

MASTERARBEIT

Personalisierung als Ursache von Search Engine Bias: eine Untersuchung personalisierter Google-Sucher- gebnisse zu kontrovers diskutierten Umweltthemen

vorgelegt im Oktober 2019 von
Manuel Reiter

1. Prüfer: Prof. Dr. Dirk Lewandowski
2. Prüferin: Prof. Dr. Ulrike Spree

**HOCHSCHULE FÜR ANGEWANDTE
WISSENSCHAFTEN HAMBURG**
Department Information
Studiengang: Information, Medien, Bibliothek (M.A.)

HOCHSCHULE FÜR ANGEWANDTE
WISSENSCHAFTEN HAMBURG
Hamburg University of Applied Sciences

Personalisierung als Ursache von Search Engine Bias: eine Untersuchung personalisierter Google-Suchergebnisse zu kontrovers diskutierten Umweltthemen

Masterarbeit vorgelegt von
Manuel Reiter

Abstract

Internet-Unternehmen haben durch die Personalisierung das Potenzial erkannt, Inhalte individuell an die Nutzer anzupassen. Die Maßnahmen dazu werden jedoch immer wieder kritisiert. Ein Vorwurf lautet, das Zustandekommen einer sogenannten Filterblase würde die Meinungsvielfalt einschränken, was besonders im Kontext kontroverser Themen als problematisch gilt. In Hinblick auf Suchmaschinen-Verzerrungen können Personalisierungsmaßnahmen potenziell eine nicht unbedeutende Rolle einnehmen. Im Rahmen dieser Masterarbeit wurden personalisierte Suchergebnisse der Suchmaschine Google untersucht, die durch ihre häufige Nutzung eine wichtige Rolle im Wissenserwerb von Gesellschaften einnimmt.

Aufgrund der hohen Aktualität wurden zur Analyse der Personalisierungseffekte kontrovers diskutierte Umweltthemen aus den Themenbereichen Energie, Verkehr und Landwirtschaft herangezogen. Das Ziel war, herauszufinden, wie stark die Personalisierung zu diesen Themen ausfällt und ob es signifikante inhaltliche Unterschiede zwischen den Ergebnissen gibt und die Meinungsvielfalt tatsächlich eingeschränkt wird. Die Ergebnisse dieser Arbeit rechtfertigen keinen Anlass, dass dies eintreffen könnte. Zwar gab es im Vergleich zwischen anonym und personalisiert erhobenen Ergebnissen durchaus Unterschiede im Ranking, signifikante inhaltliche Unterschiede traten aber nicht auf.

Das Studiendesign für diese Masterarbeit wurde so angepasst, dass es sich für das zugrundeliegende Forschungsvorhaben umsetzen lässt. So wurden Praktiken aus der Relevanzbewertung im Information Retrieval adaptiert, indem Internetdokumente auf ihren Inhalt hin bewertet wurden und nicht auf ihre Relevanz. Durch eine entsprechende Kennzahl, die auch die Rangposition eines Dokuments berücksichtigt, konnte somit ein Wert ermittelt werden, der zum einen Inhalt und zum anderen die Position des Dokuments in der Ergebnisliste beachtet.

Keywords

Personalisierung, Suchmaschine, Google, Suchmaschinen-Verzerrungen, Filterblase, Umwelt, Experiment, Search Engine Bias, Internet

Inhaltsverzeichnis

Abstract	III
Inhaltsverzeichnis	IV
Abbildungsverzeichnis	VI
Tabellenverzeichnis	VII
Glossar	VIII
1 Einleitung	1
2 Stand der Forschung	4
2.1 Search Engine Bias	4
2.2 Web-Personalisierung	12
2.3 Methodischer Ansatz zur Untersuchung von Personalisierung	16
3 Kontrovers diskutierte Umweltthemen	20
3.1 Auswahl der Themen	20
3.2 Voruntersuchung zu den Themen	24
4 Forschungsfragen und Hypothesen	26
5 Methodik	27
5.1 Experiment: Methodische Grundlagen	27
5.1.1 Zielsetzung und Merkmale eines Experiments	27
5.1.2 Validität in Experimenten: interne und externe Validität	29
5.1.3 Experimentelle Designs: Within-Subjects-Design und Between-Subjects-Design	30
5.2 Studiendesign zur praktischen Durchführung	31
5.2.1 Ziel und Aufbau des Experiments	32
5.2.2 Veränderung der UV: Manipulation des Suchverhaltens	33
5.2.3 Kontrolle von Störvariablen	40
5.2.4 Praktische Durchführung des Experiments	43
5.2.5 Inhaltliche Bewertung der Dokumente durch Juroren	46
6 Auswertung der Daten	53
6.1 Personalisierung unabhängig vom Inhalt	53
6.1.1 Kennzahlen	53
6.1.2 Deskriptive Statistik	54

6.1.3	Zusammenhang zwischen gesetzten Signalen im Profil und Personalisierungseffekten	62
6.1.3.1	Analyseverfahren: Korrelationsanalyse nach Spearman	62
6.1.3.2	Ergebnisse der Hypothesenprüfung	63
6.2	Inhaltliche Bewertung der Dokumente	65
6.2.1	Kennzahlen	65
6.2.2	Deskriptive Statistik	67
6.2.3	Überprüfung der anonymen und personalisierten Suchergebnisse auf inhaltliche Unterschiede	73
6.2.3.1	Analyseverfahren: einfaktorielle Varianzanalyse	73
6.2.3.2	Ergebnisse der Hypothesenprüfung	75
6.2.4	Auswertung der Quellen	76
7	Diskussion	80
7.1	Diskussion der Ergebnisse	80
7.2	Diskussion der Methode	85
7.3	Grenzen der Studie und Ausblick auf künftige Forschung	86
8	Fazit	88
9	Literaturverzeichnis	89
Anhang 1: Informationsdossier		A
Eidesstattliche Erklärung		J

Abbildungsverzeichnis

Abbildung 1: Verlauf der Jaccard-Koeffizienten Account 1	57
Abbildung 2: Verlauf der Jaccard-Koeffizienten Account 2	57
Abbildung 3: Verlauf der Jaccard-Koeffizienten Account 3	58
Abbildung 4: Verlauf der Jaccard-Koeffizienten Account 4	58
Abbildung 5: Verlauf der Jaccard-Koeffizienten Account 5	59
Abbildung 6: Verlauf der Jaccard-Koeffizienten Account 6	59
Abbildung 7: Streudiagramm zur Korrelationsanalyse nach Spearman	65
Abbildung 8: Aufteilung der Quellen nach Akteursgruppen- Energie	78
Abbildung 9: Aufteilung der Quellen nach Akteursgruppen- Landwirtschaft.....	79
Abbildung 10: Aufteilung der Quellen nach Akteursgruppen- Verkehr.....	79
Abbildung 11: Webseite des Europäischen Instituts für Klima & Energie	84

Tabellenverzeichnis

Tabelle 1: Typologie potenzieller Auswirkungen von Search Engine Bias (Jürgens et al., 2014, S. 109) ...	17
Tabelle 2: Beispielhafte Darstellung eines Experimentaldesigns (Jürgens et al. 2015, S. 264).....	19
Tabelle 3: Bundestags- und Landtagswahlen in Deutschland (2016-2018)	21
Tabelle 4: Identifizierte Umweltthemen aus den Wahl-O-Mat-Durchläufen	22
Tabelle 5: Charakteristiken von Informationsbedarfen im Vergleich (Frants et al., 1997, S. 38)	34
Tabelle 6: Suchanfragen-Sets - Sektor Energie	36
Tabelle 7: Suchanfragen-Sets - Sektor Verkehr	38
Tabelle 8: Suchanfragen-Sets - Sektor Landwirtschaft	39
Tabelle 9: Instruktionen für die Juroren zur Bewertung der Dokumente.....	48
Tabelle 10: Personalisierungseffekte Sektor Energie	54
Tabelle 11: Personalisierungseffekte Sektor Verkehr	55
Tabelle 12: Personalisierungseffekte Sektor Landwirtschaft.....	56
Tabelle 13: Personalisierungseffekte Top-Positionen - Sektor Energie	60
Tabelle 14: Personalisierungseffekte Top-Positionen - Sektor Verkehr	61
Tabelle 15: Personalisierungseffekte Top-Positionen - Sektor Landwirtschaft	62
Tabelle 16: Ergebnisse Korrelationsanalysen nach Spearman.....	64
Tabelle 17: Häufigkeiten von Dokumenten auf den Top-Positionen - Energie.....	68
Tabelle 18: Häufigkeiten von Dokumenten auf den Top-Positionen - Verkehr	68
Tabelle 19: Häufigkeiten von Dokumenten auf den Top-Positionen - Landwirtschaft.....	69
Tabelle 20: Durchschnittliche Bewertungen der Dokumente zu den Suchanfragen	70
Tabelle 21: Durchschnittliche Bewertung aller SERPs zu den Suchanfragen	71
Tabelle 22: Bewertungs-Scores der SERPs Account 1 - Experimentalbedingung Energie	71
Tabelle 23: Bewertungs-Scores der SERPs Account 2 - Experimentalbedingung Energie	71
Tabelle 24: Bewertungs-Scores der SERPs - Kontrollbedingung Energie	72
Tabelle 25: Bewertungs-Scores der SERPs Account 3 - Experimentalbedingung Verkehr.....	72
Tabelle 26: Bewertungs-Scores Account 4 - Experimentalbedingung Verkehr	72
Tabelle 27: Bewertungs-Scores der SERPs - Kontrollbedingung Verkehr	72
Tabelle 28: Bewertungs-Scores der SERPs Account 5 - Experimentalbedingung Landwirtschaft	72
Tabelle 29: Bewertungs-Scores der SERPs Account 6 - Experimentalbedingung Landwirtschaft	73
Tabelle 30: Bewertungs-Scores der SERPs - Kontrollbedingung Landwirtschaft	73
Tabelle 31: Die am häufigsten vorkommenden Quellen der Studie	77
Tabelle 32: Häufigkeitstabelle zu den verschiedenen Akteuren.....	78

Glossar

AGORA VERKEHRSWENDE

Thinktank zur Umsetzung der Verkehrswende

AGRAVIS RAIFFEISEN

Großes deutsches Unternehmen im Agrarhandel

ALBERT-SCHWEITZER-STIFTUNG

Tierschutzorganisation, die sich insbesondere für die Abschaffung der industriellen Nutztierhaltung einsetzt und für eine vegane Lebensweise wirbt

ALNATURA

Marke des gleichnamigen Unternehmens, das auf den Vertrieb von Bio-Lebensmitteln spezialisiert ist

BASF

Deutscher Chemiekonzern; Hersteller von Insektiziden, Herbiziden und Fungiziden;aktiv in der Biotechnologie und in der Gentechnik

BAYER

Großes deutsches Unternehmen mit Fokus auf die Pharmaindustrie und die chemische Industrie; Pflanzenschutzsparte der Bayer AG: siehe Bayer Crop Science

BAYER CROP SCIENCE

Pflanzenschutzsparte der Bayer AG; tätig in Saatgutentwicklung, Pflanzenschutz und Schädlingsbekämpfung

BAYWA

Großer deutscher Handelskonzern mit Schwerpunkt auf den Agrarsektor

BI 58

Mittel zur Bekämpfung von Schädlingen; nicht für den biologischen Landbau zugelassen

BIO COMPANY

Kette von Bio-Supermärkten mit Sitz in Berlin

BISCAYA

Markenname eines Insektizids der Firma Bayer; nicht für den biologischen Landbau zugelassen

BP

Großes Mineralölunternehmen

BUNDESVERBAND ERNEUERBARE ENERGIE

Interessensverband der Branche für erneuerbare Energien in Deutschland

BUNDESVERBAND LANDSCHAFTSSCHUTZ

Initiative gegen den Bau von Windkraftanlagen

CALYPSO

Markenname eines Insektizids der Firma; nicht erlaubt im ökologischen Landbau

CNG

Compressed Natural Gas – komprimiertes Erdgas

DEBRIV

Deutscher Braunkohlen-Industrie-Verein – Branchenverband der deutschen Braunkohlenwirtschaft

DEUTSCHER BAUERNVERBAND

Größter Interessensverband für Landwirte in Deutschland; steht in der Kritik, nur für die Interessen großer Agrarunternehmen einzutreten und nicht für kleine Landwirte

DEUTSCHES ATOMFORUM

Interessensverband von Unternehmen und Organisationen aus der Kernenergiebranche; nach Zusammenschluss mit dem Wirtschaftsverband Kernbrennstoff-Kreislauf im Mai 2019 als Kerntechnik Deutschland e. V. eingetragen

ENDE GELÄNDE

Initiative, die sich für den sofortigen Ausstieg aus der Kohleverstromung einsetzt, insbesondere aktiv an den Demonstrationen gegen den Energiekonzern RWE im Hambacher Forst; vom Verfassungsschutz als linksextremistisch beeinflusst eingestuft

ENTEKA

Energieversorger, der Strom ausschließlich aus erneuerbaren Energien bezieht

ENYWAY

Internet-Plattform, auf der Bürger ihren selbst produzierten Ökostrom anbieten können; Strom kann vom Erzeuger direkt bezogen werden

EURO 5 / EURO 6

Abgasnormen für Fahrzeuge, die Grenzen für den Ausstoß von Luftschadstoffen festlegen

E-WALD

Anbieter von Elektromobilität; zu den Unternehmensbereichen gehören Vermietung und Carsharing von Elektroautos sowie der Ausbau der Ladeinfrastruktur

EXXON MOBIL

Großer Mineralölkonzern

FORUM MODERNE LANDWIRTSCHAFT

Vereinigung diverser Organisationen aus der Agrarbranche (u. A. Bayer Crop Science, BASF, Deutscher Bauernverband, Syngenta Agro GmbH)

GEGENWIND

Initiative gegen den Bau von Windkraftanlagen

GEMÜSEKISTE

Online erwerbliche Zusammenstellungen von Gemüse aus biologischem Anbau

GLYPHOSAT

Weltweit das am meisten verbreitete Herbizid in der Landwirtschaft; Hauptwirkstoff im Produkt "Roundup" von Bayer-Monsanto; für den ökologischen Landbau nicht zugelassen

GRÜNWELT

Ökostrom-Marke des Energieanbieters Stromio

GREENPEACE ENERGY

Energieversorger, der Strom ausschließlich aus erneuerbaren Energien bezieht

HAMBI / HAMBACHER FORST

Wald in Nordrhein-Westfalen, der seit Jahrzehnten vom Energiekonzern RWE zum Abbau von Braunkohle gerodet wird; Hauptschauplatz von Demonstrationen der Anti-Kohle-Bewegung in der Auseinandersetzung mit RWE

KLIMA ALLIANZ

Bündnis diverser zivilgesellschaftlicher Organisationen, die sich für mehr Klimaschutz engagieren

MONSANTO

Unternehmen aus der Agrarbranche, 2018 übernommen von der Bayer AG; aktiv in Biotechnologie und Gentechnik; Vertreiber des umstrittenen Herbizids der Marke „Roundup“ mit dem enthaltenem Wirkstoff Glyphosat

MWV

Mineralölwirtschaftsverband; Verband, der politische und wirtschaftliche Interessen der Mineralölindustrie in Deutschland vertritt

NEEMÖL

Aus dem Niembaum gewonnenes pflanzliches Öl, das als biologisches Insektizid verwendet werden kann; für den ökologischen Landbau zugelassen

NEULAND

Marke des Vereins „Neuland-Verein für tiergerechte und umweltschonende Nutztierhaltung“, die Fleischprodukte aus artgerechter Tierhaltung zertifiziert

NORDZUCKER AGRIPORTAL

Digitales Portal des Zuckerherstellers Nordzucker, das sich an Landwirte richtet

NUCLEAR PRIDE FEST

Veranstaltung der Pro-Kernkraft-Bewegung

NUKLEARIA

Interessensverein, der sich für den Wiedereinstieg in die Kernenergie in Deutschland einsetzt

PERMACLEAN

Markenname eines Pflanzenschutzmittels der Firma Bayer; nicht erlaubt im ökologischen Landbau

PROVIEH

Tierschutzverein, der sich insbesondere gegen die industrielle Nutztierhaltung einsetzt

RETTET DIE BIENEN

Volksbegehren „Artenvielfalt & Naturschönheit in Bayern“, bekannt geworden unter „Rettet die Bienen“; Volksbegehren in Bayern, das Gesetzesänderungen im Naturschutzgesetz anstrebt, um die Artenvielfalt in Flora und Fauna zu erhalten

RING DEUTSCHER BERGINGENIEURE

Berufsständische Vereinigung zur Wahrung von Interessen im Bergbau (in diesem Kontext: Interessensvertreter u. a. für die Kohleindustrie)

ROUNDUP

Markenname eines Herbizids der Firma Monsanto mit dem umstrittenen Wirkstoff Glyphosat; für den ökologischen Landbau nicht zugelassen

ROYAL DUTCH SHELL

Großes Mineralöl- und Erdgasunternehmen

SYNGENTA

Großes Unternehmen in der Agrarbranche; gehört zu den marktführenden Unternehmen in den Bereichen chemische Pflanzenschutzmittel und kommerzielles Saatgut, aktiv in der Biotechnologie; Sitz in Basel, Schweiz

TESLA

US-amerikanisches Unternehmen, spezialisiert auf die Produktion von Elektroautos

TRANSPORT AND ENVIRONMENT

Europaweite Dachorganisation verschiedener NGOs, die eine nachhaltige Verkehrswende anstreben; Mitglied u. a. der Verkehrsclub Deutschland

VERBINDUNGSSTELLE LANDWIRTSCHAFT UND INDUSTRIE

Verein, der die Interessen zwischen Landwirtschaft und Industrie zu wahren versucht

VERKEHRSCLUB DEUTSCHLAND

Verkehrsclub, der eine umweltfreundliche und nachhaltig gestaltete Verkehrswende anstrebt

WINDWAHN

Initiative gegen den Bau von Windkraftanlagen

WIR HABEN ES SATT

Bewegung, die sich gegen die derzeitige Praxis in Landwirtschaft und Agrarindustrie richtet und sich positiv für eine Agrarwende und gegen die Massentierhaltung ausspricht

1 Einleitung

Im Jahr 2011 sprach der Internetaktivist Eli Pariser im Rahmen eines TED-Talks über das personalisierte Internet. Zu Beginn seines Vortrags benannte er sogleich die aus seiner Sicht auftretende Kernproblematik der Personalisierung und stellte sie anhand der Suchmaschine Google dar:

„If I search for something, and you search for something, even right now at the very same time, we may get very different search results. Even if you're logged out, one engineer told me, there are 57 signals that Google looks at -- everything from what kind of computer you're on to what kind of browser you're using to where you're located -- that it uses to personally tailor your query results. Think about it for a second: there is no standard Google anymore. And you know, the funny thing about this is that it's hard to see. You can't see how different your search results are from anyone else's“ (Pariser, 2011b, 01:51-02:39).

Pariser hebt hierbei die Problematik hervor, dass Internetnutzern auf Basis ihres Verhaltens im Netz, anhand ihrer Präferenzen und ihrer im Internet hinterlassenen Fußabdrücke personalisiert Inhalte zugespielt werden, und zwar ohne, dass die Nutzer davon in Kenntnis gesetzt werden. Betrachtet man nur die Suchmaschine Google, sind auch hierbei potenzielle Gefahren erkennbar. Suchmaschinen gelten als der beliebteste Dienst im Internet und besitzen eine hohe Bedeutung im Wissenserwerb der Bevölkerung (Lewandowski, 2018, S. 3). Darüber hinaus genießt die Anordnung der Ergebnisse hohes Vertrauen bei den Nutzern (Pan, Hembrooke, Joachims, Lorigo, Gay, & Granka, 2007). In Deutschland hat sich insbesondere die Suchmaschine des Unternehmens Google etabliert. Hierzulande besitzt Google einen Marktanteil von fast 95 Prozent (StatCounter, 2019). Das in Google gesetzte Vertrauen zum einen, aber auch die Tatsache, dass ein Großteil der Nutzer nur die am höchsten platzierten Ergebnisse betrachtet (Petrescu, 2014), verdeutlichen die potenziellen Auswirkungen, die durch die algorithmische Selektion von Informationen entstehen können. In seinem Buch über die Machtpotenziale des Google-Konzerns schreibt Röhle (2010, S. 12) bereits in der Einleitung:

„Die Dimensionen dieses Problems werden dann besonders deutlich, wenn es sich um kontroverse politische Inhalte handelt: Wer es hier schafft, seine Inhalte in die vordersten Plätze zu hieven, hat einen wesentlichen Schritt getan, um sich als Autorität für das Thema der Suchanfrage zu etablieren und die Meinungsbildung der Nutzer entsprechend zu prägen. Verschärft wird diese Problematik durch die zunehmende Verwendung von Suchmaschinen in Bildung, Forschung und Journalismus, mithin den zentralen Bereichen der gesellschaftlichen Wissensvermittlung.“

Suchmaschinen unterliegen gewissen Verzerrungen in der Anordnung ihrer Ergebnisse, da eine rein nach objektiven Gesichtspunkten angeordnete Reihenfolge aller potenziell relevanten Dokumente zu einer Suchanfrage nicht umzusetzen ist (Lewandowski, 2018, S. 173). Dabei muss bedacht werden, dass in der Suchmaschinenwirtschaft unterschiedliche Akteure mit verschiedenen Interessen auftreten: die Betreiber der Suchmaschine, die Anbieter der Inhalte, die Branche der Suchmaschinenoptimierung und schließlich die Nutzerschaft (Röhle, 2010, S. 80). Dieses mitunter komplexe Zusammenspiel der Akteure kann somit zur Entstehung von sogenannten Suchmaschinen-Verzerrungen, auch Search Engine Bias, führen (Lewandowski, 2018, S. 292ff.).

Die vorliegende Masterarbeit verfolgt das Ziel, durch eine angepasste Methode personalisierte Suchergebnisse zu erhalten und diese zu analysieren. Dadurch soll in Hinblick auf Suchmaschinen-Verzerrungen ein besseres Verständnis der Auswirkungen von Personalisierungstechniken gewonnen werden. Zunächst erfolgte eine Untersuchung dazu, nach welchem Zeitraum und auf welchen Positionen mit den ersten Effekten der Personalisierung zu rechnen ist. Darüber hinaus wurden Treffer auf den für die Nutzer relevanten Positionen inhaltlich bewertet, um mögliche inhaltliche Abweichungen im Vergleich zur anonymen Suche zu identifizieren. Wie bereits aufgeführt, sind mögliche Auswirkungen insbesondere im Kontext kontroverser Inhalte ernst zu nehmen. Die Untersuchung der personalisierten Ergebnisse fand daher zu kontrovers diskutierten Umweltthemen in den Themenfeldern Energie, Verkehr und Landwirtschaft statt. Die Auswahl der Themen begründet sich mit der hohen Aktualität im politischen und gesellschaftlichen Diskurs. Auch die bisher überschaubare Anzahl an Forschungsarbeiten zu den Auswirkungen der Personalisierung in Hinblick auf

kontroverse Themen war ein Anlass zur Wahl dieses Themas. So soll die Studie nicht nur exemplarisch das Ausmaß der Personalisierung analysieren, sondern auch den Zugang zu diesen personalisierten Ergebnissen im Studiendesign beinhalten.

Die Arbeit gliedert sich in mehrere Teile. Im Anschluss an die Einleitung folgt eine Darstellung des Forschungsstandes. Darin werden zunächst Suchmaschinen-Verzerrungen im Detail definiert, ehe das Zustandekommen dieser Verzerrungen im Zusammenspiel der verschiedenen Akteursverbände angerissen wird. Danach folgt eine Definition von Personalisierung im Zusammenhang mit Suchsystemen und ein Blick in die Forschungsliteratur in dieses Bereichs: Ausgehend von Forschungsarbeiten in der Informatik Mitte der 2000er-Jahre – vorrangig, um das Potenzial der Personalisierungstechniken zu erheben - wird danach ebenso auf die denkbaren Auswirkungen auf Seiten der Nutzer sowie auf methodische Zugänge zur Analyse personalisierter Suchergebnisse eingegangen. Anschließend folgen eine kurze Darstellung und Begründung der Auswahl der Umweltthemen. Aufbauend auf den Forschungsstand werden im nächsten Abschnitt die aus der Literatur abgeleiteten Forschungsfragen und Hypothesen genannt. In einem weiteren Abschnitt erfolgt im Detail die Beschreibung der in dieser Arbeit angewendeten Methodik und deren Implikation in das Studiendesign. Die letzten beiden Abschnitte befassen sich mit den Ergebnissen der Studie: Zunächst werden die Ergebnisse präsentiert, danach interpretiert. Im Rahmen der Diskussion folgt eine kritische Auseinandersetzung mit dem Forschungsdesign, ehe abschließend die potenziellen Auswirkungen von Personalisierung in den Zusammenhang mit Suchmaschinen-Verzerrungen gesetzt werden. Schlussendlich werden die Grenzen der Studie und weitere mögliche Forschungsvorhaben benannt, die im Rahmen dieser Masterarbeit nicht umgesetzt wurden, im Bereich der Thematik jedoch denkbar sind.

2 Stand der Forschung

Wie in der Einleitung bereits aufgeführt, besitzen Web-Suchmaschinen als Werkzeug zur Informationsbeschaffung einen hohen Stellenwert und nehmen deshalb eine wichtige Rolle für den Wissenserwerb einer Gesellschaft ein (Lewandowski, 2018, S. 289ff.). Die vorliegende Arbeit beschäftigt sich mit sogenannten Suchmaschinen-Verzerrungen, weswegen dieser Begriff zunächst definiert wird. Anschließend erfolgt eine Unterscheidung zwischen verzerrten, neutralen und fairen Suchergebnissen, ehe genauer auf die Forschungsliteratur zu dieser Thematik eingegangen wird. In einem weiteren Abschnitt werden abschließend wissenschaftliche Beiträge zu den Personalisierungstechniken sowie möglichen Auswirkungen dargestellt.

2.1 Search Engine Bias

Zum besseren Verständnis potenzieller Auswirkungen von Suchmaschinen-Verzerrungen ist der Definition kurz die Funktionsweise von Suchmaschinen vorangestellt. Mittels eines Algorithmus (dem sogenannten Crawler) werden systematisch Webseiten im Internet aufgesucht. Diese Webseiten bilden den Bestand (den Index) ab, den Google mit eingehenden Suchanfragen abgleicht und aus dem letztlich die Rangfolge (das Ranking) für jede einzelne dieser Suchanfragen entsteht (Weber, 2011). Hierbei wird ersichtlich, dass Google nicht sämtliche im Internet verfügbare Dokumente abdeckt, was schließlich zu verzerrten Ergebnissen führt. Lewandowski (2017, S. 7) definiert Suchmaschinen-Verzerrungen folgendermaßen:

"Search engine bias is the tendency of a search engine to prefer certain results through the assumptions inherent in its algorithms."

Um den Begriff besser einordnen zu können, ist er von fairen und neutralen Ergebnissen abzugrenzen. Verzerrte Suchergebnisse finden sich in jeder Suchmaschine wieder, da jedem Algorithmus menschliche Annahmen zugrunde liegen. Diese Annahmen spiegeln sich in der Anordnung der Ergebnisse wider. Lewandowski (2018, S. 292) spricht bei Suchmaschinen-Verzerrungen von den Unterschieden „zwischen einer idealtypischen Ergebnismenge und -reihung und der tatsächlichen Ergebnismenge und -reihung“. Eine idealtypische Ergebnismenge würde die Ansammlung aller potenziell relevanten

Dokumente bedeuten, die idealtypische Ergebnisreihung ein ausschließlich nach objektiven Kriterien angelegtes Ranking (ebd., S. 292).

Faire Suchergebnisse zeichnen sich dadurch aus, dass jedes Internet-Dokument in gleicher Art und Weise von der Suchmaschine bewertet wird. Jedes Dokument im Internet hat demnach potenziell die gleiche Chance von der Suchmaschine gefunden und bewertet zu werden und das ohne menschliche Eingriffe in die Entscheidungen des Algorithmus hinsichtlich Crawling, Indexierung und Ranking (Lewandowski, 2017, S. 7). Neutrale Suchergebnisse hingegen würden nicht nur eine Gleichbehandlung für jedes Dokument im Internet bedeuten, sondern eben diese Gleichbehandlung aller Dokumente durch die Suchmaschine tatsächlich voraussetzen. Sie wäre also im Crawling, in der Indexierung und in der Reihung der Ergebnisse verankert. Wie Lewandowski (2018, S. 301) jedoch anmerkt, können dadurch keine unverzerrten Ergebnislisten entstehen, „da jede Entscheidung über ein Ranking von Suchergebnissen schon eine Verzerrung hervorruft.“

Die Ursachen von Suchmaschinen-Verzerrungen hat Weber (2011) untersucht und er teilte sie in drei Kategorien ein. So können Verzerrungen ihren Ursprung in der Implementierung in die Algorithmen haben, von den Inhaltenanbietern ausgehen oder in der Nutzung der Suchmaschinen entstehen. Weber nennt in der Implementierung der Suchalgorithmen verschiedene Aspekte, die zu verzerrten Ergebnissen führen können. So entscheiden Suchmaschinenbetreiber selbst, welche Webseiten sie in den Index ihrer Suchsysteme aufnehmen und somit beispielsweise extremistische Inhalte ausschließen. Die Möglichkeiten, die die Betreiber in der Gestaltung ihrer Algorithmen haben, bringen ein hohes Potenzial an gezielter Beeinflussung mit sich. Doch auch im Verhalten der Inhaltenanbieter finden Manipulationen statt, das Ranking zu beeinflussen. Dabei wird oftmals versucht, Webseiten so zu gestalten, dass sie dem Suchalgorithmus angepasst werden. Hierbei ist auch von Suchmaschinenoptimierung die Rede. Ziel ist die Platzierung der eigenen Inhalte auf den vorderen Positionen im Ranking. Weber (2011, S. 280) stellt jedoch heraus, dass es keine „klare Grenze zwischen legitimer Optimierung und missbräuchlicher Gestaltung“ gebe. Der dritte Ursprung von Verzerrungen kann im Nutzerverhalten liegen, da die Anordnung von Suchergebnissen von der Qualität der Recherche abhängen kann. Betrachtet man die genannten Ursprünge von Verzerrungen, ist das mitunter komplexe Zusammenspiel der verschiedenen Akteursverbände auf dem Suchmaschinenmarkt ersichtlich. Röhle (2010) identifizierte derer vier an der Zahl: Suchmaschinenbetreiber, Inhaltenanbieter, Suchmaschinenoptimierer und Nutzer.

Die Auswahl der nachfolgenden Beiträge aus der Literatur richten sich grob nach den von Weber identifizierten Kategorien, bei denen bei genauerer Betrachtung Parallelen zu den Akteursverbänden von Röhle erkennbar sind. Zunächst erfolgt die Vorstellung von Studien, die sich mit Suchmaschinen-Verzerrungen im Bereich der Implementierung in die Algorithmen beschäftigen. Bereits Mitte der 1990er-Jahre wurden Verzerrungen in Computersystemen umfassend durch Friedman und Nissenbaum (1996) erforscht. Sie haben Computersysteme aus 17 verschiedenen Fachgebieten (u. a. Bankwesen, Bildungswesen, Computerwissenschaft) untersucht. Dabei identifizierten sie drei Kategorien von Verzerrungen: bereits vorhandene, technische und emergente. Bereits bestehende Verzerrungen haben ihre Wurzeln in sozialen Institutionen, Praktiken oder persönlichen Einstellungen, die bereits vor der Erstellung des Computersystems existieren und somit explizit und bewusst oder implizit und unbewusst durch die Entwickler in das System gelangen können. Technischem Bias hingegen liegen technische Einschränkungen zugrunde. Dies kann beispielsweise die verwendete Hard- oder Software betreffen. Emergente Verzerrungen entstehen im Kontext der Nutzung mit echten Nutzern bei neu auftretenden Phänomenen. Beispiele dafür sind neu entstandenes Wissen oder sich ändernde gesellschaftliche Werte, die sich in der Nutzung von Computersystemen herausbilden können. Auch Brey (1998) erkennt die Abbildung sozialer Strukturen einer Gesellschaft in Computersystemen. Dabei betont er die Schwierigkeit der Identifikation von Verzerrungen, da sie oftmals komplex konstruiert sind. Verzerrungen können systematisch verankert sein und gewisse Webseiten – zumeist populäre Seiten, die bereits relativ viel Traffic erhalten – gegenüber anderen hervorheben, wie Introna und Nissenbaum (2000) ausführen. Sie plädieren für eine größere Vielfalt in der Politik von Suchmaschinen, damit auch weniger populäre Inhalte Platz in diesem System finden und heben darüber hinaus ihre Zweifel an einer neutralen Suchmaschine hervor. Die Auswirkungen davon können weitreichend sein, wenn man bedenkt, dass die Bedeutung von Web-Suchmaschinen auf die Gesellschaft mit derer traditioneller Massenmedien verglichen wird, da sie eine Gatekeeper-Funktion einnehmen würden (Diaz, 2008). Anbieter von Suchmaschinen sehen sich immer wieder der Kritik ausgesetzt, den Suchprozess nicht transparent zu machen. Wieso Webseiten in den Index von Google aufgenommen bzw. ausgeschlossen werden, ist ebenso wenig im Detail bekannt wie die Erstellung des Rankings (Hinman, 2005), weswegen beispielsweise Introna und Nissenbaum (2000) eine Offenlegung der Algorithmen fordern. Wie eingangs angedeutet, können Verzerrungen bereits in der

Entwicklung von Algorithmen entstehen. Eine Studie von van Couvering (2007) beschäftigt sich daher näher mit den Ursachen. In Tiefeninterviews mit Suchmaschinenentwicklern wurde deutlich, dass die Entwicklung der Algorithmen hauptsächlich von marktwirtschaftlichen und technischen Aspekten geprägt ist, während Elemente wie Fairness und Repräsentativität nicht im Vordergrund stehen.

Neben der Ursachenforschung in der Entwicklung der Algorithmen finden sich in der Literatur zahlreiche weitere Studien, die sich mit Verzerrungen auseinandersetzen, die in der Implementierung der Suchmaschine verankert sind. Diese liegen zum einen in der Erfassung der Webseiten, zum anderen in der Darstellung der Ergebnisse. Mowshowitz und Kawaguchi (2002a, 2002b) verglichen mehrere Suchmaschinen miteinander, indem sie deren Grad an Verzerrung in verschiedenen Themenbereichen maßen. Das Ergebnis zeigte Unterschiede zwischen den von den einzelnen Suchmaschinen angebotenen Webseiten und dass auch das Maß der Verzerrung variiert. Die Autoren vermuten in den Beobachtungen einen Einfluss der unterschiedlichen Indexgrößen der Suchmaschinen. Vaughan und Zhang (2007) beschäftigten sich damit, in welchem Umfang Suchmaschinen Inhalte in verschiedenen Ländern abdecken. So ist der Anteil US-amerikanischer Webseiten in der Aufnahme durch Suchmaschinen höher als der Anteil von Webseiten aus China, Singapur oder Taiwan, was jedoch generell an einer größeren Anzahl von US-Seiten im Internet liegen könnte. Es ist jedoch auch nicht auszuschließen, dass Webseiten in gewissen Ländern gesperrt werden, wie eine Bestandsaufnahme zu geblockten Internetseiten in zehn verschiedenen Ländern deutlich macht (Mathrani & Alipour, 2010). In manchen Staaten werden Websites aktiv durch die Suchmaschine aus dem Index im jeweiligen Land ausgeschlossen. In Deutschland betrifft dies zum Beispiel Inhalte, die den Holocaust leugnen (Lewandowski, 2018, S. 44). Obwohl die genannten Studien ein gewisses Alter haben und die Funktionsweise von Suchmaschinen, wie sie Webseiten crawlen und indexieren, im Laufe der Zeit durch Entwickler vermutlich angepasst wurde, so ist der Definition zufolge davon auszugehen, dass Suchmaschinen-Verzerrungen darin auch heute implementiert sind. Suchmaschinen sind nicht in der Lage, alle Inhalte aus dem Web komplett abzudecken. Das macht schon das Beispiel mit den gesperrten Inhalten deutlich.

Einige andere Studien hatten Suchmaschinen-Verzerrungen in der Ergebnisdarstellung als Schwerpunkt. Gerhart (2004) ging der Frage nach, ob und in welchem Ausmaß Kontroversen bei bestimmten Themen in Suchmaschinenergebnissen versteckt bzw. sichtbar sind. Dabei kam sie zu dem Ergebnis, die Suchmaschinentechnologie neige zur

Darstellung der "Sonnenseite" eines Themas. Als vorrangigen Grund sieht sie in dieser Beobachtung den starken Einfluss der Linkstruktur im Ranking-Algorithmus einer Suchmaschine. An dieser Stelle sei erwähnt, dass Suchmaschinen sogenannte linktopologische Verfahren anwenden. Im Internet sind Webseiten mit Links verbunden, jeder Link hat dabei einen eigenen Wert (Lewandowski, 2018, S. 104). Ein bekannter Bestandteil der Suchtechnologie von Google ist der PageRank (Brin & Page, 1998). Dabei soll die Linkstruktur des Internets dafür eingesetzt werden, präzise Suchergebnisse zu erzielen. Der PageRank ist eine Kennzahl, die die Qualität jeder Webseite bewerten soll. Nach Lewandowski (2018, S. 106) lässt sich für ein Dokument ein hoher PageRank erreichen, wenn viele Links anderer Dokumente darauf verweisen oder wenn Dokumente mit einem hohen PageRank auf eine Webseite verlinken. Solche Zusammenhänge wurden auch im Bereich von Search Engine Bias untersucht. Die Vermutung, vorrangig der PageRank könne zu einem sogenannten Rich-get-richer-Effekt führen, in dem bereits populäre Webseiten weiter an Popularität gewinnen, wurde jedoch durch Fortunato, Flammini, Menczer und Vespignani (2006) widerlegt. Sie sehen das Ausbleiben dieses Effekts in der Heterogenität der Nutzerinteressen.

Mager (2012) führt aus, dass ein Unternehmen wie Google zwar natürlich kommerzielle Interessen verfolge, die Funktionsweise der Algorithmen in einem kapitalistischen System jedoch nicht nur in der Hand der Entwickler liege. Besonders Webseitenbetreiber und die Nutzer können, bewusst oder unbewusst, mit ihrem Verhalten im Internet die aktuell auf dem Kapitalismus beruhende Ideologie von Suchmaschinenanbietern festigen. Konkret meint sie damit die Festigung der Strategie Googles, verschiedene Parteien, also Nutzerschaft und Webseitenbetreiber, zur Kooperation zu bewegen. Bei den Nutzern drückt sich dies durch eine Nutzung der durch Google bereitgestellten Services aus. Beispiele dafür sind Google Mail, Google Maps, Google Earth, das Betriebssystem Android oder ehemals das soziale Netzwerk Google+. Die erhobenen Nutzerdaten können von Google wiederum für Werbezwecke genutzt werden. Wie Mager weiter ausführt, kooperieren auch die Inhaltenanbieter mit dem Unternehmen, indem sie beispielsweise die von Google bereitgestellten Webmaster-Tools nutzen, um die eigene Internetseite möglichst sichtbar in den Suchergebnissen zu präsentieren. In der Literatur finden sich zahlreiche Beiträge, die sich mit der als Suchmaschinenoptimierung bekannten Praxis beschäftigen. Diese richten sich zum Teil gezielt an die Betreiber von Webseiten, um sie bei der Optimierung ihrer Inhalte zu unterstützen. Die Bandbreite dabei ist groß und geht von umfassenden Handbüchern

bis hin zu knappen Tutorials, die im Kern jedoch gemeinsam haben, die grundlegenden Techniken von Suchmaschinenoptimierung der Zielgruppe näherzubringen (Beiträge dazu u. a. in Erlhofer, 2018; Killoran, 2013; Gudivada, Rao, & Paris, 2015). Welche Effekte die Optimierungsmaßnahmen mit sich bringen, ist bisher wenig erforscht. Der Einfluss der Suchmaschinenoptimierung (engl. Search Engine Optimization, abgekürzt SEO) auf die Ergebnisse wird jedoch vor allem im kommerziellen Bereich als hoch eingeschätzt, doch finden SEO-Techniken auch im Journalismus oder in NGOs Anwendung (Lewandowski, 2014, S. 237). Doch nicht nur die Hervorhebung der eigenen Inhalte durch Webseitenbetreiber muss im Zusammenhang mit Search Engine Bias genannt werden. In manchen Fällen liegt auf Seiten der Betreiber kein Interesse vor, in den Index einer Suchmaschine aufgenommen zu werden. So können durch die Verwendung von entsprechenden Meta-Elementen im Quelltext der Website Inhalte gezielt "versteckt" werden. Eine Praxis, die von Weber (2011, S. 270) als "Inselbildung" bezeichnet wird.

Suchmaschinen-Verzerrungen entstehen auch auf der Nutzerseite. Dies hängt eng mit dem Rechercheverhalten zusammen (Weber, 2011, S. 281). Hierbei lohnt sich auch ein Exkurs in die Psychologie. Exemplarisch sollen die Phänomene des sogenannten Confirmation Bias und des Myside Bias dargestellt werden, um Verzerrungen auf der Nutzerseite abzubilden. Mercier (2016) zweifelte ein Bestehen des Confirmation Bias an – der Tatsache, dass Menschen nach Argumenten suchen, die sämtliche ihrer Gedanken bestätigen würden. Mercier erkennt nach diesem Muster eine unlogische Folgerung, da damit auch Gedanken gestützt werden müssten, mit denen man nicht einverstanden ist. Er definierte daher den Begriff des Myside Bias. Menschen würden also dazu neigen, nicht generell all ihre Gedanken zu verteidigen und bestätigt sehen zu wollen, sondern nur die Positionen, die sie vertreten. Dies lässt sich auch im Online-Suchverhalten beobachten, wie eine Studie von White (2013) zeigt. Die Ergebnisse stützten die These, dass Nutzer bei der Informationssuche dazu tendieren, ihre bestehenden Überzeugungen zu bestätigen.

Im Folgenden ist eine Auswahl empirischer Untersuchungen zu Suchmaschinen-Verzerrungen im Zusammenhang mit gesellschaftlich relevanten Themen dargestellt. Ein Beispiel dafür ist eine Studie zum potenziellen Einfluss des Rankings von Suchergebnissen auf unentschlossene Wähler vor Wahlen. In Experimenten wurden unentschlossenen Wählern manipulierte SERPs (Search Engine Result Page) zu zwei zur Wahl stehenden Politikern vorgelegt. Die Ergebnislisten wurden so angefertigt, dass die darauf befindlichen Dokumente entweder den einen oder den anderen Politiker positiv

hervorheben oder neutral einzuordnen sind. Die Autoren konnten einen signifikanten Einfluss feststellen. So kann ein verzerrtes Ranking die Stimmpräferenzen von unentschlossenen Wählern erheblich beeinflussen (zu 20 Prozent oder mehr), einen Effekt, den die Autoren der Studie "Search Engine Manipulation Effect" (Epstein & Robertson, 2015) nennen. Diesen Effekt aufgreifend haben sich Kulshrestha et al. (2019) ebenfalls mit politischen Wahlen beschäftigt. Ihre Forschung hatte das Ziel, Verzerrungen in der Websuche messbar zu machen. Da Teile dieser Methodik für die vorliegende Masterarbeit relevant sind, werden sie nachfolgend etwas detaillierter beschrieben. Da die Untersuchung von Kulshrestha et al. (2019) im Vorfeld der US-Wahlen 2016 stattfand, wurden Kandidaten gesucht und geprüft, ob die Ergebnisse neutral, positiv oder negativ zur Anfrage zu bewerten sind. Die Forschungsgruppe entwickelte einen Rahmen, in dem sie das Ausmaß von Verzerrungen in drei verschiedenen Stufen des Suchprozesses bewertete. Zunächst ist hierbei der Input-Bias zu nennen. Darunter sind Verzerrungen zu verstehen, die zwischen dem Abschicken einer Suchanfrage und dem Erstellen des Rankings im Suchsystem entstehen. Das System gleicht seinen Datenbestand mit der Suchanfrage des Nutzers ab und stellt dazu relevante Ergebnisse zur Verfügung. In diesem Schritt werden also bereits potenziell relevante Dokumente herausgefiltert. Der Output Bias hingegen repräsentiert die kumulierten Verzerrungen aller Dokumente, die der Nutzer nach Eingabe seiner Suchanfrage in einer Rangliste erhält. Dabei wurde in der Bewertung auch der Ranglistenplatz berücksichtigt und nicht nur der Inhalt der Dokumente. Zwischen Input- und Output-Bias können weitere Verzerrungen entstehen, nämlich in der Anordnung der Ergebnisse. Hierbei ist von Ranking-Bias die Rede. Da die Algorithmen vieler Suchsysteme jedoch nicht vollumfänglich bekannt sind, wurde zur Berechnung des Ranking-Bias die Differenz zwischen Output- und Input-Bias genommen. Um das Ausmaß der Verzerrungen innerhalb eines Zeitraumes messen zu können, wurden die eben beschriebenen Metriken zu einzelnen Suchergebnissen ermittelt und ein Mittelwert über diese Zeit errechnet, der sogenannte Time-averaged-Bias. Mit Bezug auf die Google-Websuche (in der Studie wurde zusätzlich auch zu Verzerrungen auf Twitter geforscht) sammelte das Team innerhalb einer Woche mehrere hundert Suchergebnisse, die die Vorwahlen zur US-Präsidentenwahl 2016 thematisieren und untersuchten sie auf Verzerrungen im Output. Die Ergebnisse beinhalten eine signifikante Anzahl an Links von Nachrichtenseiten, in denen bereits Verzerrungen durch die Forschung dokumentiert wurden. Weitere populäre Ergebnisse waren Wikipedia-

Artikel oder die persönlichen Webseiten und Social-Media-Seiten der Kandidaten. Während Wikipedia-Einträge neutrale Bewertungen erhielten, wurden die persönlichen Dokumente der Kandidaten als thematisch einseitig betrachtet. Im Schnitt wies ein Großteil der Ergebnisse Verzerrungen auf, was jedoch an der Auswahl der Suchanfragen liegen könnte. Größtenteils wurde nämlich lediglich nach den Namen der Kandidaten gesucht. Im Vergleich mit den Twitter-Ergebnissen waren die Google-Dokumente insgesamt stabiler. Google neigt im Gegensatz zu Twitter auch mehr zur Anzeige von Inhalten, die durch die Kandidaten kontrolliert werden können, wie die persönlichen Social-Media-Profile oder Homepages, wie Epstein und Robertson (2015) ausführen.

Auch Abseits politischer Themen konnte in einer Reihe von Untersuchungen Verzerrungen nachgewiesen werden, wie beispielsweise eine Analyse von White und Horvitz (2009) aus dem Gesundheitssektor zeigt. Sie untersuchten Suchmaschinenergebnisse zur Suche nach eher gefahrlosen Krankheitssymptomen. Dabei zeigten sich im Ranking der Ergebnisse eine Bevorzugung von Webseiten, die die Symptome als Ursachen für schwere Krankheiten interpretieren. Ballatore (2015) analysierte die Darstellung von 15 Verschwörungstheorien in den Ergebnissen der Suchmaschinen von Google und Bing. Nutzer, die nach Verschwörungstheorien suchen, bekommen bevorzugt Webseiten angezeigt, auf denen die Theorien bestätigt werden. Auch im Kontext von Geschlechter- und Rassestereotypen wurden Suchergebnisse auf Verzerrungen hin untersucht. Noble (2018) untersuchte exemplarisch die Darstellung schwarzer Frauen in den Suchergebnissen und wies darauf hin, dass bereits bestehende Vorurteile Angehörigen dieser Personengruppen von Google wiedergegeben werden. Die Autorin führte diese Tatsache zurück auf unzureichende Kenntnisse der Entwickler hinsichtlich der Geschichte der Kultur von schwarzen Menschen in den USA. Sie fordert die Einbeziehung solcher Elemente in die Ausbildung künftiger Entwickler. Noble (2018, S. 66) wirft die Frage auf, die im Zuge der Bewertung von Search Engine Bias grundlegend erscheint: "If Google software engineers are not responsible for the design of their algorithms then who is?"

In der hier vorgestellten Literatur zu Suchmaschinen-Verzerrungen wurde der Versuch unternommen, mögliche Auswirkungen der Verzerrungen auf eine Gesellschaft abzubilden. Aufgrund der aktuellen Funktionsweise von Suchmaschinen sind viele dieser Studien als teilweise veraltet einzuschätzen, da ein Unternehmen wie Google vermutlich ständig Anpassungen an seinem Suchalgorithmus vornimmt. Dennoch liefern die ausgewählten Studien Erkenntnisse in der Diskussion darüber, dass Suchmaschinen-

Verzerrungen in verschiedenen Bereichen des Suchprozesses entstehen können und wie sie zu bewerten sind. Dabei stellt sich auch immer wieder die Frage nach der Verantwortung, die ein Unternehmen wie Google dadurch gegenüber seinen Nutzern hat. Im Laufe der Zeit haben grundlegende Veränderungen an den Algorithmen das Zustandekommen von Ergebnislisten verändert, darunter Techniken wie die Web-Personalisierung. Während Suchergebnisse zu einem Suchbegriff vor der Entwicklung hin zur Personalisierung für viele Nutzer gleich waren, können sich personalisierte Ergebnisse potenziell von Nutzer zu Nutzer unterscheiden. Suchmaschinen-Verzerrungen haben ihren Ursprung also nicht mehr primär in der Technik von Suchmaschinen, sondern können somit auch zu einem stärkeren Ausmaß im Verhalten der Nutzer zu finden sein (Jürgens, Stark, & Magin, 2014). Bevor im nächsten Abschnitt näher auf die Web-Personalisierung eingegangen wird, sind zusammenfassend die Problemfelder genannt, die Tavani (2012) mit dem Begriff des Search Engine Bias assoziiert:

- Die Suchmaschinentechnologie ist nicht neutral, stattdessen werden bestimmte Werte gegenüber anderen priorisiert
- Systematische Bevorzugung einiger Webseiten gegenüber anderen in der Anordnung der Ergebnislisten
- Suchalgorithmen verwenden in der Erstellung ihrer Ergebnislisten keine objektiven Kriterien

2.2 Web-Personalisierung

Ganz allgemein gesprochen ist Personalisierung "das Anpassen eines Objekts an die Bedürfnisse eines Subjekts" (Riemer & Brüggemann, 2009, S. 153). Im Kontext von Suchmaschinen hat die Personalisierung von Suchergebnissen das Ziel, "dass jedem Nutzer auf Basis seines Profils unterschiedliche Ergebnisse bzw. eine unterschiedliche Ergebnisreihung angezeigt werden" (Lewandowski, 2018, S. 123). Verdeutlicht wird dies durch einen simplen Vergleich (ebd., S. 124): Beim Suchwort "Jaguar" werden einem Nutzer, der in der Vergangenheit öfter nach Tieren gesucht hat, eher Ergebnisse zum Tier angezeigt, wohingegen ein Nutzer, der häufiger zum Thema Auto gesucht hat, bevorzugt Ergebnisse zur Automarke Jaguar erhält. Forschung im Bereich der Personalisierung im Internet fand zunächst überwiegend in der Informatik statt, die

implizit erstellte Nutzerprofile als eine Lösung betrachtet, dem einzelnen Nutzer relevantere Ergebnisse zu liefern und die Websuche somit zu verbessern (Gauch, Speretta, Chandramouli, & Micarelli, 2007; Teevan, Dumais, & Horvitz, 2005). Implizite Daten der Nutzer werden aus dem Surfverhalten gezogen, während explizite Daten diejenigen Angaben sind, die aktiv vom Nutzer an das System übermittelt wurden oder auf Dritte (etwa über Social-Media-Verbindungen) zurückzuführen sind (Lewandowski, 2018, S. 123). Personalisierungstechniken im Internet werden nicht nur im Bereich der Websuche angewendet. Weitere Einsatzbereiche sind beispielsweise Empfehlungssysteme oder das Ausspielen von Werbeanzeigen (Hector Garcia-Molina, Koutrika, & Parameswaran, 2011). Aufgrund der Thematik der Masterarbeit soll der Schwerpunkt der hier vorgestellten Literatur auf Personalisierungstechniken in Web-Suchmaschinen liegen.

White (2016, S. 267) unterscheidet dabei zwischen Personalisierung und Kontextualisierung. Personalisierte Ergebnisse in einem Suchsystem sind individuell auf den Nutzer zugeschnitten. Dazu wird das langfristige Nutzerverhalten hinzugezogen, um Interessen des Nutzers zu identifizieren. Merkmale, die in das Profil miteinfließen, sind beispielsweise die abgeschickten Suchanfragen und die angeklickten Dokumente über einen bestimmten Zeitraum. Liegen zu wenige Daten über den Nutzer vor, können ähnliche Daten anderer Nutzer hinzugezogen werden. Kurzfristige Verhaltensweisen sind in der Regel nicht ausreichend, um ein individuelles Nutzerprofil zu erstellen. Unter der Kontextualisierung versteht White (2016, S. 268) die Anpassung an die Suchsituation. Kontextbasierte Daten können die Klickdaten und die Suchanfragen ergänzen und ein Suchsystem verbessern. Typische kontextbasierte Daten sind beispielsweise der Standort oder Tageszeit. Google führte die personalisierte Suche bereits im Jahr 2005 ein und machte dies über Blogeinträge bekannt. In einem der Beiträge heißt es: "With the launch of Personalized Search, you can use that search history you've been building to get better results. You probably won't notice much difference at first, but as your search history grows, your personalized results will gradually improve" (Kamvar, 2005).

Bis hierher ist erkennbar, dass Verzerrungen in Suchmaschinen überwiegend kritisch betrachtet werden. Goldman (2008) nimmt dagegen eine konträre Position dazu ein. Er vergleicht Suchmaschinen mit klassischen Medienunternehmen und sieht es daher als unvermeidbar an, dass Verzerrungen entstehen, da Suchmaschinen Einfluss auf das

Zustandekommen ihrer Ergebnisse nehmen. Eine völlig neutrale Aufarbeitung der Inhalte durch die Suchmaschine sei nicht möglich. Verzerrungen entstünden dabei insbesondere durch die Verwendung von Algorithmen, die jedem Suchenden die gleichen Ergebnisse liefern. Die personalisierte Suche sieht Goldman daher als Möglichkeit, Verzerrungen zu minimieren, da verschiedene Nutzer zur gleichen Anfrage verschiedene Ergebnisse erhalten. Argumente zur Anzweiflung dieser Auffassung liefert Hinman (2005), der zunächst das Spannungsverhältnis zwischen gesellschaftlicher und wirtschaftlicher Verantwortung seitens der Suchmaschinenbetreiber hervorhebt. Durch Personalisierungstechniken sei es Unternehmen wie Google möglich, noch genauere Nutzerinformationen für werbetreibende Kunden zu nutzen.

Für Aufsehen hinsichtlich der Personalisierungstechniken großer Internet-Unternehmen wie etwa Google sorgte die Kritik von (Pariser, 2011a). In seinem Buch verwendete er den Begriff "Filter Bubble", also Filterblase, die er als ein persönliches Informationsuniversum im Internet versteht. Er nennt drei mit einer Filterblase einhergehende Dynamiken. Da wäre zum einen, dass sich Nutzer in ihrer eigenen Filterblase befinden, sie also andere Inhalte zu sehen bekommen als andere Nutzer. Als zweiten Punkt spricht er die Intransparenz der Algorithmen an, die dem Personalisierungsprozess zugrunde liegen. Sind sich Nutzer nicht im Klaren, sich in einer Filterblase zu befinden, könnten sie unter Umständen der Meinung sein, die erhaltenen Informationen seien ausschließlich unvoreingenommen und objektiv. Dass Internetnutzer in der Regel kaum entscheiden können, in welchem Umfang ihnen personalisierte Inhalte angezeigt werden, ist nach Pariser die dritte Problematik der Personalisierung.

Feuz, Fuller und Stalder (2011) prüften zunächst, ob Personalisierung überhaupt in Googles Suchergebnissen auftritt und bauten ihr Studiendesign nach den von Google getroffenen Aussagen zur Personalisierung auf (Beispiel: "Personalization is subtle"). Sie interpretierten diese Aussagen nach eigenem Ermessen und entwickelten eine Methodik zur Messbarmachung. Nach der Suche nach denselben Begriffen sollten also mit unterschiedlichem Suchverlauf für unterschiedliche Profile verschiedene Ergebnisse auftauchen. Zunächst wurden drei neue Google-Konten angelegt, denen mittels ausgewählter Suchbegriffe ein Interessensprofil gegeben wurde. Die Suchanfragen stammen aus den Indexen von Büchern dreier Philosophen. Nach Eingabe des Begriffs wurde zufällig ein Link angeklickt. Insgesamt wurden sieben solcher Durchläufe durchgeführt. Nach jedem Durchlauf wurde die Entwicklung der Personalisierung

beobachtet. Dabei wurde zusätzlich zu den Recherchen im eingeloggten Account eine anonyme Suche mit identischer Suchanfrage durchgeführt. Die Forscher fanden heraus, dass Personalisierung auftritt, eine inhaltliche Auswertung von profitierenden Webseiten blieb jedoch aus. Die Effekte waren größtenteils in den hinteren Ergebnissen festzustellen, die der durchschnittliche Nutzer nicht mehr betrachtet, und traten früher auf, als vermutet: Nach etwa 450 Suchanfragen konnten erste Effekte beobachtet werden. Diese drückten sich so aus, dass etwa jede zehnte Suchanfrage zu diesem Zeitpunkt in der personalisierten Suche ein anderes Ergebnis lieferte als die anonyme Suche.

Hannák et al. (2015) konzentrierten sich auf die Unterschiede der personalisierten Suche bei Google und Bing. Die Autoren entwickelten eine Methodik, um den Umfang von Personalisierungseffekten in Suchmaschinenergebnissen zu messen. Sie führten eine experimentelle Studie durch, in der die Ergebnisse der verschiedenen Suchmaschinen analysiert wurden. Die Themen und Suchanfragen stammten aus dem Portal Google Zeitgeist (heute Google Trends). Die Studienteilnehmer waren zunächst reale Nutzer, die ihren privaten Web-Browser so konfigurierten, dass das Forscherteam die ersten zehn Ergebnisse zu den vorgegebenen Suchanfragen beobachten konnte. Parallel dazu wurde nach jeder Suche des realen Nutzers ein Programm gestartet, das die gleiche Abfrage anonym durchführte. Anschließend erfolgte ein Vergleich der ausgegebenen Ergebnisse. Hannák et al. berechneten eine Wahrscheinlichkeit von Abweichungen personalisierter Ergebnisse gegenüber anonymen bei der Google-Nutzung von knapp 12 Prozent. Weiterhin wurde festgestellt, dass sich im Ranking weiter oben befindliche Ergebnisse weniger personalisiert sind als Ergebnisse weiter unten. Aufbauend darauf wurden neue Profile angelegt, um nach den Ursachen der Effekte zu forschen, wobei sich die aktive Anmeldung im Profil und der aktuelle Nutzerstandort am stärksten auf das Zustandekommen personalisierter Ergebnisse auswirkten.

Salehi, Du und Ashman (2015) untersuchten Personalisierungseffekte in der akademischen Websuche. Dabei haben sie zunächst 120 Studierende der University of South Australia nach ihren Präferenzen hinsichtlich der Literaturrecherche im Internet befragt. Die Suchmaschine von Google stellte sich in diesem Kontext als beliebtestes Recherche-Tool heraus. Anschließend wurden die Unterschiede von anonymen und personalisierten Ergebnissen auf der ersten SERP untersucht. Die Autoren gingen dabei experimentell vor, indem sie zunächst vier verschiedene Test-Accounts erstellten, die jeweils einem akademischen Themenspektrum zuzuordnen sind (Education, IT,

Business, Health Sciences). Ein fünfter erstellter Account fungierte als Kontroll-Account. Im nächsten Schritt wurden die Test-Accounts mit impliziten Daten angereichert. Für jedes der vier Konten wurden 30 themenbezogene, informationsorientierte Suchanfragen ausgewählt, sodass insgesamt ein Set an 120 Suchbegriffen entstand. Nach Abschicken der Suchanfragen wurde nach eigenem Ermessen entschieden, auf welchen Link auf der ersten SERP geklickt wurde. Es wurde also hingenommen, dass sich im Unterbewusstsein subjektiv getroffene Entscheidungen im Klickverhalten widerspiegeln können. Jede einzelne Suchanfrage wurde einmal anonym und einmal personalisiert abgeschickt. Im Durchschnitt unterschieden sich knapp die Hälfte der anonymen Ergebnisse von den personalisierten. Außerdem war die Mehrheit der Links auf beiden SERPs auf unterschiedlichen Positionen, was sich in dieser Studie jedoch oftmals auf den Standort zurückführen ließ.

Zu den Befürchtungen zum anfangs erwähnten Phänomen der Filterblase (Pariser, 2011a), die die Personalisierung im Web mit sich bringen soll, gab es ebenfalls Untersuchungen. (Robertson, Jiang, Joseph, Friedland, Lazer, & Wilson, 2018) untersuchten den Algorithmus in der personalisierten Google-Suche auf politische Tendenzen. Die Studie wurde in den USA durchgeführt und zeigt, dass Ergebnisse am unteren Rand der SERP in politischer Hinsicht linksgerichteter waren als Ergebnisse am oberen Rand. Hinweise auf die Hypothese der "Filterblase" konnten jedoch kaum festgestellt werden. Zuiderveen Borgesius, Trilling, Möller, Bodo, Vreese und Helmberger (2016) untersuchten Effekte der Personalisierung auf Nachrichtenseiten. Die Autoren sahen kaum berechtigten Anlass, wegen möglicher Filterblasen besorgt zu sein. Sollten personalisierte Nachrichten jedoch zur primären Informationsquelle für Individuen werden, könnte dies problematisch für eine Gesellschaft werden, behaupten die Forscher mit Hinweis auf eine Prüfung empirischer Studien im Bereich der Medienwirkung.

2.3 Methodischer Ansatz zur Untersuchung von Personalisierung

Wie aus Abschnitt 2.2 hervorgeht, ist die herkömmliche Methode zur Erzeugung personalisierter Suchergebnisse die des Experiments. Es treten dabei Herausforderungen auf, die es zu berücksichtigen gilt und die in Beiträgen von Jürgens et al. (2014; 2015), die sich explizit um den methodischen Zugang zu personalisierten

Ergebnissen im Internet drehen, diskutiert werden. Tabelle 1 zeigt die potenziellen Auswirkungen von Search Engine Bias nochmals auf. Die durch das Forscherteam erstellte Typologie greift die von Friedmann und Nissenbaum (1996) identifizierten Kategorien von Verzerrungen (bereits existierend, technisch, emergent) auf und bildet zusätzlich den methodischen Zugang ab.

Tabelle 1: Typologie potenzieller Auswirkungen von Search Engine Bias (Jürgens et al., 2014, S. 109)

	Exogener Bias (bereits existierend)	Endogener Bias (technisch)	Bias aus Interaktion (emergent)
Index	Seiten aufgrund von Zensur nicht auffindbar	Seiten nicht auffindbar, weil die Suchmaschine sie nicht indexiert	-
Sortierung (Ranking)	Repliziert bestehende Machtstrukturen	Verstärkt bestehende Machtstrukturen („rich-get-richer“-Effekt) <hr/> Bevorzugt Angebote des Suchmaschinenbetreibers	Fokussiert Ergebnisse auf Interessensprofil der Nutzer (Einengung der Ergebnisvielfalt); selbstverstärkend
Publikum	-	-	Fragmentierung der Nutzer (keine geteilte Wahrnehmung) <hr/> Digital Divide (Kompetente Nutzer profitieren stärker)

Methodik	-	-	Tests von Auswahl- und Sortierungsstrategien der Suchmaschinen nicht mehr repräsentativ
----------	---	---	---

Wie aus der Tabelle hervorgeht, wird der wissenschaftliche Zugang zu Daten der Suchmaschine durch Personalisierungsmaßnahmen erschwert. Die damit einhergehende Problematik für die Suchmaschinenforschung beschreiben Jürgens et al. (2014, S. 111) wie folgt: Der „Teil der Ergebnisse, der aufgrund von individuellen Nutzerprofilen angezeigt wird, ist der empirischen Erhebung nicht zugänglich. Alle bislang vorliegenden Erkenntnisse [...] verlieren damit aus heutiger Sicht potenziell ihre Repräsentativität.“

Doch nicht nur der Zugang zu den Analyseeinheiten ist mit Herausforderungen verbunden, sie ergeben sich auch in der Bewertung der Validität der erhobenen Daten (Jürgens et al., 2015, S. 257ff.): Während bei der Sicherung der internen Validität aufgrund fehlender Kenntnisse in der Funktionsweise der Personalisierungsalgorithmen insbesondere technische Faktoren störend wirken können, erweist sich das nicht bekannte Nutzerverhalten bei der Sicherung der externen Validität als schwierig einzuschätzen. Hierbei ist insbesondere nicht klar, wie Nutzer in den aufgerufenen Webseiten interagieren, weswegen die Simulation von realistischem Nutzerverhalten im Studiendesign empfohlen wird. In einer Fallstudie (Jürgens et al., 2014) schlägt das Forscherteam eine experimentelle Vorgehensweise vor, um personalisierte Suchergebnisse zu erhalten. In ihrer Studie gingen sie näher auf die Web-Suchmaschinen ein. Dabei wurden zwei neue Google-Konten erstellt und mithilfe eines Computerprogramms automatisch Interessensprofile aufgebaut. Die Begriffe, also das Stimulusmaterial, mit denen die Profile angereichert wurden, wurden aus Wikipedia extrahiert und sind thematisch einem Homonym mit zwei Bedeutungen zugeordnet. Das Stimulusmaterial für ein Profil war so gewählt, dass es sich klar von dem des anderen Profils unterscheidet. Auch ein zufälliges Klickverhalten gehörte zum Ablauf. Nach Eingabe des Suchbegriffs folgte ein Klick auf einen durch das Programm zufällig ausgewähltes Dokument. Die Personalisierung an sich wird nach Erstellung der

Interessenprofile mit der Recherche nach denselben Begriffen festgestellt. Die Bereiche, in denen Personalisierung in der Websuche erforscht wurde, variierten. Weiterhin heben die Autoren die Notwendigkeit hervor, Themen zu finden, die mindestens zwei klar voneinander unterscheidbare Sichtweisen auf das entsprechende Thema zulassen. Die Studie konnte zwar Personalisierungseffekte nachweisen, jedoch selten auf der ersten Ergebnisseite. Außerdem stellten sie fest, dass einige der personalisierten Links nicht eindeutig einem Thema zuzuordnen waren. Hierbei erscheint die Auswahl der Themen problematisch, wie die Autoren in ihrer methodischen Kritik selbst einräumen. Die aus der Wikipedia extrahierten Homonyme stellten vermutlich ein zu klein gewähltes Themenspektrum ab und die Personalisierungseffekte deshalb gering waren. Bekanntere Themen, über die im Internet viele Informationen zu finden sind, könnten dagegen stärkere Personalisierungseffekte in der Websuche mit sich bringen (Jürgens et al., 2014).

Das vorgeschlagene experimentelle Vorgehen ist in einfacher Weise in Tabelle 2 abgebildet. Die Methodik der vorliegenden Arbeit stützt sich auf die Beiträge von Jürgens et al. (2014; 2015) und ist in Abschnitt 5 beschrieben.

Tabelle 2: Beispielhafte Darstellung eines Experimentaldesigns (Jürgens et al. 2015, S. 264)

Faktor: Bedeutungsebenen des Wortes „Stollen“	Treatment	Messung
<i>Experimentálnutzer 1</i>	<i>Sucht nach Begriffen rund um Weihnachten</i>	<i>Suche nach „Stollen“</i>
<i>Experimentálnutzer 2</i>	<i>Sucht nach Begriffen rund um Bergbau</i>	
Kontrollsuche	<i>Keine Suche</i>	

3 Kontrovers diskutierte Umweltthemen

3.1 Auswahl der Themen

Die Identifikation kontrovers diskutierter Umweltthemen erfolgte mithilfe des Wahl-O-Mat. Der Wahl-O-Mat ist ein von der Bundeszentrale für politische Bildung zur Verfügung gestelltes Online-Tool, das vor allem junge Wähler in der Entscheidungsfindung vor Wahlen unterstützen soll. Dabei werden vor Landtags- und Bundestagswahlen insgesamt 38 Thesen zu diversen politischen Themen präsentiert, die von den Anwendern mit „stimme zu“, „stimme nicht zu“, „neutral“ oder „These überspringen“ zu beantworten sind. Die zu den Wahlen zugelassenen Parteien werden im Vorlauf mit denselben Thesen konfrontiert und können ebenfalls ihre Meinung dazu abgeben. Am Ende eines Wahl-O-Mat-Durchlaufs sieht der Nutzer, inwieweit seine politischen Standpunkte mit denen der Parteien übereinstimmen (Bundeszentrale für politische Bildung, 2013). Die Thesen werden von jungen Wählern im Alter von 16 bis 26 Jahren, Politikwissenschaftlern, Pädagogen, Statistikern sowie Wissenschaftlern und Experten zu den verschiedenen Themen ausgewählt. Grundlage der Themen stellen die Programme der zur Wahl zugelassenen Parteien dar sowie deren programmatische Aussagen. Die Themengebiete variieren, im Kern behandeln sie folgende Schwerpunkte:

- Arbeit, Soziales, Integration, Flucht und Asyl
- Energie, Umwelt, Infrastruktur, Verkehr, Gesundheit und Verbraucherschutz
- Familie, Bildung, Kultur und Religion
- Finanzen, Steuern, Währung und Wirtschaft
- Inneres, Äußeres, Demokratie, Föderalismus und EU

Die 38 Thesen werden in einem Auswahlverfahren vor jeder Wahl diskutiert. Nach Angaben der Bundeszentrale für politische Bildung (2017) werden Thesen ausgewählt, die

- die wichtigsten Themen der Wahl aufgreifen,
- von den Parteien kontrovers beantwortet werden,
- die Unterscheidbarkeit der einzelnen Parteien gewährleisten und
- ein breites thematisches Spektrum abdecken.

Aufgrund der aktuellen Debatten in der Umweltpolitik, eignen sich die von der Expertengruppe ausgewählten Themen im Sektor Umwelt daher gut, um sie für das zugrundeliegende Forschungsinteresse dieser Arbeit zu untersuchen. Um den Zeitraum möglichst aktuell zu halten, und dennoch ein umfangreiches Set an Themen zu erstellen, wurden die Thesen verschiedener Landtagswahlen in Deutschland in den Jahren 2016, 2017 und 2018 sowie der Bundestagswahl 2017 untersucht.

Tabelle 3: Bundestags- und Landtagswahlen in Deutschland (2016-2018)

Jahr	Art der Wahl
2016	Landtagswahl Baden-Württemberg
2016	Landtagswahl Rheinland-Pfalz
2016	Landtagswahl Sachsen-Anhalt
2016	Abgeordnetenhauswahl Berlin
2017	Landtagswahl Saarland
2017	Landtagswahl Schleswig-Holstein
2017	Landtagswahl Nordrhein-Westfalen
2017	Bundestagswahl
2017	Landtagswahl Niedersachsen
2018	Landtagswahl Bayern
2018	Landtagswahl Hessen

Thesen, die eindeutig nur einer bestimmten Region zugeordnet werden konnten und somit nicht verallgemeinerbar sind, wurden aus der Liste gestrichen. So zum Beispiel eine These zur Landtagswahl in Rheinland-Pfalz 2016 zum Nationalpark Hunsrück-Hochwald. Folgende Umweltthemen wurden als überregional bedeutend identifiziert:

Tabelle 4: Identifizierte Umweltthemen aus den Wahl-O-Mat-Durchläufen

Wahl	Thema der These	Sektor
Landtagswahl Baden-Württemberg 2016	Flächen für Windkraftanlagen	Energiepolitik
	Ökologische Landwirtschaft	Agrarpolitik
	Ausbau des Radwegenetzes	Verkehrspolitik
	Radioaktive Abfälle	Energiepolitik
Landtagswahl Rheinland-Pfalz 2016	Ökologischer Weinanbau	Agrarpolitik
	Standorte für Windräder	Energiepolitik
	Verbot von Gentechnik	Agrarpolitik
Landtagswahl Sachsen-Anhalt 2016	Braunkohleabbau	Energiepolitik
	Gentechnisch veränderte Pflanzen	Agrarpolitik
	Windräder	Energiepolitik
	Bienenfreundliche Landwirtschaft	Agrarpolitik
Abgeordnetenhauswahl Berlin 2016	Elektroautos	Verkehrspolitik
Landtagswahl Saarland 2017	Windkraft	Energiepolitik
	Ökologische Landwirtschaft	Agrarpolitik
	Kohlekraftwerke	Energiepolitik
Landtagswahl Schleswig-Holstein 2017	Ökologische Landwirtschaft	Agrarpolitik
	Bienenfreundliche Landwirtschaft	Agrarpolitik

	Abstand von Windkraftanlagen	Energiepolitik
Landtagswahl Nordrhein-Westfalen 2017	Ausstieg aus dem Braunkohleabbau	Energiepolitik
	Ökologische Tierhaltung	Agrarpolitik
	Windkraftanlagen im Wald	Energiepolitik
Bundestagswahl 2017	Besteuerung PKW-Diesel	Verkehrspolitik
	Ausbau erneuerbarer Energien	Energiepolitik
	Tempolimit	Verkehrspolitik
	Ökologische Landwirtschaft	Agrarpolitik
	Begrenzung der Nutztierhaltung	Agrarpolitik
	Braunkohleabbau	Energiepolitik
Landtagswahl Bayern 2018	Dieselfahrverbote	Verkehrspolitik
	Tempolimit	Verkehrspolitik
	Mindestabstand von Windrädern	Energiepolitik
	Förderung der Öko-Landwirtschaft	Agrarpolitik
Landtagswahl Hessen 2018	Windkraftanlagen in Waldgebieten	Energiepolitik
	Keine Fahrverbote für Dieselfahrzeuge	Verkehrspolitik
	Ökologische Landwirtschaft	Agrarpolitik

3.2 Voruntersuchung zu den Themen

Um mehr Erkenntnisse zu Google-Ergebnissen hinsichtlich der ausgewählten Themen zu erhalten, wurde zusätzlich zur Literatur eine kurze Voruntersuchung durchgeführt. Die Ergebnisse daraus sollen in erster Linie bei der Formulierung der Forschungsfragen und Hypothesen unterstützen. Vor Beginn der eigentlichen Datenerhebung wurden jeweils die Quellen der ersten zehn Ergebnisse von vier Suchanfragen zu umweltpolitischen Themen gesammelt. Die Suche wurde anonym bei Google durchgeführt. Bei den Suchanfragen handelte es sich um andere als die, die später in der experimentellen Studie verwendet wurden. Es wurde darauf geachtet, möglichst neutrale Begriffe zu wählen und sie sollten sich aktuellen Themen zuordnen lassen. Dazu wurde in der aktuellen Online-Berichterstattung nach kontroversen Themen aus dem Umweltbereich gesucht und als passend empfundene Suchbegriffe extrahiert. Für den Sektor Energie wurde die Debatte um eine Fortführung von Kernkraftwerken in Deutschland aufgegriffen (Der Tagesspiegel, 2019), der sich daraus ableitende Suchbegriff lautete "kernkraft deutschland". Im Sektor Verkehr fiel die Wahl auf die Suchanfrage "365 euro ticket". Damit wurde die aktuelle Forderung von Teilen aus Politik und Gesellschaft nach einem 365 Euro kostenden Jahresticket für den öffentlichen Nahverkehr aufgefasst (Süddeutsche Zeitung, 2019). Im Sektor Landwirtschaft werden derzeit verstärkt Stimmen laut, die die von der Europäischen Union praktizierten Flächensubventionen kritisieren (Schnirch, 2019). Die dritte Suchanfrage in dieser Voruntersuchung lautete deswegen "flächensubventionen". Neben Begriffen zu jedem der drei Sektoren wurde themenübergreifend die kontrovers geführte Debatte um eine Einführung eines CO₂-Preises in Deutschland aufgegriffen (Spiegel Online, 2019). Die dazugehörige Suchanfrage lautete "co2 preis".

Zu diesen Suchanfragen wurden insgesamt 40 verschiedene Dokumente gesammelt. Die Dokumente lassen sich 31 verschiedenen Quellen zuordnen. Dabei kamen fünf Quellen mehr als einmal vor: wikipedia.org (5), tagesschau.de (3), zeit.de (2), spiegel.de (2) und waz.de (2).

Auffällig war, dass mehr als die Hälfte der Dokumente (25) von Medienakteuren stammen. Darunter wurden die Online-Auftritte klassischer Medienunternehmen aus Presse und Rundfunk gezählt, Inhalte aus dem Special-Interest-Journalismus, aber auch Webseiten wie Wikipedia, die sich mit der Nutzung von neuen Medien etabliert haben. Nach den Medienakteuren waren in dieser Stichprobe Inhalte von Interessensorganisationen am zweithäufigsten vertreten (7). Beispiele dafür sind die Gewerkschaft IG Metall, der

Verein Kernenergie Deutschland e. V. oder der Naturschutzverein BUND. Die dritthäufigste Akteursgruppe stellen politische Akteure dar (5), die aus Parteien, Regierungen, Ministerien oder Politikern bestehen. Weitere identifizierte Akteure waren aus der Wirtschaft Unternehmen (2) und aus der Wissenschaft Hochschulen und andere Forschungsinstitute (1). Die Einteilung in die verschiedenen Akteursgruppen erschien plausibel, darüber hinaus waren zusätzlich Privatpersonen als Akteure denkbar, die ihre Inhalte beispielsweise über Blogs zugänglich machen.

Der relativ hohe Anteil an Wikipedia-Artikeln und journalistischen Angeboten lassen verstärkt inhaltlich neutral einzuschätzende Dokumente bei Recherchen zu kontrovers diskutierten Umweltthemen vermuten. Diese Vermutung ist in erster Linie mit der journalistischen Sorgfaltspflicht zu begründen, die Nachrichtenberichterstattung neutral zu gestalten (Schwiesau & Ohler, 2016, S. 95). Interessensorganisationen, wirtschaftliche und politische Akteure dürften dagegen eine klare Meinung zu entsprechenden Themen haben, sodass in jedem Sektor ein Dissens hinsichtlich umweltpolitischer Maßnahmen zu erwarten ist.

4 Forschungsfragen und Hypothesen

Die bereits angeführten Erkenntnisse aus der Forschungsliteratur und den Beobachtungen aus der Voruntersuchung führten zu folgenden Forschungsfragen und Hypothesen:

F1: Gibt es bei der personalisierten Google-Recherche nach kontrovers diskutierten Umweltthemen einen Zusammenhang zwischen der Anzahl gesetzter impliziter Signale und der Anordnung der Ergebnisse?

H1: Je mehr implizite Daten in Form von Suchanfragen und anschließenden Ergebnis-klicks in das persönliche Profil einfließen, umso mehr Abweichungen gibt es im Vergleich zur anonymen Suche in der Reihung der Ergebnisse.

F2: Unterscheiden sich die anonymen Suchergebnisse von den personalisierten Suchergebnissen bei Recherchen nach kontrovers diskutierten Umweltthemen inhaltlich?

H2: Zwischen den anonymen Suchergebnissen und den personalisierten Suchergebnissen in mindestens einer der beiden Experimentalgruppen gibt es signifikante inhaltliche Unterschiede, die sich durch weniger neutrale Ergebnisse in den personalisierten Ergebnislisten auszeichnen.

Zwei weitere Forschungsfragen wurden aufgrund fehlender Erkenntnisse ohne die Aufstellung von Hypothesen formuliert:

F3: Wie wirkt sich die im Suchverhalten implizite persönliche Haltung zu einer Kontroverse bei einer Umkehr auf die gegenteilige Haltung in personalisierten Google-Recherchen zu Umweltthemen aus?

F4: Welche Quellen werden Google-Nutzern in der Recherche nach kontrovers diskutierten Umweltthemen angezeigt?

5 Methodik

Nachfolgender Abschnitt beschreibt die in der Studie angewandte Methode im Detail. Das Vorgehen in der Methodik ist dabei in zwei Blöcke geteilt. Im ersten Schritt ist es notwendig, personalisierte Google-Ergebnisse zu erhalten. Im zweiten Schritt werden diese Ergebnisse auf ihren Inhalt hin bewertet. Um die für das Vorhaben benötigten personalisierten Ergebnisse zu gewinnen, wurde experimentell vorgegangen. Zunächst wird zunächst also auf die methodischen Grundlagen des Experiments in der Theorie eingegangen, im Anschluss erfolgt die Beschreibung der praktischen Durchführung. Die Beschreibung der theoretischen Grundlagen beinhaltet die Zielsetzung und Merkmale eines Experiments, die Validitätsformen im Experiment – unterschieden wird zwischen interner und externer Validität – sowie die experimentellen Designs „Between-Subjects-Design“ und „Within-Subjects-Design“.

Die personalisierten Ergebnisse wurden, um sie auf ihren Inhalt hin bewerten zu lassen, Juroren vorgelegt. Inwieweit die Bewertung in das Studiendesign miteinflusst, ist am Ende des Abschnitts beschrieben.

5.1 Experiment: Methodische Grundlagen

5.1.1 Zielsetzung und Merkmale eines Experiments

Wie erwähnt, soll in dieser Arbeit ein experimentelles Vorgehen den Zugang zu personalisierten Google-Ergebnissen gewährleisten. Experimente werden durchgeführt, um eine Beziehung zwischen Ursache und Wirkung zwischen zwei Variablen zu untersuchen. Ziel von experimentellen Untersuchungen ist also, herauszufinden, „ob und wie ein bestimmter Sachverhalt [...] einen anderen Sachverhalt beeinflusst bzw. verändert“ (Koch, Peter, & Müller, 2019, S. 3). Dieser auch als Kausalität bezeichnete Zusammenhang zwischen den Variablen ist nach Sedlmeier und Renkewitz (2018, S. 133ff.) nach Erfüllung von Voraussetzungen gegeben: Kovariation, zeitliche Präzedenz und Ausschluss von Alternativerklärungen.

Die Voraussetzung der Kovariation zwischen zwei Variablen ist bei einem bestehenden Zusammenhang gegeben. Dieser Zusammenhang zeigt sich in der Theorie folgendermaßen: Änderungen in der Variablen, in der die Ursache für einen gewissen Sachverhalt vermutet wird, führen zu Änderungen in der Ausprägung der anderen Variable. Die Überprüfung der Kovariation in einem Experiment erfolgt über den

Vergleich von wenigstens zwei unterschiedlichen Ausprägungen der als Ursache vermuteten Variable. Charakteristisch für ein Experiment ist die Manipulation der Variable, die als Ursache vermutet wird. Diese wird auch als unabhängige Variable (UV) bezeichnet. Die Auswirkungen der unabhängigen Variablen zeigen sich demnach in der abhängigen Variable (AV). Sedlmeier und Renkewitz (2018, S. 134) weisen darauf hin, dass in Experimenten im Allgemeinen mindestens zwei Bedingungen in der UV herbeigeführt werden, um mögliche Auswirkungen davon in der AV zu messen. Dabei kann es situationsbedingt vorkommen, dass die UV anwesend bzw. abwesend ist. Bei einer anwesenden unabhängigen Variablen spricht man von einer Experimentalbedingung, bei einer abwesenden von einer Kontrollbedingung.

Die zweite Voraussetzung für ein Experiment ist die zeitliche Präzedenz, denn eine bestehende Kovariation alleine reicht nicht aus, um die Kausalrichtung zu bestimmen. Für einen gültigen Kausalschluss muss die Ursache also vor der Wirkung eintreten. Somit kann eine umgekehrte Wirkung zwischen den Variablen ausgeschlossen werden. Für die Durchführung eines Experiments ist deshalb die Manipulation der unabhängigen Variablen entscheidend. Die unterschiedlichen Ausprägungen der UV werden manipuliert, die AV wird im Anschluss in den unterschiedlichen Bedingungen gemessen.

Die dritte nach Sedlmeier und Renkewitz (2018, S. 133ff.) notwendige Bedingung für gültige Kausalschlüsse ist der Ausschluss von Alternativerklärungen. Besteht eine Kovariation zwischen zwei Variablen, könnte diese potenziell auch mit dem Einfluss von Drittvariablen erklärt werden. Zur Überprüfung der Kausalität ist deshalb ein Ausschluss dieser möglichen anderen Erklärungen erforderlich. Variablen, die in einem Experiment neben der eigentlichen UV Einfluss auf die Wirkung haben, werden auch Störvariablen genannt. Besteht eine Kovariation zwischen der unabhängigen Variablen und einer potenziellen Störvariable, so sind die Auswirkungen auf die AV unter Umständen mit Einflüssen der Störvariable zu erklären. Um solche Szenarien im Experiment ausschließen zu können, ist die Kontrolle über mögliche beeinflussende Drittvariablen ein weiteres Merkmal der Methode. Die Kontrolle kann dabei über zwei Techniken erfolgen. Zum einem über das Konstanthalten von Störvariablen. Dies bringt eine Gleichhaltung aller Bedingungen in allen Ausprägungen mit sich. Störvariablen können somit nicht mit der unabhängigen respektive der abhängigen Variable kovariieren, was demzufolge eine Einflussnahme ausschließt. Wird ein Auftreten der Störvariable komplett unterbunden, ist vom Eliminieren die Rede (Sedlmeier & Renkewitz, 2018, S. 144f.). Unter der Annahme, dass es nicht möglich ist, alle zu kontrollierenden

Störvariablen konstant zu halten, spielt die zweite Kontrolltechnik eine Rolle: das Balancieren. Sedlmeier und Renkewitz (2018, S. 137) weisen auf Experimente in der Psychologie hin, in denen Studienteilnehmer unterschiedliche Merkmale mitbringen. Diese Unterschiede zeigen sich beispielsweise in der Motivation oder in der Intelligenz der Probanden und können nicht konstant gehalten werden. Das Balancieren soll gewährleisten, dass Störvariablen in den unterschiedlichen Bedingungen des Experiments den gleichen durchschnittlichen Wert annehmen. Dies würde auch den gleichen Einfluss der Störvariablen auf die AV bedeuten. Dieses Ausbalancieren erfolgt in Experimenten üblicherweise mit der sogenannten Randomisierung, also der zufälligen Anordnung von im Experiment vorkommenden Elementen. Im Beispiel von Sedlmeier und Renkewitz würden die Probanden zufällig ausgewählt an den verschiedenen Ausprägungen der Studie teilnehmen, um die unterschiedlichen Merkmale – hier aufgeführt waren Motivation und Intelligenz - im Durchschnitt gleichzuhalten.

Die Methode des Experiments kann mit folgender Definition von Koch et al. (2019, S. 6) zusammengefasst werden: „Ein Experiment ist eine Untersuchungsanordnung zur Überprüfung von Kausalannahmen, bei der unter kontrollierten Bedingungen mindestens eine unabhängige Variable aktiv variiert und deren Einfluss auf eine oder mehrere abhängige Variablen gemessen wird“.

5.1.2 Validität in Experimenten: interne und externe Validität

Die Validität von experimentellen Befunden wird unterteilt in interne und externe Validität (Koch et al., 2019). Ein Experiment erweist sich als intern valide, wenn es einen eindeutigen Kausalschluss ermöglicht und gilt laut Sedlmeier und Renkewitz (2018, S. 138) als „das wichtigste Kriterium zur Beurteilung der Güte von Experimenten“. Untermauert wird diese Aussage damit, dass es das Ziel von Experimenten sei, Zusammenhänge zwischen Ursache und Wirkung zu beweisen und im Forschungsdesign von experimentellen Studien grundlegende Elemente danach ausgerichtet seien, eine hohe interne Validität zu erzielen: die Manipulation der unabhängigen Variable und die Kontrolle der Störvariablen. Eine hohe interne Validität kann also als Indikator für einen gültigen Kausalschluss gesehen werden. Eine vollumfängliche Kontrolle über Störvariablen würde die interne Validität beispielsweise erhöhen, da der Einfluss auf die AV nur auf die manipulierte UV zurückzuführen wäre (Koch et al., 2019). In der Praxis ist es jedoch so, dass eine hohe interne Validität nicht immer gewährleistet werden kann. Das hat den einfachen Grund, dass nicht immer alle Störvariablen bekannt sind und

deshalb nicht gänzlich eliminiert werden können (Sedlmeier & Renkewitz, 2018, S. 143). Gegenstück der internen Validität ist die externe Validität. Diese Form der Validität „ist dann hoch, wenn die Ergebnisse auch auf nicht untersuchte Personen und Situationen übertragen werden können“ (ebd., S. 150). Die externe Validität ist also ein Indikator für die Generalisierbarkeit der Ergebnisse. Auch hierbei wird wieder ein Vergleich mit Untersuchungen aus der Psychologie aufgeführt. In Laborstudien, in denen eine künstliche Situation hergestellt wird, verhalten sich Probanden aus unterschiedlichen Gründen oftmals anders als im Alltag, was in der Bewertung der externen Validität berücksichtigt werden muss. Auch eine niedrige Zahl an Studienteilnehmern würde sich negativ auf die externe Validität auswirken. Es gibt verschiedene Wege, die externe Validität zu erhöhen. Bei personengebundenen Untersuchungen sollten die Studienteilnehmer repräsentativ für die Population sein, auf die generalisiert werden soll, wie Sedlmeier und Renkewitz (2018, S. 153) anmerken. Dabei erwähnen sie jedoch auch, dass die Verallgemeinbarkeit von untersuchten Situationen schwieriger zu erreichen ist als von untersuchten Personen. Als weitere Möglichkeit nennen sie deshalb eine bestimmte Form des Experiments, das Feldexperiment. Feldexperimente werden in einer realen Umgebung und nicht im Labor durchgeführt und können somit ein Ansatz sein, die externe Validität zu erhöhen. Der Haken dabei ist, dass die Kontrolle der Störvariablen erschwert wird, was eine Herabsetzung der internen Validität mit sich bringt. Die dritte genannte Option zur Erhöhung der externen Validität liegt in der Replikation von Experimenten (ebd., 2019, S. 153f.). So können die Ergebnisse vieler gleich oder ähnlich durchgeführter experimentelle Studien im Gesamten betrachtet die externe Validität erhöhen. Wie hoch die interne und die externe Validität letztlich zu bewerten sind, ist Abwägungssache. Eine höhere interne Validität kann sich negativ auf die externe Validität auswirken, ebenso verhält es sich umgekehrt (Koch et al., 2019).

5.1.3 Experimentelle Designs: Within-Subjects-Design und Between-Subjects-Design

Bei experimentellen Studien ist in der Regel auf grundlegende Unterschiede im Forschungsdesign zu achten. Dabei wird zwischen Within-Subjects-Designs und Between-Subjects-Designs unterschieden, wobei es auch gemischte Ansätze gibt, die Elemente aus beiden Designs beinhalten. Entscheidet man sich für ein Between-Subjects-Design, wird jedes zu untersuchende Element, beispielsweise eine Person, genau einer experimentellen Bedingung innerhalb der unabhängigen Variable

zugeordnet (Sedlmeier & Renkewitz, 2018). Koch et al. (2019, S. 68) ergänzen, dass es im Between-Subjects-Design mindestens zwei Untersuchungsgruppen gibt: die Experimentalgruppe und die Kontrollgruppe. Wie der Name des Between-Subjects-Designs bereits verrät, findet die Manipulation der UV zwischen den Gruppen statt. Schließlich können die Ergebnisse aus beiden Gruppen miteinander verglichen werden. Die Idee dahinter ist folgende: Nach Messung der Ergebnisse in der Kontrollgruppe und den Experimentalgruppen können die beobachteten Unterschiede als Effekte der UV auf die AV betrachtet werden (ebd., 2019, S. 68f.).

Im Within-Subjects-Design verhält es sich anders. Dort werden alle Bedingungen einer UV durchlaufen und die Ausprägung der AV bei jedem zu untersuchendem Element gemessen (Sedlmeier & Renkewitz, 2018, S. 154f.). Insbesondere bei Experimenten mit Personengruppen bringt dieses Design Vorteile mit sich. So können situationsbedingt Studien mit insgesamt weniger Versuchsteilnehmern durchgeführt werden, da die Probanden jeweils an jeder Bedingung teilnehmen und nicht nur an einer. Auch die Kontrolle personengebundener Störvariablen wird vereinfacht (ebd., S. 157). Die Autoren zählen noch weitere Vorteile des Designs auf, beantworten jedoch auch die Frage, wieso ein Within-Subjects-Design dann nicht immer in experimentellen Studien zum Einsatz kommt: Für manche Forschungsvorhaben eignet es sich nicht. Als Beispiel werden Therapiemethoden aus der Psychologie aufgeführt, die nicht beide an ein und derselben Versuchsperson untersucht werden können (ebd., S. 160f.).

Auch sogenannte gemischte Designs findet man in der Praxis. Davon ist die Rede, wenn in einem Experiment mit mehreren unabhängigen Variablen die Manipulation sowohl zwischen (Between-Subjects-Design) als auch innerhalb (Within-Subjects-Design) der Untersuchungsgegenstände stattfindet (ebd., S. 171).

5.2 Studiendesign zur praktischen Durchführung

Dieser Abschnitt befasst sich mit den Vorbereitungen des Experiments sowie mit der Durchführung der Studie an sich. Konkret wird auf die in der Theorie erläuterten Handlungsschritte eingegangen. Zunächst wird der Aufbau der experimentellen Studie dargestellt, anschließend werden die Manipulation der unabhängigen Variablen und die Kontrolle über die Störvariablen beschrieben. Der Abschnitt schließt mit den Angaben zur praktischen Durchführung

5.2.1 Ziel und Aufbau des Experiments

Ziel des experimentellen Vorgehens war der Hervorbringung personalisierter Suchergebnissen der Suchmaschine Google. Wie bereits erwähnt, sollen in einem Experiment Zusammenhänge zwischen Ursache und Wirkung erforscht werden. Für das Vorhaben dieser Masterarbeit liegen die primär für die Personalisierung herangezogenen Handlungen in der Formulierung der Suchanfrage sowie in der Auswahl der Dokumente, die ein Suchsystem für die Anfrage liefert (White, 2016, S. 269). Die Ursache in der Entstehung personalisierter Ergebnisse wird also im Suchverhalten des Nutzers vermutet, weshalb dieses die unabhängige Variable in diesem Experiment darstellt und entsprechend manipuliert werden soll. Die Manipulation ist im nächsten Abschnitt genauer beschrieben, es sei jedoch vorweggenommen, dass themenspezifische Sets an Suchanfragen gebildet wurden und die Auswahl der Dokumente für diese Anfragen festgelegt wurde. Nachdem das Suchverhalten manipuliert wurde, liegt die abhängige Variable in diesem Experiment demnach in der personalisierten SERP.

Die Generierung personalisierter Suchergebnisse erfolgt im Prinzip häufig nach dem gleichen Muster, wie bereits aus den vorgestellten Studien in der Literatur ersichtlich ist. Es werden neue Accounts angelegt, die anschließend mit Suchanfragen aus einem gewissen Themenfeld angereichert werden. Dadurch entsteht ein durch bestimmte Interessen geprägtes Profil. Profile dieser Art lassen sich beliebig oft erstellen und miteinander vergleichen. Zum Vergleich der Ausprägung der Personalisierung ist jedoch ein und derselbe Suchbegriff von Nöten. Wird dieser Begriff bei Google abgefragt, können, je nach Grad der Personalisierung, SEPRs mit verschiedenen Anordnungen der Dokumente entstehen. Die verschiedenen SERPs können im Anschluss miteinander oder mit einer in der anonymen Suche entstandenen Ergebnisliste verglichen werden, um das Ausmaß der Personalisierung zu untersuchen. Dieses gängige Vorgehen wurde auch für die vorliegende Arbeit genutzt.

Das Design entspricht daher einem Between-Subjects-Design, da jeder Account nur eine Bedingung innerhalb der UV ausgesetzt ist. In diesem Fall steht jeder Account für die Positionierung in einer Kontroverse. Zur Beantwortung Forschungsfrage 3 werden die inhaltlichen Positionen der Suchanfragen beider Accounts getauscht. Account A wird somit nun mit den Anfragen angereichert, die zuvor mit Account B in Verbindung standen und umgekehrt. Das Between-Subjects-Design bleibt, mit einem Wechsel der UV-Ausprägungen, bestehen.

5.2.2 Veränderung der UV: Manipulation des Suchverhaltens

Zur Durchführung der Studie ist es zunächst notwendig, für jedes Thema ein Set an Suchbegriffen zu erstellen. Die Herausforderung hierbei bringen Jürgens et al. (2014, S. 121) auf den Punkt:

„Eine notwendige Voraussetzung für sich durch Personalisierung unterscheidende Perspektiven in Trefferlisten ist ein Suchbegriff, zu dem es mindestens zwei unterschiedliche Sets an Treffern gibt, die sich anhand eines operationalisierbaren Kriteriums (z. B. Zugehörigkeit zu einem Thema, ähnliche Quelle o.Ä.) trennscharf unterscheiden lassen. Dazu passend muss es möglich sein, entsprechende Interessensprofile auf Seiten der Nutzer zu erzeugen. Weil die Nutzerprofile aus den vorangegangenen Suchanfragen entstehen, bedeutet das, dass zu jeder der zwei Perspektiven ein Satz an Stimulus- bzw. Trainings-Suchbegriffen vorliegen muss, der möglichst präzise mit der einen Perspektive, nicht aber mit der anderen zusammenhängt.“

In diesem Kontext bedeutet das: Für jedes der drei Themen (Energie, Verkehr, Landwirtschaft – die Herleitung der Themen ist in Abschnitt 3.1 dargestellt) müssen zwei Sets an Suchanfragen erstellt werden, mit denen die neu erstellten Google-Profile angereichert werden. Die unterschiedlichen Positionierungen sollen dabei aus den Suchanfragen ersichtlich sein. Da jedoch vorgesehen ist, das Experiment unter möglichst realistischen Bedingungen durchzuführen, sind einige Punkte zu beachten. Vor der Umsetzung lohnt es sich, einen genaueren Blick auf das Informationsverhalten der Nutzer zu werfen.

Die Intentionen der Nutzer hinter ihren Anfragen können in breiter Weise variieren. Grund für eine Suche ist jedoch, dass ein Informationsbedarf vorliegt. Der Informationsbedarf lässt sich wiederum unterscheiden in zwei verschiedene Typen, und zwar in *concrete information need* und *problem oriented information need* (Frants, Shapiro, & Voiskunskii, 1997), deren Unterschiede in Tabelle 5 ersichtlich sind.

Tabelle 5: Charakteristiken von Informationsbedarfen im Vergleich (Frants et al., 1997, S. 38)

Concrete Information Need	Problem Oriented Information Need
1. The thematic boundaries are clearly defined.	1. The thematic boundaries are not defined.
2. The request is put into exact words, that is, it corresponds exactly to the CIN thematical limits.	2. As a rule, the request does not conform to the POIN.
3. To satisfy a CIN only one good document is needed.	3. As a rule, the POIN cannot be satisfied, even with all good documents existing in the system.
4. As soon as the good document is found, the CIN disappears.	4. As soon as good documents are delivered, the thematical limits of POIN may change and the POIN itself remains for a long time.

Diese Unterscheidung lässt sich ausweiten, indem man die von Broder (2002) definierten Anfragetypen betrachtet. Dabei wurden die Absichten von Nutzern zur Informationssuche untersucht und klassifiziert: Navigationsorientierte Suchanfragen zielen auf das Auffinden einer konkreten Webseite ab, die dem Nutzer bereits bekannt ist. In der Regel gibt es bei diesem Anfragetyp aus Nutzersicht ein richtiges Ergebnis. Das unterscheidet sie grundlegend vom zweiten Anfragetyp, den informationsorientierten Suchanfragen. Diese zeichnen sich dadurch aus, dass – um ein bestimmtes Informationsbedürfnis zu befriedigen - mehrere Webseiten aufgesucht werden, z. B., um sich über ein gewisses Thema zu informieren. Während bei informationsorientierten Anfragen das Auffinden von statischen Webseiten im Vordergrund steht, also von Dokumenten, auf denen keinerlei weitere Interaktion nötig ist, außer die sich darin befindenden Informationen durchzulesen, zielen transaktionsorientierte Suchanfragen auf Interaktion ab. Die Intention von Nutzern ist das Auffinden einer Webseite, auf der sie anschließend eine Transaktion tätigen. Beispiele hierfür sind Produktkäufe oder der Download von Dateien wie etwa Musik. Lewandowski (2018, S. 72) ergänzt dazu die Unterscheidung von transaktionsorientierten Suchanfragen, die die Suche nach einer bestimmten Quelle zum Ziel haben und von Anfragen, die eine bestimmte Form der Transaktion verfolgen.

In einigen Fällen ist es jedoch nicht möglich, die Absicht zu erkennen, die hinter einer Suchanfrage steckt, weshalb bei vielen Anfragen ein gewisser Interpretationsspielraum vorliegt. Die oftmals unklare Absicht hinter einer Anfrage rührt auch daher, dass sie kurz gehalten und nicht sehr komplex in ihrer Formulierung sind (Höchstötter & Koch, 2009). Im Kontext dieser Arbeit ist jedoch folgender Punkt interessant, dass „fast jeder Nutzer seine Suche mit bestimmten subjektiven (und oft unbewussten) Annahmen startet“ (Stark, Magin, & Jürgens, 2014, S. 35). Dies wirkt sich wiederum auf die Suchergebnisse aus, die von den Formulierungen der Suchanfragen abhängig sind. So werden Suchmaschinen von den Nutzern bisweilen dazu verwendet, Behauptungen und Meinungen zu stützen. Bei entsprechenden Formulierungen der Suchanfragen können sie also als Werkzeug von Meinungsbestätigung fungieren und somit zur Entstehung von Confirmation Bias beitragen. Suchbegriffe nehmen daher einen Teil des Prozesses zur Meinungsbildung ein und deren Formulierungen können potenziell Auswirkungen zur Stärkung von Positionen haben (Haider & Sundin, 2019, S. 90f.). In ihrer Forschung zur Recherche von Nachrichten konnte Tripodi (2018) beobachten, dass Nutzer Google dazu verwenden, Meinungen zu bestätigen. Das ging bisweilen so weit, dass sie einzelne Begriffe oder ganze Zitate aus bereits bekannten Quellen wortgetreu an das Suchsystem übermittelt haben. Diesen entstehenden Kreislauf von aufgrund entsprechend formulierten Anfragen wiederkehrenden Informationen bezeichnen Haider und Sundin (2019, S. 90) auch als „information feedback loop“.

Für diese Arbeit wurden letztendlich sechs Sets mit je 30 Suchanfragen angefertigt, die jeweils den Standpunkt einer Kontroverse in den identifizierten Umweltthemen implizieren sollen. Die Herausforderung war, Begriffe zu finden, die diese implizierte Positionierung deutlich hervorheben, jedoch auch in der Realität von Google-Nutzern verwendet werden. Dazu wurde der Google Keyword-Planer zu Hilfe genommen. Der Google Keyword-Planer ist ein von Google zur Verfügung gestelltes Tool, das eigentlich auf die Verwendung im Suchmaschinenmarketing abzielt und in diesem Feld primär zur Keyword-Recherche eingesetzt wird. Neben beispielsweise Kostenabschätzungen, die in diesem Kontext nicht relevant sind, liefert der Keyword-Planer auch die durchschnittliche Anzahl der Suchen nach einem bestimmten Begriff in einen individuellen Zeitraum. So können die Suchvolumina einzelner Suchanfragen ermittelt werden (Pelzer & Gerigk, 2018). Die für diese Arbeit ausgewählten Suchanfragen richten sich jedoch nicht ausschließlich nach den Suchvolumina. Zwar werden manche Begriffe, die für die Themen geeignet sind, verhältnismäßig häufig in der Realität gesucht. Oftmals

handelt es sich dabei jedoch um einzelne Wörter, sodass die Abgrenzung zum anderen Standpunkt der Kontroverse nicht ersichtlich ist. Es wurde daher häufig mit den Wörtern „vorteil“ bzw. „nachteil“ gearbeitet. Dadurch sind die Anfragen so interpretierbar, dass sich ein Nutzer für oder gegen einen Standpunkt ausspricht. Zu erwähnen ist hierbei nochmals das von Tripodi (2018) und von Haider und Sundin (2019) beobachtete Phänomen, Suchanfragen entsprechend eigener Gesinnung zu formulieren. Auffällig war außerdem, dass häufig nach Informationen gesucht wird, die die eigenen Finanzen betreffen, weswegen einzelne Anfragen für die Studie mit den Begriffen „preis“, „preisvergleich“, „kaufen“ oder „kosten“ kombiniert sind.

Es galt bei der Erstellung der Suchanfragen-Sets also abzuwägen, welche Begriffe in der Studie Verwendung finden sollten. Hierbei sei nochmals auf die von Broder (2002) identifizierten Anfragetypen verwiesen. In jedem Satz an Suchbegriffen, die in die Studie einfließen, sind jeweils alle drei Typen interpretierbar. Wie bereits erwähnt, ist die Interpretation der Suchanfragen nicht immer eindeutig. Es ließen sich vermutlich zahlreiche verschiedene Sets bilden, die letztlich jedoch zum gleichen Effekt führen können. In den nachfolgenden Abschnitten sind alle Suchanfragen für die einzelnen Themenblöcke dargestellt.

Tabelle 6: Suchanfragen-Sets - Sektor Energie

Suchanfragen, die eine positive Haltung zu erneuerbaren Energien und eine negative Haltung zur Energiegewinnung aus fossilen Brennstoffen implizieren	Suchanfragen, die eine positive Haltung zur Energiegewinnung aus fossilen Brennstoffen und eine negative Haltung zu erneuerbaren Energien implizieren
grünwelt	vorteile fossile brennstoffe
greenpeace energy	erdgas vorteile
hambi bleibt	nachteile biomasse
solaranlage kosten	deutsches atomforum
bafa förderung wärmepumpe	erneuerbare energien nachteile
ökostrom vergleich	pro atomkraft
atomkraft nein danke	geothermie nachteile

entega ökostrom	vorteile atomkraftwerk
photovoltaik förderung	erdwärme nachteile
ökostrom anbieter	atomkraft ja bitte
solaranlage garten	pelletheizung nachteile
solarthermie kosten	erdgaspreis
einspeisevergütung photovoltaik	nuclear pride fest
bafa förderung biomasse	nachteile windenergie
pelletheizung förderung	solarenergie nachteile
bundesverband erneuerbare energien	heizöl preisvergleich
hackschnitzelheizung kosten	gasheizung kosten
vorteile windenergie	ölheizung kosten
geothermie kosten	erdgas heizung
ökostrom vergleich greenpeace	wie viele vögel sterben durch windräder
nachteile atomkraftwerk	hambi muss weg
vorteile wasserkraft	debriv
windrad kaufen	ring deutscher bergingenieure
wärmepumpe vorteile	bundesverband landschaftsschutz
braunkohle nachteile	gegenwind
ölheizung nachteile	windwahn
wärmepumpe preis	windenergie kritik
klima allianz	nuklearia
enynway	vorteile kohlekraftwerk
ende gelände	förderung kernenergie

Tabelle 7: Suchanfragen-Sets - Sektor Verkehr

Suchanfragen, die eine positive Haltung zu Maßnahmen zur Gestaltung der Verkehrswende und eine negative Haltung zu Maßnahmen zur Gestaltung der Verkehrswende implizieren	Suchanfragen, die eine positive Haltung zu Maßnahmen zur Verhinderung der Verkehrswende und eine negative Haltung zu Maßnahmen zur Gestaltung der Verkehrswende implizieren
elektroauto preisvergleich	vorteile diesel
elektroauto kaufen	contra tempolimit
vorteile elektroauto	erdgasauto kaufen
vorteile tempolimit	abgasnorm
luftverschmutzung diesel	elektroauto nachteile
elektroauto leasing	elektroauto umweltschädlich
günstige elektroautos	cng auto kaufen
elektroautos übersicht	euro 6 diesel kaufen
ladestation elektroauto	jetzt noch euro 5 diesel kaufen
bmw i3 preis	diesel kaufen
elektroauto gebraucht	ab wann lohnt sich ein diesel
e smart preis	diesel kaufen oder nicht
e auto news	exxon mobil
förderung elektroauto	gegen tempolimit
tesla leasing	aktuelle benzinpreise
kleines elektroauto	e auto kritik
e auto prämie	spritpreise
elektroauto test	dieselpreis
bafa umweltbonus	euro 6 plakette
elektroauto laden	diesel oder benziner
bmw umweltprämie	schadstoffklasse

renault zoe leasing	vorteile ottomotor
audi e tron preis	mwv
audi elektro	tankstellenpreise
agora verkehrswende	royal dutch shell
steuervorteil elektroauto	bp aktie
e wald	euro 5 auf euro 6 umrüsten
nachteile diesel	billig tanken
verkehrsclub deutschland	tankstelle in der nähe
transport and environment	cng tankstelle

Tabelle 8: Suchanfragen-Sets - Sektor Landwirtschaft

Suchanfragen, die eine positive Haltung zur ökologischen Landwirtschaft und eine negative Haltung zur konventionellen Landwirtschaft implizieren	Suchanfragen, die eine positive Haltung zur konventionellen und eine negative Haltung zur ökologischen Landwirtschaft implizieren
rettet die bienen	glyphosat kaufen
albert schweitzer stiftung	glyphosat preisvergleich
gegen massentierhaltung	massentierhaltung vorteile
gemüsebox	gentech vorteile
bio pflanzenschutzmittel	roundup kaufen
monsanto greenwashing	roundup dosierung
neuland fleisch	pestizide landwirtschaft
bio dünger	biscaya insektizid
neemöl kaufen	bi 58 kaufen
alnatura shop	syngenta landwirtschaft

bio online shop	bi 58 anwendung
wir haben es satt	bayer crop
deutscher tierschutzbund	nur fleisch essen
kein fleisch essen	massentierhaltung pro
ökologische produkte	konventionelle landwirtschaft vorteile
ökologische aufzucht	nachteile ökologische landwirtschaft
massentierhaltung nachteile	calypso pflanzenschutzmittel
zugelassene dünger ökolandbau	glyphosat anwendung
bio siegel tierhaltung	bayer glyphosat
bio hühnerhaltung richtlinien	agrar basf
fruchtfolge ökolandbau	bayer monsanto glyphosat
bio landwirtschaft förderung	forum moderne landwirtschaft
provieh	verbindungsstelle landwirtschaft industrie
vorteile ökologische landwirtschaft	deutscher bauernverband
umstellung auf bio landwirtschaft	agravis
studium ökologische landwirtschaft	permaclean anwendung
glyphosat krebserregend	bayer aktien
bio supermarkt	baywa agrar
gentechnik nachteile	nordzucker agriportal
bio company	glyphosat bienen

5.2.3 Kontrolle von Störvariablen

Die Kontrolle von Störvariablen stellt sich in diesem Experiment als herausfordernd dar, da die genaue Funktionsweise des Personalisierungsalgorithmus von Google nicht bekannt ist (Lewandowski, 2018, S. 94). Die Grundzüge dieser Algorithmen sind jedoch bekannt, weshalb sich manche Störvariablen problemlos kontrollieren lassen. Da wäre

zunächst die Auswahl der Geräte, auf denen die Studie durchgeführt wurde. Hierbei wurde darauf geachtet, gleiche Geräte vom gleichen Hersteller zu benutzen. Aus dem Department Information der HAW Hamburg wurden drei iPad mini des gleichen Modells zur Verfügung gestellt, die jeweils mit dem Betriebssystem iOS 9.3.5 ausgestattet sind. Damit keine historischen Daten in die Studie miteinfließen konnten, wurden die Geräte vor Beginn der Studie zurückgesetzt. Auch bei der Wahl des Internetbrowsers gab es keine Unterschiede, es wurde stets der auf den Geräten installierte Browser „Safari“ der Firma Apple verwendet. Trotz der Rücksetzung wurde nochmals überprüft, ob die Browsereinstellungen auf allen Geräten identisch sind, was letztlich der Fall war und so beibehalten wurde.

Nachdem die Geräte den Anforderungen für die Durchführung des Experiments entsprechend präpariert wurden, erfolgte die Erstellung neuer Google-Accounts. Neue Accounts sind in diesem Kontext unabdingbar, da sich Suchaktivitäten aus der Vergangenheit potenziell auf die Ergebnisse auswirken könnten. Bei der Erstellung eines Profils fragt Google verschiedene personenbezogene Daten ab (Google, 2019b). Neben dem Namen, dem Nutzernamen und dem Passwort gehören das Geburtsdatum und das Geschlecht zu den Pflichtfeldern bei der Kontoerstellung. Optional können eine E-Mail-Adresse zur Kontowiederherstellung und eine Telefonnummer angegeben werden. Im letzten Schritt erfolgte die Zustimmung der Datenschutzerklärung, um die Dienste überhaupt nutzen zu können. Hierbei wurde von Google auf die Nutzung der Daten hingewiesen, wobei die Verwendung der Daten zur Bereitstellung personalisierter Inhalte explizit erwähnt wird (Google, 2019a). Über einen Namensgenerator wurde online für jeden neuen Account ein zufällig generierter Nachname ausgewählt. Bei der Auswahl eines Vornamens wurde auf einen geschlechterneutralen Namen geachtet (Beispiel: Kim). Beim Geschlecht wurde die Auswahlmöglichkeit „Ich möchte dies nicht beantworten“ ausgewählt. So wurde sichergestellt, dass Google keine geschlechterspezifischen Personalisierungen in der Präsentation der Inhalte vornimmt. Die Angaben zum Geburtsdatum waren für jeden Account ähnlich. Jedem Konto wurde ein Geburtstag im April 1989 zugewiesen. Die simulierten Nutzer waren zu Beginn der Studie also alle 30 Jahre alt. Die Nutzernamen leiteten sich von den fiktiven Namen ab und wurden mit einer zufällig ausgewählten Kombination aus vier Ziffern ergänzt, die Passwörter orientierten sich wiederum am Nutzernamen. Optionale Angaben zu einer bereits bestehenden E-Mail-Adresse und einer Telefonnummer wurden keine gemacht. Erstens waren keine weiteren Angaben für diese Studie nötig. Zweitens wurde die

Gefahr, dass bereits bestehende Konten mit den neu angelegten durch die Angabe von E-Mail-Adressen und Telefonnummern verknüpft werden, als zu hoch eingeschätzt.

Weitere potenzielle Störfaktoren liegen in den Sucheinstellungen bei Google. Um diese zu kontrollieren, wurde darauf geachtet, jedes Gerät und jeden Account, sofern sinnvoll, mit denselben Sucheinstellungen auszustatten. Zunächst ist festzuhalten, dass für die Eingabe der Suchanfragen immer dieselbe Domain ausgewählt wurde: www.google.de. Die Studie wurde auf den Tablets also browserbasiert durchgeführt und nicht über die Google-App. Folgende Sucheinstellungen wurden getroffen:

- Suchverlauf: in der anonymen Suche wurden die Suchanfragen – anders als in den personalisierten Suchen - nicht gespeichert; damit soll der Einfluss vergangener Suchanfragen bei anonymen Suchen eliminiert werden
- SafeSearch-Filter: damit sollen anstößige Inhalte herausgefiltert werden; Einstellung war standardmäßig deaktiviert und auf „Relevanteste Ergebnisse anzeigen“ eingestellt, was bei allen Geräten so beibehalten wurde
- Videos: Standardeinstellung „Automatische Vorschauwiedergabe nur bei WLAN und in mobilen Netzwerken“ wurde auf allen Geräten beibehalten
- Letzte Standorte: Google hat die Speicherung der Standorte der durchgeführten Aktivitäten voreingestellt; standardmäßige Aktivierung der Speicherung wurde auf allen Geräten beibehalten
- Handschrift: Möglichkeit, Suchbegriffe nicht über die Tastatur, sondern über eine Fingerbewegung handschriftlich zu suchen; Standardeinstellung „Deaktivieren“ wurde auf allen Geräten beibehalten
- Private Ergebnisse (nur nach Login einstellbar): Standardeinstellung „Private Ergebnisse verwenden“ wurde auf allen Geräten beibehalten
- Region: Google passt die Ergebnisse an die Region an, in der sich Nutzer aktuell befinden; auch diese Einstellung wurde auf allen Geräten so beibehalten
- Sprache: die Sprache Deutsch war von Google ausgewählt und wurde auf allen Geräten übernommen

Der Standort blieb während des Untersuchungszeitraumes ebenfalls unverändert. Dazu wurde von der Forschungsgruppe Search Studies der HAW Hamburg ein Arbeitsraum zur Verfügung gestellt. Die Durchführung der gesamten Studie fand somit in den Räumlichkeiten des Departments Information der HAW Hamburg auf dem Kunst- und

Medien-campus in Hamburg statt. Wie am Beispiel der Standortdaten ersichtlich, ist die Eliminierung von Störvariablen nicht immer möglich. Die Kontrolle der Störvariablen erfolgte daher auch mit der Technik des Gleichhaltens.

5.2.4 Praktische Durchführung des Experiments

Dieser Abschnitt handelt von der praktischen Durchführung des Experiments. Die in der Theorie von Sedlmeier und Renkewitz (2018, S. 129ff.) aufgeführten Handlungsschritte sind zwar primär für Experimente in der Psychologie gedacht, an denen menschlichen Probanden teilnehmen, sie lassen sich jedoch auch für die vorliegende Studie umsetzen. Wie bereits aufgeführt, wird die unabhängige Variable im Nutzungsverhalten vermutet und die Auswirkungen zeigen sich in der Erstellung der Ergebnislisten. Die Experimentalbedingung stellt in diesem Fall eine anwesende UV dar, nämlich ein manipuliertes Suchverhalten. In der Kontrollbedingung fehlt die UV, da dort anonym gesucht wird, ohne einen Login in ein Konto. Daher wird auch kein umfangreiches Suchverhalten aufgezeichnet.

Die Suchergebnisse zu den Themen Energie, Verkehr und Landwirtschaft sollen untersucht werden, weswegen sechs neue Accounts erstellt wurden. Jedes Konto repräsentierte also einen fiktiven Nutzer, dessen Haltung zu umweltschützenden Maßnahmen aus den drei Themenbereichen jeweils entweder positiv oder negativ zu interpretieren ist. Die drei verwendeten iPads wurden folgendermaßen genutzt:

- Gerät 1: Anlegen von drei Accounts, die täglich mit Suchanfragen angereichert werden, in denen eine positive Haltung zum Umweltschutz interpretierbar ist; wöchentliche personalisierte Suche nach neutralen Begriffen
- Gerät 2: Anlegen von drei Accounts, die täglich mit Suchanfragen angereichert werden, in denen eine negative Haltung zum Umweltschutz interpretierbar ist; wöchentliche personalisierte Suche nach neutralen Begriffen
- Gerät 3: wöchentliche anonyme Suche nach neutralen Begriffen

Die Eingabe der Suchbegriffe erfolgte jeweils auf dem Gerät, auf dem der jeweilige Account angelegt wurde. Jeder neue Account muss zunächst mit einem Interessensprofil ausgestattet werden, wie von Jürgens et al. (2014) vorgeschlagen. Es wurde bereits darauf eingegangen, wie die einzelnen Sets an Suchanfragen zusammengestellt wurden. Der nächste Schritt war die Erstellung eines Zeitplans, in dem festgehalten wurde, welche Suchanfrage an welchem Tag in welchem Konto getätigt wird. Jedes Set an

Begriffen beinhaltet insgesamt 30 Suchanfragen, wovon täglich 20 in randomisierter Reihenfolge in den jeweiligen Accounts eingeloggt an Google geschickt wurden. Außerdem wurde festgelegt, dass jede Suchanfrage 30 Mal während des Untersuchungszeitraums abgeschickt wird. Der Zeitraum betrug insgesamt 90 Tage. Da jeder Account täglich mit 20 Suchanfragen angereichert wurde, waren schlussendlich 1800 Suchanfragen pro Konto gespeichert. Doch nicht nur die Formulierung der Suchanfragen trägt zum Aufbau des Interessensprofils bei, auch die Klicks auf die Internetdokumente werden, wie aus der Literatur hervorging, von Google gespeichert und für die Personalisierung verwendet. Deshalb erfolgte dieses Klickverhalten ebenfalls systematisch. Nach jeder abgeschickten Suchanfrage erfolgte zunächst ein Klick auf ein Dokument der ersten Ergebnisseite, wobei jedoch nur organische Ergebnisse berücksichtigt wurden. Ausgewählt wurden Dokumente, die aus Sicht des Versuchsleiters anhand ihrer Trefferbeschreibungen eindeutig dem Standpunkt einer Kontroverse zuzuordnen waren. Sollte in keinem Snippet eine eindeutige Positionierung erkennbar sein, wurde ein zufällig ausgewähltes Dokument angeklickt. Da die Annahme bestand, die Ergebnisse auf der ersten SERP würden im Laufe der Untersuchung wiederholend auftreten, wurden auch die Dokumente auf der zweiten Ergebnisseite berücksichtigt, da auch Treffer auf den hinteren Positionen durchaus relevant zur Suchanfrage sein können (Lewandowski, 2018, S. 298). Dazu erfolgte ein Klick auf die Schaltfläche „Weitere Ergebnisse laden“ der iPads.

Die Studie startete am 10. Juni und lief bis zum 09. September 2019. Nach jeder Woche wurden die Personalisierungseffekte gemessen. Dazu war folgender Vergleich nötig: Nachdem jedem Account entsprechend seiner Positionierung nach und nach ein Interessensprofil mit einer stetig wachsenden Anzahl an Suchanfragen zugefügt worden ist, erfolgte ein Vergleich verschiedener SERPs nach Eingabe desselben Suchbegriffs. Durch Personalisierungstechniken wird nach Eingabe derselben Suchanfrage eine anders angeordnete SERP im Vergleich zur anonymen Suche erwartet. Dazu wurden für jedes der drei übergeordneten Themen Energie, Verkehr und Landwirtschaft drei möglichst neutrale Begriffe ausgewählt, nach denen jede Woche entweder eingeloggt oder anonym gesucht wird. Bei der Auswahl dieser Begriffe wurde sich an die spezifischen Themen orientiert, die vor den Wahlen im Wahl-O-Mat mit aufgenommen wurden (siehe Tabelle 4). Darüber hinaus wurde die aktuelle Berichterstattung berücksichtigt. Hierzu sei angemerkt, dass Beiträge aus Internet, Presse und Fernsehen nicht systematisch ausgewertet wurden. Die Auswahl der Begriffe stützte sich demnach

auch auf Mutmaßungen des Verfassers. Im Folgenden sind für jeden Sektor die drei Suchanfragen aufgelistet und deren Auswahl kurz begründet.

Thema Energie

Im Sektor Energie wird insbesondere die Gestaltung der Energiewende diskutiert, die sich wie folgt definiert: „The phasing out of non-sustainable fossil fuels and nuclear energy for a phasing in of sustainable renewable energy supplies, is now called the energy transition“ (Verbruggen, 2014, S. 120).

Die Thesen im Wahl-O-Mat und Beiträge aus den Medien – einige Berichte sind beispielhaft aufgeführt – thematisieren dabei häufig den Ausbau der Windkraft (z. B. Schmidt, 2019) und den Ausstieg aus der Kohleenergie (ZDF, 2019). In beiden Themenfeldern geht es um die Art und das Tempo der Umsetzung. Für den Untersuchungssektor Energie wurden daher folgende Begriffe ausgewählt: „ausbau windkraft“, „ausstieg kohleverstromung“, „umsetzung energiewende“.

Thema Verkehr

Auch in diesem Themenfeld erfolgte die Auswahl möglichst neutraler Begriffe über die Orientierung an den Thesen im Wahl-O-Mat sowie der aktuellen Berichterstattung. Aus den im Wahl-O-Mat vorkommenden Thesen im Sektor Verkehr geht hervor, dass es insbesondere Diskussionen zur Einführung eines Tempolimits und um Fahrverbote für mit Diesel angetriebene Fahrzeuge gibt, weshalb die Begriffe „tempolimit“ und „dieselfahrverbot“ ausgewählt wurde. Die dritte Suchphrase lautet „umweltbilanz auto“, da die Umweltbilanz von Elektroautos im Vergleich zu Fahrzeugen mit Verbrennungsantrieb kontrovers diskutiert wird (Krauter, 2019; Schrader, 2017).

Thema Landwirtschaft

Bei der Auswahl der Suchbegriffe für das Thema Landwirtschaft wurde sich an den Charakteristiken der biologischen Landwirtschaft orientiert, die sich in wesentlichen Punkten von der konventionellen Landwirtschaft unterscheidet. Laut dem Bundesministerium für Ernährung und Landwirtschaft (Bundesministerium für Ernährung und Landwirtschaft, 2019, S. 4) möchte die ökologische Landwirtschaft

- „einen möglichst geschlossenen betrieblichen Nährstoffkreislauf erreichen [...]
- die Bodenfruchtbarkeit erhalten und mehren
- Tiere besonders artgemäß halten.“

Der ökologische Landbau zeichnet sich u. a. durch den Verzicht auf gentechnisch veränderte Organismen aus, strenge Regelungen bei der Haltung von Nutztvieh sowie bei zugelassenen Düngemitteln (Bundesministerium für Ernährung und Landwirtschaft, 2019). Teilweise wurden diese Themen schon als Thesen vor Wahlen im Wahl-O-Mat mitaufgenommen. Doch wie bei den beiden vorherigen Themenkomplexen wurde hier ergänzend ein Blick in Medienberichte geworfen, um Kontroversen aus dem Sektor Landwirtschaft zu identifizieren und entsprechende Suchanfragen abzuleiten. Derzeit behandeln die Debatten u. a. die Einführung einer neuen Düngeverordnung (Wulfers, 2019), die Haltungsbedingungen von Nutztieren (Mair, 2019; Norddeutscher Rundfunk, 2019) und den Einsatz von Gentechnologie in der Landwirtschaft (Kochale, 2019). Basierend darauf wurden die Begriffe „düngeverordnung“, „tierhaltung“ und „gentechnik“ für die Studie ausgewählt.

5.2.5 Inhaltliche Bewertung der Dokumente durch Juroren

Nachdem die Dokumente zu den Suchanfragen gesammelt wurden, sollten alle themenrelevanten Dokumente auf den ersten zehn Positionen auf ihren Inhalt hin bewertet werden. Dazu eignet sich als Messinstrument die bipolare Ratingskala. Diese Skalenart zeichnet sich dadurch aus, dass an jedem Skalenende ein anderes, gegensätzliches Merkmal gesetzt wird. Dies hat den Vorteil, dass durch die wechselseitige Definition der beiden gegensätzlichen Begriffe die Präzision der Urteile erhöht werden kann (Döring & Bortz, 2016, S. 245). Bezogen auf die umweltpolitische Thematik wurden für die Entwicklung der Ratingskala für jeden der drei Themensektoren Merkmalspaare gebildet, die jeweils die Skalenenden repräsentieren:

- Sektor Energie: Maßnahmen Energiewende ablehnend – Maßnahmen Energiewende befürwortend
- Sektor Verkehr: Maßnahmen Verkehrswende ablehnend – Maßnahmen Verkehrswende befürwortend
- Sektor Landwirtschaft: Maßnahmen Agrarwende ablehnend – Maßnahmen Agrarwende befürwortend

Ein weiteres Merkmal der Ratingskala ist ihre Skalierung. Trotz Diskussionen darüber, ob ordinal- oder intervallskaliert, können Messungen mit einer Ratingskala als intervallskaliert angesehen werden (Sedlmeier & Renkewitz, 2018, S. 75). Damit sind die Unterschiede zwischen den Messwerten als gleich groß zu betrachten (ebd., S. 74). Insgesamt hat die Skala fünf Merkmalsausprägungen, die von -2 bis +2 definiert sind. Döring und Bortz (2016, S. 248) führen zu diesen sogenannten numerischen Marken, also Zahlenwerten als Merkmalen, an: "Numerische Skalenbezeichnungen [...] sind knapp und eindeutig; ihre Verwendung ist jedoch nur sinnvoll, wenn die Untersuchungspersonen diese abstrakte Darstellungsform verstehen."

Aus diesem Grund bekamen die Juroren vor der Bewertung der eigentlichen Dokumente Instruktionen, wie die Inhalte zu bewerten sind. Die Idee dahinter, die Bewertung von Personen durchführen zu lassen, war einerseits mit Gründen des Aufwands verbunden. Andererseits können dadurch potenzielle subjektive Annahmen des Versuchsleiters eliminiert werden. Die teils komplexe Thematik in umweltpolitischen Fragen ist außerdem nicht immer eindeutig zu bewerten, weshalb drei Juroren für die Inhaltsbewertung akquiriert wurden. Die Instruktionen wurden den Juroren neben einer Auswahl einiger beispielhaften Dokumente zur Thematik vorab zur Verfügung gestellt. Damit konnten sie die Instruktionen auf mögliche Schwächen überprüfen und der Leitfaden ggf. angepasst werden. Die Dokumente für diesen Pretest wurden dabei so gewählt, dass sie mit großer Wahrscheinlichkeit nicht als Untersuchungsobjekte für die eigentliche Studie infrage kamen. Neben den Instruktionen sollten ein Glossar und ein Informationsdossier zur Thematik bei der Bewertung unterstützen. Beide Dokumente sind im Anhang dieser Arbeit beigefügt. In Tabelle 9 sind die Instruktionen für die Inhaltsbewertung dargestellt, wie sie auch den Juroren zur Verfügung gestellt wurden.

Tabelle 9: Instruktionen für die Juroren zur Bewertung der Dokumente

	Energie	Verkehr	Landwirtschaft
-2	<p>Dokument hebt folgenden Aspekt positiv hervor:</p> <p>Energiegewinnung aus fossilen Brennstoffen (Braunkohle, Steinkohle, Erdgas, Erdöl, Torf, Kernenergie)</p> <p>UND</p> <p>folgenden Aspekt negativ hervor:</p> <p>Energiegewinnung aus erneuerbaren Energien (Wasserkraft; Windkraft; Sonnenenergie; Meeresenergie, wie Gezeitenkraftwerke, Wellenkraftwerke, Osmosekraftwerke, Meeresströmungskraftwerke und Meereswärmekraftwerke; Bioenergie, wie Holzpellets, Biogas, Biomethan, Hackschnitzel, Strohpellets; Geothermie, wie Einsatz von Wärmepumpen oder Wärmekollektoren)</p>	<p>Dokument hebt mindestens einen der folgenden Aspekte positiv hervor: Antrieb durch fossile Kraftstoffe (Benzin, Diesel, Erdgas); Ablehnung Tempolimit</p> <p>UND</p> <p>mindestens einen der folgenden Aspekte negativ hervor: Elektromobilität; Einführung Tempolimit</p>	<p>Dokument hebt folgenden Aspekt positiv hervor: konventionelle Landwirtschaft (Einsatz chemischer Pflanzenschutzmittel wie bspw. Glyphosat; Einsatz genmanipulierter Saatgutsorten; intensive Tierhaltung; Einsatz chemischer Düngemittel)</p> <p>UND</p> <p>folgenden Aspekt negativ hervor: ökologische Landwirtschaft (Verzicht auf Gentechnik, chemische Dünger, chemische Pflanzenschutzmittel; ökologische Viehzucht)</p>
-1	<p>Dokument hebt folgenden Aspekt positiv hervor:</p> <p>Energiegewinnung aus fossilen Brennstoffen</p>	<p>Dokument hebt mindestens einen der folgenden Aspekte positiv hervor: Antrieb durch fossile Kraftstoffe; Ablehnung Tempolimit</p>	<p>Dokument hebt folgenden Aspekt positiv hervor: konventionelle Landwirtschaft</p>

	<p>Erneuerbare Energien werden im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Dokument hebt folgenden Aspekt negativ hervor: Energiegewinnung aus erneuerbaren Energien.</p> <p>Fossile Brennstoffe werden im Dokument nicht wertend erwähnt</p>	<p>Elektromobilität oder die Einführung eines Tempolimits werden im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Dokument hebt mindestens einen der folgenden Aspekte negativ hervor: Elektromobilität; Ablehnung Tempolimit</p> <p>Fossile Kraftstoffe werden im Dokument nicht erwähnt</p>	<p>Ökologische Landwirtschaft wird im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Aspekt hebt folgenden Aspekt negativ hervor: ökologische Landwirtschaft</p> <p>Konventionelle Landwirtschaft wird im Dokument nicht wertend erwähnt</p>
0	<p>Das Dokument hebt weder die Energiegewinnung aus fossilen Brennstoffen noch erneuerbare Energien positiv oder negativ hervor</p> <p>ODER</p> <p>Dokument hebt Energiegewinnung aus fossilen Brennstoffen und erneuerbare Energien in gleicher Weise positiv oder negativ hervor</p>	<p>Das Dokument steht dem Thema Verkehrswende neutral gegenüber, was sich folgendermaßen ausdrückt:</p> <p>Das Dokument hebt weder den Antrieb durch fossile Kraftstoffe noch Elektromobilität positiv oder negativ hervor; ebenso beim Thema Tempolimit</p> <p>ODER</p>	<p>Das Dokument hebt weder die konventionelle noch die ökologische Landwirtschaft positiv oder negativ hervor</p> <p>ODER</p> <p>Dokument hebt konventionelle und ökologische Landwirtschaft in gleicher Weise positiv oder negativ hervor</p>

		<p>Dokument hebt weder Antrieb durch fossile Kraftstoffe noch Elektromobilität in gleicher Weise positiv oder negativ hervor; ebenso beim Thema Tempolimit</p> <p>ODER</p> <p>Dokument hebt weder Antrieb durch fossile Kraftstoffe noch Elektromobilität positiv oder negativ hervor, sondern Hybridantriebe (Kombination aus Verbrennungs- und Elektroantrieb) oder alternative Kraftstoffe (Biodiesel, Wasserstoff, BtL-Kraftstoff)</p>	
+1	<p>1 Dokument hebt folgenden Aspekt positiv hervor: Energiegewinnung aus erneuerbaren Energien</p> <p>Fossile Brennstoffe werden im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Dokument hebt folgenden Aspekt negativ hervor: Energiegewinnung aus fossilen Brennstoffen</p>	<p>Dokument hebt mindestens einen der folgenden Aspekte positiv hervor: Elektromobilität; Einführung Tempolimit</p> <p>Fossile Kraftstoffe bzw. Ablehnung Tempolimit werden im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Dokument hebt mindestens einen der</p>	<p>Dokument hebt folgenden Aspekt positiv hervor: ökologische Landwirtschaft</p> <p>Konventionelle Landwirtschaft wird im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Dokument hebt folgenden Aspekt negativ hervor:</p>

	Erneuerbare Energien werden im Dokument nicht wertend erwähnt	folgenden Aspekte negativ hervor: Antrieb durch fossile Kraftstoffe; Ablehnung Tempolimit Elektromobilität bzw. Einführung Tempolimit werden im Dokument nicht wertend erwähnt	konventionelle Landwirtschaft Ökologische Landwirtschaft wird im Dokument nicht wertend erwähnt
+2	2 Dokument hebt folgenden Aspekt positiv hervor: erneuerbare Energien UND folgenden Aspekt negativ hervor: Energiegewinnung aus fossilen Brennstoffen	Dokument hebt mindestens einen der folgenden Aspekte positiv hervor: Elektromobilität; Einführung Tempolimit UND mindestens einen der folgenden Aspekte negativ hervor: Antrieb durch fossile Kraftstoffe; Ablehnung Tempolimit	Dokument hebt folgenden Aspekt positiv hervor: ökologische Landwirtschaft UND folgenden Aspekt negativ hervor: konventionelle Landwirtschaft

Für die Bewertung der Dokumente wurde von der Forschungsgruppe Search Studies am Department Information der HAW Hamburg ein Tool zur Verfügung gestellt, mit dem die zu bewertenden Dokumente gescraped wurden. Von den entsprechenden Webseiten wurden Screenshots angefertigt, über die die Juroren die Inhalte zu sehen bekamen. In dieses Tool inbegriffen war die entwickelte Skala. Über einen Link konnten die erfassten Dokumente somit ohne viel Aufwand an die Juroren gesendet werden. Nach einem kurzen Begrüßungstext werden die Juroren aufgefordert, die Dokumente auf ihren Inhalt hin auf der entwickelten Skala zu bewerten. Dabei wurde auch berücksichtigt, dass einzelne Dokumente nicht angezeigt werden können. Dies kann in geringem Umfang technische Probleme betreffen. In erster Linie, so die Annahme, jedoch Videoinhalte oder bezahlpflichtige Artikel. Sehen sich die Juroren nicht in der Lage, ein Dokument zu bewerten, hatten sie neben der Skala noch weitere

Auswahlmöglichkeiten ("Bewertung des Inhalts nicht möglich" – bei dieser Auswahlmöglichkeit wurde zusätzlich nach dem Grund gefragt; "Inhalte des Dokuments nicht themenrelevant"; "Ich bin mir bei der Bewertung nicht sicher"). Nicht themenrelevante Inhalte sowie Dokumente, zu denen kein Zugang gewährleistet war, wurden nicht in die Auswertung mitaufgenommen. Bei Unsicherheiten Seiten der Juroren oder Videoinhalten, übernahm der Autor der Arbeit die Bewertung. Die Häufigkeit wurde dabei als sehr gering eingeschätzt.

6 Auswertung der Daten

Im Folgenden ist die Auswertung der Daten dargestellt. Beginnend mit den Personalisierungseffekten unabhängig vom Inhalt (zu F1) folgt im Anschluss die Auswertung der inhaltlichen Bewertungen der Dokumente (F2 und F3). Beide Abschnitte sind gleich gegliedert. Zunächst werden die Kennzahlen vorgestellt, die zur jeweiligen Messung der Effekte herangezogen wurden. Die Ergebnisse sind zunächst deskriptiv dargestellt, danach erfolgt statistische Überprüfung der Hypothesen. Die Ergebnisse der Quellenauswertung (F4) bilden den Abschluss der Ergebnisdarstellung und sind rein deskriptiv abgebildet.

6.1 Personalisierung unabhängig vom Inhalt

6.1.1 Kennzahlen

Die Annahme ist, die Anzahl der gesetzten Signale im eingeloggt Zustand übt einen Einfluss auf die Anordnung der Ergebnisse aus. Unter Signalen sind Einheiten zu verstehen, die zusammengenommen als Rankingfaktor zu sehen sind (Lewandowski, 2018, S. 94). Im Rahmen dieser Arbeit werden implizite Daten im Profil als Rankingfaktor betrachtet, die Suchanfragen und die anschließenden Klicks auf ein Ergebnis als Signale. Auf den personalisierten SERPs die Anzahl der Dokumente ermittelt, die im Vergleich zur anonym erhobenen Kontroll-SERP mit identischer URL auf dem gleichen Rang platziert waren. Die Grundidee dahinter ist: Je niedriger die Anzahl derselben Dokumente auf demselben Rang, desto mehr Unterschiede gibt es in der Reihung der Ergebnisse, was als Hinweis auf Personalisierungseffekte gedeutet wird. Diese Kennzahl wurde mit den vorher gesetzten Signalen, nämlich einer Suchanfrage mit anschließendem Klick auf ein Dokument, in Beziehung gesetzt, um somit die Personalisierungseffekte zu ermitteln. Eine weitere erhobene Kennzahl, die in erster Linie einen schnellen Überblick über mögliche Personalisierungseffekte gewähren soll, ist der Jaccard-Koeffizient, der sich mit folgender Formel berechnen lässt:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Der Jaccard-Koeffizient ist ein Ähnlichkeitsmaß für Objekte (Kosub, 2016). Sein Wert liegt immer zwischen 0 und 1. Als Objekte sind hier die Suchergebnislisten zu verstehen, die miteinander verglichen werden. Unterscheiden sich die ersten zehn Dokumente einer experimentell erhobenen SERP, unabhängig von der Rangposition, nicht von denen einer anonym erhobenen, liegt der Jaccard-Koeffizient bei 1. Sind neun von zehn Dokumenten gleich, bei 0,9, usw.

6.1.2 Deskriptive Statistik

Der Zeitraum der Datenerhebung lag bei einer Dauer von 90 Tagen und fand vom 10. Juni bis zum 09. September 2019 statt. In dieser Zeit wurden an jedem Tag pro Account 40 Signale gesetzt. Nach Abschicken einer zuvor eingegebenen Suchanfrage wurde ein entsprechendes Dokument ausgewählt. Insgesamt wurden im Laufe der Untersuchungsdauer pro Account 3600 Signale gesetzt. Jede Woche wurden die personalisierten SERPs mit den anonymen verglichen und die Abweichungen ermittelt. Dabei wurden jeweils die ersten 50 organischen Ergebnisse pro Suchanfrage von jeder SERP gesammelt und in die Auswertung aufgenommen. Die Tabellen 10, 11 und 12 zeigen die Abweichungen je Themensektor und Suchanfrage im Vergleich mit den Ergebnissen der dazugehörigen anonymen Recherche. Je niedriger der Wert der Dokumente mit gleicher URL auf gleichem Rang, desto stärker wurden in diesem Vergleich die Personalisierungseffekte betrachtet.

Tabelle 10: Personalisierungseffekte Sektor Energie

	Sektor Energie					
	Account 1			Account 2		
	Dokumente mit gleicher URL auf gleichem Rang			Dokumente mit gleicher URL auf gleichem Rang		
Anzahl der Signale	Umsetzung Energiewende	Ausstieg Kohle- stromung	Ausbau Windkraft	Umsetzung Ener- giewende	Ausstieg Kohle- verstromung	Ausbau Windkraft
280	16	14	12	15	50	11
560	50	47	50	38	46	50
840	50	48	30	50	50	50
1120	50	48	40	45	50	40
1400	5	36	14	38	36	16
1680	45	50	50	36	50	50
1800	50	50	40	50	50	50

2040	33	28	50	33	50	50
2320	16	28	50	16	50	50
2600	50	5	4	42	46	32
2880	50	19	26	19	43	42
3160	25	25	8	46	48	48
3440	14	26	34	11	50	50
3600	50	49	43	47	50	50

Tabelle 11: Personalisierungseffekte Sektor Verkehr

		Sektor Verkehr					
		Account 3			Account 4		
		Dokumente mit gleicher URL auf gleichem Rang			Dokumente mit gleicher URL auf gleichem Rang		
Anzahl der Signale	<i>Umweltbilanz Auto</i>	<i>Tempolimit</i>	<i>Dieselfahrverbot</i>	<i>Umweltbilanz Auto</i>	<i>Tempolimit</i>	<i>Dieselfahrverbot</i>	
280	33	48	48	50	19	45	
560	50	33	50	50	10	41	
840	50	50	50	50	50	50	
1120	48	47	50	48	50	45	
1400	46	5	50	46	18	50	
1680	11	50	50	11	47	50	
1800	50	34	50	50	18	50	
2040	50	50	50	50	50	50	
2320	50	50	50	50	50	50	
2600	50	50	50	50	48	50	
2880	46	50	48	44	49	48	
3160	48	19	10	46	31	48	
3440	46	46	50	46	37	50	
3600	44	50	50	46	50	50	

Tabelle 12: Personalisierungseffekte Sektor Landwirtschaft

	Sektor Landwirtschaft					
	Account 5			Account 6		
	Dokumente mit gleicher URL auf gleichem Rang			Dokumente mit gleicher URL auf gleichem Rang		
Anzahl der Signale	<i>Gentechnik</i>	<i>Tierhaltung</i>	<i>Düngeverordnung</i>	<i>Gentechnik</i>	<i>Tierhaltung</i>	<i>Düngeverordnung</i>
280	50	50	49	36	32	50
560	50	38	38	33	38	50
840	32	19	44	35	33	44
1120	35	37	48	35	37	50
1400	50	50	50	50	14	50
1680	50	16	23	50	9	48
1800	35	50	50	35	50	50
2040	50	40	50	50	50	50
2320	50	50	50	50	50	50
2600	16	50	50	16	50	50
2880	32	36	50	30	37	42
3160	20	40	46	21	47	46
3440	38	23	50	13	6	48
3600	46	31	50	14	12	37

Um eine erste Orientierung zu Abweichungen auf den Top-Positionen zu erlangen, eignet sich ein Blick auf die Jaccard-Koeffizienten, die für jeden Account pro Vergleichszeitpunkt ermittelt wurden. Dabei wurden die ersten zehn organischen Ergebnisse der personalisierten und der anonymen SERPs untersucht. Sind die zehn Dokumente auf beiden Ergebnislisten dieselben, ergibt sich der maximale Wert eines Jaccard-Koeffizienten von 1, ohne daraus Hinweise auf potenzielle Unterschiede in der Reihung der Ergebnisse zu bekommen. Die Abbildungen 1-6 zeigen aber, dass bereits beim ersten Vergleich Unterschiede auf den Top-Positionen hinsichtlich differierender Dokumente zum Vorschein kamen.

Sektor Energie

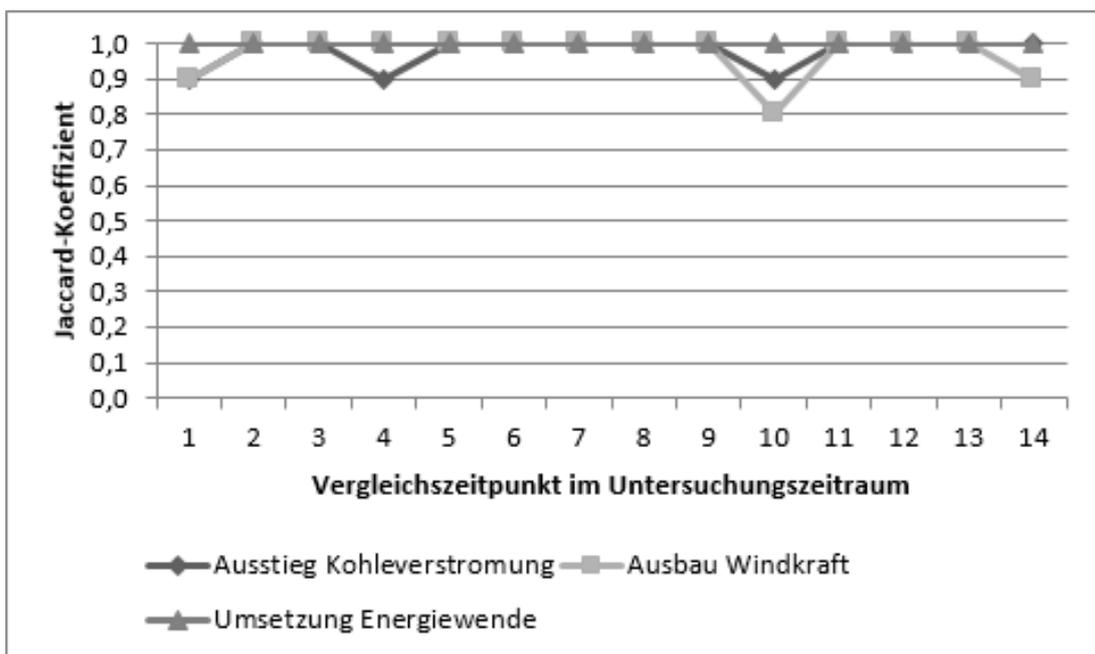


Abbildung 1: Verlauf der Jaccard-Koeffizienten Account 1

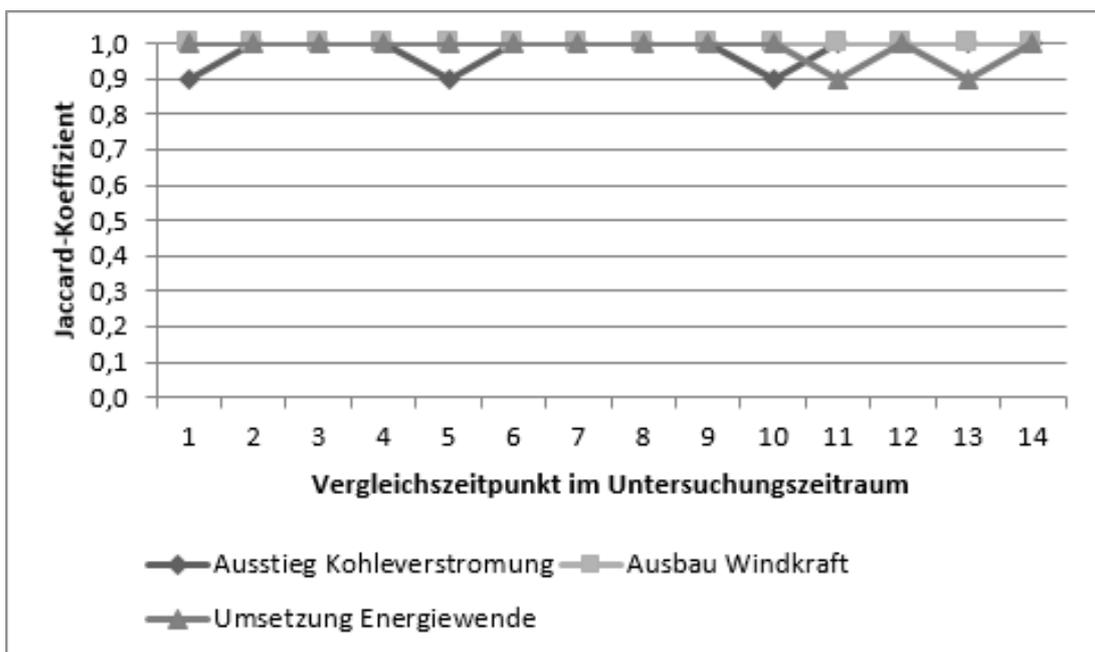


Abbildung 2: Verlauf der Jaccard-Koeffizienten Account 2

Sektor Verkehr

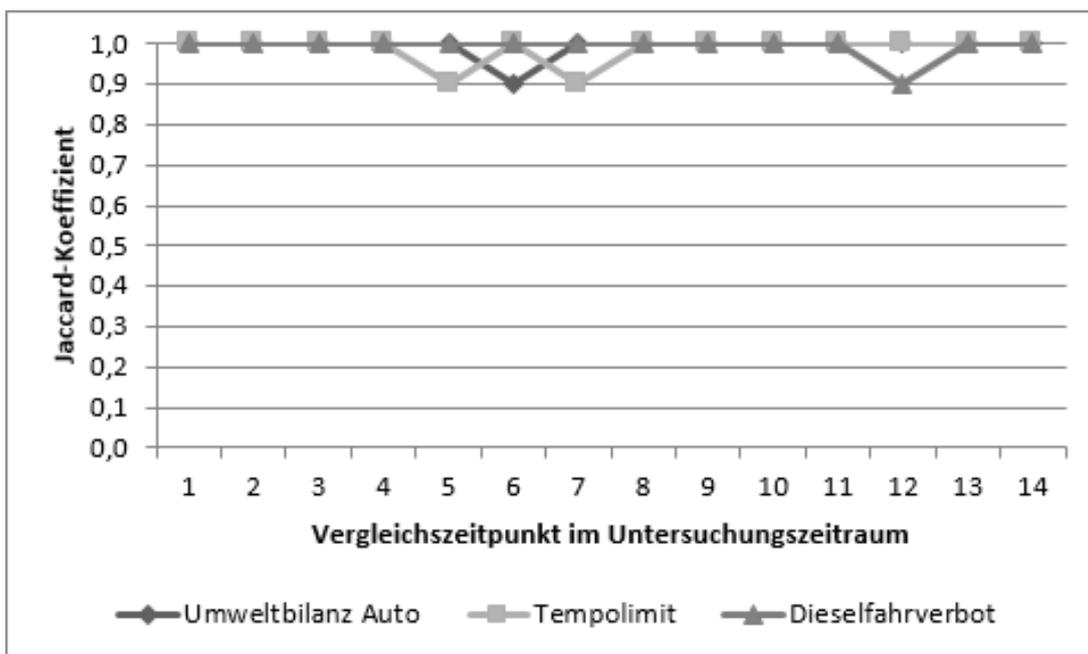


Abbildung 3: Verlauf der Jaccard-Koeffizienten Account 3

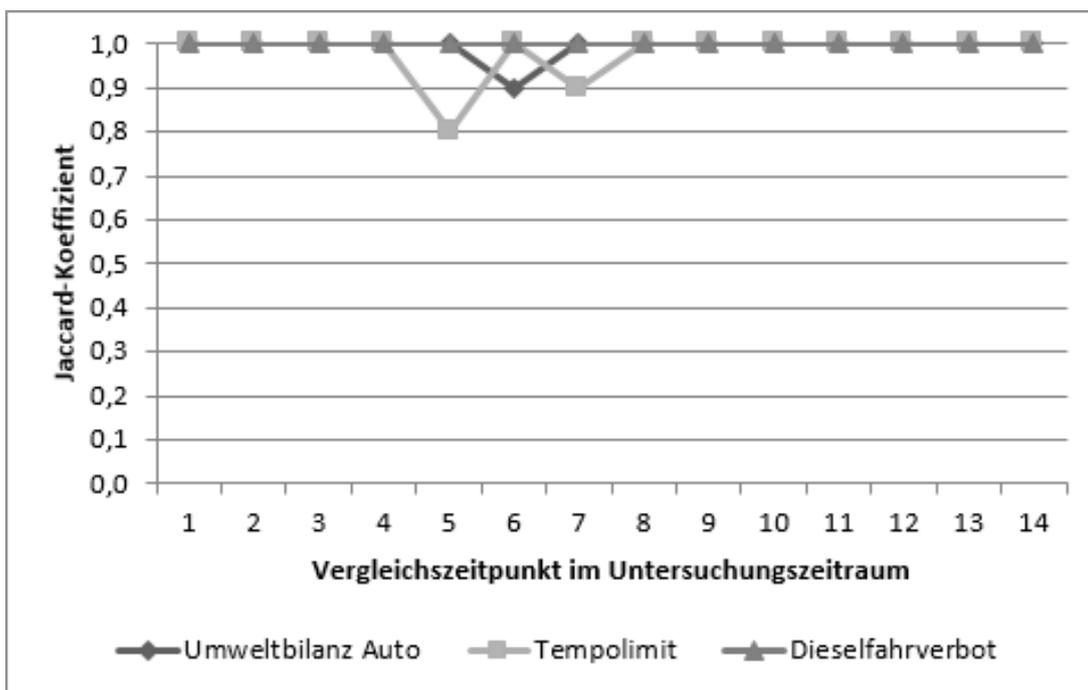


Abbildung 4: Verlauf der Jaccard-Koeffizienten Account 4

Sektor Landwirtschaft

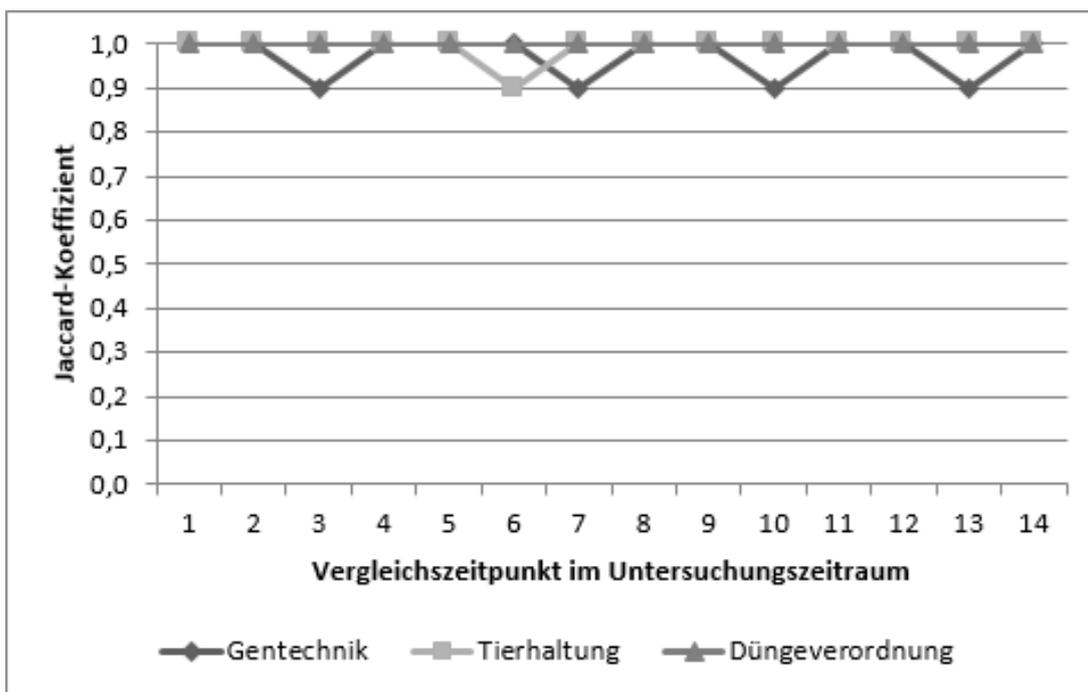


Abbildung 5: Verlauf der Jaccard-Koeffizienten Account 5

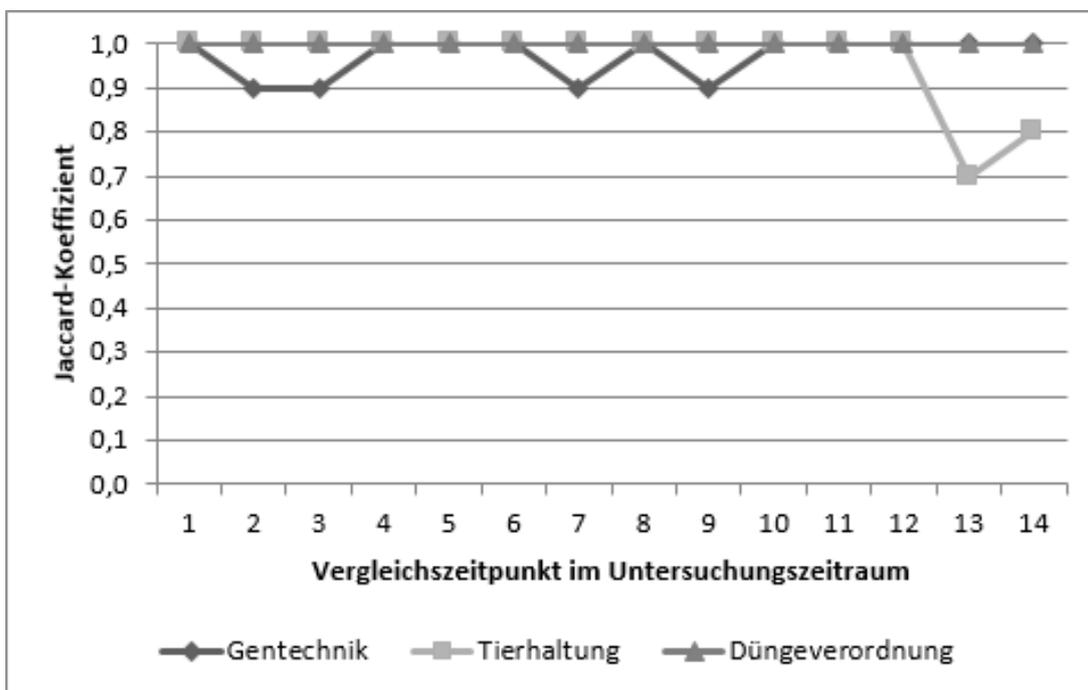


Abbildung 6: Verlauf der Jaccard-Koeffizienten Account 6

Wie bereits erwähnt, gibt der Jaccard-Koeffizient keinerlei Aufschlüsse über die Rangposition der Dokumente. Deshalb wurde, wie bei den ersten 50 Treffern, ebenfalls die Anzahl derselben Dokumente auf derselben Rangposition betrachtet (dargestellt in den Tabellen 13, 14 und 15).

Tabelle 13: Personalisierungseffekte Top-Positionen - Sektor Energie

		Sektor Energie					
		Account 1			Account 2		
		Dokumente mit gleicher URL auf gleichem Rang			Dokumente mit gleicher URL auf gleichem Rang		
Anzahl der Signale	Umsetzung	Ausstieg Kohle-	Ausbau	Umsetzung Ener-	Ausstieg Kohle-	Ausbau Wind-	
	Energie- wende	verstromung	Windkraft	giewende	verstromung	kraft	
280	6	3	9	10	4	10	
560	10	10	10	6	10	10	
840	10	8	8	10	10	10	
1120	10	10	8	10	10	10	
1400	5	4	8	10	6	8	
1680	10	10	10	6	10	10	
1800	10	10	10	10	10	10	
2040	6	10	10	6	10	10	
2320	6	10	10	6	10	10	
2600	10	4	5	10	4	8	
2880	10	10	8	6	8	10	
3160	8	8	10	10	8	10	
3440	8	10	10	7	10	10	
3600	10	8	9	10	10	10	

Tabelle 14: Personalisierungseffekte Top-Positionen - Sektor Verkehr

	Sektor Verkehr					
	Account 3			Account 4		
	Dokumente mit gleicher URL auf gleichem Rang			Dokumente mit gleicher URL auf gleichem Rang		
Anzahl der Signale	<i>Umweltbilanz Auto</i>	<i>Tempolimit</i>	<i>Dieselfahrverbot</i>	<i>Umweltbilanz Auto</i>	<i>Tempolimit</i>	<i>Dieselfahrverbot</i>
280	10	10	10	10	10	7
560	10	10	10	10	9	10
840	10	10	10	10	10	10
1120	8	4	10	8	10	10
1400	10	10	10	10	5	10
1680	5	5	10	5	10	10
1800	10	10	10	10	5	10
2040	10	10	10	10	10	10
2320	10	10	10	10	10	10
2600	10	10	10	10	10	10
2880	8	10	8	6	10	8
3160	10	10	9	10	10	10
3440	10	6	10	10	6	10
3600	10	10	10	10	10	10

Tabelle 15: Personalisierungseffekte Top-Positionen - Sektor Landwirtschaft

	Sektor Landwirtschaft					
	Account 5			Account 6		
	Dokumente mit gleicher URL auf gleichem Rang			Dokumente mit gleicher URL auf gleichem Rang		
Anzahl der Signale	<i>Gentechnik</i>	<i>Tierhaltung</i>	<i>Düngeverordnung</i>	<i>Gentechnik</i>	<i>Tierhaltung</i>	<i>Düngeverordnung</i>
280	10	10	10	10	10	10
560	10	10	6	4	10	10
840	6	7	10	7	10	10
1120	10	7	10	10	7	10
1400	10	10	10	10	4	10
1680	10	5	10	10	4	10
1800	9	10	10	9	10	10
2040	10	10	10	10	10	10
2320	10	10	10	10	10	10
2600	9	10	10	9	10	10
2880	10	10	10	10	10	8
3160	10	10	6	10	10	6
3440	10	10	10	10	2	10
3600	10	10	10	8	2	8

6.1.3 Zusammenhang zwischen gesetzten Signalen im Profil und Personalisierungseffekten

6.1.3.1 Analyseverfahren: Korrelationsanalyse nach Spearman

Nachfolgend soll der Zusammenhang zwischen den gesetzten Signalen im Profil und den Personalisierungseffekten auf der SERP statistisch überprüft werden.

Um den Einfluss des Suchverhaltens auf die Anordnung der Ergebnisse zu überprüfen, wurden Korrelationsanalysen durchgeführt. Als Analyseverfahren eignet sich in diesem Fall die Korrelationsanalyse nach Spearman, die einen linearen Zusammenhang zwischen zwei mindestens ordinalskalierten Variablen berechnet. Auf die

zugrundeliegende Hypothese bezogen wird eine negative Korrelation erwartet. Eine negative Korrelation tritt auf, wenn hohe Werte der einen Variable mit niedrigen Werten der anderen Variable einhergehen (Universität Zürich, 2018b). Zur Korrelationsanalyse nach Spearman sei jedoch angemerkt, dass damit keine Kausalbeziehungen zwischen den Variablen hergestellt werden können. Der Rangkoeffizient liefert lediglich Aussagen über Art und Intensität der Kovariation der Variablen (Döring & Bortz, 2016, S. 677). Die Rangkorrelation nach Spearman wird oftmals auch Spearman-Korrelation oder Spearman-Rho genannt. Die Berechnung des Koeffizienten ρ (rho) basiert nicht auf den tatsächlich erhobenen Messdaten. Diese werden anhand ihrer Größe durch Ränge ersetzt, nach denen letztlich gerechnet wird (Universität Zürich, 2018b).

6.1.3.2 Ergebnisse der Hypothesenprüfung

Zur statistischen Auswertung wurde die Statistik-Software SPSS verwendet. Das Signifikanzniveau α wurde dabei auf 5% fixiert (Döring & Bortz, 2016, S. 665). Zum Verständnis sind nochmals die Forschungsfrage und die Hypothese formuliert:

F1: Gibt es bei der personalisierten Google-Recherche nach kontrovers diskutierten Umweltthemen einen Zusammenhang zwischen der Anzahl gesetzter impliziter Signale und der Anordnung der Ergebnisse?

H1: Je mehr implizite Daten in Form von Suchanfragen und anschließenden Ergebnisklicks in das persönliche Profil einfließen, umso mehr Abweichungen gibt es im Vergleich zur anonymen Suche in der Reihung der Ergebnisse.

Unabhängig vom Inhalt und von den inhaltlichen Eingaben auf der unabhängigen Variablen sollten zunächst ganz allgemein Aussagen zu einem möglichen Zusammenhang zwischen den Variablen gemacht werden. Dazu wurden die Daten sämtlicher Accounts (n=252) in einer Korrelationsanalyse zusammengerechnet. Die Zahl ergab sich aus den sechs Profilen, in denen eingeloggt insgesamt 14 Mal nach drei Suchbegriffen recherchiert wurde. Zu H1 ergab die Rangkorrelation nach Spearman keinen signifikanten Zusammenhang zwischen den gesetzten Signalen und der Anzahl der gleichen Dokumente auf dem gleichen Rang ($\rho=-.047$, $p=.456$). Die Nullhypothese bleibt somit bestehen, die Alternativhypothese muss verworfen werden.

Neben der Gesamtbetrachtung wurden die Variablen noch auf Zusammenhänge in den einzelnen Sektoren (n=84) und den Suchanfragen (n=28) geprüft. Es wurden dabei

zunächst die Daten beider Accounts je Sektor und danach die Daten auf Ebene der einzelnen Suchanfragen gerechnet. Die Ergebnisse der Korrelationsanalysen sind in Tabelle 16 festgehalten.

Tabelle 16: Ergebnisse Korrelationsanalysen nach Spearman

Energie (n=84) $\rho=-.042, p=.707$			Verkehr (n=84) $\rho=.021, p=.851$			Landwirtschaft (n=84) $\rho=-.143, p=.195$		
Umsetzung Energie- wende (n=28) $\rho=-.073,$ $p=.712$	Ausstieg Kohlever- stromung (n=28) $\rho=.051,$ $p=.798$	Ausbau Wind- kraft (n=28) $\rho=-.081,$ $p=.682$	Umwelt- bilanz Auto (n=28) $\rho=-.276,$ $p=.155$	Tempoli- mit (n=28) $\rho=.230,$ $p=.668$	Diesel- fahrver- bot (n=28) $\rho=.141,$ $p=.475$	Gentech- nik (n=28) $\rho=-.443,$ $p=.018$	Tierhal- tung (n=28) $\rho=-.085,$ $p=.668$	Dünge- verord- nung (n=28) $\rho=.019,$ $p=.924$

Aus Tabelle 16 ist abzuleiten, dass auch innerhalb der einzelnen Sektoren Energie, Verkehr und Landwirtschaft keine signifikanten Zusammenhänge zwischen den Variablen festzustellen waren. Innerhalb der Suchanfragen gab es lediglich bei der Recherche nach dem Begriff „Gentechnik“ einen signifikanten Zusammenhang zwischen den gesetzten Signalen im Profil und den Personalisierungseffekten auf den erhobenen SERPs ($\rho=-.443, p=.018$). Die Korrelation ist wie erwartet negativ. Das heißt, dass die Anzahl der gleichen Dokumente auf dem gleichen Rang (im Vergleich der Experimental-SERP mit der Kontrollbedingung) anfangs noch als hoch vermutet wurde, je mehr Signale in das Profil jedoch einfließen, desto geringer wird die Anzahl. Der ermittelte Effekt ist nach einer Einteilung nach Cohen (1992, S. 80) als mittelstark einzuschätzen.

Die Verteilung der Daten der Korrelationsanalyse in der Gesamtbetrachtung ist in Abbildung 7 visualisiert. Die nicht lineare Verteilung unterstreicht das Ergebnis, dass die Variablen nicht signifikant im Zusammenhang stehen.

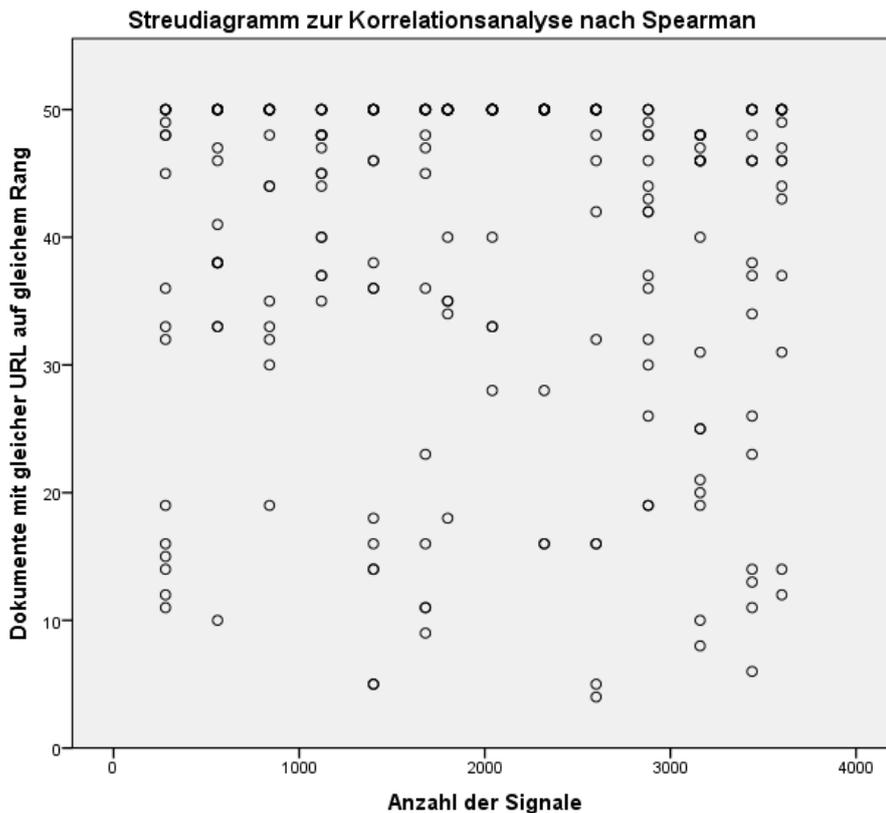


Abbildung 7: Streudiagramm zur Korrelationsanalyse nach Spearman

Zusätzlich zur Korrelationsanalyse mit dem gesamten Datenmaterial wurde außerdem die Anzahl der im Vergleich zur Kontroll-SERP ermittelten abweichenden Dokumente auf den Top-Positionen ($n=10$) und die Anzahl der gesetzten Signale auf statistisch signifikante Zusammenhänge überprüft. Mit Blick auf die bisherigen Ergebnisse konnte nach der Korrelationsanalyse nach Spearman erwartungsgemäß ebenfalls kein signifikanter Zusammenhang ermittelt werden ($\rho=.001$, $p=.984$). In den einzelnen Sektoren ergibt sich ein ähnliches Bild: In keinem Themenfeld wurden auf den ersten zehn Positionen signifikante Zusammenhänge zwischen den Variablen nachgewiesen (Energie: $\rho=.022$, $p=.843$; Verkehr: $\rho=.013$, $p=.910$; Landwirtschaft: $\rho=-.021$, $p=.846$).

6.2 Inhaltliche Bewertung der Dokumente

6.2.1 Kennzahlen

Die ersten zehn Dokumente auf den untersuchten SERPs wurden inhaltlich bewertet. Dies wurde von drei Juroren vorgenommen, die die Dokumente auf einer Skala von -2 bis +2 bewerteten. Dabei wurde bei der Entwicklung der Skala darauf geachtet, die

einzelnen Ausprägungen gemäß den Themen zu definieren, sodass die Juroren durch Instruktionen einen Bewertungsrahmen erhielten (Skala und die Definition der Ausprägungen in einer Tabelle darstellen). Da die einzelnen Dokumente nicht immer trennscharf zu bewerten waren und es vermutlich deshalb nicht immer zu identischen Bewertungen kam, wurde der Mittelwert aus den Bewertungs-Scores daraus gezogen. Um den inhaltlichen Wert einer SERP in einer Kennzahl zu ermitteln, wurde zunächst nach der Formel des Discounted Cumulative Gain (DCG) gerechnet, berechnet mit folgender Formel:

$$DCG@k = \sum_{i=1}^k \frac{2^{rel(i)} - 1}{\log_2(i + 1)}$$

Der DCG ist eine Kennzahl aus dem Information Retrieval, die zur Messung der Qualität von Ergebnislisten in Suchsystemen herangezogen wird. Entwickelt von Järvelin und Kekäläinen (2000), berücksichtigt die Berechnung des DCG neben der Relevanzbewertung eines Dokuments zu einer Suchanfrage auch die Rangposition. Jedes Dokument auf einer Ergebnisliste erhält somit einen Score, die Summe aus allen Dokumenten einer Liste ergibt den DCG:

"In this evaluation, the relevance level of each document is somehow used as a gamed value measure for its ranked position m the result and the gain is summed progressively from position 1 to n . Thus the ranked document lists (of some determined length) are turned to gamed value lists by replacing document IDs by their relevance values" (Järvelin & Kekäläinen, 2000, S. 42).

Die Berechnungsmethode des DCG wurde für die inhaltliche Bewertung der Dokumente durch die Juroren herangezogen. Statt der Relevanzurteile werden die mit der entwickelten Skala ermittelten Werte der inhaltlichen Dokumentenbewertung in die Formel übertragen. Die Berechnung des DCG wurde also auf das Vorhaben in dieser Arbeit adaptiert, sodass die entsprechenden Dokumente inhaltlich bewertet werden konnten und auch die Rangposition des Dokuments in die Kennzahl miteinfließt. Da die Werte auf der Skala in fünf Ausprägungen von -2 bis +2 definiert sind, wurden die Bewertungs-Scores entsprechend normalisiert, sodass Werte zwischen -1 und +1 entstehen. Würden alle Dokumente einer Liste mit -2 bewertet, so läge der entsprechende Score bei -1. So verhält es sich auch bei Bewertungen mit positiven Zahlen. Die negativen Werte wurden beibehalten, um die Diskrepanzen zwischen den

Positionen stärker hervorzuheben. Die normalisierten Abstufungen des Scores lassen sich gemäß der fünf Ausprägungen der Skala folgendermaßen interpretieren:

- Maßnahmen Energiewende/Verkehrswende/Agrarwende stark ablehnend: Werte zwischen -1 und $< -0,6$
- Maßnahmen Energiewende/Verkehrswende/Agrarwende ablehnend: Werte zwischen -0,6 und $< -0,2$
- Neutrale Bewertung: Werte zwischen -0,2 und $< +0,2$
- Maßnahmen Energiewende/Verkehrswende/Agrarwende befürwortend: Werte zwischen +0,2 und $< +0,6$
- Maßnahmen Energiewende/Verkehrswende/Agrarwende stark befürwortend: Werte zwischen +0,6 und < 1

Dokumente, die nicht bewertet werden konnten, wurden aus der Auswertung genommen. Dies betraf beispielsweise Dokumente, vor denen eine Bezahlschranke lag. Auch Dokumente, die offensichtlich nicht themenrelevant waren, fanden keine Berücksichtigung und wurden aus der jeweiligen SERP entfernt.

6.2.2 Deskriptive Statistik

Im Laufe der 90 Tage wurden aus allen Bedingungen 183 zu bewertende Dokumente gesammelt. Zwei davon wurden aus der Auswertung ausgenommen. Zur Suchanfrage „umsetzung energiewende“ ein Dokument der Energiate GmbH, einem Special-Interest-Verlag für energiewirtschaftliche Themen, über dem eine Bezahlschranke lag. Und im Sektor Landwirtschaft für die Suchanfrage „tierhaltung“ ein Dokument der Stadt Hamburg, das Informationen für Mieter zur Haustierhaltung bereitstellte. Der Inhalt wurde als nicht themenrelevant eingestuft. Inhaltlich bewertet wurden nur die ersten zehn Dokumente je SERP. Dabei fielen im gesamten Untersuchungszeitraum 64 Dokumente unter den Sektor Energie, 67 Dokumente unter den Sektor Verkehr und 50 Dokumente wurden dem Sektor Landwirtschaft zugerechnet. Insgesamt 43 Dokumente erschienen über den gesamten Zeitraum bei allen 14 Vergleichen auf den Top-Positionen, unabhängig von der Bedingung. Dies entspricht einem relativen Anteil von 23,6 Prozent. Diese konstanten Ergebnisse waren relativ gleichmäßig auf die Themenbereiche verteilt (Energie: $n=13$; Verkehr: $n=17$; Landwirtschaft: $n=13$). Die Tabellen 17, 18 und 19 zeigen die Häufigkeiten der Dokumente auf den Top-Positionen je Sektor. In der Spalte ganz links ist der Wert angegeben, wie oft ein Dokument in

zumindest einer Bedingung aufgetreten ist. In der Spalte rechts daneben, wie viele Dokumente dies betraf.

Tabelle 17: Häufigkeiten von Dokumenten auf den Top-Positionen - Energie

Sektor Energie					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1	14	21,9	21,9	21,9
	2	7	10,9	10,9	32,8
	3	7	10,9	10,9	43,8
	4	4	6,3	6,3	50,0
	5	1	1,6	1,6	51,6
	6	2	3,1	3,1	54,7
	7	2	3,1	3,1	57,8
	8	3	4,7	4,7	62,5
	9	3	4,7	4,7	67,2
	10	2	3,1	3,1	70,3
	11	1	1,6	1,6	71,9
	12	3	4,7	4,7	76,6
	13	2	3,1	3,1	79,7
	14	13	20,3	20,3	100,0
Gesamt		64	100,0	100,0	

Tabelle 18: Häufigkeiten von Dokumenten auf den Top-Positionen - Verkehr

Sektor Verkehr					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1	21	31,3	31,3	31,3
	2	5	7,5	7,5	38,8
	3	7	10,4	10,4	49,3
	4	1	1,5	1,5	50,7
	5	2	3,0	3,0	53,7

	6	5	7,5	7,5	61,2
	7	3	4,5	4,5	65,7
	9	3	4,5	4,5	70,1
	11	1	1,5	1,5	71,6
	13	2	3,0	3,0	74,6
	14	17	25,4	25,4	100,0
	Gesamt	67	100,0	100,0	

Tabelle 19:Häufigkeiten von Dokumenten auf den Top-Positionen - Landwirtschaft

Sektor Landwirtschaft					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1	11	22,0	22,0	22,0
	2	2	4,0	4,0	26,0
	3	4	8,0	8,0	34,0
	4	3	6,0	6,0	40,0
	6	1	2,0	2,0	42,0
	7	1	2,0	2,0	44,0
	8	1	2,0	2,0	46,0
	9	1	2,0	2,0	48,0
	10	1	2,0	2,0	50,0
	11	1	2,0	2,0	52,0
	12	5	10,0	10,0	62,0
	13	6	12,0	12,0	74,0
	14	13	26,0	26,0	100,0
		Gesamt	50	100,0	100,0

Aus den Tabellen lässt sich ableiten, dass einige Dokumente konstant auf den ersten Positionen platziert sind. Im Sektor Landwirtschaft ist beispielsweise die Hälfte der gesammelten Dokumente zehn Mal oder öfter aufgetreten. Diese konstant auftretenden Ergebnisse werden näher in der Auswertung der Quellen betrachtet.

Die Bewertungen der Inhalte wurden auf zwei Nachkommastellen gerundet. Sie zeigen im Durchschnitt eine neutrale Bewertung der Dokumente. Der Mittelwert beträgt 0,22 (n=181), wobei hier jedoch angemerkt sei, dass die Rangposition noch nicht berücksichtigt ist. Die Durchschnittswerte der einzelnen Dokumente sind nicht normalisiert, weswegen die Werte potenziell zwischen -2 und +2 liegen können. Betrachtet man die Abstufungen auf der Skala, sind die Dokumente im Mittel überwiegend neutral. Auf die Sektoren bezogen ergab sich im Bereich Energie ein Durchschnittswert von 0,30 (n=64), im Bereich Verkehr ein Wert von 0,14 (n=67) und im Bereich Landwirtschaft ein Wert von 0,21 (n=50). Zusätzlich zur Einteilung in die Sektoren wurde das arithmetische Mittel der Bewertungen der Dokumente zu den einzelnen Suchanfragen erhoben, dargestellt in Tabelle 20. Auch hier liegen die Werte um den neutralen Mittelpunkt. Die Dokumente zu den Suchanfragen „umsetzung energiewende“ und „tierhaltung“ erhielten eine durchschnittliche Bewertung von 0,45 und zeigen somit leichte Tendenzen hinsichtlich befürwortender Inhalte im Kontext der Energiewende bzw. der Agrarwende, sind aber den Abstufungen nach noch als neutral zu sehen.

Tabelle 20: Durchschnittliche Bewertungen der Dokumente zu den Suchanfragen

Energie			Verkehr			Landwirtschaft		
Umsetzung Energie- wende (n=20), Ø0,45	Ausstieg Kohlever- stromung (n=18), Ø0,35	Ausbau Windkraft (n=26), Ø0,15	Umweltbi- lanz Auto (n=15), Ø0,27	Tempoli- mit (n=27), Ø0,22	Dieselfahr- verbot (n=25), Ø-0,03	Gentech- nik (n=19), Ø0,23	Tierhal- tung (n=14), Ø0,45	Dünge- verord- nung (n=17), Ø0,00

Neben den Bewertungen für die einzelnen Dokumente wurden auch die Durchschnittswerte der gesammelten SERPs ausgewertet. Hierbei sei nochmal angemerkt, dass die Dokumentenbewertungen mit der Rangposition verrechnet und entsprechend normalisiert wurden und sich somit ein Bewertungs-Score pro SERP ergibt, der zwischen -1 und +1 liegt. Insgesamt wurden 378 SERPs betrachtet. Die Zahl ergibt sich aus den 14 Vergleichen, die mit der Zahl der Bedingungen (3) und der Zahl der Suchanfragen (9) multipliziert wurde. Je Sektor gab es 126 SERPs, die einen Wert erhielten. Je Suchanfrage also 42 SERPs und je Bedingung 14. Der durchschnittliche Bewertungs-Score aller SERPs aus allen Bedingungen liegt bei 0,10. Die Ergebnislisten in den Sektoren (n=126) erhielten die Bewertungs-Scores 0,16 (Energie), 0,05 (Verkehr)

und 0,09 (Landwirtschaft). Die durchschnittlichen Kennzahlen der SERPs zu einzelnen Suchanfragen sind in Tabelle 21 dargestellt.

Tabelle 21: Durchschnittliche Bewertung aller SERPs zu den Suchanfragen

Energie			Verkehr			Landwirtschaft		
Umsetzung Energie-wende (n=42), Ø0,30	Ausstieg Kohleverstromung (n=42), Ø0,07	Ausbau Windkraft (n=42), Ø0,10	Umweltbilanz Auto (n=42), Ø0,09	Tempolimit (n=42), Ø0,09	Dieselfahrverbot (n=42), Ø-0,02	Gentechnik (n=42), Ø0,06	Tierhaltung (n=42), Ø0,22	Düngeverordnung (n=42), Ø0,00

Auch hier sind die Kennzahlen überwiegend neutral ausgefallen. Bei den Suchanfragen „umsetzung energiewende“ und „tierhaltung“ konnten, den durchschnittlichen Scores nach, die Energiewende bzw. die Agrarwende befürwortende Inhalte für die SERPs ermittelt werden. Die nachfolgenden Tabellen zeigen die Verteilungen der Listen-Scores für jeden Vergleich pro Account und je Kontrollbedingung. Die Werte der SERPs, deren Kennzahl von der neutralen Mitte abweicht (d. h. Werte außerhalb des Raumes zwischen -0,2 und 0,2), sind gefettet dargestellt. In den Experimentalbedingungen, hierauf soll nochmals verwiesen werden, wurden die Ausprägungen der unabhängigen Variable nach der Hälfte umgekehrt.

Sektor Energie

Tabelle 22: Bewertungs-Scores der SERPs Account 1 - Experimentalbedingung Energie

Suchanfrage/Vergleich	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Umsetzung Energiewende	0,40	0,39	0,29	0,34	0,26	0,27	0,28	0,30	0,28	0,25	0,27	0,30	0,28	0,30
Ausstieg Kohleverstromung	0,04	0,08	-0,04	0,02	0,10	0,09	0,02	0,12	0,11	0,05	0,12	0,09	0,13	0,05
Ausbau Windkraft	0,12	0,18	0,08	0,01	0,05	0,01	0,04	0,08	0,08	0,15	0,16	0,11	0,17	0,14

Tabelle 23: Bewertungs-Scores der SERPs Account 2 - Experimentalbedingung Energie

Suchanfrage/Vergleich	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Umsetzung Energiewende	0,40	0,39	0,29	0,34	0,27	0,27	0,28	0,28	0,28	0,25	0,27	0,30	0,28	0,30
Ausstieg Kohleverstromung	0,04	0,04	-0,04	0,02	0,09	0,09	0,02	0,12	0,12	0,06	0,13	0,14	0,14	0,06
Ausbau Windkraft	0,23	0,19	0,14	0,00	0,04	0,01	0,04	0,07	0,07	0,13	0,15	0,11	0,16	0,14

Tabelle 29: Bewertungs-Scores der SERPs Account 6 - Experimentalbedingung Landwirtschaft

Suchanfrage/Vergleich	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Gentechnik	0,05	0,15	0,10	0,10	-0,03	0,01	0,05	0,05	0,04	0,05	0,05	0,09	0,01	0,10
Tierhaltung	0,15	0,24	0,18	0,18	0,28	0,25	0,16	0,22	0,26	0,23	0,22	0,25	0,26	0,23
Düngeverordnung	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Tabelle 30: Bewertungs-Scores der SERPs - Kontrollbedingung Landwirtschaft

Suchanfrage/Vergleich	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Gentechnik	0,05	0,10	0,15	0,10	-0,03	0,01	0,10	0,05	0,04	0,09	0,05	0,09	0,01	0,09
Tierhaltung	0,15	0,24	0,18	0,17	0,26	0,16	0,16	0,22	0,26	0,23	0,22	0,25	0,25	0,23
Düngeverordnung	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

6.2.3 Überprüfung der anonymen und personalisierten Suchergebnisse auf inhaltliche Unterschiede

6.2.3.1 Analyseverfahren: einfaktorielle Varianzanalyse

Um signifikante inhaltliche Unterschiede zwischen den beiden Experimental- und der Kontrollbedingung statistisch zu überprüfen, eignet sich das Verfahren der einfaktoriellen Varianzanalyse, auch ANOVA genannt (aus dem Englischen „Analysis of Variance“). Damit können Unterschiede zwischen den Mittelwerten aus mehr als zwei Stichproben auf Signifikanz getestet werden (Sedlmeier & Renkewitz, 2018, S. 430). Konkret bedeutet dies nach einer Definition der Methodenberatung der Universität Zürich (2018a): „Die einfaktorielle Varianzanalyse [...] testet, ob sich die Mittelwerte mehrerer unabhängiger Gruppen (oder Stichproben) unterscheiden, die durch eine kategoriale unabhängige Variable definiert werden.“ Für dieses Prüfverfahren müssen jedoch einige Voraussetzungen gegeben sein (ebd.). Zunächst ist auf die Skalierung der Variablen zu achten. Die abhängige Variable muss intervallskaliert sein. Diese Voraussetzung wurde in der vorliegenden Studie mit der Entwicklung der Skala erfüllt. Die unabhängige Variable, auch Faktor genannt, muss nominal- oder kardinalskaliert sein. Diese Voraussetzung ist ebenfalls erfüllt. Die dritte Voraussetzung, die erfüllt wurde, ist die Unabhängigkeit zwischen den einzelnen Gruppen, zwischen denen die Unterschiede ermittelt werden sollen. In dieser Arbeit hat die unabhängige Variable drei

verschiedene Bedingungen: Das manipulierte Rechercheverhalten ist so zu betrachten, dass es Nutzer simuliert, die umweltpolitischen Maßnahmen in den Sektoren Energie, Verkehr und Landwirtschaft entweder positiv (Bedingung 1) oder negativ (Bedingung 2) gegenüberstehen. In der dritten Bedingung wurde gab es vorab kein Treatment der unabhängigen Variable (Bedingung 3; Kontrollbedingung). Eine weitere Voraussetzung zur Durchführung der einfaktoriellen Varianzanalyse stellt die Normalverteilung der abhängigen Variablen dar, die aufgrund der Menge der erhobenen Daten anzunehmen ist. Die letzte zu erfüllende Voraussetzung ist die Homogenität der Varianzen. Damit soll sichergestellt werden, dass die Varianzen der abhängigen Variable der einzelnen Bedingungen etwa gleich sind. Die Methodenberatung der Universität Zürich (2018a) schlägt zur Überprüfung der Varianzhomogenität den Levene-Test vor. Dieser Test prüft, ob sich die Varianzen der Gruppen signifikant unterscheiden. Um die einfaktorielle Varianzanalyse durchführen zu können, ist ein nicht signifikanter Levene-Test die Voraussetzung, da dann von einer Varianzhomogenität ausgegangen werden kann (Universität Zürich, 2018a).

Zeigt die einfaktorielle Varianzanalyse signifikante Unterschiede zwischen den Bedingungen auf, so ist jedoch noch nicht geklärt, wo diese Unterschiede genau liegen, weshalb bei signifikanten Ergebnissen Post-hoc-Tests durchzuführen sind (Sedlmeier & Renkewitz, 2018, S. 453). Die Logik dahinter ist, dass damit sämtliche Mittelwerte einzeln miteinander verglichen werden. Somit lassen sich signifikante Unterschiede zwischen den einzelnen Bedingungen ermitteln. SPSS bietet eine Reihe solcher Post-hoc-Tests zur Durchführung an. Sedlmeier und Renkewitz (2018, S. 454) nennen in diesem Kontext die herkömmlichen Post-hoc-Tests: Scheffé-Test, Tukey-Test, Newman-Keuls-Methode und die Duncan-Methode.

6.2.3.2 Ergebnisse der Hypothesenprüfung

Zur statistischen Analyse der Suchergebnisse auf inhaltliche Unterschiede ist zunächst wieder die entsprechende Forschungsfrage mit der dazugehörigen Hypothese (F2 und H2) dargestellt.

F2: Unterscheiden sich die anonymen Suchergebnisse von den personalisierten Suchergebnissen bei Recherchen nach kontrovers diskutierten Umweltthemen inhaltlich?

H2: Zwischen den anonymen Suchergebnissen und den personalisierten Suchergebnissen in mindestens einer der beiden Experimentalgruppen gibt es signifikante inhaltliche Unterschiede, die sich durch weniger neutrale Ergebnisse in den personalisierten Ergebnislisten auszeichnen.

Die Daten zur statistischen Überprüfung der Hypothese stellen die normalisierten Kennzahlen aus den Bewertungen durch die Juroren dar. Es wurden insgesamt zwei Varianzanalysen durchgeführt. In die erste Analyse flossen die Daten aus der ersten Hälfte des Untersuchungszeitraumes mit ein, in der zweiten Analyse die Werte aus der zweiten Phase.

Die Voraussetzungen für die einfaktoriellen Varianzanalysen sind erfüllt. Dazu wurde vor der Durchführung die Varianzhomogenität jeweils mit dem Leven-Test überprüft. Die Ergebnisse der Tests waren nicht signifikant ($p=.477$; $p=.980$). Somit ist eine Varianzhomogenität gegeben.

Die Ergebnisse der einfaktoriellen Varianzanalyse in der ersten Hälfte des Untersuchungszeitraumes zeigen keine signifikanten Unterschiede zwischen den drei Bedingungen ($F(2,186)=.163$, $p=.850$). Auch für die zweite Phase konnten keine signifikanten Unterschiede zwischen den Experimentalbedingungen untereinander oder zwischen der Kontrollbedingung festgestellt werden ($F(2,186)=.016$, $p=.984$). Die Nullhypothese bleibt bestehen, die Alternativhypothese muss verworfen werden. Da keine signifikanten Unterschiede festgestellt wurden, ist die Durchführung eines Post-hoc-Tests nicht notwendig.

Mit diesen Ergebnissen kann an die dritte Forschungsfrage angeknüpft werden:

F3: Wie wirkt sich die im Suchverhalten implizite persönliche Haltung in einer Kontroverse bei einer Umkehr auf die gegenteilige Haltung in personalisierten Google-Recherchen zu Umweltthemen aus?

Da, wie angeführt, keine statistisch signifikanten Unterschiede zwischen den Bedingungen festgestellt wurden, kann die Frage damit beantwortet werden, dass eine Umkehr persönlicher Meinungen, die sich im Suchverhalten zeigen, keinen Einfluss auf die Suchergebnisse hat. Sowohl nicht im Profil eingeloggt als auch nicht während einer anonymen Recherche. Somit zeigen sich keine Tendenzen hinsichtlich potenzieller Filterblasen-Effekte. Die Ergebnisse werden ausführlich im Abschnitt der Diskussion aufgegriffen.

6.2.4 Auswertung der Quellen

Folgender Abschnitt beschäftigt sich mit der Auswertung der Quellen, die während der Studie auf den Top-Positionen platziert waren. Dazu wurde Forschungsfrage 4 formuliert:

F4: Welche Quellen werden Google-Nutzern in der Recherche nach kontrovers diskutierten Umweltthemen angezeigt?

Um die Frage zu beantworten, wurde aus den 181 bewerteten Dokumenten jeweils die Hauptdomain der URL extrahiert. Die Quellen wurden im Anschluss in Akteurskategorien eingeteilt, die sich aus der Voruntersuchung ergaben: Politische Akteure (z. B. Politiker, politische Parteien, Regierungen, staatliche Behörden), wirtschaftliche Akteure (z. B. Unternehmen), Medienakteure (klassische Medien aus Presse und Rundfunk, Internet-Portale), Interessensorganisationen (Stiftungen, Naturschutzverbände, Lobbyverbände, Vereine), wissenschaftliche Akteure (Hochschulen, private Forschungseinrichtungen) und Privatpersonen.

Die 181 Dokumente stammen insgesamt von 102 Quellen. Die am häufigsten gezeigten Quellen hatte die Hauptdomain zeit.de (9) und wird vom Medium Zeit Online herausgegeben. Dahinter hatten Wikipedia-Artikel (7) und Beiträge von Spiegel Online den zweithäufigsten Wert.

Tabelle 31: Die am häufigsten vorkommenden Quellen der Studie

Hauptdomain Quelle	Herausgeber	Akteur	Häufigkeit
adac.de	ADAC e. V.	Interessensor- ganisation	4
bmel.de	Bundesministerium für Ernährung und Landwirtschaft	Politischer Ak- teur	4
deutschland- funk.de	Deutschlandradio	Medienakteur	4
bmwi.de	Bundesministerium für Wirtschaft und Energie	Politischer Ak- teur	4
spiegel.de	Spiegel Online	Medienakteur	7
zeit.de	Zeit Online	Politischer Ak- teur	9
wikipedia.org	Wikimedia Foundation	Medienakteur	7
welt.de	Axel Springer SE	Medienakteur	4
tagesspiegel.de	Verlag der Tagesspiegel	Medienakteur	4
t-online.de	Ströer Digital Publishing GmbH	Politischer Ak- teur	4

Die am häufigsten auftretenden Quellen in der Gesamtbetrachtung sind in Tabelle 31 dargestellt. Darüber hinaus wurde wieder auf die einzelnen Sektoren geachtet. Im Themenbereich Energie waren die Quellen bmwi.de (4) und spiegel.de (4) am häufigsten auf den SERPs vertreten, gefolgt von ndr.de (3) mit dem Norddeutschen Rundfunk als Herausgeber, dem Portal windbranche.de (3) vom Internationalen Wirtschaftsforum Regenerative Energien (IWR), und mit bundestag.de (3) die Webseite des Deutschen Bundestages. Im Bereich Verkehr die bereits in Tabelle 31 genannten zeit.de (4), adac.de (4), t-online.de (4) und tagesspiegel.de (3). Im Sektor Landwirtschaft die Seiten bmel.de (4) und wikipedia.org (3).

Die Verteilung der Akteure ist nicht gleichmäßig. Wie die Häufigkeitstabelle darstellt, sind über die Hälfte aller Dokumente Medienakteuren zuzuschreiben. Auch Interessensorganisationen und politische Akteure sind häufig auf den Top-Positionen zu finden. Die anderen drei Akteursgruppen – Privatpersonen, wirtschaftliche und wissenschaftliche Akteure – treten auf den ersten zehn Plätzen bei Recherchen zu Umweltthemen relativ wenig auf. Interessensorganisationen, Medien und politische

Akteure kommen zusammen fast 90 Prozent der bewerteten Dokumente. Von Medienakteuren erstellte Inhalte haben in allen drei Sektoren die Mehrheit erreicht. Besonders im Bereich Verkehr (n=67) zeigt sich das mit fast 75 Prozent deutlich. In diesem Bereich waren demnach die anderen Akteursgruppen weniger vertreten. Besonders politische Akteure konnten ihre Inhalte im Vergleich zu den anderen Sektoren zu den gewählten Suchanfragen nicht im hohen Maße auf den Top-Positionen platzieren (Verkehr: 2,99 Prozent; Energie: 20,31 Prozent; Landwirtschaft: 24 Prozent).

Tabelle 32: Häufigkeitstabelle zu den verschiedenen Akteuren

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Interessensorganisation	37	20,4	20,4	20,4
Medien	98	54,1	54,1	74,6
Politischer Akteur	27	14,9	14,9	89,5
Privatperson	3	1,7	1,7	91,2
Wirtschaftlicher Akteur	8	4,4	4,4	95,6
Wissenschaftlicher Akteur	8	4,4	4,4	100,0
Gesamt	181	100,0	100,0	

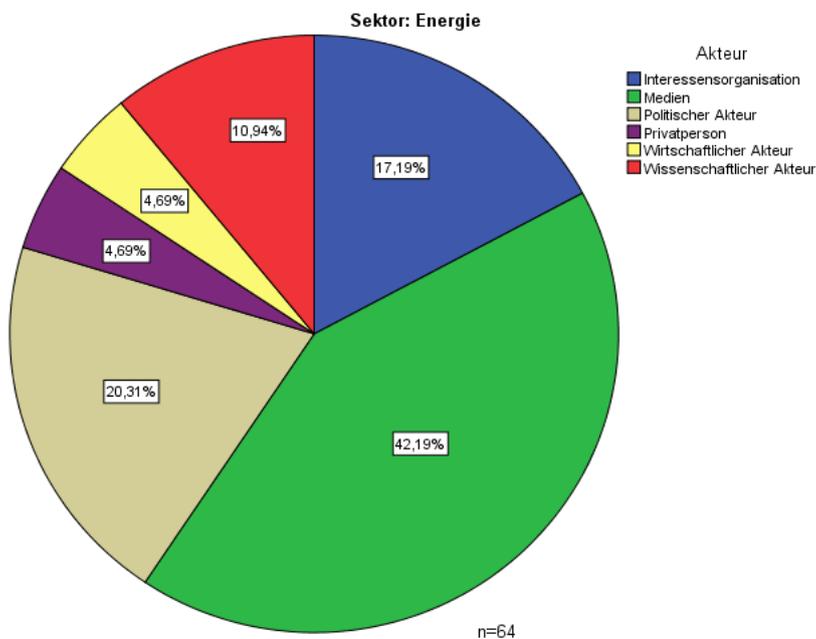


Abbildung 8: Aufteilung der Quellen nach Akteursgruppen- Energie

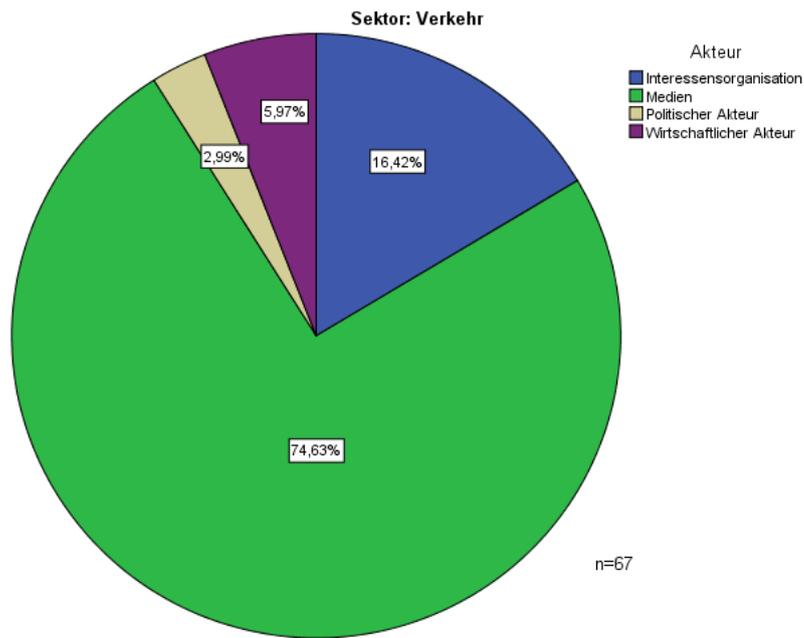


Abbildung 10: Aufteilung der Quellen nach Akteursgruppen- Verkehr

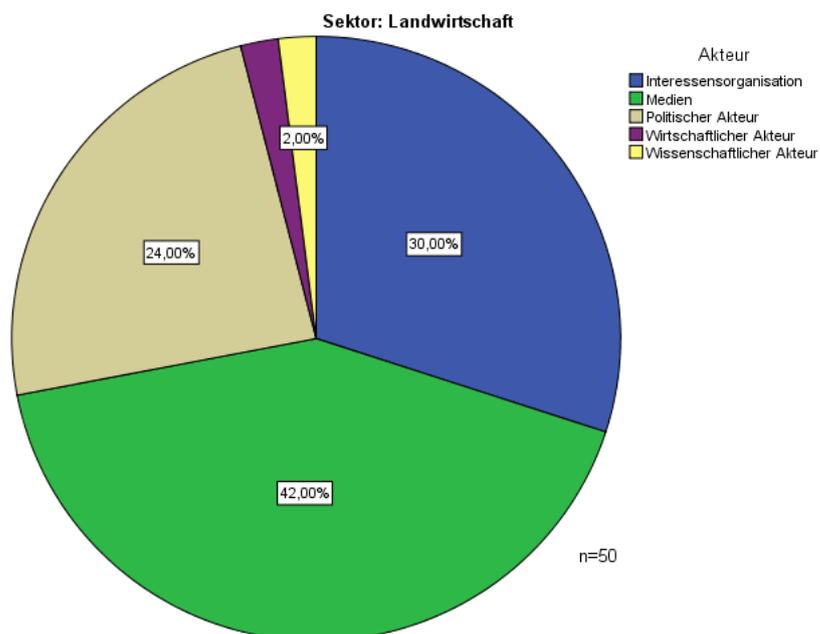


Abbildung 9: Aufteilung der Quellen nach Akteursgruppen- Landwirtschaft

7 Diskussion

In diesem Abschnitt sollen die Erkenntnisse aus dieser Arbeit diskutiert werden. Zuerst werden die im Abschnitt zuvor vorgestellten Ergebnisse interpretiert und in das Arbeitsthema eingeordnet. Im Anschluss folgt eine Kritik der Methodik und abschließend wird ein kurzer Ausblick für künftige Forschung gegeben.

7.1 Diskussion der Ergebnisse

Bestandteil der Ergebnisdiskussion ist die Beantwortung der Forschungsfragen.

F1: Gibt es bei der personalisierten Google-Recherche nach kontrovers diskutierten Umweltthemen einen Zusammenhang zwischen der Anzahl gesetzter impliziter Signale und der Anordnung der Ergebnisse?

Die Frage lässt sich unter Berücksichtigung der Ergebnisse dieser Arbeit mit Nein beantworten. Die Anzahl der gesetzten Signale im Profil - in diesem Fall die Formulierung einer Suchanfrage, eine anschließende Google-Abfrage und ein Klick auf ein Dokument – korreliert nicht negativ mit der Ergebnisreihung. Bereits beim ersten Vergleich der Ergebnislisten in den Experimental- und der Kontrollbedingung konnten entgegen der Erwartungen Effekte festgestellt werden. Im Abgleich mit weiteren SERP-Vergleichen waren die Effekte anfangs relativ hoch. Dies könnte zum einen mit der verwendeten Kennzahl zusammenhängen. Es wurde – im Vergleich der personalisierten SERPs mit den anonymen - jeweils die Anzahl der gleichen Dokumente auf dem gleichen Rang genommen. Ein einziges Dokument, das in einer Liste fehlt, kann somit potenziell einen großen Einfluss auf diese Kennzahl ausüben. Im Vergleich der SERPs würden damit alle anderen Dokumente, die dahinter angeordnet sind, um eine Position höher platziert werden. Die erhobene Kennzahl kann somit als anfällig betrachtet werden und möglicherweise zu Fehlinterpretationen führen.

Die Daten zeigen dennoch, dass eine negative Korrelation der gesetzten Signale mit der Anordnung der Dokumente auch mit anderen potenziellen Kennzahlen nicht wahrscheinlich ist. Selbst am Ende des Untersuchungszeitraumes waren die SERPs in den Experimentalbedingungen teilweise identisch mit denen aus der Kontrollbedingung, während für manche Suchanfragen am Anfang schon eine unterschiedliche Reihung der Do-

kumente selbst auf den Top-Positionen auftrat. In Anbetracht einiger Studien zur Personalisierung sowie den Aussagen von Google selbst, die Nutzer würden die Personalisierung zunächst kaum bemerken, muss dieser Umstand kritisch hinterfragt werden. Eine Erklärung dafür könnte die in der Studie von Hannák et al. (2015, S. 6) angesprochene geografische Verteilung der Rechenzentren Googles sein. In eigenen Tests haben sie herausgefunden, dass verschiedene Rechenzentren teilweise unterschiedliche Ergebnisse zur gleichen Abfrage ausgeben. Die Autoren halten es für wahrscheinlich, dass die Unterschiede durch Inkonsistenzen im Index der Suchmaschine über die Rechenzentren hinaus entstehen. Dazu sei gesagt, dass Google im Juni 2019 ein Core-Algorithmus-Update vorgenommen hat, also eine Anpassung ihres Kernalgorithmus. Google kündigte das Update zwar im Voraus an, die Informationen dazu waren jedoch nicht umfangreich. In einem Tweet dazu hieß es: “The June 2019 Core Update is now live and rolling out to our various data centers over the coming days” (Google SearchLiaison, 2019a). Am Tag davor wurden die Änderungen über Twitter bereits angekündigt, sie wurden jedoch von Google relativiert (Google SearchLiaison, 2019c). Das Update sorgte in der Online-Marketing-Branche für Aufsehen, wie einige Blogeinträge zu diesem Thema zeigen (z. B. Schwartz, 2019; Grundmann, 2019). Ob die Auswirkungen des Updates tatsächlich einen Einfluss auf die Studie hatten, ist unklar.

Somit bleibt festzuhalten, dass bei der Suche nach kontrovers diskutierten Umweltthemen zwar durchaus Personalisierungseffekte auftreten, die Anzahl der gesetzten Signale den Ergebnissen dieser Arbeit nach jedoch keine Rolle spielt. Lediglich zu einer Suchanfrage (“gentechnik”) konnte eine signifikante negative Korrelation festgestellt werden, zu den anderen acht Suchanfragen und in der Auswertung der aggregierten Daten zeigte sich kein Zusammenhang.

Darüber hinaus wurden in dieser Masterarbeit Dokumente inhaltlich bewertet. Die Forschungsfrage dazu lautete:

F2: Unterscheiden sich die anonymen Suchergebnisse von den personalisierten Suchergebnissen bei Recherchen nach kontrovers diskutierten Umweltthemen inhaltlich?

Auch diese Forschungsfrage ist mit Nein zu beantworten. So gab es nicht nur keine signifikanten Unterschiede zwischen der Kontrollbedingung und einer der beiden Experi-

mentalbedingungen. Auch zwischen diesen beiden sind über den ganzen Untersuchungszeitraum gesehen keine signifikanten inhaltlichen Unterschiede auszumachen. Somit lässt sich auch die dritte Forschungsfrage beantworten:

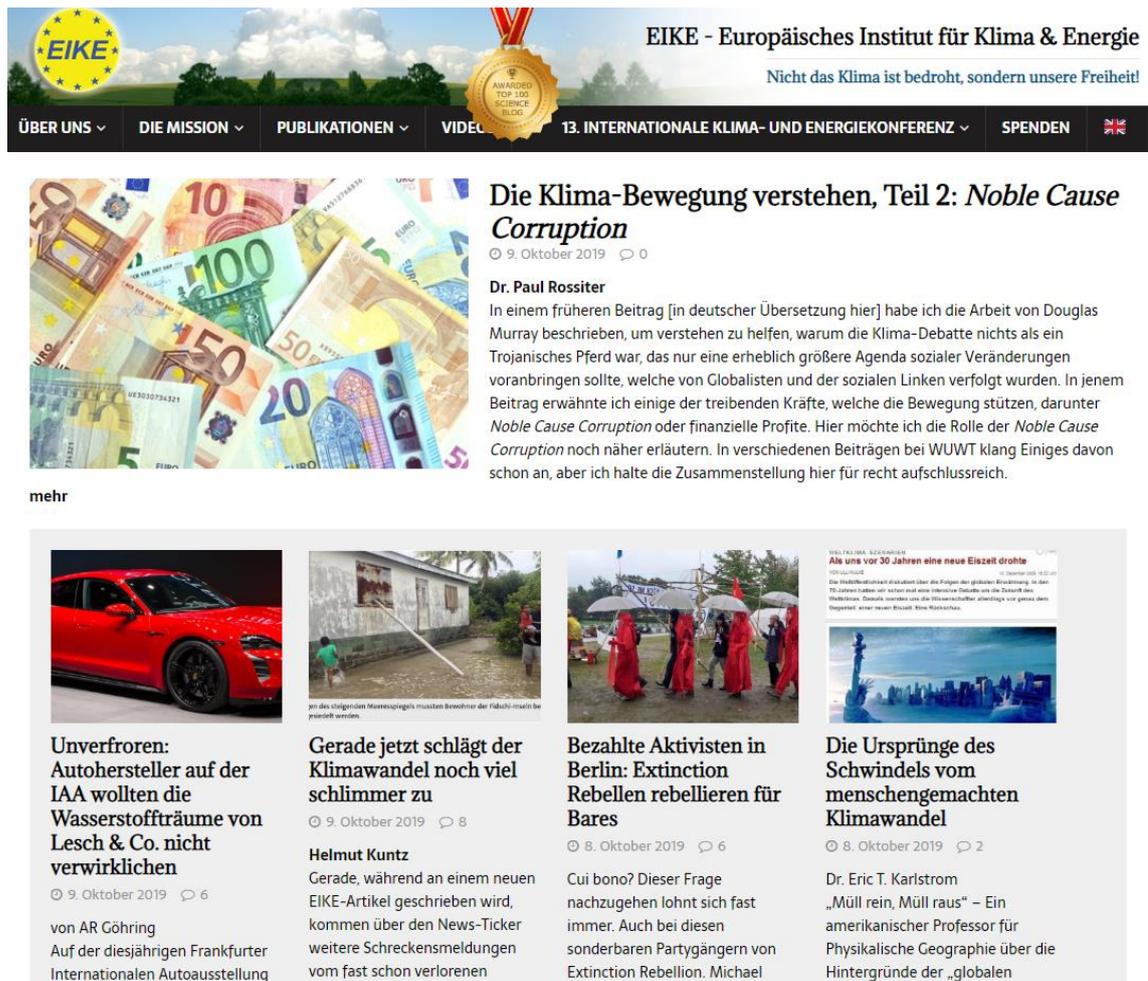
F3: Wie wirkt sich die im Suchverhalten implizite persönliche Haltung in einer Kontroverse bei einer Umkehr auf die gegenteilige Haltung in personalisierten Google-Recherchen zu Umweltthemen aus?

Eine Umkehr der persönlichen Haltung in der Diskussion um kontroverse Umweltthemen hat offenbar keinen Einfluss in der Google-Recherche. Da keine Unterschiede zwischen den Bedingungen festzustellen waren, sollen die Ergebnisse zur Beantwortung der Forschungsfragen 2 und 3 im Gesamten diskutiert werden. Die Bewertungen der Dokumente fielen überwiegend neutral aus. Ein Grund dafür könnte sein, dass relativ viele Inhalte von Medienakteuren erstellt werden, die sich oftmals an die journalistische Sorgfaltspflicht halten und - auf kontroverse Themen übertragen - verschiedene Meinungen in ihren Inhalten abbilden. Die Instruktionen für die Bewertungen waren so formuliert, dass in diesem Fall eine Neutralität besteht. Der neutrale Mittelwert aller Dokumente kann hier jedoch irreführend wirken, denn nicht alle Webseitenbetreiber veröffentlichen neutrale Inhalte. Besonders Interessensorganisationen haben eine klare Meinung zu den einzelnen Standpunkten. Zu den Suchanfragen "umsetzung energiewende", "ausstieg kohleverstromung", "tempolimit" und "tierhaltung" traten relativ häufig Dokumente auf, die, den Bewertungen nach, die umweltpolitischen Maßnahmen zur Energiewende/Verkehrswende/Agrarwende entweder stark befürworten oder stark ablehnen. Die größtenteils neutralen SERPs sind so zu erklären, dass Dokumente, in denen eine Meinung klar vertreten ist, gemischt sind mit neutralen Inhalten, und somit alle Positionen in den Ergebnislisten angezeigt werden. So ist beispielsweise bei der Recherche nach "tierhaltung" in allen SERP-Vergleichen ein Dokument der Tierrechtsorganisation Peta, die die industrielle Nutztierhaltung ablehnt, zu sehen. Jedoch auch ein Dokument des Deutschen Bauernverbandes, der die industrielle Nutztierhaltung als alternativlos ansieht. Tendenziell war jedoch eine leichte Neigung zu Inhalten festzustellen, die die Maßnahmen zur Energiewende/Verkehrswende/Agrarwende begrüßen. Ein Grund dafür könnte sein, dass der Konsens zur Anwendung solcher Maßnahmen in der Gesellschaft größer zu werden scheint. Zur anfangs von Pariser (2011a) genannten Problematik der Entstehung einer Filterblase im personalisierten Internet wurden in dieser Studie keinerlei Hinweise gesammelt, dass sich Tendenzen dazu abzeichnen würden.

Mit der vierten Forschungsfrage wurde die Frage nach den Quellen gestellt:

F4: Welche Quellen werden Google-Nutzern in der Recherche nach kontrovers diskutierten Umweltthemen angezeigt?

Hier war auffällig, dass zu einem großen Teil Medienakteure, politische Akteure und Interessensorganisationen die Ergebnislisten dominierten. Dies kann daran liegen, dass keine weitere Unterteilung dieser Akteursgruppen vorgenommen wurde. So wurden beispielsweise staatliche Behörden unter politische Akteure gefasst. Eine weitere Auffälligkeit war die Konstanz, mit der die Dokumente in den Top-Positionen vertreten waren. Von 181 bewerteten Dokumenten erschienen 90 beim ersten Vergleich, während der 13 nächsten Vergleiche kamen 91 hinzu. Die Quellen der Dokumente, die in allen 14 Vergleichen in den ersten zehn Ergebnissen landeten, sind größtenteils als seriös einzuschätzen. Bei der Suchanfrage "ausstieg kohleverstromung" fiel jedoch ein Dokument des Europäischen Instituts für Klima und Energie (EIKE) auf, das sich klar gegen den beschlossenen Kohleausstieg positioniert. Bei diesem Institut handelt es sich nicht, wie man bei diesem Namen glauben möchte, um eine wissenschaftliche Einrichtung, sondern um einen eingetragenen Verein, der nach Brunnengräber (2013, S. 27) "besonders hervorzuheben ist in der deutschen Klimaskeptiker-Szene." Ein Auszug der Webseite des Vereins ist in Abbildung 11 zu sehen. Das häufige Auftreten dieses Dokuments zur Suchanfrage "ausstieg kohleverstromung" lässt sich mit der Suchanfrage erklären. Die Überschrift des Dokuments enthält die Phrase "Ausstieg aus der Kohleverstromung", weswegen dieser Umstand als Grund für das häufige Auftreten gesehen wird. Bei Recherchen nach dem geläufigem Begriff Kohleausstieg ist der Verein nicht auf den Top-Positionen vertreten.



EIKE - Europäisches Institut für Klima & Energie
Nicht das Klima ist bedroht, sondern unsere Freiheit!

ÜBER UNS ▾ DIE MISSION ▾ PUBLIKATIONEN ▾ VIDEO ▾ 13. INTERNATIONALE KLIMA- UND ENERGIEKONFERENZ ▾ SPENDEN 

Die Klima-Bewegung verstehen, Teil 2: Noble Cause Corruption
9. Oktober 2019 0

Dr. Paul Rossiter
In einem früheren Beitrag [in deutscher Übersetzung hier] habe ich die Arbeit von Douglas Murray beschrieben, um verstehen zu helfen, warum die Klima-Debatte nichts als ein Trojanisches Pferd war, das nur eine erheblich größere Agenda sozialer Veränderungen voranbringen sollte, welche von Globalisten und der sozialen Linken verfolgt wurden. In jenem Beitrag erwähnte ich einige der treibenden Kräfte, welche die Bewegung stützen, darunter *Noble Cause Corruption* oder finanzielle Profite. Hier möchte ich die Rolle der *Noble Cause Corruption* noch näher erläutern. In verschiedenen Beiträgen bei WUWT klang Einiges davon schon an, aber ich halte die Zusammenstellung hier für recht aufschlussreich.

mehr

Unverfroren: Autohersteller auf der IAA wollten die Wasserstoffräume von Lesch & Co. nicht verwirklichen
9. Oktober 2019 6
von AR Göhring
Auf der diesjährigen Frankfurter Internationalen Autoausstellung

Gerade jetzt schlägt der Klimawandel noch viel schlimmer zu
9. Oktober 2019 8
Helmut Kuntz
Gerade, während an einem neuen EIKE-Artikel geschrieben wird, kommen über den News-Ticker weitere Schreckensmeldungen vom fast schon verlorenen

Bezahlte Aktivisten in Berlin: Extinction Rebellen rebellieren für Bares
8. Oktober 2019 6
Cui bono? Dieser Frage nachzugehen lohnt sich fast immer. Auch bei diesen sonderbaren Partygängern von Extinction Rebellion. Michael

Die Ursprünge des Schwindels vom menschengemachten Klimawandel
8. Oktober 2019 2
Dr. Eric T. Karlstrom
„Müll rein, Müll raus“ – Ein amerikanischer Professor für Physikalische Geographie über die Hintergründe der „globalen

Abbildung 11: Webseite des Europäischen Instituts für Klima & Energie

Insgesamt zeigte sich eine starke Präsenz der klassischen Medien aus Rundfunk und Presse. Dokumente der Medien “Spiegel Online” und “Zeit Online” gehörten zu den Quellen, die am häufigsten in den SERPs auftraten. Dies kann einerseits an einer hohen Reichweite und den damit einhergehenden hohen Klickzahlen liegen (Agof, 2019). Ein weiterer Grund könnte der Einfluss der Suchmaschinenoptimierung sein. Der Effekt von SEO ist bisher wenig erforscht. Die Tatsache, dass sich eine große und zahlungskräftige Branche dahinter verbirgt (Lewandowski, 2018, S. 174), ist im Kontext von Suchmaschinen-Verzerrungen vermutlich nicht zu unterschätzen.

Google selbst scheint bemüht zu sein, die Quellenvielfalt in ihren Ergebnislisten zu erhöhen. Im Juni 2019 veröffentlichte das Unternehmen das sogenannte Diversity-Update, nicht zu verwechseln mit dem Core-Update. In einem Tweet hieß es (Google SearchLiaison, 2019b)

“This site diversity change means that you usually won't see more than two listings from the same site in our top results. However, we may still show more than two in cases where our systems determine it's especially relevant to do so for a particular search....”

Zwar ist damit die Anzahl der Domains mit mehreren URLs pro Ergebnisliste begrenzt. Am Beispiel von Webseiten des Bundesministeriums für Ernährung und Landwirtschaft (BMEL) zeigt sich jedoch, dass manche Akteure ihre Inhalte über verschiedene Domains verbreiten. In dieser Studie waren je ein Dokument der Webseiten www.praxis-agrar.de und www.oekolandbau.de in allen Vergleichen in den ersten zehn Ergebnissen platziert. Beide Webseiten werden jedoch vom BMEL herausgegeben. Ob somit tatsächlich eine Diversifizierung von Inhalten erreicht werden kann, ist zumindest infrage zu stellen.

Im Allgemeinen lässt sich festhalten, dass die inhaltlichen Bewertungen neutraler und diversifizierter ausfielen, als angenommen. Suchmaschinen-Verzerrungen entstehen ganz natürlich aufgrund des Rankings, die Ergebnisse dieser Studie lassen jedoch keinen Anlass zu, die Meinungsvielfalt im Spektrum kontrovers diskutierter Umweltthemen würde durch Personalisierung eingeschränkt werden. Potenziell haben Suchmaschinen jedoch die Möglichkeit, einen Einfluss auf die Meinungsbildung auszuüben (Epstein & Robertson, 2015). Dieser Aspekt sollte auch nicht unterschätzt werden. Lewandowski (2018, S. 123) führt an, dass der Nutzer kaum Spielraum dazu hat, die Personalisierung individuell anzupassen.

7.2 Diskussion der Methode

In diesem Abschnitt wird die Methodik kritisch hinterfragt. Das experimentelle Studiendesign wurde aus anderen wissenschaftlichen Beiträgen zur Untersuchung von Personalisierung im Internet (u. a. Hanák et al., 2015) übernommen und auf das eigene Vorhaben angepasst. Das dieser Arbeit zugrundeliegende Studiendesign eignet sich zur Beschaffung personalisierter Suchergebnisse, hat jedoch im Laufe der Untersuchung Schwächen offenbart, die aufgegriffen werden sollen. Da wäre die Erstellung der Suchanfragen-Sets zur Manipulation des Suchverhaltens. Dabei wurde der Google Keyword Planer herangezogen, um real genutzte Suchanfragen zu identifizieren. Dies gestaltete sich bisweilen schwierig, da die genutzten Begriffe oftmals nicht klar dem Standpunkt in einer Kontroverse zuzuordnen waren (z. B. wenn sie nur aus einem Wort bestanden).

Auch gab es zu einzelnen Standpunkten kaum real genutzte Suchanfragen. Bei der Nutzung des Keyword Planers besteht außerdem die Gefahr, sich zu sehr von den automatisch generierten Vorschlägen des Tools leiten zu lassen. Manche Suchanfragen, die für die Studie ausgewählt wurden, hatten häufig die gleiche Bedeutung und riefen die gleichen Webseiten in der Recherche hervor. So beispielsweise bei den Suchanfragen "tankstellenpreise" und "benzinpreisvergleich". Eine systematische Auswahl der Suchanfragen basierend auf eigenen Überlegungen wäre diesem Verfahren vorzuziehen. So können z. B. gegensätzliche Begriffspaare zu systematisch erstellten Sets an Suchbegriffen und somit zu einer ähnlich gleichen Gewichtung dieser Sets führen.

Ein weiterer Kritikpunkt ist die Eliminierung von Dokumenten in der inhaltlichen Bewertung aufgrund Bezahlschranken und nicht gegebener Themenrelevanz. Die ermittelten Bewertungs-Scores sind somit als leicht verfälscht anzusehen. Zwar ist die Rangposition je Dokument in die Kennzahl miteingeflossen. Durch die Entfernung von Dokumenten auf den Top-Positionen flossen jedoch die Bewertungen von Inhalten ein, die die Nutzer aufgrund ihrer Platzierung auf der zweiten Ergebnisseite möglicherweise nicht betrachtet hätten. Da die Anzahl der entfernten Dokumente jedoch sehr gering war (2), wird der Einfluss auf die Ergebnisse in dieser Studie als vernachlässigbar eingeschätzt.

Im Allgemeinen sind die Kernelemente des Studiendesigns zur Untersuchung personalisierter Suchergebnisse geeignet. Auch die ermittelte Kennzahl zur Bewertung der Inhalte wird insbesondere wegen der Berücksichtigung der Rangposition als passend betrachtet. Die Grenzen der Studie sowie weitere potenzielle Forschungsarbeiten in diesem Bereich werden im nächsten Abschnitt behandelt.

7.3 Grenzen der Studie und Ausblick auf künftige Forschung

Diese Masterarbeit hatte zwar personalisierte Suchergebnisse als Untersuchungsgegenstand, es wurden jedoch nur Teile des Prozesses untersucht. So wurde der Zusammenhang zwischen der Anzahl der implizit gesetzten Signale im Profil auf die Personalisierungseffekte untersucht. Diese impliziten Daten sorgen jedoch nicht ausschließlich für personalisierte Ergebnisse. Auch explizite Daten spielen als Faktoren für die Personalisierung eine Rolle (Lewandowski, 2018, S. 123). Diese wurden in der vorliegenden Arbeit nicht berücksichtigt. Hier könnte die Berücksichtigung von sogenannten sozialen Signalen, also Empfehlungen in sozialen Medien (bspw. durch eine Gefällt-mir-Angabe oder

das Teilen eines Links), für die weitere Forschung interessant sein. Hildebrandt, Schumann und Wolling (2015) konnten einen Einfluss solcher Signale auf das organische Ranking von Google feststellen. Dies wurde jedoch nicht im personalisierten Umfeld untersucht.

Das Studiendesign ist außerdem nicht nur für Untersuchungen der Suchmaschine Google geeignet. Das Gerüst lässt sich auf andere Suchsysteme adaptieren, die ihre Ergebnisse an den individuellen Nutzer anpassen. In diesem Kontext ist beispielsweise die Untersuchung des Edge-Algorithmus von Facebook vorstellbar. Ein entsprechendes Profil könnte explizit mit Angaben zu den Interessen relativ einfach aufgebaut werden.

Für die weitere Forschung wird jedoch die automatische Durchführung der Studie empfohlen. Im Rahmen dieser Masterarbeit wurden die Profile rein mit manuellen Eingaben auf den Tablets aufgebaut. Dies hatte den Vorteil, die Intensität der Eingaben von Suchanfragen real zu halten. Dies müsste auch bei automatisch generiertem Suchverhalten beibehalten werden. Der zeitliche Aufwand würde sich so jedoch erheblich reduzieren. Berücksichtigt man das reale Nutzerverhalten, ließen sich so über einen gewissen Zeitraum mehr Daten sammeln, die zu valideren Ergebnissen führen könnten.

8 Fazit

Personalisierungsmaßnahmen im Internet werden immer wieder kritisch hinterfragt. Ein bekannter Kritikpunkt ist die Theorie der dadurch entstehenden Filterblase (Pariser, 2011a). In der vorliegenden Masterarbeit wurden personalisierte Suchergebnisse der Suchmaschine Google zu kontrovers diskutierten Umweltthemen untersucht. Mit einem experimentellen Vorgehen konnten entsprechende Suchergebnisse beschafft werden. Diese wurden von Juroren auf ihren Inhalt hin bewertet. Die Ergebnisse zeigen, dass es war durchaus zu unterschiedlichen Ergebnisreihungen kommt, die Häufigkeit von Recherchen im eingeloggten Zustand und die damit verbundenen impliziten Daten scheinen darauf jedoch keinen Einfluss zu haben. Eine statistische Überprüfung konnte keinen signifikanten Einfluss zwischen der Anzahl der impliziten Daten im Profil und den Effekten in der Ergebnisreihung feststellen.

Auch inhaltlich gibt es in dieser Studie keine Ansatzpunkte, die die Bildung einer Filterblase rechtfertigen würden. Zwischen den gebildeten Bedingungen, die sich durch ein aufgebautes Interessensprofil zu Umweltthemen auszeichnen, konnten keine signifikanten Zusammenhänge festgestellt werden.

Personalisierte Ergebnislisten scheinen den Ergebnissen dieser Arbeit nach nicht "verzerrter" zu sein als welche in der anonymen Suche. Diese Aussage trifft jedoch nur in Hinblick auf die impliziten Angaben im Profil zu. Die potenziellen Auswirkungen anderer Faktoren wurden nicht untersucht, eignen sich jedoch für künftige Vorhaben ebenso wie die Untersuchung anderer Suchsysteme mit personalisierten Inhalten. Dazu wurde im Zuge dieser Masterarbeit ein Studiendesign mit entsprechenden Kennzahlen entwickelt, das sich auf andere mögliche Forschungen anpassen lässt.

9 Literaturverzeichnis

- Agof (2019). daily digital facts 02.09.2019 / Rangreihe. Abgerufen 8. Oktober 2019, von https://www.agof.de/download/Downloads_daily_digital_facts/Downloads_ddf_08_2019/ddf_August_2019_Angebote_Ranking_Internet.pdf
- Ballatore, A. (2015). Google chemtrails: A methodology to analyze topic representation in search engine results. *First Monday*, 20(7). Abgerufen von <http://dx.doi.org/10.5210/fm.v20i7.5597>
- Brey, P. (1998). The Politics of Computer Systems and the Ethics of Design. In *Proceeding ACM/SIGCAS Conference* (S. 64–75).
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1–7), 107–117. [https://doi.org/10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X)
- Broder, A. (2002). A taxonomy of web search. *ACM SIGIR Forum*, 36(2), 3. <https://doi.org/10.1145/792550.792552>
- Brunnengräber, A. (2013). *Klimaskeptiker in Deutschland und ihr Kampf gegen die Energiewende*. Abgerufen von: http://refubium.fu-berlin.de/bitstream/handle/fub188/20150/FFU-Report_Achim_Brunnengraeber_endgueltige_Version.pdf?sequence=1&isAllowed=y
- Bundesministerium für Ernährung und Landwirtschaft. (2019). Ökologischer Landbau in Deutschland. abgerufen 1. Mai 2019, von https://www.bmel.de/SharedDocs/Downloads/Landwirtschaft/Oekologischer_Landbau/OekolandbauDeutschland.pdf?jsessionid=CE2ACF250ED0D49CD0961461D2407BBB.2_cid288?__blob=publicationFile
- Bundeszentrale für politische Bildung. (2017). Die Entstehung eines Wahl-O-Mat. Abgerufen von <http://www.bpb.de/politik/wahlen/wahl-o-mat/45292/die-entstehung-eines-wahl-o-mat>
- Bundeszentrale für politische Bildung. (2013). Wie funktioniert der Wahl-O-Mat? Abgerufen von <http://www.bpb.de/politik/wahlen/wahl-o-mat/166945/wie-funktioniert-der-wahl-o-mat>

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. Abgerufen von <http://www.utstat.toronto.edu/~brunner/oldclass/378f16/readings/CohenPower.pdf>
- Der Tagesspiegel. (2019). Konervative CDU-Abgeordnete wollen Kernkraftwerke länger laufen lassen. Abgerufen 5. Juni 2019, von <https://www.tagesspiegel.de/politik/kein-falscher-respekt-vor-oekopopulisten-konservative-cdu-abgeordnete-wollen-kernkraftwerke-laenger-laufen-lassen/24418002.html>
- Diaz, A. (2008). Through the Google Goggles: Sociopolitical Bias in Search Engine Design. In A. Spink & M. Zimmer (Hrsg.), *Web Search* (S. 11–34). Berlin Heidelberg: Springer-Verlag.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-41089-5>
- Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(33), E4512-21. <https://doi.org/10.1073/pnas.1419828112>
- Erlhofer, S. (2018). *Suchmaschinen-Optimierung: das umfassende Handbuch* (9. Aufl.). Bonn: Rheinwerk Verlag.
- Feuz, M., Fuller, M., & Stalder, F. (2011). Personal Web searching in the age of semantic capitalism: Diagnosing the mechanisms of personalisation. *First Monday*, *16*(2). Abgerufen von <https://journals.uic.edu/ojs/index.php/fm/article/view/3344/2766>
- Fortunato, S., Flammini, A., Menczer, F., & Vespignani, A. (2006). Topical interests and the mitigation of search engine bias. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(34), 12684–12689. <https://doi.org/10.1073/pnas.0605525103>

- Frants, V. J., Shapiro, J., & Voiskunskii, V. G. (1997). Vital Activities and Needs. In V. J. Frants, J. Shapiro, & V. G. Voiskunskii (Hrsg.), *Automated Information Retrieval* (Vol. 97, S. 27–50). Emerald Group Publishing Limited.
[https://doi.org/10.1108/S1876-0562\(1997\)000097A003](https://doi.org/10.1108/S1876-0562(1997)000097A003)
- Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems*, 14(3), 330–347. <https://doi.org/10.1145/230538.230561>
- Garcia-Molina, H., Koutrika, G., & Parameswaran, A. (2011). Information seeking: Convergence of search, recommendations, and advertising. *Communications of the ACM*, 54(11), 121–130. <https://doi.org/10.1145/2018396.2018423>
- Gauch, S., Speretta, M., Chandramouli, A., & Micarelli, A. (2007). User Profiles for Personalized Information Access. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Hrsg.), *The Adaptive Web* (Vol. 4321, S. 54–89). Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-72079-9>
- Gerhart, S. (2004). Do web search engines suppress controversy? *First Monday*, 9(1). Abgerufen von <https://ojphi.org/ojs/index.php/fm/article/view/1111/1031>
- Goldman, E. (2008). Search Engine Bias and the Demise of Search Engine Utopianism. In A. Spink & M. Zimmer (Hrsg.), *Web Search* (S. 121–133). Berlin Heidelberg: Springer-Verlag.
- Google. (2019a). Warum fragt Google bei der Erstellung eines Kontos nach bestimmten Informationen? Abgerufen 1. Mai 2019, von <https://support.google.com/accounts/answer/1333913?hl=de>
- Google. (2019b). Datenschutzerklärung & Nutzungsbedingungen. Abgerufen 2. Mai 2019, von <https://policies.google.com/privacy?hl=de>
- Google SearchLiaison (2019a, 3. Juni). The Juni 2019 Core Update is now live. [Tweet]. Abgerufen von <https://twitter.com/searchliaison/status/1135586621204525060>
- Google SearchLiaison (2019b, 6. Juni). This site diversity change means that you usually won't see more than two listings from the same site in our top results. [Tweet] Abgerufen von <https://twitter.com/searchliaison/status/1136739064193998848>
- Google SearchLiaison (2019c, 2. Juni). Tomorrow, we are releasing a broad core algorithm update, as we do several times per year. [Tweet]. Abgerufen von <https://twitter.com/searchliaison/status/1135586621204525060>

- Grundmann, J. (2019). June 2019 Core Update: Alle Infos zum aktuellen Google Update. Abgerufen October 7, 2019, from <https://blog.searchmetrics.com/de/google-update-june-2019-core/>
- Gudivada, V. N., Rao, D., & Paris, J. (2015). Understanding Search-Engine Optimization. *Computer*, 48(10), 43–52. <https://doi.org/10.1109/MC.2015.297>
- Haider, J., & Sundin, O. (2019). *Invisible Search and Online Search Engines: The Ubiquity of Search in Everyday Life*. London, New York: Routledge.
- Hannák, A., Sapiezłyński, P., Khaki, A. M., Lazer, D., Mislove, A., & Wilson, C. (n.d.). Measuring Personalization of Web Search. Abgerufen von <http://arxiv.org/pdf/1706.05011v1>
- Hildebrandt, C., Schumann, C., & Wolling, J. (2015). Licht in die Black Box?! Experimente zum Einfluss von Social Signals auf das organische Ranking von Google. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft*, (S. 225–249). Berlin <https://doi.org/10.17174/dcr.v2.10>
- Hinman, L. M. (2005). Esse est indicato in Google: Ethical and Political Issues in Search Engines. *International Review of Information Ethics*, 3(06), 19–25.
- Höchstötter, N., & Koch, M. (2009). Standard parameters for searching behaviour in search engines and their empirical evaluation. *Journal of Information Science*, 35(1), 45–65. <https://doi.org/10.1177/0165551508091311>
- Introna, L. D., & Nissenbaum, H. (2000). Shaping the Web: Why the Politics of Search Engines Matters. *The Information Society*, 16(3), 169–185. <https://doi.org/10.1080/01972240050133634>
- Jarvelin, K., & Kekalainen, J. (2000). IR evaluation methods for retrieving highly relevant documents. *SIGIR Forum (ACM Special Interest Group on Information Retrieval)*, 41–48.
- Jürgens, P., Stark, B., & Magin, M. (2015). Messung von Personalisierung in computervermittelter Kommunikation. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 251–270). Berlin. <https://doi.org/10.17174/dcr.v2.11>

- Jürgens, P., Stark, B., & Magin, M. (2014). Gefangen in der Filter Bubble? Search Engine Bias und Personalisierungsprozesse bei Suchmaschinen. In S. Stark, Birgit; Dörr, D.; Aufenanger (Hrsg.), *Die Googleisierung der Informationssuche. Suchmaschinen zwischen Nutzung und Regulierung* (S. 98–135). Berlin, Boston: Walter de Gruyter GmbH. <https://doi.org/10.13109/9783666252884.17>
- Kamvar, S. (2005). Search gets personal. Abgerufen von <https://googleblog.blogspot.com/2005/06/search-gets-personal.html>
- Killoran, J. B. (2013). How to Use Search Engine Optimization Techniques to Increase Website Visibility. *IEEE Transactions on Professional Communication*, 56(1), 50–66. <https://doi.org/10.1109/TPC.2012.2237255>
- Koch, T., Peter, C., & Müller, P. (2019). *Das Experiment in der Kommunikations- und Medienwissenschaft*. Wiesbaden: Springer VS. <https://doi.org/https://doi.org/10.1007/978-3-658-19754-4>
- Kochale, S. (2019). Gentechnik und das Glaubwürdigkeits-Problem. Abgerufen 1. Mai 2019, von <https://www.mdr.de/wissen/umwelt/bayer-monsanto-wissenschaftler-gentechnik-100.html>
- Kosub, S. (2019). A note on the triangle inequality for the Jaccard distance. *Pattern Recognition Letters*, 120(1), 36–38. <https://doi.org/10.1016/j.patrec.2018.12.007>
- Krauter, R. (2019). “Elektroautos sind nicht die sofortige Wunderwaffe.” Abgerufen 2. Mai 2019, von https://www.deutschlandfunk.de/co2-reduzieren-elektroautos-sind-nicht-die-sofortige.676.de.html?dram:article_id=450643
- Kulshrestha, J., Eslami, M., Messias, J., Zafar, M. B., Ghosh, S., Gummadi, K. P., & Karahalios, K. (2019). Search bias quantification: investigating political bias in social media and web search. *Information Retrieval Journal*, 22(1–2), 188–227. <https://doi.org/10.1007/s10791-018-9341-2>
- Lewandowski, D. (2014). Die Macht der Suchmaschinen und ihr Einfluss auf unsere Entscheidungen. *Information - Wissenschaft & Praxis*, 65(4–5). <https://doi.org/10.1515/iwp-2014-0050>
- Lewandowski, D. (2018). *Suchmaschinen verstehen* (2. Aufl.). Heidelberg: Springer Vieweg. Abgerufen von <https://doi.org/10.1007/978-3-662-56411-0>

- Lewandowski, D. (2017). Is Google Responsible for Providing Fair and Unbiased Results? In M. Taddeo & L. Floridi (Hrsg.), *The Responsibilities of Online Service Providers* (Vol. 31, S. 61–77). Cham: Springer International Publishing.
<https://doi.org/10.1007/978-3-319-47852-4>
- Mager, A. (2012). Algorithmic Ideology: How capitalist society shapes search engines. *Information, Communication & Society*, 15(5), 769–787.
<https://doi.org/10.1080/1369118X.2012.676056>
- Mair, M. (2019). Ein bisschen Tierwohl. Abgerufen 2. Mai 2019, von <https://www.tagesschau.de/inland/tierwohllabel-101.html>
- Mathrani, A., & Alipour, M. (2010). Website Blocking Across Ten Countries: A Snapshot. In *PACIS 2010 Proceeding* (S. 152). Abgerufen von <https://aisel.aisnet.org/pacis2010/152>
- Mercier, H. (2016). Confirmation bias myside bias. *Cognitive Illusions: Intriguing Phenomena in Judgement, Thinking and Memory*, 99–114.
<https://doi.org/10.4324/9781315696935>
- Mowshowitz, A., & Kawaguchi, A. (2002a). Bias on the web. *Communications of the ACM*, 45(9). <https://doi.org/10.1145/567498.567527>
- Mowshowitz, A., & Kawaguchi, A. (2002b). Assessing bias in search engines. *Information Processing & Management*, 38(1), 141–156.
[https://doi.org/10.1016/S0306-4573\(01\)00020-6](https://doi.org/10.1016/S0306-4573(01)00020-6)
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.
- Norddeutscher Rundfunk. (2019). Landwirte ärgern sich über Feinstaub-Studie. Abgerufen 2. Mai 2019, von <https://www.ndr.de/nachrichten/niedersachsen/Landwirte-aergern-sich-ueber-Feinstaub-Studie,feinstaubbelastung102.html>
- Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google We Trust: Users' Decisions on Rank, Position, and Relevance. *Journal of Computer-Mediated Communication*, 12(3), 801–823.
<https://doi.org/10.1111/j.1083-6101.2007.00351.x>

- Pariser, E. (2011a). *The Filter Bubble: What the Internet is Hiding from You*. New York: The Penguin Press.
- Pariser, E. (2011b). *Beware online "filter bubbles."* Abgerufen von https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?language=de
- Pelzer, G., & Gerigk, D. (2018). *Google Adwords: Das umfassende Handbuch* (2nd ed.). Bonn: Rheinwerk Verlag.
- Petrescu, P. (2014). Google Organic Click-Through Rates in 2014. Abgerufen von https://moz.com/blog/google-organic-click-through-rates-in-2014?fbclid=IwAR2Wd0aaZ8nONnk5LMAhJHHVf1iDIhmIbG6TNf663wKZR8ddqK_obhsFp0g
- Riemer, K., & Brüggemann, F. (2009). Personalisierung der Internetsuche - Lösungstechniken und Marktüberblick. In D. Lewandowski (Hrsg.), *Handbuch Internet-Suchmaschinen* (S. 148–174). Heidelberg: Akademische Verlagsgesellschaft Aka GmbH.
- Robertson, R. E., Jiang, S., Joseph, K., Friedland, L., Lazer, D., & Wilson, C. (2018). Auditing Partisan Audience Bias within Google Search. In *Proceedings of the ACM on Human-Computer Interaction* (p. 148). New York: ACM.
- Röhle, T. (2010). *Der Google-Komplex: Über Macht im Zeitalter des Internets*. Bielefeld: transcript Verlag.
- Salehi, S., Du, J. T., & Ashman, H. (2015). Examining Personalization in Academic Web Search. In Y. Yesilada, R. Farzan, & G.-J. Houben (Hrsg.), *Proceedings of the 26th ACM Conference on Hypertext & Social Media - HT '15* (S. 103–111). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2700171.2791039>
- Schmidt, W. (2019). Wie es beim Windkraft-Ausbau weitergeht. Abgerufen 2. Mai 2019 von <https://www.welt.de/regionales/hamburg/article187563840/Schleswig-Holstein-Hitzige-Debatte-zum-Thema-Windkraft-Ausbau.html>
- Schnirch, P. (2019). Der Frust sitzt tief. Abgerufen 5. Juni 2019, von <https://www.sueddeutsche.de/muenchen/freising/landwirte-der-frust-sitzt-tief-1.4451857>

- Schrader, C. (2019). Ein kritischer Blick. Abgerufen 2. Mai 2019, von <https://www.spektrum.de/news/wie-ist-die-umweltbilanz-von-elektroautos/1514423>
- Schwartz, B. (2019). Early data around the Google June 2019 core update shows some winners, losers. Abgerufen 7. Oktober 2019, von <https://searchengineland.com/early-data-around-the-google-june-2019-core-update-shows-some-winners-losers-317861>
- Schwiesau, D., & Ohler, J. (2016). *Nachrichten – klassisch und multimedial: Ein Handbuch für Ausbildung und Praxis*. Wiesbaden: Springer VS.
<https://doi.org/10.1007/978-3-658-08717-3>
- Sedlmeier, P., & Renkewitz, F. (2018). *Forschungsmethoden und Statistik in der Psychologie* (3. Aufl.). Hallbergmoos: Pearson Deutschland.
- Spiegel Online. (2019). "Verbindliche Preisempfehlung" statt CO2-Steuer. Abgerufen von <https://www.spiegel.de/politik/deutschland/co2-steuer-bernd-althusmann-und-thomas-strobl-fordern-preisempfehlung-a-1270339.html>
- Stark, B., Magin, M., & Jürgens, P. (2014). Navigieren im Netz: Befunde einer qualitativen und quantitativen Nutzerbefragung. In B. Stark, D. Dörr, & S. Aufenanger (Hrsg.), *Die Googleisierung der Informationssuche. Suchmaschinen zwischen Nutzung und Regulierung* (S. 20–74). Berlin: de Gruyter.
- StatCounter. (2019). Search Engine Market Share Germany. Abgerufen 9. Oktober 2019, von <https://gs.statcounter.com/search-engine-market-share/all/germany>
- Süddeutsche Zeitung. (2019). DGB fordert 365-Euro-Ticket für Auszubildende. Abgerufen 5. Juni 2019, von <https://www.sueddeutsche.de/karriere/ausbildung-hannover-dgb-fordert-365-euro-ticket-fuer-auszubildende-dpa.urn-newsml-dpa-com-20090101-190604-99-506768>
- Tavani, H. (2012). Search Engines and Ethics. Abgerufen from <https://plato.stanford.edu/archives/fall2016/entries/ethics-search/>
- Teevan, J., Dumais, S. T., & Horvitz, E. (2005, August). Personalizing search via automated analysis of interests and activities. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval* (S. 449-456). ACM.

- Tripodi, F. (2018). Searching for Alternative Facts: Analyzing Scriptural Inference in Conservative News Practices, 1–64. Abgerufen von https://datasociety.net/wp-content/uploads/2018/05/Data_Society_Searching-for-Alternative-Facts_FINAL-5.pdf
- Universität Zürich. (2018a). Einfaktorielle Varianzanalyse (ohne Messwiederholung). Abgerufen 5. Oktober 2019, von https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/unterschiede/zentral/evarianz.html
- Universität Zürich. (2018b). Rangkorrelation nach Spearman. Abgerufen 5. Oktober 2019, von https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/zusammenhaeng_e/rangkorrelation.html
- Van Couvering, E. (2007). Is Relevance Relevant? Market, Science, and War: Discourses of Search Engine Quality. *Journal of Computer-Mediated Communication*, 12(3), 866–887. <https://doi.org/10.1111/j.1083-6101.2007.00354.x>
- Vaughan, L., & Zhang, Y. (2007). Equal Representation by Search Engines? A Comparison of Websites across Countries and Domains. *Journal of Computer-Mediated Communication*, 12(3), 888–909. <https://doi.org/10.1111/j.1083-6101.2007.00355.x>
- Verbruggen, A. (2014). Could it be that Stock-Stake Holders Rule Transition Arenas? In A. Brunnengräber & M. R. Di Nucci (Hrsg.), *Im Hürdenlauf zur Energiewende* (S. 119–132). Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-06788-5>
- Weber, K. (2011). Search Engine Bias. In D. Lewandowski (Hrsg.), *Handbuch Internet-Suchmaschinen 2* (S. 265–285). Heidelberg: Akademische Verlagsgesellschaft AKA GmbH.
- White, R. W. (2016). *Interactions with Search Systems*. New York: Cambridge University Press.
- White, R. W. (2013). Beliefs and biases in web search. *SIGIR 2013 – Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 3–12. <https://doi.org/10.1145/2484028.2484053>

- White, R. W., & Horvitz, E. (2009). Cyberchondria. *ACM Transactions on Information Systems*, 27(4), 1–37. <https://doi.org/10.1145/1629096.1629101>
- Wulfers, A. (2019). Kampf um die Gülle. Abgerufen von <https://www.faz.net/aktuell/wirtschaft/landwirte-protestieren-gegen-neue-guelleregeln-16124210.html>
- ZDF. (2019). Union schließt längere Kohlenutzung nicht aus. Abgerufen von <https://www.zdf.de/nachrichten/heute/carsten-linnemann-ralph-brinkhaus-plaene-zum-kohleausstieg-stossen-bei-der-cdu-auf-skepsis-100.html>
- Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., Helberger, N., & Internet Policy Review. (2016). Should we worry about filter bubbles? *Internet Policy Review*, 5(1). <https://doi.org/10.14763/2016.1.401>

Anhang 1: Informationsdossier

Projektbeschreibung

Dieses Informationsdossier soll bei der Bewertung von Internetdokumenten zu umweltpolitischen Themen unterstützen. Im Rahmen einer Masterarbeit sollen Ergebnisse aus einer Google-Suche auf ihren Inhalt hin bewertet werden. Dazu wurden drei Juroren rekrutiert, die ihre Bewertung auf einer Skala abgeben. Die Instruktionen dazu finden sich im letzten Abschnitt dieses Dokuments.

Die in der Arbeit zu untersuchenden Internetdokumente drehen sich in erster Linie um die Themenfelder Energie, Verkehr und Landwirtschaft. Ziel ist es, herauszufinden, in welchem Ausmaß Google die Ergebnisse zu kontrovers diskutierten Umweltthemen personalisiert. Zu jedem der drei genannten Sektoren wurden mehrere Internetdokumente gesammelt. Nachfolgend finden sich zu jedem Sektor wesentliche Informationen. Das Dossier schließt mit den Instruktionen für die Juroren.

Informationen zum Sektor Energie

Im Sektor Energie sind in den Vorschlägen zur Umsetzung der Energiewende Diskrepanzen erkennbar. Der vollumfängliche Umstieg auf erneuerbare Energien zur Energieproduktion ist notwendig, um die Auswirkungen des Klimawandels so gering wie möglich zu halten. Auch die Bundesregierung bekennt sich zur Energiewende, die eine grundlegende Umstellung der Energieversorgung mit sich bringt: die Abkehr von nuklearen und fossilen Energieträgern hin zu erneuerbaren Energien. In der politischen und gesellschaftlichen Debatte werden dabei besonders die Vor- und Nachteile der verschiedenen Energieträger, sowie der Zeitrahmen, in dem die Energiewende umgesetzt werden soll, diskutiert.

Fossile Energien

Zu den fossilen Energieträgern zählen Erdöl, Erdgas, Kohle und Torf. Kernenergie, in der Regel gewonnen durch Uran und später im Prozess durch Plutonium, nimmt unter den

fossilen Energieträgern eine Sonderstellung ein und wird deshalb im Kontext dieser Arbeit zu den fossilen Brennstoffen gezählt.

Nachfolgend sind stichpunktartig die Eigenschaften der verschiedenen fossilen Energieträger aufgelistet.

Kohle

Nutzung der durch die Verbrennung von Kohle frei gewordenen Energie. In Deutschland in der Regel durch Braunkohle, seltener durch Steinkohle.

Erdöl

Erdöl befindet sich in unter der Erdoberfläche liegenden sogenannten Ölfeldern. Das dort geförderte Öl wird u. a. als flüssiger Brennstoff in Form von Heizöl genutzt.

Erdgas

Unterirdisch vorkommendes Gas, das zum größten Teil aus Methan besteht. Wird zum Heizen und zur Stromerzeugung benutzt.

Torf

Energieerzeugung durch die Verbrennung von Torf.

Kernenergie/Atomkraft

Durch nukleare Reaktionen freigesetzte Energie, die hauptsächlich zur Stromerzeugung verwendet wird. Die Energie wird durch Kernspaltung gewonnen (Spaltung des Atomkerns von Uran oder Plutonium). Durch die Kernspaltung wird Hitze freigesetzt, die sich wiederum in elektrische Energie umwandeln lässt. Perspektivisch ist die Energiegewinnung mittels Kernfusion denkbar.

Erneuerbare Energien

Nachfolgend sind stichpunktartig die Eigenschaften der verschiedenen erneuerbaren Energieträger aufgelistet.

Windenergie

Energiegewinnung durch mittels Windkraft angetriebene Windräder in sogenannten Windparks oder Windkraftanlagen.

Wasserkraft

Üblicherweise erfolgt die Energieproduktion durch mit Wasserkraft angetriebene Generatoren. Entsprechende Kraftwerke stehen meistens an Gewässern mit hoher Fließkraft. Eine neue Form der Nutzung der Wasserkraft liegt in der Energiegewinnung aus dem Meer. Dabei werden beispielsweise Wellenströmungen oder Gezeitenströmungen genutzt.

Solarenergie

Sonnenenergie oder Solarenergie wird in erster Linie durch zwei verschiedene Verfahren gewonnen: in Photovoltaikanlagen durch Solarzellen und durch Solarthermie mittels Sonnenkollektoren.

Biomasse

Energiegewinnung aus tierischen und pflanzlichen Abfällen (Gülle, Ernteabfälle, Biomüll) oder aus nachwachsenden Rohstoffen (Holz, extra angebaute Energiepflanzen wie Raps oder Mais); zu den Energieträgern der Bioenergie gehören Holzpellets, Biogas, Biomethan, Hackschnitzel und Strohpellets.

Geothermie/Erdwärme

Energiegewinnung aus der in der Erdkruste gespeicherten Wärmeenergie, die zum Heizen, Kühlen oder zur Stromproduktion verwendet werden kann. Bei der oberflächennahen Geothermie findet die Übertragung der Energie in der Regel in einer sogenannten Wärmepumpe oder mittels Wärmekollektoren statt, die tiefe Geothermie wird vor allem von Geothermiekraftwerken genutzt.

[Informationen zum Sektor Verkehr](#)

Das Themenfeld Verkehr weist Parallelen zum Sektor Energie auf. Die Debatte dreht sich hierbei insbesondere um die Zukunft der Mobilität und den Einsatz von klimaneutralen Antriebsformen für Fahrzeuge. Die Verkehrswende ist unterteilt in eine Mobilitätswende und eine Energiewende im Verkehrswesen. Während die Mobilitätswende die Reduzierung des Energieverbrauchs ohne eine zusätzliche

Einschränkung der Mobilität anstrebt, soll die Energiewende im Verkehr zu einer klimaneutralen Energieerzeugung zur Deckung des verbleibenden Bedarfs führen.

Bei der Energiewende im Verkehrswesen steht die Abkehr von Fahrzeugantriebsformen, die auf fossile Brennstoffe angewiesen sind (Benzin, Diesel, Erdgas), hin zur Elektromobilität an vorderster Stelle. In den gesellschaftlichen und politischen Diskussionen ist vor allem die Umweltbilanz der verschiedenen Fahrzeugtypen ein Thema. Während Fahrzeuge mit Verbrennungsmotor in ihrer Nutzung Emissionen ausstoßen, sind Elektrofahrzeuge klimaneutral unterwegs. Aus einem ökologischen Standpunkt betrachtet sind sie jedoch in ihrer Herstellung umstritten, da die Produktion der in Elektrofahrzeugen verbauten Akkus als sehr energieaufwendig gilt. Auch der Abbau von seltenen Erden, die zum Bau der Batterien benötigt werden, gilt vielerorts als zweifelhaft.

Die Mobilitätswende hingegen soll vor allem durch den Ausbau des öffentlichen Nahverkehrs, den Ausbau von Radwegen und Innovationen wie Carsharing oder Ridesharing umgesetzt werden. Andere Maßnahmen, wie die Einführung eines Tempolimits, spielen in der Verkehrswende zwar eine untergeordnete Rolle, werden in der Gesellschaft mitunter aber emotional diskutiert.

Informationen zum Sektor Landwirtschaft

Die Kontroversen im Sektor Landwirtschaft sind größtenteils im Spannungsfeld zwischen ökologischer und konventioneller Landwirtschaft zu finden. Die Unterschiede zwischen beiden Betriebsformen sind nachfolgend in kurzer Form aufgelistet, bestehen aber hauptsächlich in strengeren Verordnungen für den biologischen Landbau, an die sich konventionelle Landwirte nicht halten müssen.

Unterschiede gibt es zum einem im Ackerbau. Bio-Landwirte verzichten auf chemisch hergestellte Dünge- und Pflanzenschutzmittel, die im konventionellen Landbau erlaubt sind. Beim Einsatz von Gentechnologie, beispielsweise in der Herstellung von Saatgut, verhält es sich ähnlich: in der ökologischen Landwirtschaft verboten, in der konventionellen Landwirtschaft erlaubt.

Ein weiterer Unterschied liegt in der Tierhaltung. Auch dort gibt es strengere Auflagen für Bio-Landwirte. Das Platzangebot für die Tiere ist in der ökologischen Landwirtschaft höher. Außerdem muss das Futter für das Vieh dort zu mindestens 95 Prozent aus

biologischem Anbau stammen, der Einsatz von Gentechnologie ist auch im Futter verboten.

Der Verkauf von Bio-Lebensmitteln sieht außerdem vor, auf künstliche Stoffe wie Geschmacksverstärker oder Konservierungsmittel zu verzichten.

Bewertungsinstruktionen für Juroren

Die erhaltenen Informationen sollen nun die Bewertung von Internetdokumenten vereinfachen. Die Dokumente sind auf ihren Inhalt hin auf einer Skala von -2 bis 2 zu bewerten und drehen sich um die genannten Themenfelder Energie, Verkehr und Landwirtschaft. Die einzelnen Bewertungspunkte für jeden Sektor sind wie folgt zu verstehen:

	Energie	Verkehr	Landwirtschaft
--2	<p>Dokument hebt folgenden Aspekt positiv hervor:</p> <p>Energiegewinnung aus fossilen Brennstoffen (Braunkohle, Steinkohle, Erdgas, Erdöl, Torf, Kernenergie)</p> <p>UND</p> <p>folgenden Aspekt negativ hervor:</p> <p>Energiegewinnung aus erneuerbaren Energien (Wasserkraft; Windkraft; Sonnenenergie; Meeresenergie, wie Gezeitenkraftwerke, Wellenkraftwerke,</p>	<p>Dokument hebt mindestens einen der folgenden Aspekte positiv hervor: Antrieb durch fossile Kraftstoffe (Benzin, Diesel, Erdgas); Ablehnung Tempolimit</p> <p>UND</p> <p>mindestens einen der folgenden Aspekte negativ hervor: Elektromobilität; Einführung Tempolimit</p>	<p>Dokument hebt folgenden Aspekt positiv hervor: konventionelle Landwirtschaft (Einsatz chemischer Pflanzenschutzmittel wie bspw. Glyphosat; Einsatz genmanipulierter Saatgutsorten; intensive Tierhaltung; Einsatz chemischer Düngemittel)</p> <p>UND</p> <p>folgenden Aspekt negativ hervor: ökologische Landwirtschaft (Verzicht auf Gentechnik,</p>

	<p>Osmosekraftwerke, Meeresströmungskraftwerke und Meereswärmekraftwerke; Bioenergie, wie Holzpellets, Biogas, Biomethan, Hackschnitzel, Strohpellets; Geothermie, wie Einsatz von Wärmepumpen oder Wärmekollektoren)</p>		<p>chemische Dünger, chemische Pflanzenschutzmittel; ökologische Viehzucht)</p>
--1	<p>Dokument hebt folgenden Aspekt positiv hervor: Energiegewinnung aus fossilen Brennstoffen</p> <p>Erneuerbare Energien werden im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Dokument hebt folgenden Aspekt negativ hervor: Energiegewinnung aus erneuerbaren Energien.</p> <p>Fossile Brennstoffe werden im Dokument nicht wertend erwähnt</p>	<p>Dokument hebt mindestens einen der folgenden Aspekte positiv hervor: Antrieb durch fossile Kraftstoffe; Ablehnung Tempolimit</p> <p>Elektromobilität oder die Einführung eines Tempolimits werden im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Dokument hebt mindestens einen der folgenden Aspekte negativ hervor: Elektromobilität; Ablehnung Tempolimit</p>	<p>Dokument hebt folgenden Aspekt positiv hervor: konventionelle Landwirtschaft</p> <p>Ökologische Landwirtschaft wird im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Aspekt hebt folgenden Aspekt negativ hervor: ökologische Landwirtschaft</p> <p>Konventionelle Landwirtschaft wird im Dokument nicht wertend erwähnt</p>

		Fossile Kraftstoffe werden im Dokument nicht erwähnt	
0 0	<p>Das Dokument hebt weder die Energiegewinnung aus fossilen Brennstoffen noch erneuerbare Energien positiv oder negativ hervor</p> <p style="text-align: center;">ODER</p> <p>Dokument hebt Energiegewinnung aus fossilen Brennstoffen und erneuerbare Energien in gleicher Weise positiv oder negativ hervor</p>	<p>Das Dokument steht dem Thema Verkehrswende neutral gegenüber, was sich folgendermaßen ausdrückt:</p> <p>Das Dokument hebt weder den Antrieb durch fossile Kraftstoffe noch Elektromobilität positiv oder negativ hervor; ebenso beim Thema Tempolimit</p> <p style="text-align: center;">ODER</p> <p>Dokument hebt weder Antrieb durch fossile Kraftstoffe noch Elektromobilität in gleicher Weise positiv oder negativ hervor; ebenso beim Thema Tempolimit</p> <p style="text-align: center;">ODER</p>	<p>Das Dokument hebt weder die konventionelle noch die ökologische Landwirtschaft positiv oder negativ hervor</p> <p style="text-align: center;">ODER</p> <p>Dokument hebt konventionelle und ökologische Landwirtschaft in gleicher Weise positiv oder negativ hervor</p>

		Dokument hebt weder Antrieb durch fossile Kraftstoffe noch Elektromobilität positiv oder negativ hervor, sondern Hybridantriebe (Kombination aus Verbrennungs- und Elektroantrieb) oder alternative Kraftstoffe (Biodiesel, Wasserstoff, BtL-Kraftstoff)	
1 +1	<p>Dokument hebt folgenden Aspekt positiv hervor: Energiegewinnung aus erneuerbaren Energien</p> <p>Fossile Brennstoffe werden im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Dokument hebt folgenden Aspekt negativ hervor: Energiegewinnung aus fossilen Brennstoffen</p> <p>Erneuerbare Energien werden im Dokument nicht wertend erwähnt</p>	<p>Dokument hebt mindestens einen der folgenden Aspekte positive hervor: Elektromobilität; Einführung Tempolimit</p> <p>Fossile Kraftstoffe bzw. Ablehnung Tempolimit werden im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Dokument hebt mindestens einen der</p>	<p>Dokument hebt folgenden Aspekt positiv hervor: ökologische Landwirtschaft</p> <p>Konventionelle Landwirtschaft wird im Dokument nicht wertend erwähnt</p> <p>ODER</p> <p>Dokument hebt folgenden Aspekt negativ hervor: konventionelle Landwirtschaft</p>

		<p>folgenden Aspekte negativ hervor: Antrieb durch fossile Kraftstoffe; Ablehnung Tempolimit</p> <p>Elektromobilität bzw. Einführung Tempolimit werden im Dokument nicht wertend erwähnt</p>	<p>Ökologische Landwirtschaft wird im Dokument nicht wertend erwähnt</p>
2 +2	<p>Dokument hebt folgenden Aspekt positiv hervor: erneuerbare Energien</p> <p>UND</p> <p>folgenden Aspekt negativ hervor: Energiegewinnung aus fossilen Brennstoffen</p>	<p>Dokument hebt mindestens einen der folgenden Aspekte positiv hervor: Elektromobilität; Einführung Tempolimit</p> <p>UND</p> <p>mindestens einen der folgenden Aspekte negativ hervor: Antrieb durch fossile Kraftstoffe; Ablehnung Tempolimit</p>	<p>Dokument hebt folgenden Aspekt positiv hervor: ökologische Landwirtschaft</p> <p>UND</p> <p>folgenden Aspekt negativ hervor: konventionelle Landwirtschaft</p>

Eidesstattliche Erklärung

Ich versichere, die vorliegende Arbeit selbstständig ohne fremde Hilfe verfasst und keine anderen Quellen und Hilfsmittel als die angegebenen benutzt zu haben. Die aus anderen Werken wörtlich entnommenen Stellen oder dem Sinn nach entlehnten Passagen sind durch Quellenangaben kenntlich gemacht.

Hamburg, 11. Oktober 2019

Ort, Datum

Unterschrift

(Manuel Reiter)