

Reflecting Bubbles – Ein IT-Artefakt zur Visualisierung der eigenen Echokammer

Wie sollte ein IT-Artefakt designt sein, um Twitter-Nutzer:innen die eigene Echokammer reflektieren zu lassen und Entscheidungsautonomie zu stärken?

Marie Schmidt
marie.schmidt@hu-berlin.de

Julia Heuritsch
heuritsj@hu-berlin.de

Klara Podkowik
podkowik@hu-berlin.de

Nina Schlömer
n.schloemer@fu-berlin.de

Laura von Welzeck
laura.welzeck@fu-berlin.de

Abstract

Demokratiegefährdende Polarisierungstendenzen werden im Kontext sozialer Medien durch den oftmals unreflektiert limitierten Meinungs Austausch in Filterblasen befördert. Vor diesem Hintergrund zielt die vorliegende Arbeit auf die Entwicklung eines nutzerzentrierten Artefakts ab, das Twitter-Nutzer:innen zur Reflexion über die eigene Filterblase anregt. Dies soll Ambiguitätstoleranz und Entscheidungsautonomie fördern und dadurch Polarisierungsprozessen in sozialen Medien entgegenwirken. Hierzu wurden innerhalb eines Design Science Research-Ansatzes anhand von qualitativen Interviews mit Twitter-Nutzer:innen Designprinzipien entwickelt und in einem prototypischen Artefakt umgesetzt und evaluiert. Die Ergebnisse zeigen, dass die Visualisierung der eigenen Filterblase Reflexionsprozesse anstoßen kann. Ausschlaggebend dafür ist eine mehrdimensionale, transparente und wertneutrale Darstellung der eigenen Filterblase, um Nutzer:innen die Zusammensetzung von Filterblasen und die Wirkungsmechanismen von Filteralgorithmen zu verdeutlichen und die positive Auseinandersetzung mit der eigenen Filterblase anzuregen.

Einleitung

Pluralismus bildet eines der höchsten Güter und zentrales Leitbild liberaler Demokratien (u.a. Dahl 2006) und gleichzeitig eine ihrer größten Herausforderungen (Foroutan 2019). Die legitime Vielfalt von Meinungen und Interessen sowie Austausch und Auseinandersetzung mit diesen ist

für demokratische Prozesse unentbehrlich (vgl. Sunstein 2018: 46; Habermas 1968). In diesem Zusammenhang gewinnen soziale Medien zunehmend an gesellschaftlicher Bedeutung. Digitale Plattformen werden immer stärker als Medium für politische Diskurse verwendet und geben individuellen Nutzer:innen die Möglichkeit zur Partizipation (vgl. Burbach et al. 2019: 442). Mit einem wachsenden Anteil der Bevölkerung, der soziale Medien als Nachrichtenquelle nutzt, wächst die Relevanz sozialer Medien für Prozesse der (politischen) Meinungsbildung (Conover et al. 2011; Bozdag & van den Hoven 2015). Diese globale digitale Vernetzung kann einerseits neue Wege des Austausches und diversere Informations- und Interessensspektren eröffnen, andererseits verweist eine zunehmende Anzahl an Studien auf das Gefahrenpotenzial sozialer Medien für die Demokratie (vgl. Barberá 2020: 2; Burbach et al. 2019: 444; Sunstein 2018).

Ein als potenziell demokratiegefährdend diskutierter Mechanismus ist die als Homophilie bezeichnete Tendenz, mit gleichgesinnten Individuen soziale Cluster zu bilden. Dieser Prozess führt sowohl im analogen als auch im digitalen Raum zur Entstehung sogenannter Echokammern, welche einem Austausch unterschiedlicher Perspektiven im Sinne einer pluralistischen Gesellschaft entgegenwirken (vgl. Justwan et al. 2018: 428f.; vgl. Sunstein 2018: 68), da hier die Fähigkeit, „Vieldeutigkeit und Unsicherheit zur Kenntnis zu nehmen und ertragen zu können“ (Häcker & Stapf 2004: 33), nicht zum Tragen kommt. Die Fähigkeit, Mehrdeutigkeit auszuhalten wird im Folgenden als Ambiguitätstoleranz bezeichnet (vgl. Habermas 1968: 128). Echokammern funktionieren dementsprechend als Katalysator für politische Polarisierungsprozesse, wenn Personen mit homogenen Standpunkten und weniger glaubwürdigen Gegenmeinungen konfrontiert sind (vgl. Conover et al. 2011: 95). In sozialen Medien wird dieses Phänomen durch die personalisierte Selektion angezeigter Inhalte durch Filter-Algorithmen verstärkt und sogenannte Filterblasen entstehen, wobei die Auswahl der Inhalte für Nutzer:innen meist intransparent ist (Pariser 2011) und die Eingeschränktheit der eigenen Entscheidungsautonomie verschleiert bleibt (Bozdag & van den Hoven 2015: 251ff.). Verschiedene wissenschaftliche Beiträge haben die Effekte von Filterblasen auf Meinungsbildungsprozesse untersucht und Lösungsansätze und -strategien für Nutzer:innen entwickelt, um aus diesen auszubrechen (vgl. Burbach et al. 2019: 445; Bozdag & van den Hoven 2015; Pariser 2011). Oftmals bleibt allerdings die Perspektive der Nutzer:innen unterrepräsentiert (vgl. Jürgens et al. 2014: 106), welche von den Filter-Algorithmen auch profitieren, da diese bei der Suche nach persönlich relevanten Inhalten unterstützen.

Da insbesondere Twitter als soziales Medium zum Austausch politischer Perspektiven eine zentrale Rolle einnimmt (vgl. Conover et al. 2011: 90; Justwan 2018: 425), soll die folgende Arbeit

durch die Entwicklung eines IT-Artefakts einen Beitrag dazu leisten, Twitter-Nutzer:innen für die eigene Echokammer sowie den eigenen (selbst- und fremdbestimmt limitierten) Bewegungsradius zu sensibilisieren, um somit Ambiguitätstoleranz sowie die Reflektion diverser Perspektiven von Twitter-Nutzer:innen zu fördern und damit Polarisierungstendenzen entgegenzuwirken.

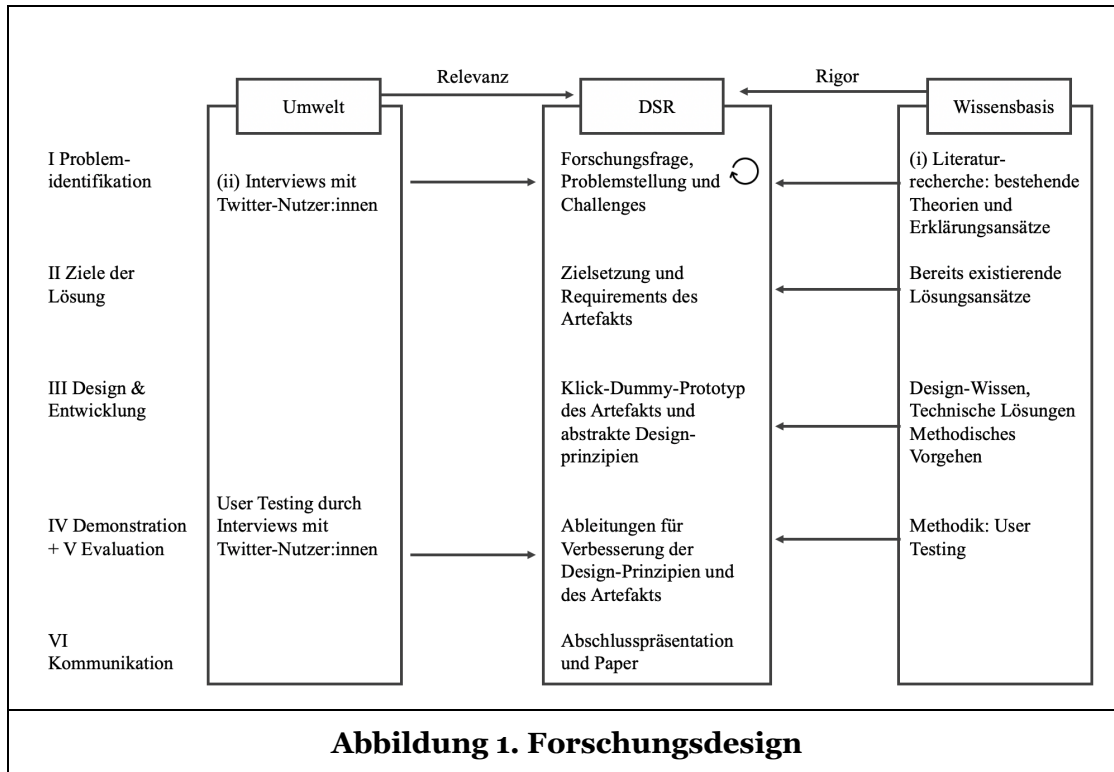
Daher lautet die Forschungsfrage dieses Papers:

Wie sollte ein IT-Artefakt designt sein, um Twitter-Nutzer:innen die eigene Echokammer reflektieren zu lassen und Entscheidungsautonomie zu stärken?

Zur Beantwortung dieser Forschungsfrage wurden nach einem Design Science Research (DSR)-Ansatz Designprinzipien für ein Dashboard, das Twitter-Nutzer:innen deren Echokammer visualisiert, entwickelt und in einem nutzerorientierten Artefakt prototypisch umgesetzt. Das vorliegende Paper bildet dabei den ersten Design-Zyklus ab und orientiert sich in seiner Struktur an den Phasen des DSR-Forschungsprozesses nach Peffers et al. (2007). Nach der Beschreibung des Forschungsdesigns werden Problemstellung und Motivation dargelegt, bevor die Zielsetzung, sowie Entwicklung der Designprinzipien und die konzeptionelle Ausarbeitung des Artefakts beschrieben werden. Anschließend wird die Evaluation des Artefakts ausgewertet, bevor abschließend die Ergebnisse des Forschungsprozesses diskutiert werden.

Forschungsdesign

Durch die Entwicklung eines spezifischen sozio-technischen IT-Artefakts anhand von präskriptiven Designprinzipien trägt diese Arbeit zum Aufbau von Design-Wissen bei (Gregor et al. 2020: 2f.). Die dabei entwickelten Designprinzipien sind generisch und demnach auch für spätere DSR nutzbar (Gregor & Jones 2007: 322). DSR bietet neue, innovative Lösungsansätze für bisher nicht gelöste Probleme und verbindet dabei den Problemraum mit der Lösungssuche im Lösungsraum. Hierbei werden in der Umwelt identifizierte Probleme und Bedürfnisse mit schon existierenden Methoden und Erklärungsansätzen aus der Wissensbasis zusammengebracht, um sowohl relevante und nützliche als auch wissenschaftlich fundierte Lösungen für ein real existierendes Problem zu entwickeln (vgl. Hevner et al. 2004: 79ff.; Gregor et al. 2020: 2f.; Maedche et al. 2019: 18), wie unser Vorgehen verdeutlicht (*Abb. 1*).



Dieses Vorgehen stimmt mit den Designprinzipien der ISO-Norm für Human-Centered-Design überein, welche u.a. die Involvierung von Nutzer:innen, ein iteratives Vorgehen und die nutzerzentrierte Evaluation des Artefakts als essentiell ansehen, um ein nutzerzentriertes Artefakt zu designen, was ein Kernanliegen unserer Arbeit darstellt (International Organization for Standardization 2019: 13f.)

In der ersten DSR-Stufe, der Problem-Identifizierung, wurden nach einer ersten Konkretisierung der Problemstellung durch Auseinandersetzung mit der Literatur zur weiteren Erkundung des Problemraums qualitative Interviews mit Twitter-Nutzer:innen durchgeführt, um tiefere Einblicke in das Nutzungsverhalten und den Umgang mit Filterblasen auf Twitter zu erhalten. Anhand der Interviewergebnisse wurde anschließend erneut in der Wissensbasis nach Erklärungen und bestehenden Lösungsansätzen gesucht. Dieses iterative Vorgehen ermöglicht eine tiefgreifende Auseinandersetzung mit dem Problemraum, der in den einzelnen Iterationen immer weiter konkretisiert wird. Somit soll vermieden werden, dass die Problem-Identifizierung überstürzt wird und dabei relevante Dimensionen des Problems übersehen werden (Maedche et al. 2019: 29). Ergebnis dieser ersten DSR-Phase sind *Challenges*, durch die das Problem konzeptionell zerlegt und dadurch Komplexität reduziert werden kann (Peppers et al. 2007: 12), aus denen die Forschungsfrage abgeleitet wird.

Anschließend wurden die Ziele konkretisiert, die unsere Lösung erreichen soll. Um realisierbare Anforderungen zu entwickeln, wurden hierzu bestehende Lösungsansätze untersucht und aus diesen, sowie aus den Challenges, qualitative *Requirements* für das zu entwickelnde Artefakt abgeleitet (ebd.: 12), welche durch die Kooperation mit potentiellen Nutzer:innen die Entwicklung eines bedürfnisorientierten Artefakts ermöglichen (Maedche et al. 2019: 19).

In Stufe III des DSR-Prozesses wurden aus diesen Requirements Designprinzipien abgeleitet, welche in Features, die die Funktionalitäten des umzusetzenden Artefakts beschreiben, übersetzt wurden. Anhand dieser wurde das Artefakt dann in Form eines Prototypen umgesetzt, was eine typische DSR-Methode des Artefakt-Designs ist (Österle et al. 2011: 3).

Anschließend wurde das Artefakt hinsichtlich seiner Zielerreichung durch qualitative Interviews mit dem gleichen Nutzer:innen-Sample wie in der ersten DSR-Phase evaluiert. Dies erfolgte nach einem partizipatorischen Ansatz, um besser zu verstehen was die Nutzer:in bei der Nutzung des Prototypen denkt, wie sie das Artefakt bewertet und wie dieses verbessert werden könnte (Maguire 2001: 614). Die Interviewergebnisse wurden danach in die Designprinzipien eingearbeitet.

Die letzte Phase des DSR-Prozesses ist die Kommunikation der Forschungsergebnisse. Diese wurden bereits im Rahmen des Seminars präsentiert und anhand dieses Papers verschriftlicht.

Problem-Identifizierung und Motivation

Digitale Plattformen, welche zunehmend zum Schauplatz politischer Diskurse werden, könnten Nutzer:innen die Möglichkeit der selbstbestimmten Partizipation ermöglichen, um einem liberal-demokratischen Verständnis von freier Meinungsbildung und einem deliberativ-demokratischen Verständnis von Diskursfähigkeit zu entsprechen. Die Entscheidungsautonomie von Nutzer:innen wird jedoch innerhalb des digitalen Raums durch die intransparente Vorselektion angezeigter Inhalte durch Algorithmen eingeschränkt (vgl. Bozdag & van de Hoven 2015: 251ff.). Obwohl wissenschaftliche Forschung bisher nicht eindeutig und abschließend klären kann in welcher Art und in welchem Umfang diese Vorselektion von Informationen politische Meinungsbildungsprozesse beeinflusst, steht die Intransparenz der Informationsselektion zunehmend unter Kritik (Jürgens et al. 2014). Das marktwirtschaftliche Interesse der Social Media-Plattformen, Nutzer:innen möglichst oft und lange zur Nutzung der Plattform zu animieren und somit mehr Daten über die Nutzer:innen sammeln und verwerten zu können, führt zu einer individualisierten Auswahl von angezeigten Informationen (ebd.: 104f.). Dies hat für Nutzer:innen potenziell positive Effekte, da

eine gelungene Vorselektion passende Informationen anzubieten und überflüssige bzw. irrelevante Informationen auszusortieren vermag (ebd.) und somit einen wichtigen Beitrag zur Alltagsbewältigung von Nutzer:innen durch Komplexitätsreduktion leisten kann (Burbach et al. 2019, Haken & Schiepek 2006). Gleichzeitig vermittelt die individualisierte Vorselektion eine eingeschränkte Informations- und Meinungsvielfalt (vgl. Sunstein 2018: 6; Jürgens et al. 2014: 98f.), die von Nutzer:innen nicht unbedingt als solche reflektiert wird (Epstein & Robertson 2015) und zu einer eingeschränkten Entscheidungsautonomie von Nutzer:innen führt (Bozdog & van de Hoven 2015). Obwohl sich auch im analogen Raum das Phänomen der Homophilie als Zusammenfinden von Menschen in Gruppen mit ähnlichen Weltanschauungen und Wertesystemen beobachten lässt, scheint die individualisierte Vorselektion von Informationen und Zusammenstellung von Startseiten oder Threads im digitalen Raum zusätzliche Effekte auf Gruppenbildungsprozesse zu haben. Dadurch radikalieren sich Positionen innerhalb von Filterblasen stärker als in analogen Räumen, was als negativ für demokratische Diskurse bewertet werden kann (vgl. Justwan 2018: 428f.; Zuiderveen Borgesius et al. 2016; Sunstein 2018).

Außerdem bedroht ein Austausch, der sich lediglich innerhalb der eigenen Filterblase abspielt die Ambiguitätstoleranz von Nutzer:innen, indem automatisch Meinungen aussortiert werden, die der eigenen widersprechen (vgl. Bakshy et al. 2015: 1130). Damit nimmt die Heterogenität des eigenen sozialen Netzwerks ab, was sich wiederum negativ auf die individuelle Ambivalenz-Fähigkeit auswirkt (vgl. Rudolph 2011: 561).

Werden Filterblasen als den öffentlichen Diskurs einschränkendes Phänomen anerkannt, ist es im Sinne eines liberalen Demokratieverständnisses dieses Phänomen sichtbar zu machen. Des Weiteren sollten Nutzer:innen, wenn sie dies wünschen, durch das Aufzeigen von Entscheidungs- und Handlungsoptionen dazu befähigt werden, diesen entgegenzuwirken.

“The possible negative consequences of algorithms would be eliminated if all people were aware that filter bubbles do exist and whether they are caught inside of one.” (Burbach et al. 2019: 444)

Da wissenschaftlichen Studien bisher wenig Aufschluss darüber bieten, welchen Umgang Nutzer:innen mit Filterblasen wünschen (vgl. Burbach et al. 2019: 442), werden zur Klärung dieser Fragen und zur weiteren Erkundung des Problemraums zunächst qualitative Interviews mit neun Twitter-Nutzer:innen anhand halbstrukturierter Interviewleitfäden (*Appendix I.*) geführt. Dabei

bedienen wir uns einem *Convenience-Sample*, d.h. wir baten Twitter-Nutzer:innen in unserem Umfeld um ein kurzes Interview¹.

Unsere Interviewees nutzen Twitter sehr unterschiedlich - von einem professionellen "engaging people" (IP#8), über einen Austausch mit Gleich- oder Andersgesinnten bis zur Nutzung als reiner Informationsquelle. Acht von neun Interviewees wussten um die Existenz von Filterblasen, allerdings war die Auseinandersetzung mit der Thematik unterschiedlich. Der Großteil wünschte sich mehr konstruktiven Diskurs, was die Erweiterung des eigenen Horizonts und das Verlassen der Komfortzone inkludiert. Einige schätzen jedoch die Komfortzone der Filterblase sowie die einhergehenden Bestätigungen der eigenen Perspektiven, weil diese dabei helfe, den Alltag emotional zu bewältigen (Homophilie).

"Für mich funktioniert Twitter tatsächlich als Meinungsbestätigung." (IP#3)

Gleichzeitig wünschen sich die Interviewees eine bessere Aufklärung über die Existenz, Entstehungsprozesse und möglichen Konsequenzen von Filterblasen². Visualisierungen unterschiedlicher Dimensionen einer Filterblase wie zum Beispiel Netzwerkanalysen und im Feed repräsentierte Themen wurden von einigen Interviewten als reflektionsfördernd eingeschätzt. In Bezug auf die Verlässlichkeit der konsumierten Daten wurden Vertrauens-Checks (in Inhalte und Personen) von allen als wichtig erachtet und selbst durchgeführt, indem beispielsweise auf die Verifizierung von Accounts geachtet wird oder Informationen von vertrauenswürdigen Nachrichtenquellen verglichen werden.

Zusammenfassend lässt sich erkennen, dass Filterblasen, trotz ihrer potenziellen negativen Auswirkungen auf die liberale Demokratie auch Vorteile für die Nutzer:innen haben, welche sich teilweise bewusst in ihren Filterblasen bewegen und wenig Interesse haben sich dem konträren Meinungsspektrum auszusetzen ("Ich glaube, es würde mich zu sehr aufregen, wenn ich dann jetzt oft so AfD Leuten folgen würde"; IP#6). Der Wunsch nach mehr Transparenz der Filter-Algorithmen zeigt allerdings auch, dass Interesse an der Auseinandersetzung mit der eigenen Filterblase vorhanden ist. Auch wünschen sich die Interviewees eine gewisse Diversität der Meinungen in ihrem Feed, welche sich allerdings nicht zu sehr von dem eigenen Wertesystem entfernen sollte ("Auch innerhalb meiner Blase gibt es Auseinandersetzungen und Aushandlungsprozesse, aber

¹ Bei der Interviewauswertung wurde berücksichtigt, dass die geringe Größe des Samples sowie die Methode des Convenience Samplings dazu beigetragen haben, dass die Reflektion über Filterblasen nicht als repräsentativ eingeordnet werden kann, da die Interviewten überdurchschnittlich gebildet und politisch engagiert sind.

² Die positiven Effekte von Transparenz auf User:innen Experience wird in der Literatur auch beispielsweise von Nagulendra und Vassileva (2016: 146f.) hervorgehoben.

ich will mich nicht in meinem Feed mit Menschen auseinandersetzen, die menschenverachtende Meinungen vertreten”; IP#3).

Zielsetzung

Um für die beschriebene Problemstellung - der Ambivalenz von Filterblasen als einerseits demokratiegefährdende Treiber der Polarisierung und andererseits notwendige Hilfsmittel zur Komplexitätsbewältigung für Nutzer:innen - ein problemlösungsorientiertes Artefakt entwickeln zu können, wird im Folgenden die Zielsetzung der Arbeit konkretisiert. Dazu werden bestehende Lösungsansätze aus der Literatur untersucht und anschließend Requirements für das Artefakt abgeleitet. Requirements beschreiben Konditionen, die ein Artefakt besitzen sollte, damit Nutzer:innen ein beschriebenes Problem lösen oder ein Ziel erreichen kann (Maedche et al. 2019: 25).

Um der eingeschränkten Autonomie von Nutzer:innen innerhalb von Filterblasen zu begegnen, gibt es bereits verschiedene Ideen und Tools, welche das gemeinsame Ziel verfolgen, Nutzer:innen ein Gefühl von Kontrolle über ihre angezeigten Inhalte zu vermitteln und angewandte Filterblasen transparent zu machen.

Pariser (2011), der den Begriff der Filterblase begründet hat, nimmt Nutzer:innen in die Pflicht, um der eigenen Filterblase entgegenzuwirken. Er empfiehlt, den eigenen Interessensraum in sozialen Netzwerken aktiv zu vergrößern, Cookies regelmäßig zu löschen und ein Grundverständnis für Algorithmen zu entwickeln (ebd.: 232ff.). Diese vorgeschlagenen Maßnahmen sind zwar konstruktiv, berücksichtigen aber kaum die positiven Effekte von Filterblasen und sind mit einem hohen Aufwand für die Nutzer:in verbunden.

Andere Lösungsansätze sind nutzerzentrierter. So haben beispielsweise Munson et al. (2013) das Browser-Tool “Balancer” kreiert, welches User:innen die Varianz an Leseinhalten im Browser spiegelt, um eine Balance zwischen liberalen und konservativen gelesenen Artikeln zu fördern. Die Wirksamkeit des Tools muss jedoch als gering eingestuft werden, da ein überwiegender Anteil an User:innen trotz der grafischen Übersetzung der eigenen Filterblasen keine Verhaltensänderungen zeigen.

Nagulendra und Vassileva (2014; 2016) hingegen fokussieren in ihrem “control and visualization tool” die Transparenz und Reflektion der eigenen Filterblase. Dieses soll aufzeigen, welche Themen und Menschen innerhalb und außerhalb der eigenen Blase verortet sind. Durch die anschließende Option, die Filterblase dementsprechend zu erweitern, wird die eigene Kontrolle über die

Beschaffenheit und Konsistenz der Filterblase gesteigert. Allerdings ist diese Anwendung für das eher unbekannte Netzwerk MADMICA entwickelt worden.

Twitter selbst bietet mit Twitter Analytics auch eine Analyse-Tool an (Twitter, o.J.), welches zur Analyse der eigenen Wirkung auf Twitter und dem Erfolg von Werbeschaltungen dient. Indem beispielsweise die Reichweite des eigenen Tweets angezeigt wird, veranschaulicht Twitter Analytics jedoch lediglich die eigene Wirkung auf die Follower und nicht die Wirkung der Filterblase auf die Nutzer:in selbst.

Die Analyse bestehender Lösungsansätze betont die Relevanz des Themas erneut. Es wurden in den letzten Jahren verschiedene Artefakte entwickelt, die versuchen Filterblasen aufzuzeigen, aufzubrechen oder Nutzer:innen einer größeren Meinungs-Diversität auszusetzen und dabei teilweise nutzerorientiert vorgehen (zusammenfassend in Bozdag & van den Hoven 2015). Dabei konzentrieren sich allerdings nur wenige direkt auf soziale Netzwerke und kaum eines auf Twitter als meinungsbildende Plattform. Außerdem lässt sich eine Fokussierung auf die US-amerikanische Politik erkennen und ein Großteil der in der Literatur erwähnten Tools sind weitestgehend unbekannt (vgl. ebd.: 261ff.) oder nicht mehr auffindbar.

Unser Artefakt soll auf diesen Ergebnissen aufbauen, dabei aber durch hohe Nutzerorientierung und Fokussierung auf Twitter sowie den deutschsprachigen Raum auf die identifizierten Problem-Challenges und Schwachstellen der bisherigen Lösungsansätze eingehen. Die *Causa finalis* (Heidegger 1997: 6) unseres Artefakt ist es also:

- über die Existenz von Filterblasen und deren Effekte aufzuklären und die eigene Filterblase zu visualisieren, um
- Reflexion ohne Zwang zu schaffen und das Interesse am Diskurs und diversen Meinungen zu wecken und damit
- Ambiguitätstoleranz und Entscheidungsautonomie zu fördern und somit schlussendlich
- Polarisierung entgegenzuwirken und demokratische Meinungsbildungsprozesse zu unterstützen.

Um diese Ziele zu erreichen und damit Nutzer:innen bei der Erfüllung ihrer Bedürfnisse zu unterstützen, muss unser Artefakt folgende Requirements erfüllen:

- DR1: Ein Echokammer-Dashboard sollte gesichertes Wissen über Filterblasen und Echokammern im Allgemeinen bereitstellen.

Burbach et al. (2019: 449) zeigen auf, dass das individuelle Wissen über Filterblasen und deren Auswirkungen das eigene Verhalten im Bezug auf diese Filterblasen beeinflusst. Da unsere Interviews ergeben haben, dass der Kenntnisstand über Filterblasen bei den Interviewees variiert, sollte ein Echokammer-Dashboard allgemeines Wissen über Filterblasen vermitteln, um der Nutzer:in Reflexion und potentielle Verhaltensänderungen zu ermöglichen.

- DR2: Ein Echokammer-Dashboard sollte Inhalte wertfrei vermitteln.

Die Ambivalenz von Filterblasen war eine grundlegende Erkenntnis aus dem Problemraum. Demnach sollten die im Artefakt dargestellten Inhalte so neutral wie möglich sein, um eine Bewertung der Filterblase der Nutzer:in zu vermeiden und ihr die Interpretation der Ergebnisse selbst zu überlassen. Wie Bozdogan und van den Hoven (2015: 245) aus Studien über Technologie-Ethik schließen, manifestieren sich die Werte und Biases der Designer:innen oftmals in den Produkten. Dem soll durch dieses Requirement entgegengewirkt werden.

- DR3: Ein Echokammer-Dashboard sollte die eigene Filterblase anhand unterschiedlicher Dimensionen darstellen.

Wie die Ergebnisse unserer Interviews zeigen, sind die Befragten an der Darstellung ihrer eigenen Filterblase und an Transparenz der Filter-Algorithmen interessiert. Da der Filter-Algorithmus von Twitter (sowie von anderen sozialen Netzwerken) nicht bekannt ist, soll mehr Transparenz und Verständnis für die eigene Filterblase und die Wirkungsweisen der Filter-Algorithmen durch die Darstellung der Filterblase in mehreren Dimensionen erfolgen. Dies führt zu einer Komplexitätsreduktion und soll damit die Verständlichkeit verbessern. Grundsätzlich erhöht die Darstellung der Filterblase das Bewusstsein von Nutzer:innen für diese Blase (vgl. Nagulendra & Vassileva 2016: 146).

- DR4: Ein Echokammer-Dashboard sollte Filterblasen-externe Meinungen aufzeigen.

Filterblasen zeichnen sich dadurch aus, dass Meinungen, die nicht dem vorherigen Nutzungsverhalten der Nutzer:innen entsprechen automatisch "weggefiltert" werden (vgl. Bakshy et al. 2015: 1130). Um die Dimensionen der eigenen Filterblase aufzuzeigen und damit Bewusstsein und Reflexion anzuregen, sollten die diese "weggefilterten" Themen und Meinungen dargestellt werden (vgl. Nagulendra & Vassileva 2016: 148). Das geht mit unseren Interview-Ergebnissen einher, in welchen die Befragten tendenziell an anderen Meinungen interessiert waren, solange diese nicht zu sehr vom eigenen Wertesystem abweichen. Demnach sollten Themen und Meinungen innerhalb eines gesellschaftlich-gerechtfertigten Rahmens und keine menschenverachtenden Aussagen, dargestellt werden.

Die letzten zwei Requirements beziehen sich auf Konditionen der Usability. Auch wenn die Nutzer:innen die Beschäftigung mit der eigenen Echokammer prinzipiell als nützlich erachtet haben, stellt dies dennoch vorerst Mehraufwand für die Nutzer:innen dar. Daher sollte das Artefakt den Usability-Anforderungen des Human-Centered-Designs, wie z.B. Attraktivität des Artefakts, entsprechen (u.a. Maguire 2001: 603), um die Akzeptanz des Produkts und dementsprechend auch die Nutzung zu fördern. Hierbei sind für unser Artefakt zwei Aspekte als alleinstehende Requirements zu fokussieren: Da die Freiwilligkeit der Beschäftigung mit Filterblasen ein zentrales Ziel unseres Artefakts ist, sollte dieses nicht aufdringlich, sondern schnell wegzuklicken sein. Außerdem sollte das Artefakt intuitiv zu nutzen, also einfach zugänglich und leicht zu verstehen sein, um die Hürden zur Nutzung möglichst gering zu halten. Das steht im Einklang mit dem Technology-Acceptance-Modell, in dem die wahrgenommene Einfachheit der Nutzung ein zentraler Einflussfaktor für die Nutzungsakzeptanz ist (Davis 1989: 320) und auch Maguire (2001: 603) hebt die Verständlichkeit als zentrales Usability-Kriterium hervor.

- DR5: Ein Echokammer-Dashboard sollte freiwillig zu nutzen sein.
- DR6: Ein Echokammer-Dashboard sollte einfach zu nutzen sein.

Abb. 2 fasst unsere Ergebnisse jeder DSR-Stufe noch einmal zusammen.

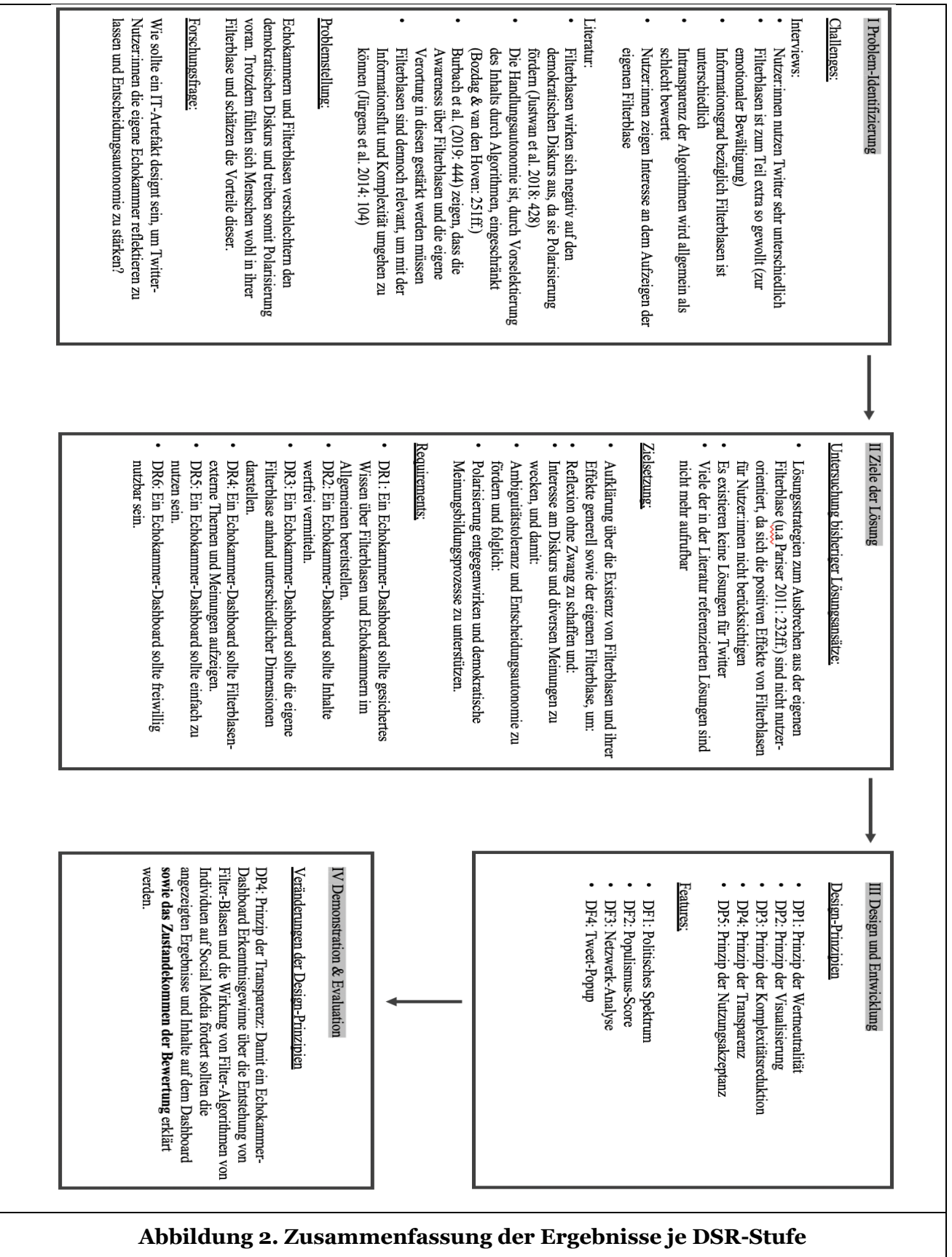


Abbildung 2. Zusammenfassung der Ergebnisse je DSR-Stufe

Design und Entwicklung

Basierend auf den zuvor ausgearbeiteten Requirements leiten sich unsere abstrakten Designprinzipien ab (vgl. Peffers et al. 2004). Hierbei haben wir uns schematisch an Gregor et al. (2020: 22) orientiert und unsere Designprinzipien in Anlehnung an deren dritte Kategorie von Designprinzipien (“design principles about user activity and an artifact”; ebd.: 14) formuliert.

Wir leiten aus den Requirements fünf Designprinzipien ab:

- *DP1: Prinzip der Wertneutralität:* Damit ein Echokammer-Dashboard die Ambiguitätstoleranz von Individuen auf Social Media fördert, sollten Nutzer:innen verschiedene gegensätzliche Meinungen und Themen wertneutral aufgezeigt werden.
- *DP2: Prinzip der Visualisierung:* Damit ein Echokammer-Dashboard das Interesse und die Awareness von Individuen auf Social Media gegenüber Filterblasen und deren Effekten steigert, sollte die eigene Filterblase dargestellt werden.
- *DP3: Prinzip der Komplexitätsreduktion:* Damit ein Echokammer-Dashboard das Verständnis über die eigene Filterblase und die Entwicklung möglicher zielgerichteter Gegenstrategien unterstützt, sollte die Filterblase in verschiedenen Dimensionen zerlegt und visualisiert werden.
- *DP4: Prinzip der Transparenz:* Damit ein Echokammer-Dashboard Erkenntnisgewinne über die Entstehung von Filterblasen und die Wirkung von Filteralgorithmen von Individuen aus Social Media fördert, sollten die angezeigten Ergebnisse und Inhalte auf dem Dashboard erklärt werden.
- *DP5: Prinzip der Nutzungsabsicht:* Damit ein Echokammer-Dashboard die Reflexion über Filterblasen anstößt, sollte es nach Kriterien der Usability designt sein, um so die Akzeptanz der Nutzer:innen und die tatsächliche Nutzung zu fördern.

In Abb. 3 ist dargestellt, wie wir die Designprinzipien aus den Designanforderungen abgeleitet und dadurch die Designanforderungen zusammengefasst und abstrahiert haben.

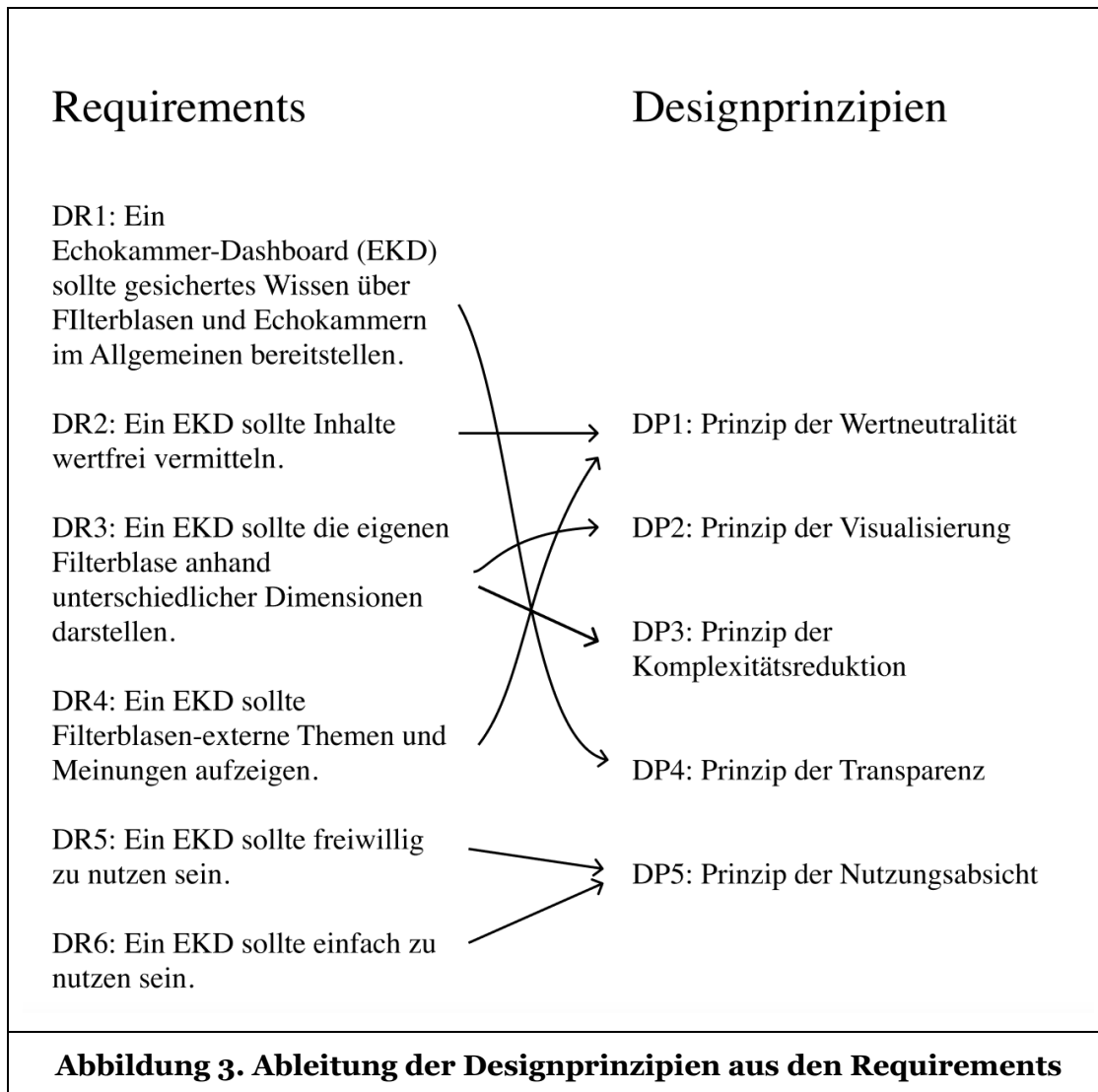


Abbildung 3. Ableitung der Designprinzipien aus den Requirements

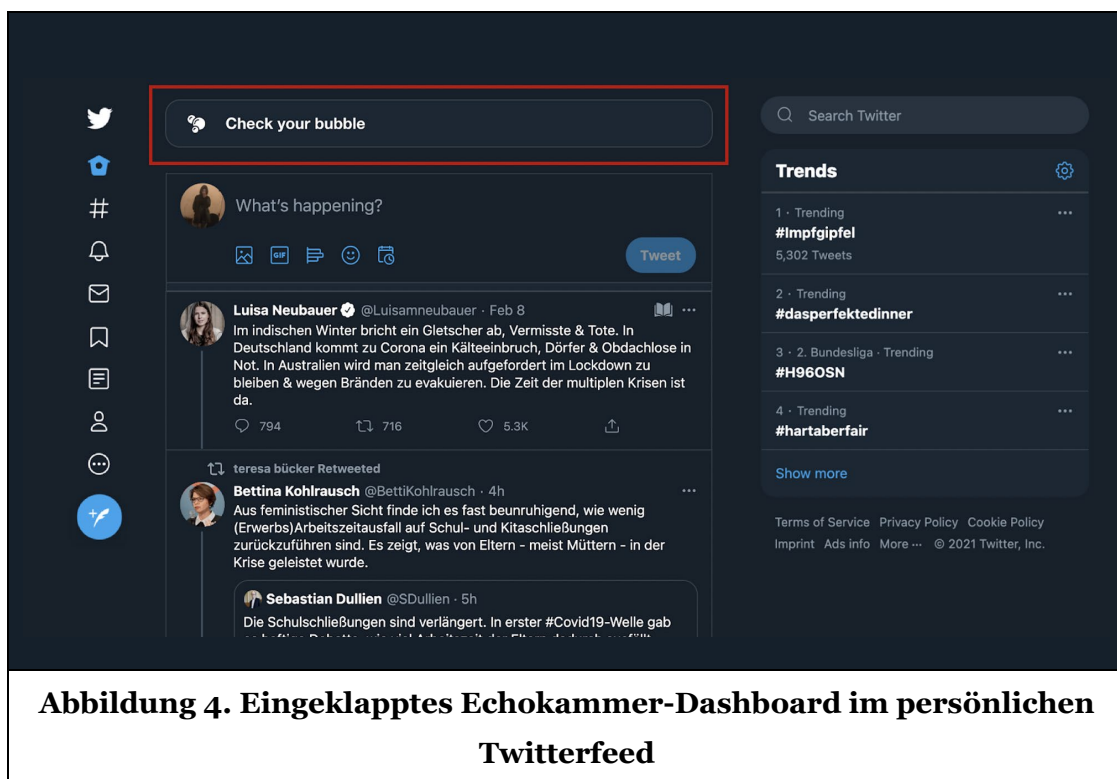
Aus den fünf Designprinzipien werden im nächsten Schritt konkrete Design-Features abgeleitet, wobei ein Feature mehrere Prinzipien enthält. Diese Features bilden in ihrer Kombination unser Artefakt, welches wir “Echokammer-Dashboard” nennen.

In der ersten Design-Iteration unseres Artefakts haben wir einen konzeptionellen Prototypen erstellt, also eine Vorab-Darstellung einiger Aspekte unseres Konzepts (Camburn et al. 2017: 1), um die Designprinzipien möglichst gut bezüglich ihrer Zielerfüllung evaluieren zu können. Außerdem erlaubt die Prototyp-Entwicklung eine idealtypische Gestaltung, die sehr nah an unseren Designprinzipien ist, ohne bereits technischen Einschränkungen folgen zu müssen. Wir entschieden uns dafür, einen Klickdummy-Prototypen mit Figma zu erstellen, der entsprechend des Prinzips

der Nutzungsabsicht ein möglichst ansprechendes und Twitter-nahes Design aufweist. Gleichzeitig entwickeln wir erste Ideen für eine konkrete technische Umsetzung (*siehe Appendix II.* für eine detaillierte Erklärung).

Grundlegendes Designkonzept und Homescreen

Das grundlegende Design-Konzept des Echokammer-Dashboards orientiert sich an DP5, aufgrund dessen das Artefakt nach Kriterien der Usability designt sein sollte. Zurückgeführt auf die Designanforderungen bedeutet dies, dass das Echokammer-Dashboard einfach und freiwillig zu nutzen sein sollte, um unserer Problemstellung möglichst gut zu begegnen. Daher haben wir uns dafür entschieden, das Artefakt “nahtlos” in die Twitter-Umgebung einzubinden, indem die Farben und Schriftarten des Corporate Designs von Twitter verwendet und eine einfache Ein- und Ausblendung des Artefakts durch Klicken ermöglicht werden (*Abb. 4*).



Nach dem Öffnen des Artefakts zeigt sich der Homescreen (*Abb. 5*). Dieser besteht aus verschiedenen, individuell anpassbaren Widgets, die sich aus den anderen Tabs *Political Spectrum*, *Populism* und *Network* zusammensetzen. Diese Funktionen gehören zusammen mit dem *Tweet-Popup* zu den insgesamt vier Features, die wir aus den Designprinzipien abgeleitet und in der

ersten Iteration unseres Artefakts umgesetzt haben. Die mehrdimensionale Darstellung der eigenen Echokammer basiert auf dem Prinzip der Komplexitätsreduktion, nach welchem die Filterblase in verschiedene Dimensionen zerlegt und visualisiert werden soll (vgl. auch DR3). Durch den Homescreen lassen sich diese verschiedenen Dimensionen nach individueller Präferenz anordnen und das Artefakt dadurch personalisieren (vgl. DR5).

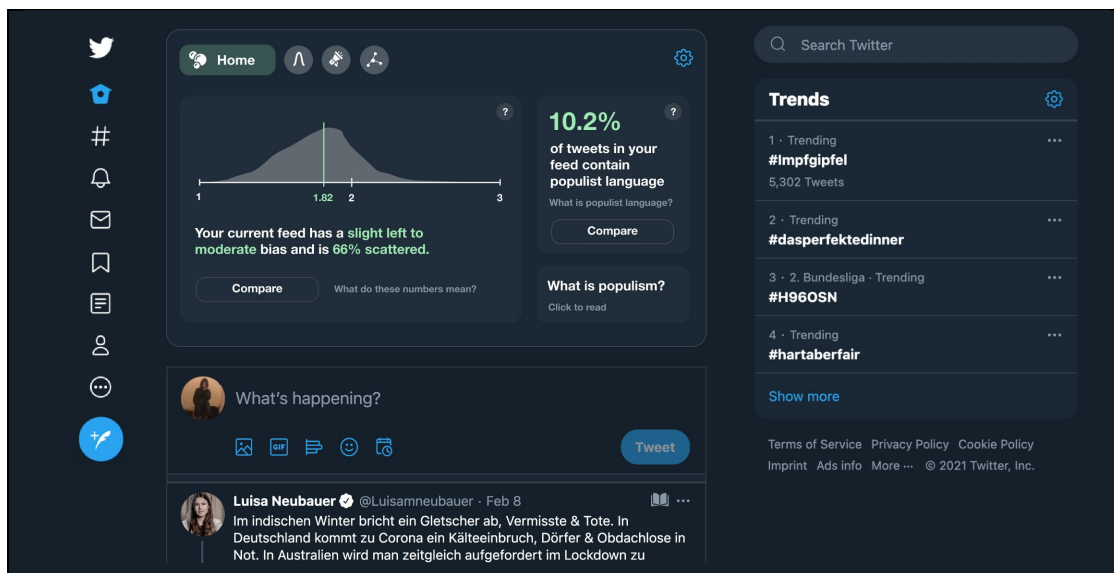


Abbildung 5. Aufgeklapptes Echokammer-Dashboard im persönlichen Feed, welches den mit Widgets individualisierbaren Homescreen zeigt

Feature 1: Politisches Spektrum (Links-Rechts-Barometer und Streuungsdiagramm)

Als erstes Feature, anhand dessen die eigene Filterblase visualisiert werden soll, stellt unser Prototyp die politische Rechts-Links-Einordnung und Streuung des eigenen Feeds unter dem Tab *Political Spectrum* dar (Abb. 6). Diese Analyse basiert auf den Tweets, die der Nutzer:in in den letzten 24 Stunden im Feed angezeigt wurden, wobei nicht bekannt ist, wie diese angeordnet sind. Das Links-Rechts-Spektrum stellt nach wie vor die wichtigste Kategorisierung in der vergleichenden politischen Analyse dar (vgl. Jahn 2011: 746) und wird häufig zur Messung von ideologischer Polarisierung verwendet (vgl. Huber 1989: 600). Unsere Interviews unterstützen die Relevanz des Links-Rechts-Spektrums, da zur Beschreibung des eigenen Feeds ohne vorherige Einführung oftmals auf die Begriffe “links” und “rechts” zurückgegriffen wurde. Daher verwenden wir in unserem Prototypen eine eindimensionale Links-Rechts-Skala, wobei die Zahl 1 das extrem linke und die Zahl 3 das extrem rechte Ende der Skala darstellt (Eilders 2004). Für die Berechnung wollen

wir die Wordscore-Analyse anwenden, die neue Texte anhand von Referenztexten, deren politische Positionierung bekannt ist, einordnet (Laver et al. 2003; siehe Appendix II. für eine detaillierte Erklärung).

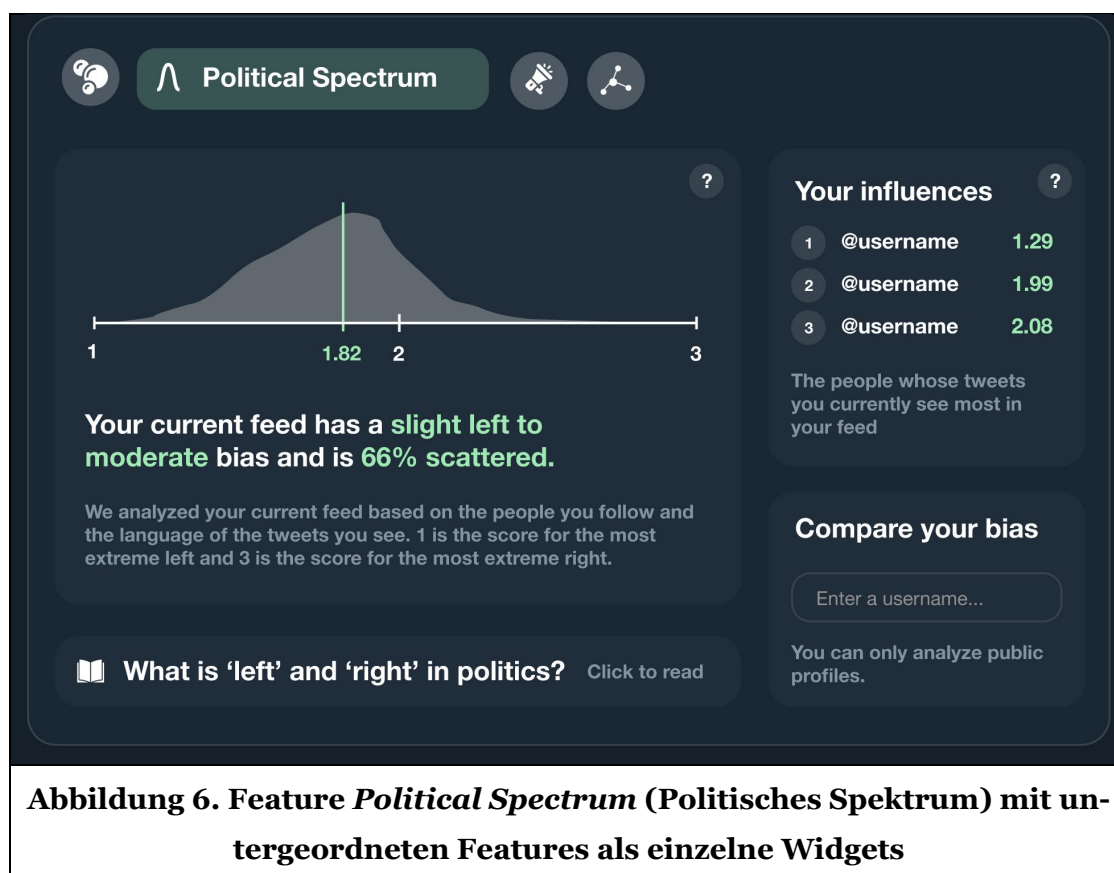


Abbildung 6. Feature *Political Spectrum* (Politisches Spektrum) mit untergeordneten Features als einzelne Widgets

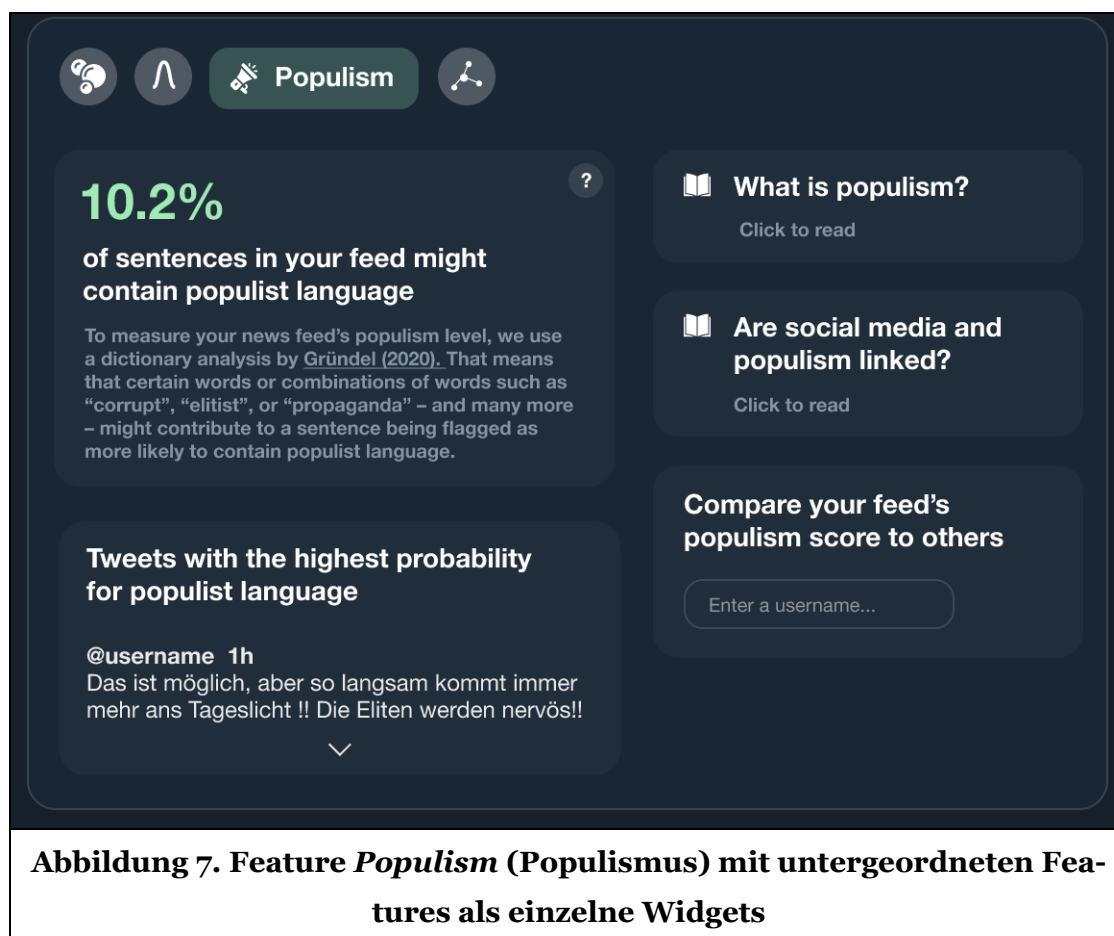
Als Streuungsmaß verwenden wir den Variationskoeffizienten, also das prozentuale Verhältnis der Standardabweichung zum arithmetischen Mittel, da wir eine relative Angabe in Prozenten für verständlicher einschätzen. Als zusätzliche Visualisierung der Streuung ist die absolute Auftretenshäufigkeit der in die Links-Rechts-Skala eingeordneten Tweets in dem Widget hinterlegt, an der sich auch Polarisierungen oder Konvergenzen erkennen lassen.

Weiterhin finden die Nutzer:innen unter diesem Tab Informationen zu den Nutzer:innen, deren Tweets sie aktuell am Häufigsten in ihrem Feed vorfinden ("Your influences") sowie deren Einordnung auf dem Links-Rechts-Spektrum, was das Zustandekommen der politischen Ausrichtung des aktuellen Feeds teilweise erklären kann. Verwandt dazu ist die Möglichkeit auch die Feeds anderer (öffentlicher) User:innen zu analysieren. Insbesondere im Hinblick auf das *Prinzip der Transparenz* beinhaltet dieser Tab Erklärung der Messungen über „?“-Schaltflächen sowie

Verlinkungen von Artikeln, die gesichertes Wissen über die Einordnung in das politische Spektrum bereitstellen (vgl. DR1, DP4).

Feature 2: Populismus-Score

Populismus beschreibt eine Ideologie, die die Gesellschaft in zwei homogene und antagonistische Gruppen aufgeteilt betrachtet und argumentiert, dass Politik ein Ausdruck des allgemeinen Willens des Volkes sein sollte (Mudde 2007: 23). Ein zentraler Mechanismus – zumindest in rechts-populistischen Bewegungen – ist die Politik der Angst, die auf der Polarisierung der Beziehung der eigenen Gruppe zur Außengruppe baut (ebd.: 89). Das Zusammenspiel von Polarisierung und Populismus versuchen wir daher durch die Interaktion der Features *Political Spectrum* und *Populism* darzustellen (Abb. 7) und dadurch wertneutral zur Reflektion einer möglichen Gefahr durch Populismus auf persönlicher Ebene anzuregen.



Aufgrund dieser Demokratiegefährdung haben wir das *Populism*-Feature als weitere Dimension in unseren Prototypen eingebettet. Zur Ermittlung des Populismuslevels des Newsfeeds möchten

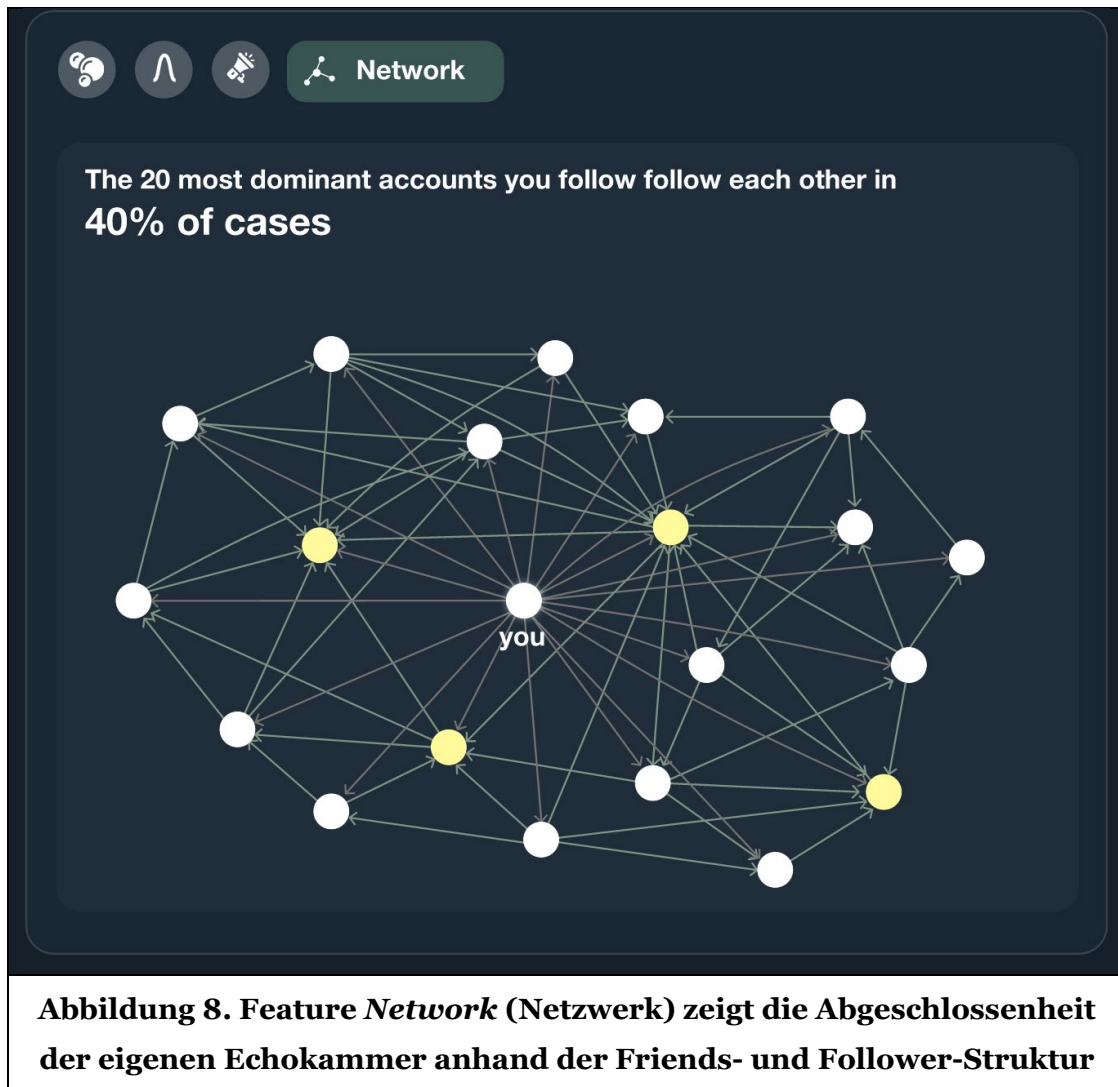
wir ein deutsches Wörterbuch verwenden, welches zur Messung von populistischer Rhetorik in Sozialen Medien entworfen wurde (Gründl 2020). Wird dieses auf den Textkorpus angewandt, werden diejenigen Worte bzw. Wortkombinationen gezählt, die als Indikatoren für Populismus klassifiziert sind. Der aggregierte Populismuswert zeigt an, wieviel Prozent der Sätze in einem Feed populistische Rhetorik enthalten.

Auch das Feature *Populism* wird entsprechend des Prinzips der Komplexitätsreduktion durch weitere Funktionen in Widget-Form unterstützt. Den Nutzer:innen werden als populistisch eingordnete Tweets angezeigt und sie können ihren Populismus-Score mit dem anderer (öffentlicher) Profile vergleichen. Auch bezüglich des Prinzips der Transparenz finden sich weiterführende Artikel und Erklärungen.

Feature 3: Netzwerkanalyse

Politischer Diskurs auf Twitter weist eine stark segregierte Clusterstruktur mit begrenzter Konnektivität zwischen links- und rechtsgerichteten Nutzern auf³, wobei der ausgetauschte Inhalt innerhalb dieser Echokammern homogen ist (Conover et al. 2011: 89ff.). Aufgrund des engen Zusammenhangs zwischen Polarisierung und Clusterbildung trägt das Feature *Network* ebenfalls zu einem erweiterten Blick auf die eigene Echokammer anhand unterschiedlicher Dimensionen bei (vgl. DR3, DP2, DP3). Die Visualisierung der Netzwerkanalyse eignet sich besonders gut dafür, die Abgeschlossenheit von Filterblasen zu illustrieren. Da sich der persönliche Twitter-Feed hauptsächlich aus den Beiträgen gefolgter Personen ergibt, stellen wir das Netzwerk als Netzwerk von "Follows" dar. Unser Prototyp veranschaulicht, inwieweit sich die Friends und Follower einer bestimmten Nutzer:in auch untereinander gegenseitig folgen (*Abb. 8*).

³ Diese Erkenntnis ergibt sich aus der Analyse der *Retweets* auf Twitter. Die *Erwähnung* von anderen Nutzer:innen in Tweets hingegen weist deutlich weniger starke Polarisierung auf (Conover et al. 2011).



Feature 4: Tweet-Popup

Das vierte Feature bildet das *Tweet-Popup*, welches oben rechts von einem Tweet erscheint. Um Meinungen außerhalb der eigenen Filterblase aufzuzeigen (DR4), blenden wir passende Artikel verschiedener Zeitungen zu Tweets ein (Abb. 9). Dabei orientieren wir uns an Thornhill et al. (2019), die ein Artefakt für denselben Zweck entwickelt haben (vgl. *Appendix II.*). Die Zeitungen der angezeigten Artikel sind, ähnlich dem *Political Spectrum*, von 1 bis 3 geordnet, wobei wir diese Ordnung von Eilders (2004: 145) übernommen haben. Entsprechend den Interview-Ergebnissen können wir daher filterblasen-externe Meinungen aufzeigen, die qualitativ abgesichert sind und innerhalb unseres demokratischen Wertekanons liegen und somit die Konfrontation mit “menschenverachtenden Meinungen” (IP#3) vermeiden.



Abbildung 9. Aufgeklapptes Echokammer-Dashboard im persönlichen Feed, welches den mit Widgets individualisierbaren Homescreen zeigt

Evaluation

Für die Evaluation unseres Mockups führten wir einen “Participatory Approach” (Maguire 2001: 614) mit den selben Twitter-Nutzer:innen wie in der ersten Interviewphase durch. Dabei konnten die Nutzer:innen das Mockup selbstständig testen und die einzelnen Features ausprobieren. Ziel der leitfadengestützten Interviews (*Appendix III.*) war es nachzuvollziehen, wie sich Nutzer:innen dem Prototyp annähern und wie nützlich die einzelnen Features im Hinblick auf unsere Objectives eingeschätzt werden.

Die Interviewpartner:innen haben das Artefakt insgesamt als nützlich zur Reflektion der eigenen Filterblase empfunden, was Kern unserer Zielsetzung war und bewerteten es als verständlich und intuitiv. Allerdings gab es auch Verbesserungsvorschläge zu einzelnen Feature-Details. So wurde beispielsweise bei der Netzwerkanalyse angestoßen, die Analyse um weitere Informationskategorien und Visualisierungen zu erweitern. Es wurden detaillierte Erklärungen dazu erwünscht, wie genau die Visualisierung in Bezug auf Details wie die Pfeilrichtungen, -längen etc. konzipiert ist und vor allem, nach welchen Kriterien visualisierte Accounts ausgewählt wurden. “*Most dominant account doesn't mean much just yet.*” (IP#8) betont diesen Wunsch nach klar nachvollziehbaren Definitionen und Funktionsmechanismen des Features. In Bezug auf alle Features wurde reflektiert, dass sich eine größtmögliche Transparenz über das Zustandekommen und die Bedeutung der einzelnen Ergebnisse positiv auf die wahrgenommene Aussagekraft des Artefakts auswirken würde. Umsetzbar wäre eine solche Transparenz-Steigerung beispielsweise anhand eines einführenden Informations-Clips, der eventuell auch Nutzer:innen abholen und involvieren könnte, die durch die Darstellung von Scatterplots und politikwissenschaftlichen Begriffen eher abgeschreckt sind.

Das Links-Rechts-Barometer wurde von den Teilnehmenden als intuitiv anwendbar und spannend, aber auch als ausbaufähig empfunden. So wurde beispielsweise vorgeschlagen, die Darstellung im Sinne einer größeren Aussagekraft nach verschiedenen Themenbereichen wie Klima, Migration und Wirtschaft zu differenzieren.

Während das Prinzip der Wertneutralität insofern als positiv eingeordnet werden kann, als dass sich Nutzer:innen grundsätzlich nicht bewertet empfunden haben, äußerten zwei der Interviewpartner:innen, dass sie die wertende Einordnung von Tweets anhand eines Populismusscores als kritisch betrachten und eine höhere Transparenz des angewandten Bewertungsprozesses befürworten.

"Tweets leben von zugespitzten Formulierungen" (IP#3)

Weiterhin lässt sich vermuten, dass ein sich mit der individuellen Nutzung entwickelndes Artefakt zur langfristigen Anwendung führt. Daher sollten weiterführende Informationen zu einzelnen Features mit jeder Nutzung vertieft und erweitert werden, damit die User:innen in einem kontinuierlichen Lern- und Reflektionsprozess unterstützt werden.

Damit lassen sich die entwickelten Designprinzipien grundsätzlich bestätigen. Allerdings wurde die Bedeutung des Prinzips der Transparenz hervorgehoben, welches wir daher umfassender formulieren möchten. So soll durch das DP4 nicht nur die Transparenz im Bezug auf die Entstehung von Filterblasen und den Wirkungsmechanismen von Filteralgorithmen gestärkt werden, sondern auch die Bewertungsprozesse transparenter dargestellt werden, um einer Skepsis gegenüber dem Echokammer-Dashboard entgegenzuwirken und damit die Auseinandersetzung mit der eigenen Filterblase zu fördern. Demnach möchten wir das Designprinzip wie folgt anpassen:

*Prinzip der Transparenz: Damit ein Echokammer-Dashboard Erkenntnisgewinne über die Entstehung von Filter-Blasen und die Wirkung von Filter-Algorithmen von Individuen auf Social Media fördert, sollten die angezeigten Ergebnisse und Inhalte auf dem Dashboard **sowie das Zustandekommen der Bewertung** erklärt werden.*

Diskussion

In unserem DSR-Projekt konnten wir anhand von Designprinzipien ein prototypisches Artefakt in einem ersten Design-Zyklus umsetzen und evaluieren. Allerdings war unser Forschungsprozess einigen Limitationen unterworfen, auf welche wir folgend eingehen. Da die Umsetzung entwickel-

ter Ideen im Rahmen des Seminars zeitlich beschränkt war, sind sowohl die technische Entwicklung als auch die Evaluierung pragmatisch umgesetzt worden. Daher sollten bei Weiterführung der Forschung Detailfragen stärker fokussiert werden. So sollten bei der technischen Umsetzung des Artefakts mehr theoretische Überlegungen dazu angestellt werden, wie die Visualisierung eine maximale Intuitivität der Anwendung garantiert. Optionale Informationen zu den einzelnen Features, deren Entstehungsmechanismen und Interpretationsmöglichkeiten sowie visuelle Details sollten darüber hinaus nicht nur umfassend zur Verfügung gestellt werden, sondern an individueller Nutzung orientiert vertieft und erweitert werden.

Auch die von uns angedachte technische Umsetzung in späteren Design-Zyklen muss kritisch beleuchtet werden. In Bezug auf die politische Verordnung des Newsfeed können invalide Messungen erfolgen, wenn beispielsweise ein Tweet keinerlei politische Aussage enthält und sich somit nicht politisch verorten lässt, was durch einen Algorithmus dennoch geschehen würde. Auch ist die Qualität der Messung populistischer Rhetorik durch ein Wörterbuch fraglich, da in sozialen Medien eine spezifische Sprache mit vielen Abkürzungen verwendet wird, wodurch eine Messung der populistischen Rhetorik enorm erschwert wird⁴. Da die Features nur prototypisch umgesetzt werden konnten und daher noch nicht vollumfänglich ausgearbeitet waren, muss davon ausgegangen werden, dass in der Evaluation keine abschließende Kritik erhoben werden konnte, zumal das Convenience Sampling eine nicht-repräsentative Stichprobe bedingt hat. Die Nutzung des Artefakts sollten daher anhand einer heterogenen Fokusgruppe über einen längeren Zeitraum wiederholt erhoben werden. Hierbei wäre auch eine Bewertung der Zielerreichung nachgelagerter Ziele, wie z.B. dem Entgegenwirken von Polarisierung, möglich, was in dem kurzen Erhebungszeitraum bisher nicht umsetzbar war. Um erhobene Kritikpunkte umsetzen und erneut testen zu können, sollten außerdem mehrere Iterationen des DSR-Prozesses durchgeführt werden. Letztendlich ist die Implementierung des Artefakts in dem Twitter-nativen Design von der Kooperation mit Twitter abhängig, sodass die Weiterführung dieses Projekts eine solche Kooperation voraussetzt.

⁴Weitere (nicht erschöpfende) Kritikpunkte, die hier aus Kapazitätsgründen nicht weiter ausgeführt werden können: 1. Die Vorauswahl der vorgeschlagenen Zeitungsartikel stellt eine Herausforderung in Bezug auf die links-rechts-Einstufung als auch auf die Qualitätssicherung dar. 2. Wenn User:innen nur einer begrenzte Anzahl an anderen User:innen folgen, kann sich dies auf die Aussagekraft auswirken 3. Dem Artefakt stehen nur Informationen über die User:in in den letzten 24 Stunden angezeigten Tweets zur Verfügung. Es kann nicht ausgewertet werden, in welcher Reihenfolge der User:innen Tweets angezeigt werden.

Fazit

Ziel dieser Arbeit war die Entwicklung eines nutzerzentrierten Artefakts, das Nutzer:innen zur Reflexion über die eigene Filterblase anregt, ohne sie aus ihrer eigenen Blase zu zwingen, um somit Ambiguitätstoleranz und Entscheidungsautonomie zu fördern und dadurch Polarisierungsprozessen in sozialen Medien entgegenzuwirken. Hinsichtlich dieser Zielerreichung kann das entwickelte Artefakt als nutzbares Produkt für Nutzer:innen bewertet werden und stellt somit eine Lösung für das real-existierende Problem, welches im Problemraum identifiziert wurde, dar. Durch die nutzerorientierte und wertneutrale Darstellung der verschiedenen Echokammer-Dimensionen, die Umsetzung im Kontext eines der relevantesten sozialen Netzwerke und die leichte Übertragbarkeit in andere Sprach- und Länderkontexte, bietet unser Artefakt einen relevanten praktischen Beitrag, dessen verschiedene Aspekte noch kein anderes Artefakt vereinen konnte (vgl. Munson et al. 2013, Nagulendra & Vassileva 2016, Bozdog & van den Hoven 2015). Die vorliegende Arbeit trägt darüber hinaus auf zwei verschiedene Arten zur Wissensbasis bei: Sowohl durch die Umsetzung eines realisierbaren Artefakts als auch auf einem abstrakten Level durch die Entwicklung von Designprinzipien, welche nach Gregor und Hevner (2013: 341f.) als erste Stufe hin zur Entwicklung einer Design-Theorie verstanden werden können. Obwohl der Prototyp des vorliegenden Artefakts noch verschiedene Entwicklungsschritte durchlaufen muss, um diskutierte Kritikpunkte umsetzen zu können, kann er als wertvoller wissenschaftlicher und praktischer Beitrag betrachtet werden.

Referenzen

- Bakshy, E., Messing, S., & Adamic, L.A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6329), 1130-1132. <https://doi.org/10.1126/science.aaa1160>
- Barberá, P. (2020). *Social Media and Democracy: The State of the Field, Prospects for Reform*. Cambridge University Press. <https://doi.org/10.1017/9781108890960>
- Bernhard, L., Kriesi, H., & Weber, E. (2015). The Populist Discourse of the Swiss People's Party. In Kriesi, H. & Pappas, T. (Hrsg.), *European Populism in the Shadow of the Great Recession* (125-139). ECPR Press.

- Bozdag, E., & Van Den Hoven, J. (2015). Breaking the filter bubble: democracy and design. *Ethics and information technology*, 17(4), 249-265. <https://doi.org/10.1007/s10676-015-9380-y>
- Burbach, L., Halbach, P., Ziefle, M., & Valdez, A. C. (2019). "Bubble trouble: strategies against filter bubbles in online social networks." In *International Conference on Human-Computer Interaction* (441-456). Springer. https://doi.org/10.1007/978-3-030-22219-2_33
- Camburn, B., Viswanathan, V., Linsey, J., Anderson, D., Jensen, D., Crawford, R., ... & Wood, K. (2017). Design prototyping methods: state of the art in strategies, techniques, and guidelines. *Design Science*, 3. <https://doi.org/10.1017/dsj.2017.10>
- Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political Polarization on Twitter. *Proceedings of the Fifth International Conference on Weblogs and Social Media*, 5(1), 89-96.
- Dahl, R. A. (2006). *On political equality*. Yale University Press.
- Davis, F.D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Eilders, C. (2004). Von Links bis Rechts - Deutung und Meinung in Pressekommentaren. In C. Eilders, F. Neidhardt, B. Pfetsch (Hrsg.), *Die Stimme der Medien. Pressekommentare und politische Öffentlichkeit in der Bundesrepublik* (1. Auflage, 129-166), VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-322-80557-7>
- Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. In *Proceedings of the National Academy of Sciences*, 112(33), E4512-E4521. <https://doi.org/10.1073/pnas.1419828112>
- Foroutan, N. (2019), *Die postmigrantische Gesellschaft: ein Versprechen der pluralen Demokratie*, transcript Verlag.
- Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, 37(2), 337-355. <https://doi.org/10.25300/MISQ/2013/37.2.01>
- Gregor, S., & Jones, D. (2007). The anatomy of a design theory. *Association for Information Systems*, 8(5), 312-335. <https://doi.org/10.17705/ijais.00129>

- Gregor, S., Chandra Kruse, L., & Seidel, S. (2020). Research Perspectives: The Anatomy of a Design Principle. *Journal of the Association for Information Systems*, 21(6), 2. <https://doi.org/10.17705/1jais.00649>
- Gründl, J. (2020). Populist ideas on social media: A dictionary-based measurement of populist communication. *New Media & Society*. <https://doi.org/10.1177/1461444820976970>
- Habermas, J. (1968): Stichworte zu einer Theorie der Sozialisation. In J. Habermas (1973) *ders., Kultur und Kritik* (118-194). Suhrkamp.
- Häcker, H.O., & Stapf, K. (2004): *Dorsch Psychologisches Wörterbuch* (14. Auflage). Huber.
- Haken, H., & Schiepek, G. (2006, [2010]). *Synergetik in der Psychologie – Selbstorganisation verstehen und gestalten*. Hogrefe.
- Heidegger, M. (1977). *The Question Concerning Technology and other Essays*. Garland Publishing.
- Hevner, A., March, S.T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75-105. <https://doi.org/10.2307/25148625>
- Huber, J. D. (1989). Values and partisanship in left-right orientations: Measuring ideology. *European Journal of Political Research*, 17(5), 599-621. <https://doi.org/10.1111/j.1475-6765.1989.tb00209.x>
- International Organization for Standardization. (2019). *Ergonomie der Mensch-System-Interaktion - Teil 210: Menschzentrierte Gestaltung interaktiver Systeme* (ISO 9241-210:2019; Deutsche Fassung EN ISO 9241-210:2019).
- Jahn, D. (2011). Conceptualizing Left and Right in comparative politics: Towards a deductive approach. *Party Politics*, 17(6), 745–765. <https://doi.org/10.1177/1354068810380091>
- Jürgens, P., Stark, B., & Magin, M. (2014). “Gefangen in der Filter Bubble? Search Engine Bias und Personalisierungsprozesse bei Suchmaschinen.” In B. Stark, D. Dörr, & S. Aufenanger (Hrsg.), *Die Googleisierung der Informationssuche* (98-135). De Gruyter. <https://doi.org/10.1515/9783110338218>
- Justwan, F., Baumgaertner, B., Carlisle, J. E., Clark, A. K., & Clark, M. (2018). Social media echo chambers and satisfaction with democracy among Democrats and Republicans in the aftermath of the 2016 US elections. *Journal of elections, public opinion and parties*, 28(4), 424-442. <https://doi.org/10.1515/9783110338218>

- Laver, M., Benoit, K., & Garry, J. (2003). Extracting Policy Positions from Political Texts Using Words as Data. *The American Political Science Review*, 97(2), 311-331.
<https://doi.org/10.1017/S0003055403000698>
- Maedche, A., Gregor, S., Morana, S., & Feine, J. (2019). Conceptualization of the Problem Space in Design Science Research. In B. Tulu, S. Djamasbi, G. Leroy (Hrsg.), *Extending the Boundaries of Design Science Theory and Practice* (18-31), Springer.
- Maguire, M. (2001). Methods to support human-centred design. *International Journal of Human-Computer Studies*, 55(4), 587-634. <https://doi.org/10.1006/ijhc.2001.0503>
- Mihalcea, R., & Tarau, P. (2004). TextRank: bringing order into texts. *Conference on empirical methods in natural language processing*, 404-411.
- Mudde, C. (2007). *Populist Radical Right Parties in Europe*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511492037>
- Munson, S., Lee, S., & Resnick, P. (2013). Encouraging reading of diverse political viewpoints with a browser widget. In *Proceedings of The International AAAI Conference on Web and Social Media*, 7(1).
- Nagulendra, S., & Vassileva, J. (2014). Understanding and controlling the filter bubble through interactive visualization: a user study. *Proceedings of the 25th ACM conference on Hypertext and social media*, 107-115. <http://dx.doi.org/10.1145/2631775.2631811>
- Nagulendra, S., & Vassileva, J. (2016). Providing awareness, explanation and control of personalized filtering in a social networking site. *Information Systems Frontiers*, 18(1), 145-158. <https://doi.org/10.1007/s10796-015-9577-y>
- Österle, H.; Becker, J., Frank, U., Hess, T., Karagiannis, D., Krcmar, H., Loos, P., Mertens, P., Oberweis, A., & Sinz, E.J. (2011). Memorandum on design-oriented information systems research. *European Journal on Information Systems*, 20(1), 7-10.
<https://doi.org/10.1057/ejis.2010.55>
- Otte, E., & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences, *Journal of Information Science*, 28 (6), 441-453.
<https://doi.org/10.1177/016555150202800601>
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin, London.

- Pauwels, T. (2011). Measuring populism: A quantitative text analysis of party literature in Belgium. *Journal of Elections, Public Opinion and Parties*, 21(1), 97-119.
<https://doi.org/10.1080/17457289.2011.539483>
- Peppers, K., Tuunanen, T., Rothenberger, M.A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45-78. <https://doi.org/10.2753/MIS0742-1222240302>
- Rooduijn, M., & Pauwels, T. (2011). Measuring populism: Comparing two methods of content analysis. *West European Politics*, 34(6), 1272-1283.
<https://doi.org/10.1080/01402382.2011.616665>
- Rudolph, T.J. (2011). The Dynamics of Ambivalence. *American Journal of Political Science*, 55(3), 561-573. <https://doi.org/10.1111/j.1540-5907.2010.00505.x>
- Sunstein, C. R. (2018). *# Republic: Divided democracy in the age of social media*. Princeton University Press.
- Thornhill, C., Meeus, Q., Peperkamp, J., & Berendt, B. (2019). A digital nudge to counter confirmation bias. *Frontiers in big data*, 2, 1-9. <https://doi.org/10.3389/fdata.2019.00011>
- Twitter. (o.J.). Twitter Analytics. Test, learn, and optimize with Twitter Analytics.
- Zuiderveen Borgesius, F.J., Trilling, D., Möller, J., Bodó, B., de Vreese, C.H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, 5(1).
<https://doi.org/10.14763/2016.1.401>

Appendix

Appendix.I.– Interviewleitfaden für die Erkundung des Problemraums

Einstieg

Vielen Dank, dass du dir die Zeit für dieses Interview nimmst. Es ist jetzt wahrscheinlich keine Überraschung für dich, dass es in diesem Interview um Twitter und deine Nutzung von Twitter gehen wird, nachdem ich dich ja in meiner Anfrage schon gefragt habe, ob du Twitter nutzt.

Exploration: Twitter use

Warum nutzt du Twitter?

Und wie sieht deine Nutzung aus - wie oft nutzt du Twitter und hast du dabei bestimmte Routinen?

Welche Art Tweets bevorzugst du in deinem Feed?

Trust & Criteria

Inwieweit vertraust du Tweets? Wovon ist dein Vertrauen in Tweets abhängig?

Überprüfst du die faktische Korrektheit von Tweets mithilfe von anderen Quellen? Wenn ja: Bitte beschreibe, wie du das machst/ welche Quellen du benutzt.

Bevor du etwas retweetest - überprüfst du das Profil desjenigen, von dem der Tweet stammt?

Was sind deine Kriterien um anderen zu folgen?

Tendence to echo-chamber

Wenn du an die Tweets in deinem Feed denkst entsprechen sie zum Großteil deinen Meinungen?

Wenn du an die Tweets in deinem Feed denkst, wie fühlt es sich für dich an, wenn die Tweets nicht deiner Meinung entsprechen?

Reflexion on Filter-Bubbles

Wie denkst du, dass dein Tweet sich zusammensetzt?

Meinst du, dass deine Tweets ausbalanciert sind?

Hast du schonmal von dem Begriff Filter Bubbles gehört? Wenn ja: Was weißt du darüber?

Wish for change

Meinst du, dass du dich in einer Filter Bubble befindest?

Wenn ja, Inwiefern?

Wenn ja: Unternimmst du Versuche, eine Filter Bubble zu vermeiden?

Würdest du gerne vielfältigere Tweets lesen?

Sind Filter Algorithmen etwas, was dich stört?

Wenn ja: Wie wichtig ist die Transparenz solcher Algorithmen für dich?

Wenn ja: Was wäre hilfreich für dich, vielfältigere Tweets zu erhalten?

Abschlussfragen

Hast du noch irgendwelche Gedanken, die du mit uns teilen willst, bevor wir zum Schluss kommen?

Appendix.II.– Erste Ideen für eine konkrete technische Umsetzung

Um zu untersuchen, welche Form von Analysen in Twitter möglich wären, haben wir uns zuerst mit RTweet⁵, einem R Package, das zur Analyse von Twitter-Daten entwickelt wurde, beschäftigt. Wenn man einen Developer Account auf Twitter anlegt, kann man mit den meisten der RTweet Funktionen auf die Tweets sämtlicher Twitter-Nutzer:innen zugreifen. Als Inspiration für unser Artefakt fielen uns besonders jene Funktionen auf, die die Nutzerdaten von Friends/ Followers von bestimmten Nutzer:innen sammeln⁶ und jene, die die Timeline der letzten Tweets, inklusive deren Content und wie oft andere Nutzer:innen in diesen erwähnt werden, ausgeben⁷. Diese Daten ermöglichen es theoretisch, mithilfe von Text- und Netzwerkanalysen (SECTIONS) jene Aspekte von Filterblasen zu untersuchen, die wir in unseren Design-Features herausgearbeitet haben.

Messung

Politisches Spektrum (DF1)

Zur Ermittlung der politischen Ausrichtung des jeweiligen Newsfeeds planen wir die Anwendung der *wordscore* Analyse (Laver et al. 2003). Dafür werden mindestens zwei Referenztexte benötigt. Deren (am besten extreme) Position auf der jeweiligen policy-Dimension muss bekannt sein. Außerdem sollten sie den zu untersuchenden Texten (*virgin texts*) sprachlich ähneln sowie ein großes Vokabular verwenden (ebd.). Anhand der Häufigkeit bzw. Einzigartigkeit der verwendeten Begriffe werden allen Worten in den Referenztexten Gewichte zugeteilt und dadurch, in einem nächsten Schritt, die Position der *virgin texts* berechnet. Die Anwendbarkeit dieser Methode ist also durch die Existenz geeigneter Referenztexte bedingt. Dies stellt bei der Analyse von Twitter ein Problem da, weshalb geeignete Referenztexte zunächst selbst erstellt werden müssen: Hierfür planen wir die Kategorisierung von ca. 5000 bis 10000 Tweets auf einer Links-Rechts-Skala durch menschliche Kodierer:innen. Im besten Fall sollte jeder Tweet zwei bis drei Mal von verschiedenen Menschen kodiert werden, um die Validität der Ergebnisse zu erhöhen. Durch diese neuen Referenzwerte kann dann die beschriebene *wordscore* Methode angewandt und die Newsfeeds im politischen Raum verordnet werden – die policy Positionierung dieser Texte kann errechnet werden.

⁵ <https://cran.r-project.org/web/packages/rtweet/rtweet.pdf>

⁶ `get_friends()` & `get_followers()`

⁷ `get_timeline()`

Populismus (DF2)

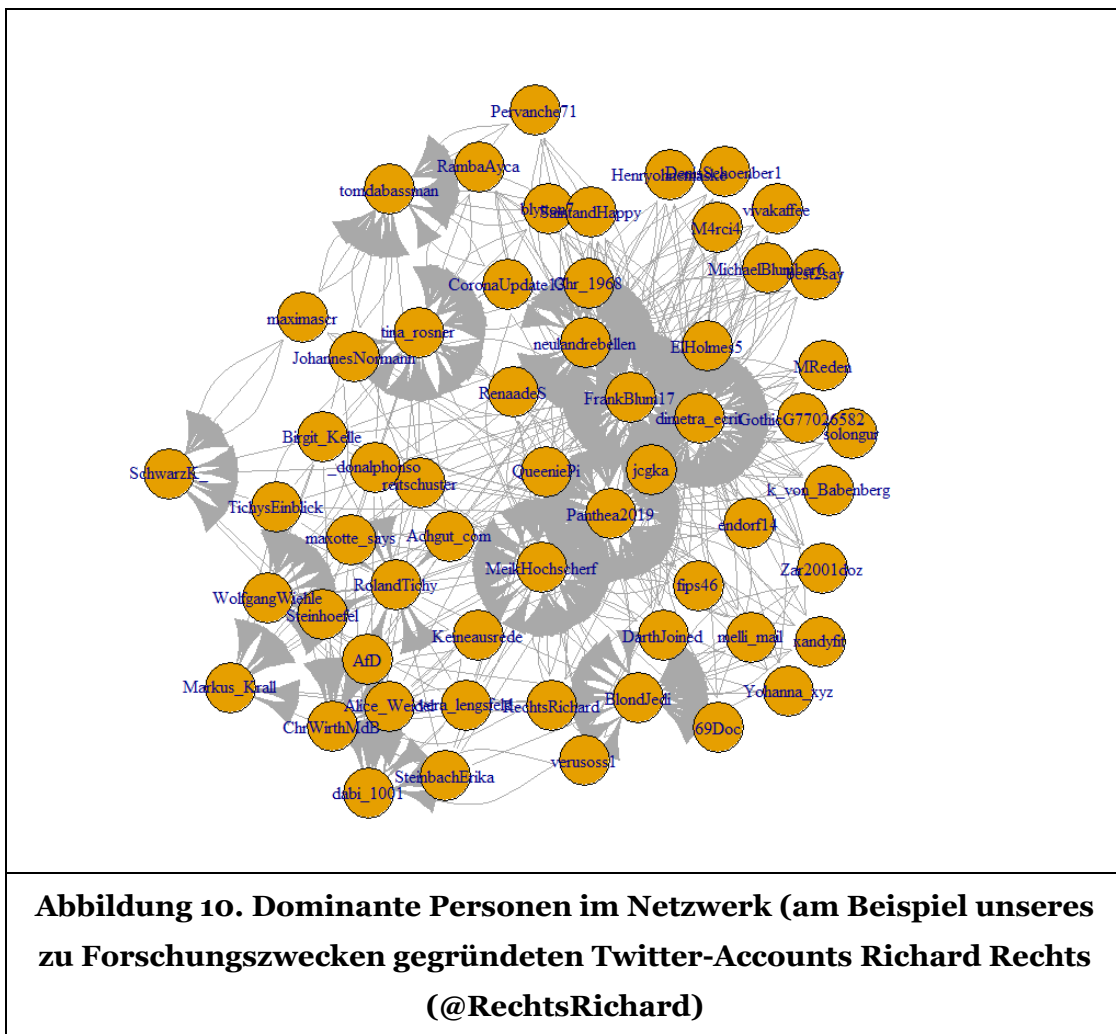
Um das Populismuslevel des jeweiligen Newsfeeds zu berechnen, verwenden wir eine Wörterbuchanalyse. Während es zwar viel Forschung zu quantitativer Textanalyse im Allgemeinen gibt, existiert deutlich weniger Forschung bezüglich des Konzepts Populismus. Die wenigen existierenden Studien erzielen zwar relativ valide Ergebnisse (Bernhard et al. 2015; Pauwels 2011; Rooduijn & Pauwels 2011), doch existieren insgesamt kaum Wörterbücher auf Deutsch und diese sind alle für andere Kontexte entworfen worden (z.B. zur Analyse von Parteiprogrammen oder politischen Reden). Die in Sozialen Medien verwendete Sprache ist jedoch grundlegend verschieden, was ein Wörterbuch beachten sollte. Hier hat Gründl (2020) in seiner Forschung angesetzt und ein Deutsches Wörterbuch zur Messung von populistischer Rhetorik in Sozialen Medien entworfen und dieses auch mit sehr guten Ergebnissen validiert. Abgesehen von dem spezifischen Kontext, hat dieses Wörterbuch auch noch weitere Vorteile: es beachtet mehrere Dimensionen von Populismus und nicht nur *anti-elitism* (wie z.B. Rooduijn & Pauwels, 2011) – misst das Konzept also deutlich umfangreicher und theoriegetriebener – und verwendet eine deutlich größere Anzahl von Begriffen. Dies ist vor allem für die Analyse von sehr kurzen Texteinheiten wie Tweets von großer Relevanz.

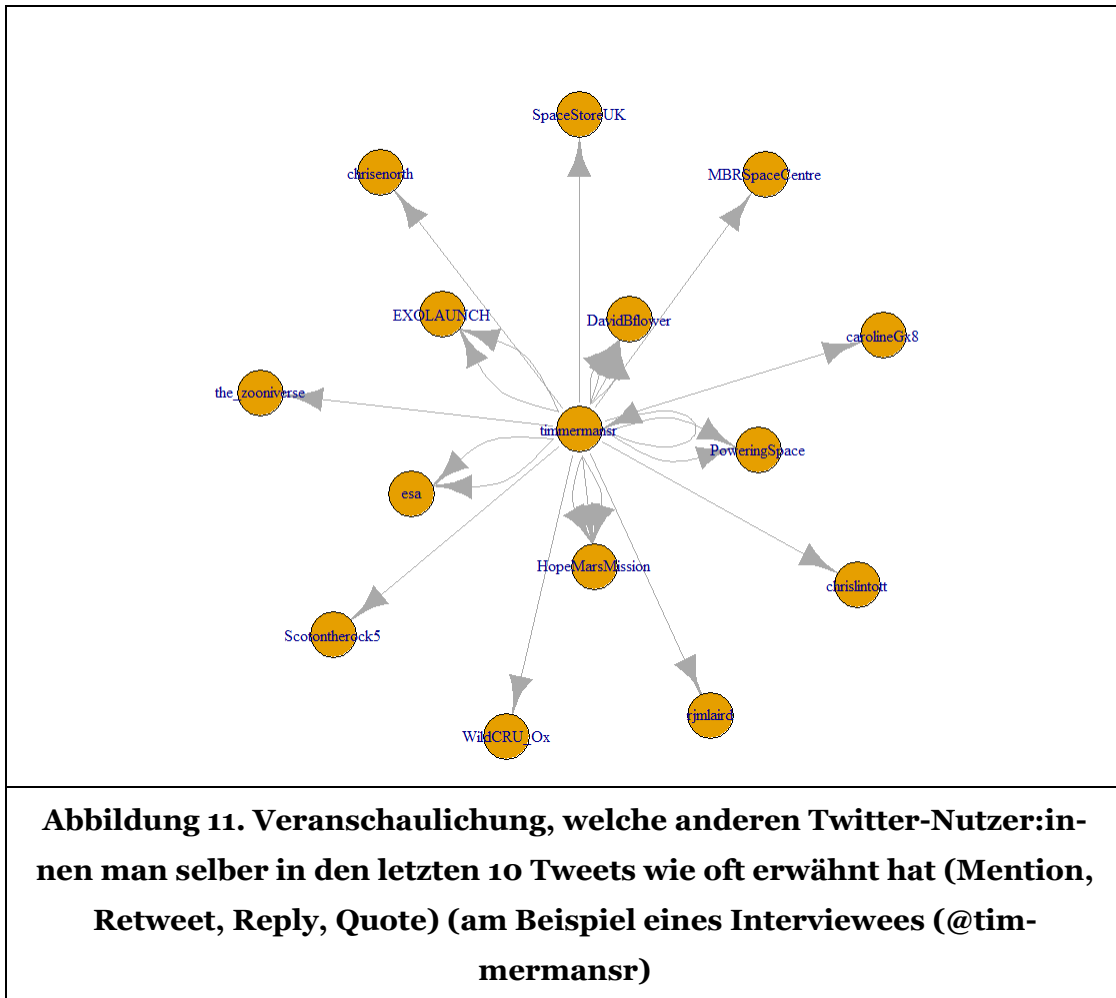
Um dieses Wörterbuch auf den Newsfeed anzuwenden, müssen die Tweets zunächst vorbereitet werden, indem beispielsweise Nummern, Symbole oder Emojis entfernt werden sowie der gesamte Text in Kleinbuchstaben formatiert wird. Da Gründl (2020) nicht nur einzelne Wörter beachtet, sondern auch teilweise die Abfolge oder Kombination von Begriffen, werden die Tweets nicht *tokenized*. Wenn nun das Wörterbuch auf den vorbereiteten Textkorpus angewandt wird, zählt der Computer die Worte, die als Indikatoren für Populismus angesehen werden. Gründl (2020) empfiehlt hier die Untersuchung auf Satzebene statt auf Tweetebene, da somit eine grobe mit Einbeziehung der Länge der Texte passieren kann (da teilweise die Abfolge/Kombination der Wörter von Bedeutung ist, kann nicht auf Wortebene agiert werden). Der daraus resultierende Wert ist eine binäre Messung, ob ein Satz einen populistischen Begriff beinhaltet. Der aggregierte Populismuswert zeigt dann an, wie viel Prozent der Sätze in einem Feed populistische Rhetorik enthalten.

Allerdings müssen auch hier einige Einschränkungen beachtet werden. Eine automatisierte quantitative Textanalyse ist insofern problematisch, als dass Ironie, Sarkasmus, Humor, Zitate oder Redewendungen kaum als solche erkannt werden können (hierfür würde es menschliche Kodierer:innen benötigen) und somit verzerrte Ergebnisse entstehen können.

Netzwerkanalyse (DF3)

Zusätzlich zur Abgeschlossenheit der Filterbubble, die wir in unserem Prototypen veranschaulichen, können durch eine Netzwerkanalyse auch “dominante” Personen im Netzwerk identifiziert werden, also jene Nutzer:innen, denen viele andere Personen im Netzwerk folgen (Abb. 10). Eine dritte Form der Visualisierung kann darstellen, welche Nutzer:Innen ein bestimmter User in dessen letzten Tweets wie oft erwähnt hat (Abb. 11). Zusätzlich zur Veranschaulichung dieser unterschiedlichen Aspekte einer Filterbubble, können diese auch berechnet werden (Otte & Rousseau 2002). So eignet sich das Zentralitätsmaß für die Identifikation von dominanten Personen im Netzwerk und die Maßzahl Dichte um die Abgeschlossenheit des Netzwerks zu berechnen. Mit Hilfe der Cliquenanalyse können Teilgruppen im Netzwerk identifiziert werden.





Tweet-Popup (DF4)

Um den Nutzer:innen passende Artikel zu Tweets im Newsfeed anzeigen können, verwenden wir die Methode von Thornhill et al. (2019). Diese haben das Artefakt *BalancedView* für eben diesen Zweck entwickelt. Bevor geeignete Artikel gesucht werden können, müssen zunächst die relevanten Informationen aus einem Tweet gezogen werden. Hierfür kann der *TextRank* Algorithmus (Mihalcea & Tarau 2004) verwendet werden, der in fünf Schritten die wichtigsten Schlagwörter extrahiert: (1) Der Text wird zunächst bereinigt. Dafür wird er *tokenized*⁸ sowie Stoppwörter, Zahlen, Satzzeichen und twitterspezifische Aspekte (z.B. Hashtags) werden entfernt. (2) Der Algorithmus analysiert die Wortart und es werden nur Nomen, Adjektive und Verben behalten, (3) von welchen jeweils nur der Wortstamm verwendet wird. (4) Als nächstes wird Worten, die sehr oft

⁸ Der Text wird in einzelne tokens – meist Worte, es kann sich aber auch um Sätze oder Satzfragmente handeln – aufgeteilt.

vorkommen bzw. Wörter, die nach einem solchen Wort stehen, ein deutlich höherer Wert zugeschrieben als beispielsweise Begriffen, die nur einmal am Satzende vorkommen. Der Algorithmus geht hier wie folgt vor: Er baut eine Grafik, in der jeder Token ein Knotenpunkt ist und die Kanten die Beziehung zwischen ihnen darstellen. Eine Kante zwischen zwei Wörtern bedeutet, dass diese beiden Wörter im Text aufeinander folgen. Ein Scoring-Faktor weist jedem Knoten Scores zu, basierend auf den Knoten, die vom ersten Wort des Tweets aus erreichbar sind. (5) als letztes werden diese Begriffe nach ihrem Score sortiert und nur die drei bis fünf Schlagwörter mit dem höchsten Score werden behalten. Diese Schlagwörter werden dann dazu verwendet relevante Newsartikel zu finden. Hierzu wird eine Anfrage an NewsAPI⁹ geschickt. Das gute an NewsAPI ist, dass hier simultan mehrere Newsoutlets durchsucht werden können – wenn auch nur aus dem letzten Monat.

⁹ <https://newsapi.org>

Appendix.III.– Interviewleitfaden für das Evaluationsgespräch

Einstieg

Vielen Dank für das letzte Interview - dein Beitrag war sehr wertvoll für das Design unseres Artefakts für Twitter-User. Nun haben wir erstmal ein Mockup designed und würden dich bitten, ein bisschen damit rumzuspielen. Bitte teile uns dabei gleich **laut** deine Gedanken dabei mit, wenn du dich mit den unterschiedlichen Features befasst.

Participatory & Begleitung

Diese Fragen werden für jedes der Features gestellt.

Was denkst du ist die Intentionen hinter diesem Features?

Anweisung für die Interviewerinnen: Wenn sie es nicht wissen, erkläre, was mit dem Feature erreicht werden soll! Setze die Erklärungen immer in Verbindung mit unseren Zwecken "Bewusstseinschaffung" oder "Entscheidungsautonomie"

Wie wahrscheinlich ist es, dass du dieses Feature selbst verwenden wirst?

Hast du Gedanken dazu, wie wir dieses Feature verbessern könnten?

Abschlussfrage

Wenn du nochmal an unser Ziel zurückdenkst - Bewusstseinschaffung über Filterblasen und Entscheidungsautonomie stärken - welche Art von Features fehlen noch?