu gestalten (Morris 2015, S. 447f.) Ihr Mehrwert liegt dabei gerade in der selektiven Bereitstellung von Information. Deshalb ist es

fraglich, ob die dort verhandelten Diskurse die Bevölkerungsmeinung tatsächlich oder lediglich stark verzerrt abbilden.

4 Entwicklung von Qualitätskriterien zur Bestimmung der Aussagekraft von Medienquellen

Der öffentliche Diskurs und die freie Meinungsäußerung stellen in einer demokratischen Gesellschaft einen wesentlichen Modus der Konsensfindung dar (z. B. Dubiel 1994), weshalb sich die Qualität der Demokratie im Wesentlichen von der Qualität der öffentlichen Kommunikation ableiten lässt. Als Ort zur Bestimmung und Diskussion der intersubjektiv erachteten und allgemeinverbindlich zu lösenden Probleme sind soziale Medien durch ihre freie Zugänglichkeit durchaus voraussetzungsvoll, denn sie ermöglichen viel stärker als früher eine breite Teilhabe möglichst vieler BürgerInnen an den kollektiven Bewusstseinsbildungs- und Entscheidungsprozessen (vgl. Seethaler 2015, S. 12). Eine nicht zu vernachlässigende Gefahr in sozialen Medien stellen jedoch aktuell Fake News und Hass-Postings und sogenannte Filter-Bubbles¹⁰ dar. Aufgrund der hohen NutzerInnen-Zahlen von Facebook können sich Falschmeldungen extrem schnell verbreiten. Der Echokammer-Effekt ist ein damit zusammenhängendes Konzept, das die eingeschränkte Sichtweise auf Sachverhalte in sozialen Medien beschreibt, da Inhalte, die dem Konsens einer Gruppe/eines Freundeskreises widersprechen, ausgefiltert werden. Durch den Umgang mit Gleichgesinnten verstärken sich die eigenen Positionen gegenseitig und es werden hauptsächlich oder ausschließlich mediale Inhalte geteilt, die die eigenen Ansichten stützen. Dies kann in sozialen Netzwerken zu einer gefährlichen Dynamik führen, wie die Debatten rund um die Flüchtlingskrise und die damit verbundenen Integrationsherausforderungen zeigen.

Die Qualität der Medienberichterstattung, die Aussagekraft von Medienquellen und der Einfluss verschiedener Online-Medien lassen sich jedoch nur schwer messen. „Qualität“ ist ein vieldimensionales und relationales Konstrukt. Ein verbreiteter Ansatz in Bezug auf die Analyse der Qualität in den Medien sind die Dimensionen Relevanz, Vielfalt, Einordnungsleistung und Professionalität, die auf Kurt Imhof zurückgehen (föG – Forschungsinstitut Öffentlichkeit und Gesellschaft 2016, S. 46). Vieldimensional meint jedoch auch, dass es unterschiedliche Qualitätsanforderungen gibt, die zueinander auch in einem Spannungsverhältnis stehen können; relational verweist darauf, dass Qualität nicht ein absolutes, inhärentes Merkmal ist, sondern durch die beabsichtigte Verwendung bestimmt wird – Qualität im Sinne von *fitness for use* (vgl. Weichbold 2009). Die Bewertung der Qualität der verwendeten Medien-

¹⁰ Sogenannte Filterblasen (Pariser 2011) entstehen auf Facebook und auch anderen Seiten aufgrund von Algorithmen. Diese Computerprogramme zeigen Inhalte entsprechend der Informationen, die sie über eine Person haben an und dienen damit der Personalisierung. Zu diesen Informationen zählen u. a. Standort, Geschlecht und Alter, v. a. aber auch das bisherige Verhalten (Suchen, Likes etc.) bzw. Interesse. Folgt man auf Facebook nun beispielsweise einer bestimmten politischen Partei und vielleicht noch einigen PolitikerInnen dieses Lagers und ist auch der Freundeskreis auf Facebook in etwa gleich gesinnt, entsteht eine so genannte Filterblase, da einem nur noch Inhalte und Berichte angezeigt werden, die mit der eigenen politischen Ausrichtung konform gehen.

Tab. 5 Qualitätskriterien zur Bestimmung der Aussagekraft von Medienquellen. (Quelle: eigene Darstellung)

Rating	Reichweite	Differenziertheit der Behandlung der Themen	Abbildung des Meinungsspektrums
*****	Höchste Reichweite (umfassende Rezeption in der Bevölkerung)	Sehr hohe Aussagekraft/Seriosität der Berichterstattung	Breite Abdeckung der Meinungs- und Wertevielfalt in der Gesamtbevölkerung
****	Hohe Reichweite (hohe Auflage und breite Rezeption der Mediendiskurse)	Hohe inhaltliche Qualität der Berichterstattung, geringe Einschränkungen	Generalisierbare Schlussfolgerungen zu relevanten öffentlichen Diskursen durch aggregierte Datenanalyse möglich
***	Mittlere Reichweite (zumindest für einzelne Milieus in der Bevölkerung von Interesse)	Eingeschränkte Aussagekraft, Oberflächlichkeit in der Behandlung der Themen	Darstellung milieuspezifischer Diskurse (spezifische inhaltliche Ausrichtungen erkennbar)
**	Geringe Reichweite (für spezifisch interessierte NutzerInnen)	Deutlich eingeschränkte Aussagekraft, fragwürdige Art der Meinungsbildung	Einseitige Diskursdarstellung, erkennbar durch spezifisches NutzerInnenprofil
*	Äußerst geringe Reichweite („Nischenprodukte“, „hochspezialisiert“, „abseits vom Mainstream“)	Äußerst geringe Aussagekraft (Transport von Einzelmeinungen, Social Bots, Fake News)	Gefahr der Manipulation gegeben (z. B. durch „Filter Bubbles“ und „Echokammern“)

quellen erfolgte deshalb auf Basis eigens erstellter Dimensionen – in Bezug auf die Anforderungen des Weblyzards – im interdisziplinären Team der Universität Salzburg, das sich aus zwei Kommunikationswissenschaftlern und vier SoziologInnen zusammensetzt.

Wir beurteilen die Qualität der eingebetteten Medienquellen (s. Tab. 5) in Relation zur *Reichweite*, also wie umfassend der Einfluss des Mediendiskurses auf das Stimmungsbild der österreichischen Bevölkerung einzuschätzen ist. Zusätzlich versuchen wir aufgrund kommunikationswissenschaftlicher Erkenntnisse, Aussagen zur *Differenziertheit der Behandlung der Themen* abzuleiten. Diese Dimensionen lässt Schlüsse auf die *Seriosität* der jeweiligen Medienquelle zu. Die letzte Dimension, die wir zur Beurteilung der Aussagekraft der jeweiligen Medienquellen verwenden, ist die *Abdeckung des Meinungsspektrums* in der österreichischen Bevölkerung. Hier beurteilen wir, wie breit die Akzeptanz einzelner Medien in der Bevölkerung ist und inwiefern einzelne Medien einen vielschichtigen öffentlichen Diskurs ermöglichen. Die Spannweite reicht von einer unabhängigen Positionierung bis hin zu einer manipulativen Berichterstattung, die auf die Gefahr der Produktion von „Filterbubbles“ und „Echokammern“ hindeutet und einer postfaktischen Meinungsbildung (erstmal Keyes 2004) Vorschub leisten könnte. Die Tab. 6 fasst die Bewertungen der Medienquellen nochmals zusammen und ermöglicht eine vergleichende Analyse der Relevanz einzelner Quellen.

Tab. 6 Einschätzung der Qualität der Quellen für die Ableitung von medialen und öffentlichen Diskursen. (Quelle: eigene Darstellung)

Medium	Reichweite	Differenziertheit der Behandlung der Themen	Abbildung des Meinungsspektrums
Nachrichten	****	****	****
Kommentare	*	**	**
Facebook	***	**	***
Twitter-Medien	**	**	**
Twitter-Accounts	**	*	**

Die Gesamtheit der Quellen der Nachrichten-Kategorie im Weblyzard liefert ein relativ aussagekräftiges Bild des medialen Diskurses in Österreich. Insgesamt lässt sich zur Internetnutzung in Österreich sagen, dass 83 % der Bevölkerung online sind und davon 73 % das Internet regelmäßig zum Lesen von Nachrichten nutzen (vgl. Fletcher et al. 2016, S. 63, 2017, S. 6). Dies unterstreicht nochmals die Reichweite der Nachrichten-Kategorie, zumal die österreichische Bevölkerung jüngsten Erhebungen zufolge sehr an Nachrichten interessiert ist. Laut Erhebungen der ÖWA (Österreichische Webanalyse) stellen für 77,3 % der österreichischen Internet-UserInnen „Nachrichten und Politik“ einen wesentlichen Nutzungsschwerpunkt dar (vgl. ÖWA 2017, S. 9). Zwei Drittel der NutzerInnen lesen laut Digital Media Report 2017 mindestens zweimal täglich Online-Nachrichten. Zur Beurteilung der Differenziertheit der Berichterstattung von Online-Nachrichten kann die Seethaler-Studie (2015) als Gradmesser herangezogen werden. Diese Studie befasste sich mit vier Qualitätskriterien, der Objektivität der Berichterstattung, der Urheber-Transparenz, der Quellentransparenz und der diskursfördernden analytischen Qualität. Die Studie kam zu dem allgemeinen Ergebnis, dass die Angebote der Qualitätsmedien durchwegs als qualitativ hochwertig anzusehen sind, wohingegen die Boulevard-Medien bei allen Qualitätsindikatoren im unteren Drittel bis Viertel angesiedelt sind (ebd., S. 78). Die Abbildung des Meinungsspektrums ist durch den breiten Korpus an Medienquellen auch weitgehend gewährleistet. Die im Weblyzard integrierten Medienquellen reichen von den Online-Versionen der österreichischen Tages-, Wochen- und Regionalzeitungen über Online-Nachrichtenseiten aus ganz Österreich bis hin zu den Portalen von ORF und anderer TV- und Radioanstalten. Zudem sind neben einer Vielzahl von Magazinen und Regionalmedien auch Nischenprodukte für spezielle Milieus in der Bevölkerung enthalten. Mit Ausnahme extremer Positionen (z. B. rechtsextreme Quellen) werden in dieser Kategorie alle Gesinnungen sämtlicher Bevölkerungsgruppen repräsentiert. Es wurde ebenfalls darauf geachtet, auch Sprachrohre von Minderheiten in die Auswahl der Quellen zu inkludieren (z. B. Biber Magazin).

In Bezug auf die Stimmungslage der ÖsterreicherInnen kann die Medienberichterstattung jedoch nicht mit den Positionen der Bevölkerung gleichgesetzt werden. Durch die hohen NutzerInnen-Zahlen wird deshalb immer wieder Facebook als zentrale Informationsquelle genannt. Hinsichtlich der Nachrichtennutzung über Facebook liegt die Zahl dem Digital News Report 2017 zufolge bei 34 % der österreichischen UserInnen (vgl. Fletcher et al. 2017, S. 57). Die Geschlechterverteilung ist

mit 51 % männlichen und 49 % weiblichen UserInnen äußerst ausgewogen und mit knapp 1,2 Mio. ist die Altersgruppe der 20- bis 29-Jährigen am stärksten vertreten, doch auch die NutzerInnen-Zahlen der Generation 60+ wachsen jährlich weiter und liegen mittlerweile zwischen zwei und drei Prozent (Stand: August 2016). Trotz der relativ hohen Reichweite werden die Differenziertheit der Behandlung der Themen und die Seriosität der behandelten Diskurse nur mit einer geringen Qualitätsbeurteilung versehen. Die Abbildung des gesamten Meinungsspektrums ist über Facebook nur schwer möglich. Zwar finden sich auf dieser Plattform verstärkt gesellschaftliche Milieus, die sich von klassischen Mainstream-Medien verabschiedet haben und alternative Kanäle der Informationsvermittlung suchen und nutzen. Andererseits ist die vollständige Einbindung von Facebook-Sites nur begrenzt möglich und deshalb ist eine einseitige Nutzung von Facebook zur Ableitung des öffentlichen Diskurses als kritisch zu betrachten.

Im Gegensatz zu den USA weist Twitter in Österreich nur eine sehr geringe Reichweite auf und findet hier v. a. bei Unternehmen, PolitikerInnen und bekannten Persönlichkeiten Anklang.¹¹ Die Plattform ermöglicht ihren registrierten NutzerInnen das Verfassen und Veröffentlichen telegrammartiger Kurznachrichten, die auf 140 Zeichen beschränkt sind und „Tweets“ (engl. „to tweet“ = „zitschern“) genannt werden. Dem Digital News Report 2017 des Reuters Institute ist zu entnehmen, dass nur vier Prozent der österreichischen UserInnen zum Lesen von Nachrichten auf Twitter zurückgreifen. Stellt man diese Zahl den 73 % der Online-NutzerInnen gegenüber, die im Internet Nachrichten lesen, wird nochmals deutlich, dass Twitter in Österreich nur eine Nischen-Plattform darstellt (vgl. Fletcher et al. 2016, S. 63, 2017, S. 57). Selbst wenn breitere Bevölkerungsgruppen die Plattform nutzen würden, verhindert ein starkes Gefälle zwischen aktiven (sprechenden) und passiven (hörenden) UserInnen die Aussagekraft der erhobenen Daten in Bezug auf öffentliche Meinung. Letztere machen laut Boyd und Crawford (2012, S. 669) 40 % der GesamtnutzerInnen aus. Zudem besteht die Gefahr, dass Einzelpersonen durch mehrere Profile aktiv zum Diskurs beitragen bzw. computergestützt durch Bots den Diskurs manipulieren (Boyd und Crawford 2012, S. 669).

Die Kategorie „Twitter-Accounts“ im Webyzard greift zwar eine große Zahl offizieller österreichischer Twitter-Profile von Unternehmen, Zeitungen, Magazinen, NGOs, Initiativen, Organisationen, Vereinen, JournalistInnen, PolitikerInnen und Persönlichkeiten des öffentlichen Lebens ab, dennoch ist sowohl die Reichweite (durch die geringe Nutzung in der österreichischen Bevölkerung), die Differenziertheit der Behandlung der Themen (durch die vorgegebene Beschränkung auf wenige Zeichen) und die Abbildung des Meinungsspektrums (durch die fehlende Berücksichtigung breiter Bevölkerungsschichten) als sehr begrenzt einzustufen. Es wird nicht empfohlen, Twitter-Accounts als einzige Social-Media-Quelle zu berücksichtigen. Am ehesten stellt diese Quelle eine Ergänzung dar, wenn Diskurse in Online-Nachrichtenmedien beleuchtet werden.

¹¹ Zu den Top-5-Twitter-Profilen in Österreich zählen laut OTS-Twitterlist (Stand: KW 30/2017): 1. Armin Wolf, 2. Der Standard, 3. Sebastian Kurz, 4. Ingrid Thurnher und 5. Florian Klenk (vgl. APA-OTS 2017, online).

Die Kategorie der „Twitter-UserInnen“ umfasst im Gegensatz zu den offiziellen Profilen der „Twitter-Accounts“ sämtliche Privatpersonen, die auf der Microblogging-Plattform angemeldet sind. Im Unterschied zu den Twitter-Accounts liegt bei den Twitter-UserInnen im Weblyard eine sehr umfassende Datenbasis vor, die sich prinzipiell zur Analyse von Social-Media-Diskursen eignet. Es gelingt über die Ebene der Twitter-UserInnen sicherlich, relevante Diskurse in Österreich abzubilden, jedoch sind diese nicht als repräsentativ für die österreichische Gesamtbevölkerung einzustufen. Die Differenziertheit der Behandlung der Themen bleibt auf oberflächlichem Niveau, zudem ist die Gefahr von Falschmeldungen (Fake News) als gravierend einzustufen. Es ist durchaus denkbar, dass einzelne einflussreiche Twitter-UserInnen Agenda-Setting (z. B. McCombs und Shaw 1972) betreiben, das nur bedingt für Österreich relevant ist.

Die Kategorie der Kommentare weist schließlich die mit Abstand geringste Reichweite auf, da beim Weblyard in der derzeitigen bestehenden Form ausschließlich die Kommentare der LeserInnen zu den publizierten Beiträgen auf www.derStandard.at, www.debatte.orf.at und www.krone.at abgegriffen werden. Es dürfte sich auch hier bei den PosterInnen um eine spezifische Klientel handeln, die ein erhöhtes Bedürfnis nach (Medien-)Aufmerksamkeit aufweist. Auch wenn eine umfangreiche Datenbasis zu Postings bestehen würde, ist nicht davon auszugehen, dass derartige Diskussionsforen in sozialen Medien für breite Bevölkerungsschichten meinungsbildend wirken.

5 Resümee – Die theorieorientierte Forschung als Gegengewicht zu Big Data

Die quantitative Forschung befindet sich durch den Siegeszug neuer Technologien der Datenverarbeitung zweifellos im Umbruch und ist ambivalenten Entwicklungen ausgesetzt. Die kostengünstigen Möglichkeiten von automatisierten Datenanalysen gewährleisten, dass sozialwissenschaftlich relevante Informationen teils in beliebig hoher „Auflösung“ sowohl auf der Mikro- als auch auf der Makroebene gewonnen werden können, und es können mittlerweile auch beliebig kleine Aggregate empirisch beleuchtet werden, um korrelative oder netzwerkartige Strukturen aufzudecken. Die interdisziplinär orientierte „Computational Social Science“ als aufstrebende Disziplin fokussiert in ihren Analysen vorwiegend auf diese neuen Arten von Massendaten. Dabei kommt eine Vielzahl von rechenintensiven Verfahren aus unterschiedlichen Disziplinen zum Einsatz, und die interdisziplinäre Zusammenarbeit ist oftmals eine Vorbedingung (vgl. Alvarez 2016). Das wesentliche Merkmal der Computational Social Science ist der flexible Einsatz von Computern zur Erkenntnisgewinnung. Cioffi-Revilla (2010, S. 260) vergleicht es passend mit Galileis Teleskop:

Just like Galileo exploited the telescope as the key instrument for observing and gaining a deeper and empirically truthful understanding of the physical universe, computational social scientists are learning to exploit the advanced

and increasingly powerful instruments of computation to see beyond the visible spectrum of more traditional disciplinary analyses.

Meist werden deshalb gewaltige Datenmengen (volume) mit hoher Geschwindigkeit (velocity) ausgewertet und auch unterschiedliche Daten (variety) miteinander kombiniert (Ekbja et al. 2015). Dementsprechend ist Computational Social Science eine verfahrensorientierte Disziplin, die in einem Spannungsverhältnis zu theorieorientierten Ansätzen steht. Gerade der Projektzusammenhang und die Evaluationsstudie zeigen, dass durch den Aufstieg von Big Data die Gefahr besteht, dass die theorieorientierte, deduktive Sozialforschung zu stark ins Hintertreffen gerät. Im Wettstreit klassischer und neuer Ansätze der Datengewinnung und -analyse sollte das Gewicht der sozialwissenschaftlichen Forschung stärker auf die Dateninterpretation gelegt werden, denn häufig beginnt die sozialwissenschaftliche Interpretation erst dort, wo die Big-Data-Analyse endet. Schließlich wird bei Algorithmen primär explorativ und induktiv gearbeitet, ohne auf Stichprobencharakteristika und Messfehler ausreichend Bezug zu nehmen. Hinzu kommt das Problem, dass die Daten zwar öffentlich zugänglich gemacht werden, aber die Information von den Untersuchten nicht aktiv zur Verfügung gestellt wird. Daraus resultiert, dass jeglicher Kontext in Bezug auf die Entstehung der untersuchten Aussagen für die Interpretation fehlt (vgl. Matzner 2016, S. 201 f.). Somit ist – auch im Kontext unseres Fallbeispiels – die Frage nach einem tiefgreifenden Systemwandel zu stellen, weil gesellschaftliche Subsysteme (wie beispielsweise Politik, Wirtschaft, Medien und Wissenschaften) immer mehr in den Bann von Algorithmen – bei Monahan (2018) „algorithmic fetishism“ – geraten und politische Akteure durch das Credo einer evidence-based policy stärker zu schnellen Entscheidungen gedrängt werden.¹² Die automatisierten Auswertungen großer Datenmengen bergen die Gefahr, dass weitreichende Maßnahmen an Artificial Intelligence delegiert werden und damit Reflexions- und Kontrollverluste einhergehen. Diese Reduzierung auf „calculated publics“ (Crawford 2016, S. 78) kann beispielsweise zu algorithmischer Diskriminierung führen (Monahan 2018, S. 2). Zudem steht die ökonomische Vermarktung neuer Softwarelösungen im Vordergrund, womit die angebotenen Tools weniger am Wahrheitsgehalt der erzielten Aussagen, sondern stärker an den Verkaufszahlen orientiert bleiben dürften. Dass Erfolge nach außen inszeniert und „Baustellen“ intern ausgehandelt werden müssen, trifft wohl auf zahlreiche Forschungsprojekte zu, wobei die Kommunikation über die Projekthürden viel zu selten erfolgt. Gerade an dieser Stelle könnte ein wissenschaftlicher Partner eine tragfähige und wertvolle Rolle in anwendungsbezogenen Forschungsprojekten vertreten, weil eine sachliche und wertfreie Analyse unabhängig von ökonomischen und politischen Verwertungszusammenhängen einer am Wahrheitsanspruch orientierten Wissenschaft zweckdienlich sein sollte. In diesem Geflecht sollte sich die Soziologie auf ihre Stärke der theorieorientierten Sozialforschung besinnen. Auch die sozialwissenschaftliche Methodologie kann ihre Rolle darin sehen, hinter die Fassade des Impression Managements und der rei-

¹² Siehe dazu auch die jüngste Konferenzinitiative der OECD (2017): <http://www.oecd.org/gov/evidence-informed-policy-making.htm>.

nen Deskriptionslogik von Big-Data-Analysen zu blicken, und somit im Idealfall wirkungsvoll zur Vermeidung vorschneller Schlussfolgerungen beitragen.

Funding Die beiden Projekte, auf die im Artikel näher verwiesen wird, wurden von der Österreichischen Forschungsförderungsgesellschaft (FFG) im Rahmen des KIRAS-Sicherheitsforschungsprogramms gefördert. Open access funding provided by Paris Lodron University of Salzburg.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Literatur

- Alvarez, R.M. 2016. Introduction. In *Computational Social Science*, Hrsg. R.M. Alvarez, 1–24. Cambridge: Cambridge University Press.
- Aschauer, W. 2017. *Das gesellschaftliche Unbehagen in der EU. Ursachen, Dimensionen, Folgen*. Wiesbaden: Springer VS.
- Beck, U. 2003. Weltrisikogesellschaft revisited: Die terroristische Bedrohung. In *Irritierte Ordnung. Die gesellschaftliche Verarbeitung von Terror*, Hrsg. R. Hitzler, J. Reicherts, 275–299. Konstanz: UVK.
- Bigo, D. 2008. Globalized (in)Security: the field and the Ban-opticon. In *Terror, Insecurity and Liberty. Illiberal practices of liberal regimes after 9/11*, Hrsg. D. Bigo, A. Tsoukala, 10–48. Abingdon: Routledge.
- Blühdorn, I. 2013. *Simulative Demokratie. Politik nach der postdemokratischen Wende*. Berlin: Suhrkamp.
- Bohle, H.H., et al. 1997. Anomie in der modernen Gesellschaft. Bestandsaufnahme und Kritik eines klassischen Ansatzes soziologischer Analyse. In *Was treibt die Gesellschaft auseinander?*, Hrsg. W. Heitmeyer, 29–68. Frankfurt am Main: Suhrkamp.
- Boyd, D., und K. Crawford. 2012. Critical Questions for Big Data. *Information, Communication & Society* 15(5):662–679.
- Bude, H. 2014. *Gesellschaft der Angst*. Hamburg: Hamburger Edition.
- Burzan, N.P.A., und Berger. 2010. *Dynamiken (in) der gesellschaftlichen Mitte*. Sozialstrukturanalyse. Wiesbaden: VS.
- Cioffi-Revilla, C. 2010. Computational Social Science. *Wiley Interdisciplinary Reviews: Computational Statistics* 2(3):259–271.
- Crawford, K. 2016. Can an Algorithm be Agonistic? Ten Scenes from Life in Calculated Publics. *Science, Technology, & Human Values* 41(1):77–92.
- Crouch, C. 2008. *Postdemokratie*. Frankfurt am Main: Suhrkamp.
- Daase, C. 2010. National, social and human security: on the transformation of political language. *Historical Social Research* 4:22–37.
- Dubiel, H. 1994. *Politik und Ungewißheit*. Frankfurt am Main: Suhrkamp.
- Ekbia, H., et al. 2015. Big Data, Bigger Dilemmas: A Critical Review. *Journal of the Association for Information Science and Technology* 66(8):1523–1545.
- Fletcher, R., A. Kalogeropoulos, D. Levy, N. Newman, und R.K. Nielsen. 2017. Digital News Report 2017. Oxford: Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital%20News%20Report%202017%20web_0.pdf.
- Fletcher, R., D. Levy, N. Newman, und R.K. Nielsen. 2016. Digital News Report 2016. Oxford: Reuters Institute for the Study of Journalism. <http://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital-News-Report-2016.pdf>.
- fög – Forschungsinstitut Öffentlichkeit und Gesellschaft. 2016. *Jahrbuch 2017. Qualität der Medien: Schweiz–Suisse–Svizzera*. Basel: Schwabe.
- Fredriksen, K. 2012. Income Inequality in the European Union, OECD Economics Department Working Papers, No. 952, OECD Publishing. <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ECO/WKP%282012%2929&docLanguage=En>.
- Galtung, J., und M.H. Ruge. 1965. The structure of Foreign News. *Journal of Peace Research* 2(1):64–91.
- Greve, W., und D. Wentura. 1997. *Wissenschaftliche Beobachtung: Eine Einführung*. Weinheim: Beltz.
- Heitmeyer, W. 2010. Disparate Entwicklungen in Krisenzeiten, Entsolidarisierung und Gruppenbezogene Menschenfeindlichkeit. In *Deutsche Zustände*, Bd. 9, Hrsg. W. Heitmeyer, 13–33. Frankfurt am Main: Suhrkamp.

- Hirtenlehner, H. 2009. Kriminalitätsangst – klar abgrenzbare Furcht vor Straftaten oder Projektionsfläche sozialer Unsicherheitslagen? *Journal für Rechtspolitik* 17:13–22.
- Huth, I. 2004. *Politische Verdrossenheit: Erscheinungsformen und Ursachen als Herausforderungen für das politische System und die politische Kultur der Bundesrepublik Deutschland im 21. Jahrhundert*. Münster: LIT.
- Keyes, Ralph. 2004. *The post-truth era: Dishonesty and deception in contemporary life*. New York: St. Martin's Press.
- Lischka, K., und C. Stöcker. 2017. *Digitale Öffentlichkeit. Wie algorithmische Prozesse den gesellschaftlichen Diskurs beeinflussen. Arbeitspapier*. Gütersloh: Bertelsmann Stiftung.
- Lombard, M., J. Snyder-Duch, und C.C. Bracken. 2002. Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability. *Human Communication Research* 28(4):587–604.
- Lüdemann, C. 2006. Kriminalitätsfurcht im urbanen Raum. Eine Mehrebenenanalyse zu individuellen und sozialräumlichen Determinanten verschiedener Dimensionen von Kriminalitätsfurcht. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 58(2):285–306.
- Magin, M., und B. Stark. 2011. Österreich – Land ohne Leuchttürme? Qualitätszeitungen im Spannungsfeld zwischen publizistischer Leistung und strukturellen Zwängen. In *Krise der Leuchttürme öffentlicher Kommunikation*, Hrsg. R. Blum, 97–114. Heidelberg: VS.
- Matzner, T. 2016. Beyond data as representation. The performativity of Big Data in surveillance. *Surveillance & Society* 14(2):197–210.
- McCombs, M.E., und D.L. Shaw. 1972. The agenda-setting function of mass media. *Public opinion quarterly* 36(2):176–187.
- Monahan, T. 2018. Algorithmic Fetishism. *Surveillance & Society* 16(1):2–5.
- Morris, J.W. 2015. Curation by code: Infomediaries and the data mining of taste. *European Journal of Cultural Studies* 18(4–5):446–463.
- ÖWA. 2017. ÖWA Plus. Ergebnisse der Studie 2016 – IV. http://www.oewa.at/fileadmin/Documents/documents/oewa_plus_16q4.pdf.
- Pariser, E. 2011. *The filter bubble: What the Internet is hiding from you*. UK: Penguin.
- Pollach, I., A. Scharl, und A. Weichselbraun. 2009. Web content mining for comparing corporate and third-party online reporting: a case study on solid waste management. *Business Strategy and the Environment* 18(3):137–148.
- Scharl, A., und A. Weichselbraun. 2008. An automated approach to investigating the online media coverage of US presidential elections. *Journal of Information Technology & Politics* 5(1):121–132.
- Schulz, W. 1976. *Die Konstruktion von Realität in den Nachrichtenmedien: Analyse der aktuellen Berichterstattung*. Freiburg: Alber.
- Schulz, W., und K. Dankert. 2017. Informationsintermediäre – Anknüpfungspunkte für rechtliche Regulierung. *Informatik-Spektrum* 40(4):351–354.
- Seethaler, J. 2015. Qualität des tagesaktuellen Informationsangebots in den österreichischen Medien. Eine crossmediale Untersuchung. Wien: Rundfunk und Telekom Regulierungs-GmbH (RTR). https://www.rtr.at/de/inf/SchriftenreiheNr12015/Band_1-2015.pdf.
- Standing, G. 2011. *The Precariat. The new dangerous class*. London: Bloomsbury.
- van de Vijver, F.J.R., und N.K. Tanzer. 2004. Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology* 54(2):119–135.
- Weichbold, M. 2009. Zur Bestimmung und Sicherung der Qualität von Umfragen. In *Umfrageforschung. Herausforderungen und Grenzen*, Hrsg. M. Weichbold, J. Bacher, und C. Wolf, 553–570. Heidelberg: Springer VS.
- Weichselbraun, A., S. Gindl, F. Fischer, S. Vakulenko, und A. Scharl. 2017. Aspect-based extraction and analysis of affective knowledge from social media streams. *IEEE Intelligent Systems* 32(3):80–88.
- Weichselbraun, A., S. Gindl, und A. Scharl. 2013. Extracting and grounding contextualized sentiment lexicons. *IEEE Intelligent Systems* 28(2):39–46.
- Weichselbraun, A., S. Gindl, und A. Scharl. 2014. Enriching semantic knowledge bases for opinion mining in big data applications. *Knowledge-Based Systems* 69:78–85.

Wolfgang Aschauer ist seit 2016 Assoziierter Professor an der Abteilung Soziologie und Kulturwissenschaft der Universität Salzburg. 2015 hat er sich mit der Monografie *Das gesellschaftliche Unbehagen in der EU. Ursachen, Dimensionen, Folgen* habilitiert, das Buch ist 2017 bei Springer VS erschienen. Seine Forschungsbereiche sind quantitative Methoden (insbesondere ländervergleichende Survey-Forschung) sowie Mobilität und sozialer Wandel (Rassismusforschung, Migrationsforschung, Tourismusforschung). Wolfgang Aschauer is Associate Professor in the Department of Political Science and Sociology at the University of Salzburg. In 2015, he completed his habilitation project on *The societal malaise of EU citizens. Causes, characteristics, consequences*, which was published as a monograph by Springer VS in 2017. His main research fields are quantitative methods (particularly cross-national survey research) and mobility and social change (particularly racism and migration research, tourism research).

Alexander Seymer Dr., ist Senior Lecturer an der Abteilung für Soziologie und Kulturosoziologie der Universität Salzburg. Er promovierte 2013 an der Universität Salzburg zum Thema Werte und politische Einstellungen im europäischen Ländervergleich. Seine Forschungsinteressen umfassen die Werte- und Einstellungsforschung, politische Soziologie, quantitative Methoden der Sozialwissenschaften und hier im Besonderen die Verwendung latenter Konstrukte. Dr. Alexander Seymer is Senior Lecturer in the Department of Sociology and Cultural Sociology at the University of Salzburg. He received his doctorate in 2013 from the University of Salzburg on the subject of values and political attitudes in the European Union. His research interests include research on values and attitudes, political sociology, quantitative methods of the social sciences and, in particular, the application of latent variables.

Martin Weichbold Dr., ist A.O. Univ.Prof. für Soziologie und empirische Sozialforschung an der Universität Salzburg. Nach dem Studium der Soziologie, Politikwissenschaft und Psychologie und der Habilitation über computergestützte Formen der Befragung beschäftigt er sich insbesondere mit Qualitätskriterien der Sozialforschung, mit interkulturellen Vergleichen sowie methodischen und methodologischen Fragen der Datenkonstruktion. Dr. Martin Weichbold is university professor for sociology and empirical social research at the University of Salzburg. After studying sociology, political science and psychology and completing his habilitation on computer-assisted forms of questioning, he is now particularly concerned with quality criteria of social research, intercultural comparisons and methodical and methodological questions of data construction.

Thomas Herdin Dr., ist Assoz. Prof. für Kommunikationswissenschaft an der Universität Salzburg und Leiter der Abteilung für Transkulturelle Kommunikation. Seine Forschungsschwerpunkte liegen in den Bereichen interkulturelle Kompetenz und kulturelle Resilienz, Wertewandel und transkulturelle Transformationsprozesse in Asien, interkulturelles Management, De-Westernisierungsansätze in der Kommunikationswissenschaft und Tourismusforschung. Dr. Thomas Herdin is Assoc. Prof. in the Department of Communication Studies at the University of Salzburg and head of the Transcultural Communication Division. Field of research: intercultural competence and cultural resilience, value changes and cultural transformation processes in Asia, intercultural management, de-westernization in communication studies and tourism research.

Andreas Röser B.A., ist am Fachbereich Kommunikationswissenschaft der Universität Salzburg in der Abteilung Transkulturelle Kommunikation als Studienassistent tätig und unterstützt den Fachbereich bei der Methodenausbildung. Seine Forschungsschwerpunkte sind die transkulturelle Kommunikation, Kommunikation für Entwicklung sowie die Methodenforschung. Andreas Röser, B.A. is working as a teaching assistant in the Communications department of the University of Salzburg (Transcultural Communication). He supports the department in courses concerning methods and statistical training. His research focuses on transcultural communication, communication for development and survey research.