



DIE ÖFFENTLICHE VERHANDLUNG VON BIG DATA IN POLITISCHEN KAMPAGNEN

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

01IS15016A-F



Universität Bremen*



ZeMKI

Zentrum für
Medien-, Kommunikations- und
Informationsforschung

Universität Bremen, Zentrum für Medien-, Kommunikations-
und Informationsforschung (ZeMKI)

Prof. Dr. Christian Pentzold

Lena Fölsche

ABIDA – ASSESSING BIG DATA

PROJEKTLAUFZEIT 01.03.2015-28.02.2019



Westfälische Wilhelms-Universität Münster,
Institut für Informations-, Telekommunikations- und
Medienrecht (ITM), Zivilrechtliche Abteilung



Karlsruher Institut für Technologie,
Institut für Technikfolgenabschätzung
und Systemanalyse (ITAS)



Leibniz Universität Hannover
Institut für Rechtsinformatik
(IRI)



Technische Universität Dortmund,
Wirtschafts- und Sozialwissenschaftliche
Fakultät (WiSo) Techniksoziologie



Ludwig-Maximilians-Universität München,
Forschungsstelle für Information, Organisation
und Management (IOM)



Wissenschaftszentrum Berlin
für Sozialforschung

Wissenschaftszentrum
Berlin für Sozialforschung



ABIDA – Assessing Big Data
Über das Gutachten

Das Gutachten wurde im Rahmen des ABIDA-Projekts mit Mitteln des Bundesministeriums für Bildung und Forschung erstellt. Der Inhalt des Gutachtens gibt ausschließlich die Auffassungen der Autoren wieder. Diese decken sich nicht automatisch mit denen des Ministeriums und/oder der einzelnen Projektpartner.

ABIDA lotet gesellschaftliche Chancen und Risiken der Erzeugung, Verknüpfung und Auswertung großer Datenmengen aus und entwirft Handlungsoptionen für Politik, Forschung und Entwicklung.

www.abida.de

© 2018 – Alle Rechte vorbehalten

Inhaltsverzeichnis

| | | |
|-------|--|----|
| 1 | Zusammenfassung | 2 |
| 2 | Konzeptueller Rahmen | 4 |
| 2.1 | Big Data im öffentlichen Diskurs | 4 |
| 2.2 | Datengetriebene Wahlkämpfe als ‚computational politics‘ | 6 |
| 3 | Forschungslücken | 7 |
| 4 | Ziele..... | 10 |
| 5 | Arbeitsprogramm | 13 |
| 6 | Methode..... | 15 |
| 6.1 | Datensammlung und Datenorganisation | 15 |
| 6.2 | Inhaltsanalytische Aufbereitung | 21 |
| 7 | Ergebnisse..... | 22 |
| 7.1 | Makroanalytische Diskursstrukturierung | 22 |
| 7.2 | Diskurs in der Nusschale: Der Fall Cambridge Analytica..... | 27 |
| 7.3 | Datenbasierte Abbildungsverhältnisse, Repräsentationsbeziehungen und Regulierungsabsichten..... | 31 |
| 7.3.1 | Datenbasierte Abbildungsverhältnisse | 31 |
| 7.3.2 | Datenbasierte Repräsentationsbeziehungen..... | 36 |
| 7.3.3 | Datenbasierte Regulierungsabsichten | 49 |
| 8 | Ergebnisdiskussion | 57 |
| 9 | Handlungsempfehlungen | 65 |
| 10 | Literatur..... | 69 |
| 11 | Anhang..... | 75 |

DIE ÖFFENTLICHE VERHANDLUNG VON BIG DATA IN POLITISCHEN KAMPAGNEN

DATENBASIERTE ABBILDUNGSVERHÄLTNISSE,
REGULIERUNGSABSICHTEN UND
REPRÄSENTATIONSBEZIEHUNGEN IN DER WAHL-
KAMPFBERICHTERSTATTUNG IN DEUTSCHLAND,
GROSSBRITANNIEN UND DEN USA

1 ZUSAMMENFASSUNG

In dem vorliegenden Gutachten werden die datenbasierten Abbildungsverhältnisse, Regulierungsabsichten und Repräsentationsbeziehungen in der Wahlkampfberichterstattung in Deutschland, Großbritannien und den USA rekonstruiert. Die Verfügbarkeit und Auswertung großer Datensätze, das Entdecken neuer Muster und das Verfolgen sozialer Prozesse in Echtzeit kann beeinflussen, so die ambivalenten Erwartungen in diesem exponierten Anwendungs- und Forschungsfeld politischer Kampagnen, wie das Politische und das Soziale gedacht, verstanden und gesteuert werden können.

Entsprechend bearbeitet das Gutachten **Big Data-basierte politische Praktiken und diskursive Zuschreibungen in ländervergleichender Perspektive**. Genutzt wird dabei ein qualitativ-verstehendes methodisches Design. Es erfasst, welche Bedeutung und welche Bedeutsamkeit Daten in Wahlkampfaktivitäten zugeschrieben werden. Die Ergebnisse werden im Gutachten zusammengefasst, konzeptionell bewertet und es werden auf dieser Basis Handlungsempfehlungen aufgestellt.

Im Kern erfasst das empirisch fundierte Gutachten die **medienöffentliche Debatte um datenbasierte Wahlkämpfe** – die tatsächliche Anwendungspraxis aggregierter Daten und darauf basierender Analyseverfahren für politische Strategien ist nicht Gegenstand der Studie. Untersucht wird vielmehr die gesellschaftlich zentrale Dimension der

öffentlichen Verhandlung und Bewertung datenbasierter Politik, die wesentlich deren Akzeptanz, Legitimierung und Reflexion bestimmt.

Zentrale Erkenntnisse der Studie sind:

- Die Verhandlung der Bedeutungen von Big Data in Wahlkämpfen erfolgte aus **vier grundlegenden Diskurspositionen**: Optimisten, Skeptiker, Pessimisten und Warnende; sie zeigen an, dass sich bisher keine hegemoniale Vorstellung darüber herausgebildet hat, welche Bedeutung Big Data in und für Wahlkämpfe besitzt.
- Die **zentralen Themen** kreisen um Fragen der Erkenntnisgewinne durch Big Data-Analysen, um Fragen des Datenschutzes, der Datenklüfte sowie der Manipulation, Überwachung und Parzellierung der Wählerschaft.
- Die **Repräsentativität der Daten** wird im untersuchten Diskursmaterial kaum kritisch hinterfragt; meist gehen die Kommunikate implizit von einer Ubiquität datenintensiver Praktiken aus, von der alle gleichermaßen betroffen seien – nämlich durch Praktiken des Datensammelns und -auswertens oder der darauf basierenden Microtargeting-Maßnahmen.
- Explizit thematisiert werden im Diskurs unterschiedliche **datenschutzrechtliche Regulierungsansätze** in Europa gegenüber den USA, wo datenintensive Wahlkampfpraktiken nicht zuletzt aus diesem Grund weiter vorangeschritten sind als in europäischen Ländern.
- **Datenintensive Wahlkampfmaßnahmen** werden im Diskurs als einseitig wahrgenommen: Bürger*innen treten im Diskurs nicht aktiv in Erscheinung, sie werden von den auftretenden Sprecher*innen – Journalist*innen, Datenanalyt*innen, Kampagnenleiter*innen und Wissenschaftler*innen – als Ressourcen und Ziele der datenintensiven Praktiken dargestellt, nicht als handelnde Subjekte.

2 KONZEPTUELLER RAHMEN

2.1 BIG DATA IM ÖFFENTLICHEN DISKURS

Gesellschaftliche Prozesse wie auch Formen persönlicher Lebensführung gehen zusehends einher mit der Generierung großer Datenmengen, die gespeichert und weiterverwendet werden (boyd/Crawford, 2012; Dorschel, 2015; Mayer-Schönberger/Cukier, 2013). Der in diesem Zusammenhang aufgekommene mehrdeutige Begriff **Big Data** verweist auf eine Reihe von Sachverhalten (Beer, 2016; Kennedy/Moss, 2015; Kitchin, 2014; Puschmann/Burgess, 2014).

(1) Zum einen meint **Big Data als Technologie** Datensätze, zu deren Bearbeitung aufgrund ihres Volumens Rechenzentren nötig sind, sowie Praktiken, um soziale oder technische Gegebenheiten in einzelne Datenpunkte zu überführen und diese maschinell zu aggregieren, zu analysieren und zu kombinieren (Hilbert/López, 2011). ‚Big‘ Data bezieht sich dabei nicht nur auf die Größe der Datenmengen, sondern auch auf die Vielfalt der Daten, die Geschwindigkeit, mit der sie ausgewertet werden können, sowie die zu ihrer Verwertung verfügbare analytische Kapazität. Integral damit verbunden sind Algorithmen, „sets of defined steps structured to computationally process instructions/data to produce an output“ (Kitchin, 2016: 1).

(2) Zum anderen verbindet sich mit **Big Data als analytischer Ambition** die Hoffnung nach präziseren Einsichten und verlässlicheren Prognosen (Anderson, 2008; Burrows/Savage, 2014; Ruppert/Law/Savage, 2013).

(3) Außerdem macht **Big Data als regulatorische Herausforderung** die Überwachung und Kontrolle durch staatliche Institutionen und privatwirtschaftliche Unternehmen möglich, welche Anlass für Bedenken bezüglich des Schutzes von Privatheit, Sicherheit, informationeller Selbstbestimmung und bürgerlichen Freiheiten gibt (Andrejevic, 2014; van Dijck, 2014; Lane et al., 2014; Lyon, 2014; O’Neil, 2016).

Ein grundlegender Aspekt datenbasierter Praktiken, darauf bezogener Erwartungen und Deutungsrahmen sowie der sich verändernden gesellschaftlichen Verhältnisse ist die Annahme, dass **Daten Realität abbilden** (Couldry/Hepp, 2017; Ulbricht/von Grafenstein, 2016). Das heißt, ihre Analyse durch Firmen, Politikberater*innen, Verwaltungen, Journalist*innen oder Forschende geht einher mit Repräsentationsansprüchen.

Diese begründen weitreichende unternehmerische und politische Maßnahmen bzw. Interventionen und gerade deshalb ist die Reflexion des Umgangs mit Big-Data-bezogenen Praktiken fundamental notwendig und aktuell dringend geboten.

Das Gutachten greift auf gegenwärtige Diskussionen in den *critical data studies* zurück. Sie verweisen zum einen auf die **Selektivität von Datenbeständen, auf methodische Unzulänglichkeiten und Probleme durch mangelnde Datenqualität** (Boellstorff, 2015; Halavais, 2015; Hoeren, 2017). Zum anderen hinterfragen sie die epistemologische Haltung, Daten eine objektive Faktizität und abstrakte Neutralität zuzusprechen, statt deren Komplexität und mögliche epistemische Widersprüche zu betrachten. Im Gegensatz dazu weisen sie darauf hin, dass Daten stets interpretationsbedürftig sind und der Umgang mit Daten das voraussetzungsreiche Zusammenwirken von Akteuren, Wissensordnungen und Technologien benötigt (Bowker, 2013, Gitelman/Jackson, 2013; Star/Bowker, 1999).

Doch jenseits einer naiven Annahme eines realitätsabbildenden und neutralen Charakters von Daten in der Forschung ist die **mediale Verhandlung von Daten** und die ihnen im öffentlichen Diskurs zugeschriebenen Potentiale zur Veränderung von sozialer Realität bzw. deren Verständnis in der alltagsweltlichen Praxis der Wahlkampfberichterstattung kritisch zu untersuchen.

Vor diesem Hintergrund führt das Gutachten eine ländervergleichende Studie durch. Es werden die datenbasierten Abbildungsverhältnisse, Regulierungsabsichten und Repräsentationsbeziehungen in der Wahlkampfberichterstattung in Deutschland, Großbritannien und den USA rekonstruiert.

Politische Wahlkämpfe sind aus zwei Gründen ein exponiertes Anwendungs- und Forschungsfeld.

- a. In der politischen Praxis sind sie als zeitlich befristete Phasen des konzentrierten, intensiv und öffentlich geführten Wettbewerbs um politische Zustimmung, Mobilisierung und der Entwicklung politischer Inhalte gekennzeichnet, verknüpft mit der Gestaltung und Implementierung neuartiger Technologien der Meinungsforschung und kommunikativen Ansprache. Entsprechend bauen auch die momentanen Strategien datenintensiver Kampagnen auf einer Tradition fortlaufender Innovationen in den Wahlkampftechniken auf (Schulz, 2014).

b. Analytisch sind Wahlkämpfe zentrale Gegenstände der politik- und kommunikationswissenschaftlichen empirischen Forschung und Theoriebildung. Untersucht werden dabei u.a. der Einsatz verschiedener Medien, Inszenierungsstrategien, die Art und der Inhalt der Berichterstattung, der Gebrauch von Umfragen und die Kommunikationsdynamiken zwischen Politikbetrieb und Wählerschaft (Übersicht bei Kenski/Jamieson, 2014).

2.2 DATENGETRIEBENE WAHLKÄMPFE ALS 'COMPUTATIONAL POLITICS'

Ungeachtet der Unschärfe des Begriffes verbinden sich mit ‚Big Data‘ konkrete Erwartungen eines **Wandels politischer Akteurskonstellationen, Strukturen und Prozesse** (Jungherr, 2015 u. Shorey/Howard, 2016 m.w.N.). Sie stehen in unterschiedlichen Bezügen zu bestehenden Nutzungsweisen von Metriken und statistischen Informationen in Administration und Verwaltung (Behnke et al., i.Dr.; Schnapp et al., 2009; Scott, 1999). Die Verfügbarkeit und Auswertung großer Datensätze, das Entdecken neuer Muster und das Verfolgen sozialer Prozesse in Echtzeit kann beeinflussen, so die ambivalente Erwartung, wie das Politische und das Soziale gedacht, verstanden und gesteuert werden können. Das Novum der *computational politics* definiert Tufekci (2014: 2) demgemäß so: „applying computational methods to large datasets derived from online and off-line data-sources for conducting outreach, persuasion and mobilization“ (auch Römmele/von Schneidmesser, 2016).

Im Blick auf datengetriebene Wahlkämpfe ist Big Data in der **affirmativen Perspektive** ein wirkungsvolles Instrument, diese kampagnenförmigen politischen Prozesse planbarer, kontrollierbarer und letztlich erfolgreicher zu führen – so zumindest die durch die Siege der datenintensiv geführten Kampagnen von Obama begründete Erwartung (Gülden-zopf, 2011, Jungherr/Schoen, 2013). Die Nutzung von Daten wird als entscheidendes Element professioneller Kampagnen gesehen. So diene das *Microtargeting* der Identifizierung und Ansprache potentiell überzeugbarer und mobilisierbarer Wahlberechtigter oder Spender*innen (Broockman/Kalla, 2016; Green/Gerbner, 2015; Nickerson/Rogers, 2013). Zudem könnten vermittelt „personalized political communication“ (Nielsen, 2012: 7) jenseits von Massenmedien und postalischen Sendungen un-

entschlossene Wähler*innen überzeugt bzw. politisch überzeugte Unterstützer*innen (gemessen am *persuadability score*) auch zur Urne gebracht werden oder aber davon abgebracht werden, wählen zu gehen (Castelman, 2016; Lundry, 2016). Durch Daten würden Wahlkämpfe schließlich effizienter, da sie die Zuwendung knapper Ressourcen durch intensivere individuellere Kontakte vereinfachen und die Kosten vergeblicher Kontaktversuche verringern würden. Kampagnen könnten außerdem effektiver geführt werden, da Big Data Streueffekte in der Kommunikation vermeiden helfe, etwa durch die detaillierte Anpassung der Kommunikationsinhalte (Arceneaux/Nickerson, 2009). Darüber hinaus soll Big Data helfen, die Kampagnenhelfer*innen vor Ort zu steuern, einen Unterstützer*innenkreis jenseits von Parteizugehörigkeit zu organisieren und die Aktivitäten der Kandidat*innenteams zu kontrollieren (Nielsen, 2012).

Im Fokus der **kritischen Perspektive** steht einmal die Frage nach der Parzellierung der Wählerschaft entlang persönlich zugeschnittener Botschaften, die nicht dazu beitragen, den Austausch zwischen Wähler*innen zu fördern. Damit einher geht, so Tufekci (2014), die Betonung trennender Themen, womit die Ausbildung gemeinschaftlicher Formen der politischen Willensbildung behindert und die ‚Balkanisierung‘ der Wählerschaft und ihrer politischen Orientierung unterstützt werden würde. In anderer Hinsicht steht der Schutz der Privatsphäre in Frage, sowohl im Rahmen der massiven Sammlung und Kombination von Daten als auch bezüglich ihrer Sicherung und Weiternutzung (Kreiss/Howard, 2010). Nicht zuletzt ist die Validität der datenbasierten Wähleranalysen und die Effektivität der darauf aufgesetzten politischen Strategien in Frage zu stellen (Jungherr, 2017).

3 FORSCHUNGSLÜCKEN

Die Potentiale und Risiken der Nutzung großer Datensätze und statistischer Modelle durch politische Organisationen werden in intensiven, bisher aber nur in Ansätzen auch wissenschaftlich aufgearbeiteten, Diskursen verhandelt. Deutlich wird, dass datenbasierte Wahlkämpfe nicht nur erhöhte Anforderung an die Rechen- und Verarbeitungskapazität der Kampagnenmacher stellen. Zugleich ändert sich die Art der Kommunikation selbst, weg von areal-basierten und hin zu individuums-basierten Formen der Ansprache (Jungherr, 2016; Stromer-Galley, 2014). Damit gehen datengetriebene

Kampagnen über bestehende Formen der Adressierung von Wahlberechtigten in Stimmbezirken, in denen vorherige Wahlergebnisse eine Kampagne vielversprechend erscheinen ließen, hinaus (Gerber et al., 2011; Issenberg, 2012), und sie berücksichtigen in größerem Ausmaß Einstellungen, Werte und Verhalten der Wahlberechtigten (Bleamire, 2012; Strauss, 2012).

Ausgehend von diesem Stand der Praxis und der wissenschaftlichen Reflexion werden im Folgenden **zwei Forschungslücken** betrachtet:

Big Data-basierte politische Praktiken und diskursive Zuschreibungen. Die bisherige Auseinandersetzung mit Big Data beachtet kaum, mit welchen Abbildungs-, Repräsentations-, und Regulierungsansprüchen datenintensive Strategien einhergehen und wie deren Rationalitäten hervorgebracht werden (Beer, 2016; Kitchin, 2014). Es fehlt also die konsequente Rekonstruktion der Bedeutungszuweisungen an Big Data in öffentlichen Diskursen. Unbeachtet bleibt damit die interpretative Dimension von Big Data, in der Akteure, Rechtfertigungsmuster und Technologien über die Hervorbringung epistemischer Annahmen in Beziehung gesetzt werden (Michael/Lupton, 2016; Portmess/Tower, 2015). Diese Lücke ist problematisch, weil damit übersehen wird, dass „data need to be imagined as data to exist and to function as such, and the imagination of data entails an interpretative base“, so Gitelman und Jackson (2013: 3). Die Beschäftigung mit Big Data als ‚Mythologie‘, wie Boyd und Crawford (2012: 663) schreiben, und folglich dem „widespread belief that large data sets offer a higher form of intelligence and knowledge [...] with the aura of truth, objectivity, and accuracy“ ist im Blick auf Kampagnen noch aus einem weiteren Grund geboten. Und zwar weist Jungherr (2013) darauf hin, dass die Kommunikation des Gebrauchs neuer Ansätze und Technologien in Wahlkämpfen ein bisher unerforschtes rhetorisches Mittel zur Inszenierung der Professionalität und des Innovationswillens einer hochfortschrittlichen Kampagne sein kann. Bezeichnend für die „campaign tech bullshit season“ (Karpf, 2016) ist die öffentliche Faszination über den Umfang der Informationen und der vermuteten Präzision der Modelle in US-amerikanischen Wahlkämpfen (Grassegger/Krogerus, 2016; Issenberg, 2012; Madrigal, 2012; Sides/Vavreck, 2014). Sie scheint geprägt von der Marketing-Rhetorik kommerzieller Analytiken zu sein (z. B. Cambridge Analytics; Strategic Communications Laboratories; Civis Analytics; Catalist; Blue Labs; Echelon Insights; Optimus Consulting; Giles-Parscale).

Zu fragen ist also:

- (1) Wie wird das Verhältnis von Daten und gesellschaftlicher Realität in öffentlichen Diskursen verhandelt?**
- (2) Inwiefern wird die Repräsentativität der in Daten abgebildeten und aus diesen resultierenden Willensbekundungen, Aktivitäten oder Relationen thematisiert bzw. problematisiert?**

Big Data-basierte politische Praktiken in vergleichender Perspektive. Die Integration datengestützter Entscheidungen in die politische Arbeit ist ein anspruchsvoller Prozess, bei dem rechtliche, technische und polit-kulturelle Kontextbedingungen sowie unterschiedliche Mediensysteme nicht ignoriert werden können (Anstead, 2016; Esser/Strömbäck, 2012; Hallin/Mancini, 2004; Pfetsch, 2014). Gerade weil aber die „digital opportunity structure“, wie Kreiss (2016: 109) sie nennt, – „features of the political environment and candidates and campaigns’ symbolic material [...] which shape the possibilities for using digital media for a strategic ends“ – sehr verschieden ausgebildet sein kann, ist die bisherige Fokussierung auf die USA zu eng (Hersh, 2015; Howard, 2006; Issenberg, 2012; Kreiss, 2012, 2016; Nickerson/Rogers, 2014; Nielsen, 2012; Tufekci, 2014). Sie muss durch eine komparative Perspektive erweitert werden, um zu klären, wie sich datengetriebene Wahlkämpfe und damit der Einsatz von Daten in politischen Prozessen in anderen Ländern ausgestalten (Anstead, 2017). So kann vergleichend untersucht werden, wie sich Kampagnenformen in spezifischen sozio-politischen Kontexten herausbilden, wie sie mit bestehenden Kontexten zusammenwirken und wie datenbasierte Wahlkampfpraktiken in den jeweiligen Öffentlichkeiten reflektiert werden.

Zu fragen ist also:

- (3) Inwiefern werden im öffentlichen Diskurs die analytischen Praktiken in Bezug gesetzt zum politischen System, insbesondere Wahlsystem, sowie zur politischen Kultur, vor allem im Blick auf das Parteienverhältnis und die Formen des Wahlkampfes?**

4 ZIELE

Das Projekt analysierte die öffentliche Verhandlung von Big Data im Kontext politischer Kampagnen. Hierdurch wurde es möglich, an einem exponierten Feld politischer Praxis die Generierung von und kontroverse Debatte um Big Data-basierte Abbildungsverhältnisse, Regulierungsabsichten und Repräsentationsbeziehungen zu rekonstruieren. Im Fokus stehen **drei miteinander verknüpfte Themen**:

- **Datenbasierte Abbildungsverhältnisse:** Datenbasierten Praktiken liegt die Vorstellung einer adäquaten Abbildbeziehung zwischen Daten und Realität zugrunde. Sie beruhen auf der Annahme, dass ausreichende, verlässliche und für eine Kampagne interessante Informationen über das Verhalten und die Einstellungen von Wahlberechtigten verfügbar sind und in Daten abgebildet werden können. Es wird überprüft, ob die Repräsentativität der verwendeten Daten kritisch diskutiert wird. Außerdem soll die diskursive Verhandlung von theoretisch möglichen und tatsächlich eingesetzten Mitteln datenbasierter Wahlkampfmethoden vor dem Hintergrund ihrer technischen, rechtlichen und moralischen Grenzen hin untersucht werden.
- **Datenbasierte Repräsentationsbeziehungen.** Ausgehend von personenbezogenen Daten zu bisherigem Wahlverhalten und damit korrelierenden soziodemographischen Merkmalen seien Rückschlüsse möglich – so die wahlkampfpraktische Annahme –, die stabile Prognosen für zukünftige Wahlhandlungen ermöglichen (Jungherr, 2013). Es wird geprüft, inwiefern dieses Abbildungsverhältnis und die behauptete ‚epistemische Autorität‘ von Daten diskursiv thematisiert und instrumentalisiert werden. Neben konkreten Problemen der Selektivität und Qualität der Kampagnen zugrunde gelegten Daten, ihrer Prognosefähigkeit für Handlungen und Intentionen und der auf ihnen basierenden Modelle geht es dabei grundlegender um das epistemische Potential von Big Data und Big-Data Analysen in Abbildungsverhältnissen. Zu überlegen ist, inwiefern sie die Wahrnehmung von Akteuren und Anspruchsgruppen prägen, und es ist zu überprüfen, mit welchen kategorialen Annahmen bzw. epistemischen Modellen sie einhergehen und damit

die politische Wissensproduktion informieren (Ulbricht et al., 2017). In diesem Zusammenhang liegt es zudem nahe zu prüfen, welche Funktionen dem Sammeln und Auswerten digitaler Daten als Instrument der politischen Meinungsforschung im öffentlichen Diskurs zugeschrieben werden. So wird es einerseits möglich, das Verhältnis der sich herausbildenden analytischen Praktiken zu traditionellen Formen der Wählerforschung, Fokusgruppen und Meinungsforschung (*polling*), zu beleuchten, sollte im Diskurs ein solcher Vergleich hergestellt werden. Weiterhin ist zu fragen, welche Funktionen dem Sammeln und Auswerten digitaler Daten als Instrument der politischen Werbung und Ansprache von Bezugsgruppen in öffentlichen Diskursen zugeschrieben werden – und in welchem Verhältnis sie zu traditionellen Formen der Kommunikation mit Bürger*innen stehen (Bimber, 2014).

- **Datenbasierte Regulierungsabsichten:** Bei der Untersuchung der Repräsentationsverhältnisse kommt zugleich zum Tragen, dass Big Data nicht nur Muster gesellschaftlicher Wirklichkeit abbildet (bzw. Big Data diese Repräsentanz zugeschrieben wird), sondern es diese reziprok aktiv gestaltet. In dieser Richtung untersucht das Gutachten die Frage, inwiefern gesellschaftliche Wirklichkeit in Daten nicht aufgefunden, sondern vielmehr erst mittels datenbasierter Analysen erzeugt wird, und diese datenbasiert generierten, mit Repräsentativitätsansprüchen verknüpften Modelle als Ressourcen politischer Autorität mobilisiert werden (Disch, 2008; Rosanvallon, 2006). In diesem Sinn entfaltet die Sammlung, Aggregation, Auswertung und Interpretation von Daten in politischen Kampagnen eine performative Wirkung, indem sie sinnstiftend Identitäten oder Willensbekundungen hervorbringen, praktisch handlungsleitend eingesetzt werden und Regulierungsabsichten fundieren (Callon, 2007; Law, 2009). Entsprechend ist zu untersuchen, auf welche Weise Daten in „representative claims“ (Saward, 2010) artikuliert werden, um beispielsweise darzustellen, was die ‚Mehrheit der Leute‘ denkt, was ein ‚kollektives Interesse‘ darstellt und welches politische Programm daraus abgeleitet werden sollte. Dadurch wird aber, so ist anzunehmen, nicht der Wille der Wähler*innen gespiegelt, sondern im Prozess des Herstellens (datenbasierter) Repräsentationen hervorgebracht. Konsequenterweise geht es nicht nur darum, den Wählerwillen zu ermessen. Vielmehr kommen Bürger*innen selbst reflexiv in

die Lage, sich im politischen Prozess zu verstehen und zu positionieren. Geht man von prinzipiell interaktiv bewerkstelligten Repräsentationen aus, dann ist andererseits zu überlegen, auf welche Weise Bürger*innen aggregierte Daten einsetzen, um ihrerseits Themen zu setzen, politische Entscheidungsträger*innen zu adressieren und die politische Agenda zu bestimmen, wie etwa im *data activism* (Karpf, 2017; Milan, 2013; Schrock, 2016).

Die Bearbeitung dieser drei Themen wird durch folgende **drei Thesen** orientiert:

- **These 1:** Es wird erstens angenommen, dass sich nicht eine solitäre hegemoniale Vorstellung darüber herausbildet, welche Bedeutung Big Data in und für Wahlkämpfe hat. Vielmehr werden der Begriff und die damit verknüpften datenbasierten Praktiken mit variablen Definitionen erfasst, in Bezug auf verschiedene historische, literarische oder räumliche Analogien konkretisiert und mit unterschiedlichen Bedingungen, Konsequenzen und Bewertungen in Verbindung gebracht. Big Data als „Herausforderungslandschaft“ (Geiselberger/Moorstedt, 2013: 9) wird so gesehen gesellschaftlich relevant, indem es als Phänomen und Problem diskursiv gerahmt wird (Keller/Knoblauch/Reichertz, 2013). Die Vielzahl an Formen des diskursiven Bezugs trägt hingegen nicht dazu bei, das Verständnis für den Sachverhalt zu schärfen – vielmehr wird in den Diskursen ‚Big Data‘ als semantisch zu füllender ‚empty signifier‘ gebraucht (Barassi, 2016; Laclau, 1996).
- **These 2:** Es wird zweitens angenommen, dass die Formulierung von Abbildungs-, Repräsentations-, und Regulierungsansprüchen im Zustand der fundamentalen Unsicherheit probabilistischer Schlüsse geschieht. Entsprechend stehen sie in der Spannung zwischen dem Definieren retrospektiver Gewissheiten, dem Prognostizieren prospektiver Wahrscheinlichkeiten und dem Diskutieren ihres Eintretens bzw. Nicht-Eintretens.
- **These 3:** Es wird drittens angenommen, dass in den diskursiv konstituierten Repräsentationsbeziehungen zum einen die Akteure nicht implizit bleiben, sondern sowohl Diskursakteure in Sprecherpositionen aktiv mit Wortmeldungen auftreten als auch datenbasiert Subjektpositionen (etwa als Wählerschaft, Mütter, Arbeitnehmende) formuliert werden. Zum anderen wird angenommen, dass sich das

Spektrum um neue Akteure bzw. Akteurskonstellationen (z.B. Datenanalyt*innen, Programmierer*innen) erweitert.

5 ARBEITSPROGRAMM

Um diese Themen zu bearbeiten, wurde ein **qualitativ-verstehendes methodisches Design in sieben Modulen** angewandt. Die überwiegende Zahl der Studien zu Big Data bedient sich ihres analytischen Gegenstandes als Instrument und nutzt somit selbst das Erkenntnispotential, das in der Aggregation und Analyse von Daten, insbesondere digitalen Daten, liegt. Entsprechend entstehen derzeit neue Felder an der Schnittstelle verschiedener Disziplinen wie die Digital Humanities oder die Computational Social Sciences (Lazer et al., 2009). Im Zuge dieser Forschungsdynamiken sind hingegen qualitativ-verstehend vorgehende Studien bisher marginal (aber: Couldry/Powell, 2014; Ribes/Jackson, 2013). Dies ist problematisch, weil gerade Ansätze, die sich mit Big Data als Gegenstand befassen, in der Lage sind, auch die Varianz im Umgang mit Big Data und die epistemischen Widersprüche des Verstehens von Big Data offenzulegen. Im Blick auf datenbasierte Wahlkämpfe kann folglich eine qualitativ-verstehend vorgehende Analyse von Diskursen bestehende Ansätze sinnvoll ergänzen. Bisherige Ansätze in der Big-Data-Forschung arbeiten mit kommerziell verfügbaren Datensätzen (Hersh, 2015), ethnografischen Methoden (Nielsen, 2012) oder Interviews (Anstead, 2017; Kreiss, 2012, 2016). Auf diesem komplementären Weg wird es möglich zu untersuchen, welche Bedeutung und welche Bedeutsamkeit Daten in Wahlkampfaktivitäten zugeschrieben werden.

Die Studie konzentriert sich auf Mediendiskurse, die sich in Artikeln aus etablierten, professionellen journalistischen Quellen sowie Kommunikaten in sozialen Medien konstituieren. Diskurse werden verstanden als strukturierte und regulierte Komplexe von Aussagen, die kulturelle Bedeutungen und soziale Realitäten konstituieren und kollektiv vermitteln und die ihre symbolische Gestalt zumeist in verbal-kommunikativen Äußerungen erhalten (zur politikwissenschaftlichen Diskursforschung vgl. Herschinger/Nonhoff, 2014).

Die Studie ist komparatistisch angelegt und vergleicht Deutschland, Großbritannien und die USA. Die USA sind aus verschiedenen Gründen der ‚incubation space‘ (Anstead, 2017) datenbasierter politischer Praxis, während Großbritannien und Deutsch-

land zwar als G8/G7-Staaten technisch und wirtschaftlich entwickelte Demokratien sind, doch in der Implementierung und öffentlichen Aufarbeitung der Rolle und Funktion von Daten im Wahlkampf noch am Anfang stehen. Die Gründe für diese Unterschiede, die mit dem Gutachten aufgearbeitet werden, sind vielfältig (Anstead/Chadwick, 2008). Die USA waren bereits Vorreiter in computergestützten Direktmailings. In ihrem de facto Zweiparteiensystem finden sich nur schwach institutionalisierte Parteistrukturen mit stratachischer Struktur (gestuftes Herrschaftssystem, das eine Anzahl von innerparteilichen Gruppen und Personen umfasst) (McKibbin, 2016). Es existiert ein umfangreicher politischer Beratungssektor und es werden aufgrund der extensiven Wahlkampf- und Parteienfinanzierung deutlich mehr finanzielle Ressourcen für Wahlkämpfe aufgewandt als in europäischen Staaten (Falguera et al., 2014; Thompson, 2012). Das hohe Kampagnenbudget (Anstead, 2008; Stier, 2016) geht einher mit einer langfristigen Zusammenarbeit mit Dienstleistern zur Datensammlung, -aufbereitung und -modellierung sowie der Entwicklung kampagnen- und parteiinterner datenbezogener Infrastrukturen und Kompetenzen (Kreiss, 2016). Diese Praxis wird befördert durch offenere Bestimmungen zum Datenschutz und zur Datenweitergabe, sodass in den USA die Bundesstaaten Informationen verfügbar machen, beispielsweise Kontaktinformationen aller Wahlberechtigten, Informationen zu demografischen Merkmalen, zur Wahlbeteiligung in der Vergangenheit, auch zur selbsterklärten politischen Zugehörigkeit und zur Registrierung zur demokratischen bzw. republikanischen Vorwahl (sog. *voter files*; Pearlman, 2012; Sullivan, 2012). Hinzu kommen von Datenhändlern (Acxiom, Experian etc.) gekaufte Informationen (Hersh, 2015). Außerdem finden sich im Blick auf die Parteienkultur eine grundsätzliche Bereitschaft, quantitative Prozesse einzusetzen (u.a. wegen der Bedeutung online gesammelter Spenden und der Mobilisierung freiwilliger Wahlkampfhelfer*innen; Kreiss, 2012; Saatchi, 2012) sowie eine quantitative Evaluations- und Kampagnenkultur im Umfeld der Parteien (Gerber/Green, 2012).

In Großbritannien und Deutschland als den beiden Vergleichsländern dürfen deutlich weniger Informationen an Kampagnenorganisatoren weitergegeben werden (für GB: Cowley/Kavanagh, 2015; Crabtree, 2010; Ross, 2015; für D: Jungherr, 2016). Datenschutzrechtliche Bestimmungen schränken die Möglichkeit ein, ergänzende Informationen einzukaufen bzw. diese zusammenzuführen. Entsprechend steht hier den Parteien eine vergleichsweise unzuverlässige und rechtlich eingeschränkte Datenlage zur

Verfügung (Richel, 2016). In beiden Staaten haben die politischen Parteien in jeweils Mehrparteiensystemen einen höheren Grad an formaler Institutionalisierung, wobei Großbritannien eine relative Mehrheitswahl zum Unterhaus hat, Deutschland ein personalisiertes Verhältniswahlrecht (sowie Direktkandidat*innen, die nach relativem Mehrheitswahlrecht gewählt werden). In Deutschland wird der Wahlkampf nicht durch einzelne Kandidat*innen, sondern zentral durch die Parteien geführt; in Großbritannien liegen viele Entscheidungen auf Ebene lokaler politischer Vertretungen (Zittel, 2010). Durch den Vergleich Deutschland und Großbritannien wird es zudem möglich, europäische Spezifika und Unterschiede zu erfassen, etwa im Hinblick auf das Verhältnis zum Datenschutz. Hinzu kommt, dass der Einbezug einer dritten Sprache den Aufwand des Projekts enorm vergrößern würde und somit entfällt.

Der Vergleich soll aufzeigen, inwiefern in öffentlichen Diskursen die analytischen Praktiken in Bezug zum politischen System gesetzt werden, insbesondere zum Wahlsystem sowie zur politischen Kultur, vor allem im Blick auf das Parteienverhältnis und die Formen des Wahlkampfes. Möglich ist auch, dass durch den Vergleich weitere kulturelle oder rechtliche Faktoren wie etwa das Verständnis von Datenschutz und die entsprechenden rechtlichen Regelungen zutage treten, die die diskursive und handlungspraktische Verwendung von Big Data prägen. Dabei ist davon auszugehen, dass es sich nicht um technologisch determinierte, wohl aber um technologisch geprägte datenintensive Praktiken handelt. Weiterhin wird angenommen, dass die Wahlkampftechniken nicht länderübergreifend repliziert werden, sondern in die jeweiligen politischen Kontexte zu übersetzen sind.

6 METHODE

6.1 DATENSAMMLUNG UND DATENORGANISATION

Im ersten Schritt wurde das Korpus an Quellen zusammengestellt und für die weitere Analyse aufbereitet. Die Auswahl erfolgte als Kombination aus einer Erhebung aller relevanten Artikel in sechs überregionalen Tageszeitungen, die sich auf die jeweiligen Kampagnen beziehen und auf Datenpraktiken bzw. deren Implikationen abheben, er-

gänzt durch eine weiterführende Auswahl mittels Schneeballsampling und Suchmaschinenrecherchen.

Im Fokus der Analyse stehen drei Zeiträume: die 58. Präsidentschaftswahl in den USA (Vorwahlen und Präsidentschaftswahlkampf 2016), die Wahl zum britischen Unterhaus (2017) sowie die 19. Bundestagswahl (2017).

Die Erhebungszeiträume beginnen mit der Erklärung der offiziellen Kandidaturen durch Spitzenkandidat*innen (USA: 04/2015 durch H. Clinton; D 11/2016 durch A. Merkel) bzw. der Bekanntgabe der geplanten Neuwahl (GB: 04/2017). Die Erhebungszeiträume enden 100 Tage nach Wahltag (USA: 03/2017; UK: 09/2017; D 12/2017, hier wird Material fortlaufend einbezogen).

In Anlehnung an die Verfahrensweise der Grounded Theory wurde eine Strategie des *theoretical sampling* gewählt, die es ermöglicht, das Korpus schrittweise zu erweitern und eine „konzeptuelle Repräsentativität“ zu erreichen (Glaser/Strauss, 1998), also keine statistische, sondern eine Repräsentativität hinsichtlich der Auswahl aller relevanten Fälle für eine konzeptuelle Sättigung. Es handelt sich um ein zirkuläres Auswahlverfahren, bei dem aus dem bereits erhobenen Material weitere Kriterien für die Auswahl supplementärer Analyseeinheiten gewonnen werden. Während die ersten Erhebungen noch von theoretischen und praktischen Vorkenntnissen geleitet sind (sog. „sensibilisierende Konzepte“, vgl. Strauss/Corbin, 1996: 25ff.), werden die weiteren Selektionsentscheidungen auf Basis der Analysen des bereits erhobenen Materials getroffen (Pentzold/Meier 2010). Die Erhebung endet, wenn keine neuen Aspekte des Untersuchungsgegenstandes mehr gefunden werden, also eine theoretische Sättigung eintritt (vgl. Glaser/Strauss, [1967] 2005: 69f).

Phase 1. Offenes Sampling und themenbezogene Strukturierung

Als Einstiegsmaterial dienten zwölf thematisch einschlägige und aus einer früheren Erhebung zum Thema ‚Big Data‘ bekannte Beispielartikel aus *New York Times*, *Washington Post* und *Süddeutscher Zeitung*. Aus dem Material wurden teilweise trun-kierte Schlüsselbegriffe generiert, die sich im medialen Diskurs als zentral für die Beschreibung datengetriebener Praktiken in politischen Kampagnen erwiesen: ‚**election**‘

bzw. ‚campaign‘ in Verbindung mit ‚big data‘, ‚data‘, ‚algorithm‘, ‚microtargeting‘, ‚micro-targeting‘, ‚bot‘ bzw. deren deutschsprachige Pendanten.

Als Basis der Erhebung wurden in jedem Land zwei Tageszeitungen ausgewählt, die hinsichtlich ihrer redaktionellen Ausrichtung einmal eher dem linken und einmal eher dem konservativen politischen Spektrum zugeordnet werden (Blake, 2014; Maurer/Reinemann, 2007; Smith, 2017):

- für USA: *Washington Post*, *New York Times*
- für GB: *The Guardian*, *The Times*
- für D: *Süddeutsche Zeitung*, *Frankfurter Allgemeine Zeitung*

Die Gesamtausgaben der sechs Zeitungen wurden mithilfe der Recherchedatenbank *factiva* bzw. für die Frankfurter Allgemeine Zeitung mittels des *F.A.Z-Bibliotheksportals* nach den Schlüsselbegriffen durchsucht. Ergänzend wurde nach ‚big data‘ allein sowie nach ‚Cambridge Analytica‘ als angenommenen zentralem Akteur in den untersuchten anglo-amerikanischen Wahlkämpfen gesucht, um sicherzustellen, dass alle relevanten Beiträge erfasst wurden. Aufgreifkriterium der Artikel war die Thematisierung datenbasierter Praktiken im Zusammenhang mit politischen Wahlen. Anhand der Artikelüberschriften wurden zunächst diejenigen Artikel aussortiert, die offensichtlich keinen Bezug zu politischen Wahlen aufwiesen. Im nächsten Schritt wurden die Absätze gelesen, die die jeweiligen Suchbegriffe enthielten. Im Zweifel wurden Artikel komplett gelesen, um über die Aufnahme ins Sample zu entscheiden. Eine erste Analyse des so gewonnenen Materials erbrachte die Identifikation von und Zuordnung der Artikel zu drei Themenkomplexen, die den medialen Diskurs zu Big-Data-Praktiken und Wahlen inhaltlich strukturierten: a. Datenbasierte Prognosen, b. Datenbasierte Kommunikationskampagnen, c. Datenbasierte Öffentlichkeit.

a. Datenbasierte Prognosen. Big Data ist mit der Hoffnung auf neue, erweiterte Erkenntnispotentiale verbunden (vgl. Geiselberger/Moorstedt, 2013; Mayer-Schönberger/Cukier, 2013; Mayer-Schönberger, 2015), die sich nicht nur auf den Zustand des *realis* beschränken, sondern auch die Sphäre des *potentialis* umfassen, also prognostische Schlüsse erlauben. Dieses Diskurscluster bildet den Kernbereich der Analyse: Es enthält Beiträge, die sich mit datenintensiven Methoden der Wähleranalyse befassen – einmal vonseiten der Parteien und Kandidat*innen und einmal vonseiten der Meinungsforscher*innen und Demoskop*innen. Politische Akteure bedienen sich

der unter dem Begriff des Microtargeting oder Psychotargeting bekannt gewordenen Methode der Auswertung verschiedener personenbezogener Daten zur Identifikation und gezielten Ansprache von potentiellen Wähler*innen. Wahlforscher*innen, Demoskop*innen und Statistiker*innen verwenden teilweise ebenfalls personenbezogene Daten, um Wahlausgänge zu prognostizieren. Die aus unterschiedlichen Datensets gewonnenen Prognosen werden hinsichtlich ihrer Validität, Belastbarkeit, Nützlichkeit, ethischen Zulässigkeit und individuellen und gesellschaftlichen Implikationen diskutiert. Das Sample umfasst insgesamt 52 Artikel, die mehrheitlich im Politikressort der jeweiligen Zeitungen erschienen sind.

b. Datenbasierte Kommunikationskampagnen. Ein weiteres Diskurscluster befasst sich mit der Anwendung datenintensiver Methoden in der politischen Kommunikation. Ein zentrales Thema ist der Einsatz von *social bots*, die automatisiert Nachrichten in sozialen Netzwerken absetzen – entweder in Form von Spam oder aber als automatisierte Antworten auf Nutzerposts. Ebenfalls prominent vertreten waren Beiträge, die sich mit dem Phänomen viraler Prozesse im Fall der Verbreitung von Fake News, also von Falschmeldungen, die vor allem Online verbreitet werden, als Propagandastrategie auseinandersetzen. Zwar basieren diese Praktiken entweder auf der Anwendung von Algorithmen, die durch Big Data gespeist werden, oder sie erzeugen selbst große Datenmengen (wie im Falle von Spam-Attacken oder der Verbreitung von Fake News), doch stehen Verwendung oder Erzeugung von Big Data nicht im Zentrum des Diskurses. Die epistemische Dimension der zugrundeliegenden oder produzierten Daten wird nicht thematisiert. Vielmehr geht es um Produktion und Wirkung neuer Kommunikationsstrategien in Online-Umgebungen.

c. Datenbasierte Öffentlichkeit. Das dritte Diskurscluster thematisiert die algorithmisierten Bedingungen, in denen heute politische Information und Meinungsbildungsprozesse stattfinden. Besonders Suchmaschinen und soziale Plattformen wie Facebook und Twitter stehen im Verdacht, die Wählerschaft zu parzellieren, indem sie durch Personalisierung und auf Ähnlichkeiten basierenden Rankings thematisch-weltanschaulich separierte, inhärent homogene und verbindungsschwache Echokammern oder Filterblasen erzeugen, die die Herausbildung einer partizipativen, bürgerlichen Öffentlichkeit verhindern. Zwar handelt es sich bei diesen strukturellen Bedingungen der algorithmisierten Öffentlichkeit ebenfalls um Ergebnisse datenbasierter

Prozesse, die Bedeutung von Big Data für Ergebnisse von Suchmaschinenrankings oder News Feeds werden aber nicht explizit diskutiert.

Phase 2. Selektives Sampling

Nach der Systematisierung des Materials in drei Diskurscluster und einer Festlegung auf den Kernbereich der Untersuchung, der sich auf datenbasierte Prognosen konzentriert, konnte gezielt nach Material gesucht werden, welches das Sample nach dem Prinzip der minimalen und maximalen Kontrastierung vervollständigt: Anhand zentraler Begriffe aus den Zeitungsartikeln wurden Kommunikate recherchiert, die die Anwendung von prognostischen Analysen im Wahlkampf, deren Bedingungen, Bedeutungen und Implikationen thematisierten.

Dabei wurden lediglich solche Kommunikate aufgenommen, die neue Aspekte oder Argumente bezüglich des Untersuchungsgegenstands enthielten. Zum einen wurden nach dem Schneeballprinzip Quellenverweise in den vorliegenden Artikeln verfolgt und geprüft, zum anderen eine Suchmaschinenrecherche auf Basis einer Liste mit 53 Suchbegriffen durchgeführt, die den Artikeln der sechs Tageszeitungen entnommen wurden. Hierbei handelte es sich um zentrale Akteure (Personen, Firmen, Softwareprogramme) oder Praktiken datenbasierter Prognosen, über die weitere Informationen gewonnen werden sollten: **1. Datenquellen** (z. B. *voter registration records, voter files*), **2. Methoden** (z. B. *OCEAN, microtargeting / micro-targeting, predictive models, A/B testing, postcode-targeting*), **3. Firmen, Agenturen, Meinungsforschungsunternehmen** (z. B. Cambridge Analytica, Proofpoint, Aggregate IQ, Civis Analytics, Targeted Victory, YouGov), **4. Akteure** (z. B. Alexander Nix, Steve Bannon, Robert Mercer, Nate Silver, Sasha Issenberg, Vincent Harris, Yvonne Hofstetter, Michal Kosinski), **5. Soziale Plattformen und Apps** (z. B. Twitter, Snapchat, Facebook).

Die Suchanfragen wurden im privaten Surfmodus bei Google eingegeben, die sowohl in Deutschland als auch in Großbritannien und den USA die mit Abstand meistgenutzte Suchmaschine ist (Alexa.com, 2017). Es sollten alle im öffentlichen Diskurs relevanten und durch die Suche sichtbaren Kommunikate erfasst werden, daher wurde angesichts der jeweils sechs- bis siebenstelligen Trefferzahlen zu den Suchanfragen eine Beschränkung auf die Auswertung der ersten fünf Ergebnisseiten – also die ersten 50 Suchergebnisse – auf thematisch relevante Treffer hin vorgenommen.

Zum einen erschienen auf hinteren Ergebnisseiten in den Stichproben keine relevanten Ergebnisse mehr, zum anderen entspricht es dem alltäglichen Google-Nutzungsverhalten, die Suche nach den ersten, wenn nicht schon nach der ersten Trefferseite zu beenden (van Deursen/van Dijk, 2009). Enthielten die gefundenen Kommunikate wiederum Hinweise auf weitere relevante Beiträge, wurden diese ebenfalls verfolgt. Bei einer Gesamtanzahl von 121 Kommunikaten wurde die Erhebung beendet, als keine neuen Aspekte des Gegenstandsbereichs mehr identifiziert wurden. Die Suchmaschinenrecherche wurde von zwei Personen ausgeführt, die wiederum trotz unterschiedlicher Suchanfragen in einigen Fällen die gleichen Artikel archivierten (insgesamt zwanzig). Die dem theoretischen Sampling innewohnende Zirkularität, die bei einer entsprechenden Sättigung des Samples dazu führt, dass kein neues Material bzw. keine neuen Aspekte des Untersuchungsgegenstandes mehr gefunden werden, konnte durch sich doppelnde Archivierungen, also Doubletten, als erfüllt betrachtet werden. Hauptsächlich handelt es sich bei den Beiträgen um Artikel aus Zeitungen und Nachrichtenmagazinen. Ergänzt werden sie von Blog- und Microblog-Einträgen (z. B. von David Carroll, der die Herausgabe der von Cambridge Analytica über ihn gesammelten Daten erwirkt und dies öffentlich dokumentiert hat).

Weiterhin ließ sich bereits in der Phase der Diskurseingrenzung ein für den Diskurs zentraler Beitrag über die Verwendung von Big-Data-Analysen im Wahlkampf identifizieren: Der von Hannes Grassegger und Mikael Krogerus (3.12.2016) im Schweizer *Magazin* veröffentlichte Artikel „Ich habe nur gezeigt, dass es die Bombe gibt“ über den Einsatz von aus Facebook-Tests generierten Persönlichkeitsprofilen für politisches Microtargeting. Dieser wurde in verschiedenen Publikationen reproduziert, teilweise auch in übersetzter Fassung (z. B. am 28. Januar 2017 auf *motherboard.com* oder am 3. Dezember 2016 im *Tages-Anzeiger*), sodass er mehrmals im Sampling-Prozess erfasst wurde. Für die spätere Analyse der Kommunikate wurde lediglich die Erstveröffentlichung in *Das Magazin* beibehalten. Zusammen mit den Ergebnissen der Zeitungsrecherche ergibt sich ein **Gesamtsample von 153 Analyseeinheiten**.

Dass die Strategie des *theoretical sampling* (Schneeballprinzip und Suche nach zentralen aus dem Einstiegsmaterial gewonnenen Begriffen) für Online-Diskurse gut geeignet ist, belegt die lexikalische Suche in den Kommunikaten des Gesamtsamples: Lediglich die Hälfte der Kommunikate (52,9%) enthalten die Begriffe ‚Big Data‘, ‚Microtargeting‘.

ting' oder ‚Psychotargeting‘ (oder verwandte Schreibweisen), sodass bei einer konventionellen Stichwortsuche nach diesen Begriffen ein Großteil des Materials nicht gefunden worden wäre.

6.2 INHALTSANALYTISCHE AUFBEREITUNG

Die inhaltsanalytische Aufbereitung des Materials erfolgte in einem qualitativen Verfahren, bei dem deduktive Kategorien während des Codierprozesses um induktive Kategorien nach dem Prinzip der Subsumption erweitert werden (vgl. Mayring, 2015; Schreier, 2014). Dieses Verfahren erlaubt sowohl den Abgleich zwischen den Korpusdokumenten und somit die Feststellung von Homologien in der Argumentationsstruktur als auch die Markierung textualisierter Differenzen.

Die Aufbereitung orientiert sich dabei an drei Schritten: Erstens der **Zusammenfassung inhaltlich gleicher Passagen**; zweitens der **Explication von indexikalischen, inhaltlich zu erläuternden Passagen**; drittens der **inhaltlichen Strukturierung** gemäß der drei Schwerpunktthemen des Gutachtens (Abbildungsverhältnisse, Regulierungsabsichten und Repräsentationsbeziehungen). Um das Korpus von 153 Diskursbeiträgen möglichst detailliert und gleichzeitig ergebnisorientiert im Hinblick auf die Forschungsfragen zu analysieren, wurde zunächst auf Basis von 15 Beiträgen (zehn Prozent des Samples) ein Kategoriensystem erarbeitet. In einem zweiten Schritt erfolgte der Pretest: Das Kategoriensystem wurde von drei Codierern auf neun besonders argumentativ gehaltvolle Kommunikate angewandt. Im Verfahren der kommunikativen Validierung (Steinke, 2008) wurde das Codesystem um zwei Codes erweitert und die Segmentierung des Diskursmaterials sowie die Vergabe der Codings in der Forschungsgruppe evaluiert, sodass eine hohe Übereinstimmung der Codierungen gewährleistet werden konnte. Zur Absicherung versahen die Codierer strittige Stellen mit Memos, die in der abschließenden Auswertung der inhaltsanalytischen Aufbereitung gelöst wurden. Das verwendete Codesystem bestand schließlich aus 15 Obercodes und 19 Subcodes. Insgesamt wurde in den 153 diskursiven Beiträgen 4784 Codings vergeben (Codebuch siehe Anhang).

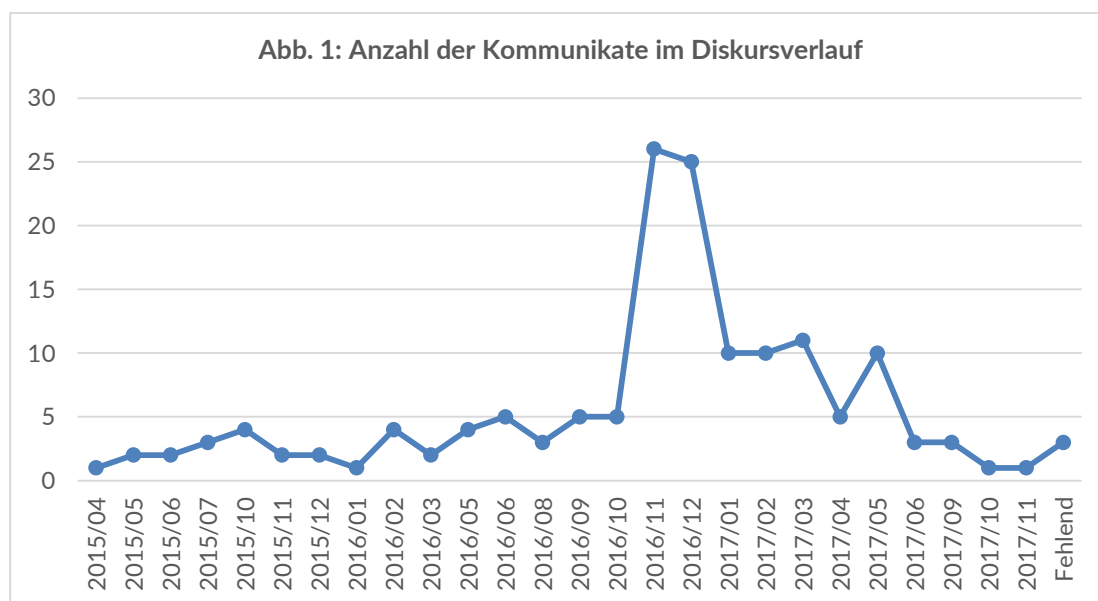
7 ERGEBNISSE

Im Folgenden werden die empirisch gewonnenen Ergebnisse in **drei Schritten** dargestellt und diskutiert. Erstens wird ein **Überblick über die grobe inhaltliche Struktur und den thematischen Verlauf des Diskurses** gegeben (Makro-Ebene). Zweitens werden mit **Cambridge Analytica** eine wesentliche Akteurskonstellation und damit verbundene Aushandlungsprozesse im Detail betrachtet (Mikro-Ebene). Schließlich werden drittens die Ergebnisse hinsichtlich der drei Dimensionen von **Abbildungsverhältnissen, Repräsentationsbeziehungen und Regulierungsabsichten** erläutert (Meso-Ebene).

Die Ergebnisdarstellung enthält eine Reihe von Belegziten aus dem Material, längere Passagen sind eingerückt.

7.1 MAKROANALYTISCHE DISKURSSTRUKTURIERUNG

In der Betrachtung der Menge der Kommunikate im Zeitverlauf des Untersuchungszeitraums lässt sich ein **Diskurshöhepunkt** in den Monaten November und Dezember 2016 erkennen (siehe Abb. 1). Allein in diesen beiden Monaten wurde ein Drittel (33,3%) des diskursiven Materials veröffentlicht, rund ein weiteres Viertel danach bis Mai 2017, sodass 57 Prozent aller Kommunikate innerhalb von sieben Monaten kulminieren.



Vier **Ereignisse** bzw. Ereignisketten konnten identifiziert werden, die eine entsprechende Diskursverdichtung mit einem höheren Kommunikationsvolumen nach sich zogen:

1) Im näheren zeitlichen **Vorfeld der Präsidentschaftswahl** am 8. November 2016 in den USA beschäftigen sich vermehrt Veröffentlichungen mit datenintensiven Wahlkampagnen der Kandidatin Hillary Clinton und des Kandidaten Donald Trump. Sie thematisieren beispielsweise Budgets, Methoden und Nutzen der angewandten Strategien und diskutieren die datenbasiert gemachten Prognosen, auch im Verhältnis zu herkömmlichen Formen der Wahlkommunikation und Abschätzung von Wahlausgängen.

2) Nach dem **Wahlsieg Trumps** am 8. November wendet sich der Diskurs dem Eintreten bzw. hier häufiger dem Nicht-Eintreten der vorab gemachten Wahlprognosen zu, die in der Mehrheit einen Sieg Clintons vorausgesagt hatten. Hier wird dann auch die epistemische Autorität durch Big-Data-Analysen diskutiert. Die entsprechende Grundfrage wird in einem Artikel der *Süddeutschen Zeitung* mit dem Titel „Voll daneben: Warum Prognosen trotz Big Data immer unzuverlässiger werden“ (6.2.2017) anschaulich umrissen: „Die Ära von Big Data ist von einem Paradox geplagt: So viele Informationen wie nie zuvor stehen frei zur Verfügung, und dennoch sind verlässliche Aussagen über die Zukunft rar. Welcher Demoskop hätte vor einem Jahr gewusst, dass Großbritannien aus der EU austritt und Donald Trump US-Präsident wird? In beiden Fällen lagen selbst ausgefeilte mathematische Modelle grob daneben.“

3) Anders gelagert ist das dritte Moment, in dem nicht politisches Geschehen diskurssetzend war, sondern publizistische Tätigkeit: Das **Erscheinen des Artikels** „Ich habe nur gezeigt, dass es die Bombe gibt“ von Hannes Grassegger und Mikael Krogerus im Schweizer *Das Magazin* am 3. Dezember 2016, also in einem zeitlichen Abstand zur Wahl, löst eine starke Resonanz aus. Viele Kommunikate nehmen explizit Bezug auf die darin geäußerten Aussagen des Cambridge-Psychologen Michal Kosinski, des Cambridge Analytica-CEOs Alexander Nix und der Autoren selbst über den Einsatz und die Wirksamkeit von Microtargeting-Maßnahmen auf der Basis von Psychogrammen, die von Cambridge Analytica anhand von Facebook-Quizen erstellt und in Trumps Wahlkampf zum Einsatz gekommen seien. Es erscheinen Zusammenfassungen, Kommentare, Kritiken, Hintergrundberichte und Interviews mit den im *Magazin*-Artikel

genannten Akteuren sowie Kommunikate, die Bezüge zu den anstehenden Wahlen in Deutschland und Großbritannien herstellen.

4) Schließlich wird im Frühjahr 2017 über die Arbeit von Cambridge Analytica für verschiedene **UK Leave-Kampagnen** sowie die Datensammelpraktiken mehrerer Datenfirmen (u.a. Cambridge Analytica, Aggregate IQ) während des Brexit-Referendums 2016 ebenfalls medienöffentlich diskutiert, deren Rechtmäßigkeit in Untersuchungen durch die Electoral Commission festgestellt werden soll. Hier werden häufig Bezüge zum US-Wahlkampf hergestellt, wo Cambridge Analytica für die Trump-Kampagne gearbeitet hat. Die Diskussionen stehen ebenfalls im Zusammenhang mit den Neuwahlen zum britischen Unterhaus am 23. Juni 2017 und den Einsatz von datenintensiven Maßnahmen im Zuge der Kampagnen.

Für die beiden weniger diskursintensiven Zeiträume, April 2015 bis Oktober 2016 und Juni bis Dezember 2017, sollen hier ebenfalls kurz die Hauptthemen rekonstruiert werden:

Von Frühjahr 2015 bis Spätsommer 2016 beschäftigen sich die Kommunikate hauptsächlich mit den datenintensiven Methoden verschiedener Kandidat*innen und Parteien, die zur **britischen Unterhauswahl am 7. Mai 2015** antraten, sowie der Kandidat*innen der **Vorwahlen in den USA** in der ersten Jahreshälfte 2016, insbesondere von Ted Cruz, Marco Rubio, Bernie Sanders, aber auch Donald Trump und Hillary Clinton. In diesem Zuge erscheinen einige Hintergrundberichte, die die Methoden politischen Microtargetings vorstellen, zum Beispiel über die Kooperationen während des US-Wahlkampfs mit Facebook, das Targeting-Maßnahmen der Parteien unterstützt:

„Facebook, which told investors on Wednesday it was ‚excited about the targeting‘, does not let candidates track individual users. But it does now allow presidential campaigns to upload their massive email lists and voter files – which contain political habits, real names, home addresses and phone numbers – to the company’s advertising network. The company will then match real-life voters with their Facebook accounts, which follow individuals as they move across congressional districts and are filled with insightful data. [...] Facebook will find users who like lots of political content and share it with their friends, mark them as ‘political influencers’ and allow campaigns to target them specifically. As one of Facebook’s monetization executives explained to Wired after

the quiet launch of influencer tracking in November: 'People are more likely to trust information that their friends share.'" (Guardian, 28.1.2016)

Auch einige Kommunikate, die neue **datenintensive Prognosemodelle** diskutieren, erscheinen in dieser Zeit:

"A traditional poll is basically a kind of spot check—a dipstick dropped into one part of an engine at one particular moment in time. But even back in 2007, sampling errors and nonresponse rates were beginning to make those spot checks chronically inaccurate. Now the dipstick wasn't just a momentary reading; it didn't even tell you how much oil you had left. The rise of data analytics in campaigning suggests a model that's more like an engine that is monitored continuously, with sensors collecting a record of performance over time. Getting to that kind of continuous monitoring, however, means building long-term databases of information about voters that can be refreshed and crunched a bunch of different ways." (Wired, 6.6.2016)

Es fällt auf, dass in der gesamten eineinhalbjährigen Zeitspanne vor der diskursiven Klimax nur ein einziger deutschsprachiger Artikel im Sample zu finden ist (Heinrich-Böll-Stiftung, 12.5.2016). Der Diskurs über den Einsatz von datenintensiven Methoden in Wahlkämpfen erreichte die deutschen Debatten also erst mit deutlicher Verzögerung.

Insgesamt liegt der **diskursive Schwerpunkt auf und in den USA**: Mehr als die Hälfte der Kommunikate (52,9%) stammen aus US-amerikanischen Quellen. Großbritannien und der deutschsprachige Raum (Deutschland, Österreich und der deutschsprachige Teil der Schweiz) sind mit jeweils 22 Prozent der Kommunikate im Gesamtsample vertreten. Mehrheitlich (zu 70%) rekurrieren britische und deutsche bzw. deutschsprachige Kommunikate auf Entwicklungen und Methoden in den USA.

Weiterhin konnte beobachtet werden, dass einige Informationen die bereits im Diskurs benannt wurden, nicht sofort zirkulierten, sondern erst zu einem späteren Zeitpunkt Aufmerksamkeit erhalten. Beispielsweise erscheint bereits am 7. Juli 2015 ein Bericht über Cambridge Analytica in der Online-Ausgabe von *Politico*, der die Dienste des Unternehmens für die Kampagne von Ted Cruz in den US-amerikanischen Vorwahlen thematisiert. Er erwähnt auch die Verwendung von Psychogrammen, fokus-

siert aber hauptsächlich die personalen Verbindungen zwischen Cambridge Analytica, Ted Cruz und der Mercer-Familie, Großspendern der Konservativen, die später, nach Cruz' Ausscheiden aus dem Wahlkampf, auch Donald Trump unterstützen. Vier Monate später veröffentlicht Sascha Issenberg bei *Bloomberg* einen detaillierten Hintergrundbericht über die Methoden und Rolle von Cambridge Analytica im Wahlkampf von Ted Cruz unter dem Titel „Cruz-Connected Data Miner Aims to Get Inside U.S. Voters' Heads“ (Bloomberg, 12.11.2015), in dem er das Potential des psychographischen Profiling kritisch diskutiert. Am 11. Dezember 2015 schließlich veröffentlicht der *Guardian* einen Beitrag mit dem Titel „Ted Cruz using firm that harvested data on millions of unwitting Facebook users“. Alle drei Artikel publizieren weitestgehend dieselben Informationen über Cambridge Analytica wie der über ein Jahr später veröffentlichte *Magazin*-Artikel (3.12.2016): Die Firma erstelle anhand von Facebook-Daten Persönlichkeitsprofile von Nutzer*innen, die wiederum als Basis für Microtargeting-Strategien dienen. Anders als dieser Beitrag lösen die drei englischsprachigen Veröffentlichungen allerdings keine große Resonanz im öffentlichen Diskurs aus.

Erst der *Magazin*-Artikel, der nach dem Sieg Donald Trumps bei den US-Wahlen erscheint, initiiert eine breite Debatte über Methoden, Chancen, Risiken und Rechtsgrundlagen des politischen Micro- und Psychotargeting, insbesondere auch im deutschsprachigen Raum. Eine mögliche Erklärung hierfür könnte darin liegen, dass Cruz sich in den Vorwahlen nicht durchsetzen konnte, sein datengetriebener Wahlkampf ihm letztlich also keinen Vorsprung vor seinen Mitbewerber*innen um die Präsidentschaftskandidatur sichern konnte. Erst mit dem Wahlsieg Trumps, der den Erfolg von Cambridge Analytica Methoden zu belegen schien, ist ein Interesse an den datenintensiven Methoden der Kampagnenführung entstanden, weil nun Erklärungen für den vermeintlichen Überraschungssieg gesucht wurden.

In diesem Sinne argumentiert auch Dennis Horn in einem Kommentar (WDR, 5.12.2016), nämlich dass Schuldige für den Sieg Trumps gesucht würden – in diesem Fall Datenanalysten, die durch Microtargeting Wähler*innen manipuliert hätten:

„Mir ist selten ein Artikel so häufig in die Timelines gespült worden wie der mit dem nachdenklichen Mann im Karohemd, von dem behauptet wird, Donald Trump mit zum Sieg verholfen zu haben. Michal Kosinski heißt dieser Mann. ‚Ich habe nur gezeigt, dass es die Bombe gibt‘, heißt der Artikel. Er erzählt von

Kosinskis Erfindung in der Psychometrie: einer Form der Big-Data-Analyse, mit der sich einzelne Menschen so durchleuchten lassen, dass man mit diesem Wissen am Ende ganze Wahlen manipulieren könne. [...] Soweit dieser Artikel aus dem Schweizer ‚Magazin‘. Und nun mein Widerspruch: Etwas, das so gut ins Bild passt, das als tragische Heldengeschichte daherkommt und endlich einen Sündenbock und einen Grund für den Wahlsieg Donald Trumps liefert – etwas, das gerade jeder hören möchte, das fast schon zu perfekt klingt, um wahr zu sein, sollte man vielleicht zweifach, dreifach, vierfach auseinandernehmen. Schon grundlegende journalistische Reflexe müssten da greifen.“

Nach Abflauen des Diskurses im Juni 2017 beziehen sich die Kommunikate hauptsächlich auf die **Wahlen in Deutschland** oder **Targetingmöglichkeiten in neuen Online-Umgebungen** (*Snapchat*), nur einzelne Berichte lieferten noch Hintergründe zu den Wahlen in den USA oder Großbritannien.

Der überwiegende Teil des diskursiven Materials stammt aus **professionellen journalistischen Quellen (81,1%)**. Korrespondierend mit dem Diskursschwerpunkt in den USA ist die *New York Times* mit insgesamt 21 Artikeln als häufigste Quelle vertreten. Weitere zehn Prozent der Quellen sind Unternehmenswebsites und -blogs. Eine genaue Quellenaufschlüsselung befindet sich im Anhang.

7.2 DISKURS IN DER NUSSSCHALE: DER FALL CAMBRIDGE ANALYTICA IN DER ANALYTISCHEN MIKROPERSPEKTIVE

Cambridge Analytica und deren Mutterkonzern Strategic Communication Laboratories (SCL) werden als **zentrale Akteure im öffentlichen Diskurs über Big-Data-Anwendungen in Wahlkämpfen** verhandelt: 38 Prozent des diskursiven Materials referieren auf die Unternehmensgruppe oder Cambridge Analytica im Speziellen, keine andere Datenfirma wird derart häufig thematisiert. Es folgt Civis Analytics mit knapp neun Prozent an Kommunikaten. Dies ist eine Big-Data-Firma, die von ehemaligen Mitarbeiter*innen der Obama-Kampagne 2012 gegründet wurde, und auf deren Einschätzungen etwa zu Wahlprognosen häufig als Expertenmeinungen in den journalistischen Berichten zurückgegriffen wird. Andere Datenfirmen wie Targeted Victory oder Aggregate IQ werden dagegen nur vereinzelt thematisiert. Diese Fokussierung auf

Cambridge Analytica und deren Mutterkonzern SCL spiegelt sich ebenfalls in der Konzentration des Diskurses auf die datenintensiven Praktiken derjenigen Kampagnen, die von dieser Firma unterstützt wurden: Trumps Wahlkampf und die Leave-Kampagnen im Zuge des Brexit-Votums 2016. Datenbasierte Wahlkampfpraktiken der anderen Parteien werden weitaus seltener thematisiert.

Im Zuge der großen Aufmerksamkeit, die Cambridge Analytica nach Erscheinen des *Magazin*-Artikels (3.12.2016) im deutsch- und im englischsprachigen Raum zuteil wird, wird auch vermehrt über das vermeintlich weltweite Agieren des Unternehmens berichtet: „The firm has worked on campaigns in Argentina, Kenya, Ghana, Indonesia and Thailand; the Pentagon has used it to conduct surveys in Iran and Afghanistan.“ (The Huffington Post, 17.3.2017) Andere Kommunikate berichten von Big-Data-Kampagnen in Trinidad, Lettland, Nigeria, Moldawien, Iran: „Cambridge Analytica – a U.S.-incorporated affiliate of SCL Group, a British firm that has worked on campaigns in 22 countries – says it is working on 50 campaigns in the United States, all of them Republican.“ (WP, 27.10.2016)

In den ersten Wochen vor und nach der Trump-Wahl befinden sich Cambridge Analytica und deren verschiedene Sprecher, insbesondere der CEO Alexander Nix, auf dem Höhepunkt ihrer diskursiven Präsenz. Bis zur Wende der öffentlichen Meinung nach der Trump-Wahl, als Cambridge Analytica für seine Methoden und den reklamierten Anspruch auf den Beitrag am Wahlsieg Trumps heftig kritisiert wird, treten die Firma und ihre Sprecher mit Aussagen im Diskurs hervor, die das revolutionäre Potential der genutzten Methoden hervorheben: „Your behavior is driven by your personality and actually the more you can understand about people’s personality as psychological drivers, the more you can actually start to really tap in to why and how they make their decisions,’ says Nix. ‘We call this behavioral microtargeting and this is really our secret sauce, if you like. This is what we’re bringing to America.’“ (Bloomberg, 12.11.2015)

Es lässt sich beobachten, dass der Cambridge Analytica zugesprochene Einfluss auf Wahlen und Kampagnenausgängen im ersten Quartal des Jahres 2017 zunehmend diskursiv relativiert und angezweifelt wird, sodass die Firma im März schließlich von den eigenen Aussagen zu Aktivitäten und Einfluss abrückt. Im Zentrum dieser Verhandlung steht ein Artikel der *New York Times* (6.3.2017) mit dem Titel „Data Firm Says ‘Secret Sauce’ Aided Trump; Many Scoff“, der den Einfluss der Firma auf mehreren

Ebenen anzweifelt: die Auftragslage und die Involvierung der Firma in internationale Wahlkampagnen oder Regierungsprojekte, den Einsatz psychographischer Methoden im Trump-Wahlkampf, die Beteiligung an der Brexit-Kampagne. Hatte Cambridge Analytica zunächst mit seiner Rolle als Präsidentenmacher und Brexit-Herbeiführer geworben, dementiert es diese Aussagen später und bestätigt öffentlich, weder von den Leave-Kampagnen offiziell engagiert gewesen zu sein noch im Trump-Wahlkampf psychographisches Profiling eingesetzt zu haben. Hierzu ein längerer Ausschnitt aus besagtem Artikel:

„The firm’s claims [Cambridge Analytica] about its client base have also shifted. As recently as October, the firm said it had 50 clients in the 2016 elections. But a company spokesman said federal elections records showing just a dozen were correct. [...] The spokesman also said neither Cambridge nor SCL had done any work, paid or unpaid, with the pro-‘Brexit’ Leave.eu campaign last year, although Mr. Nix once claimed that Cambridge had helped ‘supercharge’ Leave.eu’s social media campaign. British authorities are now investigating the company’s exact role with Leave.eu and whether Cambridge’s techniques violated British and European privacy laws. At a conference in Munich last month, Alexander Tayler, Cambridge’s chief data officer, dodged a question about whether Cambridge would work with far-right parties in European elections this year. He also played down the role of psychological profiling in the company’s work, much of which, Mr. Tayler suggested, is still based on traditional data analytics and marketing.” (NYT, 6.3.2017)

Es kann ein Zusammenhang angenommen werden zwischen der Selbstrelativierung der eigenen Rolle und der massiven Kritik an Cambridge Analytica Methoden. Es geht indes nicht aus dem Diskurs hervor, ob die Relativierungen durch die Firma dem eigentlichen Sachstand näherkommen oder ob, im Falle der Beteiligung an den Leave-Kampagnen, steuer- und datenschutzrechtliche Untersuchungen durch die UK Electoral Commission zu einer Dementierung führten, bzw. im Falle der US-Kampagne die massive Kritik an der psychographischen Profiling-Methode in Zusammenhang mit einem Sieg von Donald Trump (Wired, 15.8.2016, Bloomberg Businessweek, 23.3.2017). Trotz der offiziellen Dementi durch Cambridge Analytica beharren Dritte darauf, z. B. Andy Wigmore, Kommunikationschef von Leave.EU, dass die Firma in Wahrheit doch

für sie tätig gewesen sei und Facebook-Profiling zum Einsatz gekommen sei. Hierzu ein Ausschnitt aus einem Artikel des *Guardian*:

„A few weeks later, the Observer received a letter. Cambridge Analytica was not employed by the Leave campaign, it said. Cambridge Analytica ‘is a US company based in the US. It hasn’t worked in British politics.’ [...] Cambridge Analytica had worked for them, he [Andy Wigmore] said. It had taught them how to build profiles, how to target people and how to scoop up masses of data from people’s Facebook profiles. A video on YouTube shows one of Cambridge Analytica’s and SCL’s employees, Brittany Kaiser, sitting on the panel at Leave.EU’s launch event. [...] They hadn’t ‘employed’ Cambridge Analytica, he said. No money changed hands. ‘They were happy to help.’ Why? ‘Because Nigel [Nigel Oakes, CEO der SCL Group, dem Mutterkonzern von Cambridge Analytica] is a good friend of the Mercers. And Robert Mercer introduced them to us. He said, ‘Here’s this company we think may be useful to you.’ What they were trying to do in the US and what we were trying to do had massive parallels. We shared a lot of information. Why wouldn’t you?’ Behind Trump’s campaign and Cambridge Analytica, he said, were ‘the same people. It’s the same family.’” (The Guardian, 28.2.2017)

Cambridge Analytica und die SCL-Gruppe haben aufgrund dieses Artikels Klage gegen den *Guardian* eingereicht, das Verfahren läuft noch. Ein ehemaliger Mitarbeiter von Cambridge Analytica berichtet vom Wandel der öffentlichen Wahrnehmung des Unternehmens nach Brexit und Trump-Sieg: „That was before we became this dark, dystopian data company that gave the world Trump,’ a former Cambridge Analytica employee who I’ll call Paul tells me.“ (The Observer, 7.5.2017)

Nach den ersten alarmierten Reaktionen auf die Veröffentlichung des *Magazin*-Artikels werden auch Kommunikate veröffentlicht, die Zweifel am weitreichenden Einfluss der psychographischen Methoden von Cambridge Analytica anmelden. Den Autoren wird teilweise Übertreibung und Sensationslust hinsichtlich der in den Brexit- und Trump-Kampagnen verwendeten Daten und Modelle und deren Einsatz im politischen Microtargeting vorgeworfen. Daraufhin räumen sie einige Tage später im Tages-Anzeiger (10.12.2016) ein, dass sie „überspitzt formuliert“ hätten.

Als die öffentliche Debatte um Cambridge Analytica allmählich abflaut, herrscht weitgehend die publizistische Meinung vor, dass deren Beitrag zu Trumps Wahlsieg marginal bis nicht zu erkennen gewesen sei. So nennt beispielsweise Julius van de Laar, ehemaliger Kampagnenmitarbeiter von Obama, die Vorstellung, Trump sei wegen seiner Microtargeting-Kampagne, die wiederum auf psychographischem Profiling beruhe, gewählt worden, eine „Verschwörungstheorie“ (FAZ, 14.5.2017). „Mit extrem offensiver Selbstvermarktung hat sie [Cambridge Analytica] erfolgreich heiße Luft verkauft. So konnte sie davon ablenken, dass ihre Microtargeting-Methoden fragwürdig und die versprochenen Effekte nicht nachgewiesen sind.“ (Zeit online, 7.3.2017) Allerdings finden sich auch in dieser Zeit noch starke Positionen, die vor den Methoden des Unternehmens warnen, z. B. die Journalistin Carole Cadwalladr, die im *Guardian* beziehungsweise *Observer* zwei ausführliche Artikel zum Thema veröffentlicht, die in der weiteren Analyse noch genauer vorgestellt werden (The Guardian, 26.2.2017; The Observer, 7.5.2017).

7.3 DATENBASIERTE ABBILDUNGSVERHÄLTNISSE, REPRÄSENTATIONSBEZIEHUNGEN UND REGULIERUNGSABSICHTEN IN MESOANALYTISCHER PERSPEKTIVE

Im Folgenden werden die Ergebnisse der Diskursanalyse anhand der drei Themenblöcke **datenbasierte Abbildungsverhältnisse**, **datenbasierte Repräsentationsbeziehungen** und **datenbasierte Regulierungsabsichten** dargestellt.

7.3.1 DATENBASIERTE ABBILDUNGSVERHÄLTNISSE

„Die Quellen sind vielfältig: Behörden, Verlage, Banken, Social-Media-Plattformen, Cookies. Selbst welche Magazine die Menschen lesen und welche Fernsehsendungen sie schauen, kann erfasst werden.“ (MEEDIA, 3.3.2017)

Im diskursiven Material werden eine breite Palette an **Quellen aggregierter Datenmengen** benannt, wobei die Aufzählungen oft pars pro toto für die Varianz und Potenz datenbasierter Abbildungsverhältnisse verwendet wurden: „Big Data bedeutet auch, dass alles, was wir treiben, ob im Netz oder ausserhalb [sic], digitale Spuren hinter-

lässt.“ (Das Magazin, 3.12.2016) Explizit gemacht wird ebenfalls, dass nicht nur Online-Aktivitäten in Daten abgebildet werden, sondern auch (vermeintliche) Offline-Merkmale und -Aktivitäten: soziodemographische Merkmale wie Alter, Geschlecht, Adresse, Telefonnummer, Konsumentendaten wie Mitgliedschaften, Abonnements und Einkäufe mit Treue- oder Payback-Karten, Spenden, Wahlregister-Einträge (*voter files*), medizinische und polizeiliche Daten. Hinzu kommen genuin online erzeugte Daten: Surfverhalten (z. B. Cookies), Online-Käufe, GPS-Daten von mobilen Geräten, E-Mail-Adresse, Aktivitäten und Kontakte auf sozialen Netzwerkplattformen, von Online-Dienste-Anbietern angelegte Profile über Nutzer (z. B. bei Netflix oder Amazon) sowie Selbstauskünfte etwa aus online (bzw. auch offline/telefonisch) ausgefüllten Persönlichkeitstests und Meinungsumfragen. Die mit Abstand am häufigsten genannte Referenz im Diskurs sind Facebook-Daten, 63 Prozent der Kommunikate thematisieren die Rolle des sozialen Netzwerkes.

Dabei wird die **Repräsentativität der Datensammlungen** in der überwiegenden Zahl der Berichte nicht kritisch hinterfragt. Dass ein *Bias* durch unterschiedlich intensive Nutzungsaktivitäten entsteht, mit den äußeren Polen des Typs des Nicht-Nutzers, der jegliche Nutzung von Online-Praktiken verweigert, und dem Typ des Always-On-Nutzers, der viele Dienste frequentiert, die persönliche Daten aggregieren, speichern und / oder anderen Akteuren zur Verfügung stellen, spielt im Diskurs kaum eine Rolle.

Lediglich in Bezug auf Online-Umfragen wird auf Probleme hinsichtlich der Repräsentativität der erzeugten Daten verwiesen: „Internet-polling companies try to sidestep sampling bias by recruiting large, stable ‘panels’ made up of the right numbers of the educated, the young and so on, from which they pick representative samples each time they run a poll. But this can still produce poor results. After finding that its internet polling in 2015 oversampled politically engaged voters, who tend to be leftish, YouGov tried hard to recruit less-engaged voters to its panel.“ (The Economist, 17.6.2017) Ein Eintrag auf dem Firmenblog von Spark Cognition Inc., einem KI-Unternehmen, räumt in einem Nebensatz ein, dass Daten aus sozialen Netzwerken aufgrund der Altersstruktur der Nutzer*innen eigentlich keine Rückschlüsse auf die Gesamtbevölkerung zulassen. Dieser Einwand sei allerdings aufgrund der Erkenntnisgewinne von Big-Data-Analysen in Echtzeit zu vernachlässigen:

„Even since the first election of President Obama in 2008, the percentage of adults who use at least one social media site has skyrocketed from 25% to 65%, and social media grew from being just another distraction in our lives to an extension of our personalities. A useful side effect of all this growth has been an exponential increase in the amount of data people create, share, and enjoy with each other using social media websites such as Google, Facebook, and Twitter. Political discourse wasn't left out of this growth, making social media has a potentially ripe source for real time data on how people think and feel about certain politicians at any given time.! [sic] Although social media's demographical bias towards younger people makes it a challenge to extrapolate trends to the general population, it can provide a window into how people's thoughts, beliefs, and actions are changing in real time in a way polls never could.” (sparcognition, 20.10.2016)

Ein anderer Einwand besteht darin, dass politische Information nicht nur über soziale Medien stattfindet, wo die Microtargeting-Botschaften platziert werden, sondern über viele mediale Kanäle, sodass die Fokussierung auf eine Plattform weder der datenbasierten Wahlkommunikation noch dem Nutzer*innenverhalten entsprechen: „Sharvani Mukherjee, 27, opened the door to his knock. She planned to vote, she said, but was still undecided. Despite being an avid user of Instagram and Facebook, her political views are shaped from watching television debates and reading the news, she said, 'not what people talk about on Twitter.'” (NYT, 3.5.2015).

Insgesamt zwei kritische Aussagen zur Repräsentativität von Big-Data-Kontexten bei Stichprobenbildungen und eine kritische Aussage zur Anwendung darauf basierender Maßnahmen sind die einzigen geäußerten Bedenken im untersuchten diskursiven Material, die sich konkret auf die Frage der Repräsentativität und damit der datenbasierten Abbildbarkeit von sozialen Sachverhalten und persönlichen Merkmalen in ihrer gesellschaftlichen Gesamtheit beziehen. Die Reichweite datenintensiver Maßnahmen – sowohl die Datensammlung als auch den Einsatz von Microtargeting-Maßnahmen betreffend – wird in der Darstellung nicht weiter differenziert, sondern es wird vielmehr der Eindruck erweckt, Datensammlung und darauf basierende Targeting-Maßnahmen betreffen alle Bürger*innen gleichermaßen und es würden für alle ähnliche Daten erhoben, ohne Möglichkeit von Interventionen.

Auf die Option, Daten aus unterschiedlichen Quellen zusammenzuführen, das sogenannte *Onboarding*, wird ebenfalls verwiesen:

„In the age of Facebook, it has become far easier for campaigners or marketers to combine our online personas with our offline selves, a process that was once controversial but is now so commonplace that there's a term for it, 'onboarding.' Cambridge Analytica says it has as many as 3,000 to 5,000 data points on each of us, be it voting histories or full-spectrum demographics -- age, income, debt, hobbies, criminal histories, purchase histories, religious leanings, health concerns, gun ownership, car ownership, homeownership -- from consumer-data giants.” (NYT, 20.11.2016)

Es werden sowohl die unterschiedlichen zur Verfügung stehenden Datenquellen als auch deren Verwendungsmöglichkeiten in staaten- und ländervergleichender Perspektive thematisiert. Mehrere Beiträge befassen sich beispielsweise mit den rechtlichen Bestimmungen zu den sogenannten *voter files*, personenbezogene Akten, die in den USA von Bundesstaat zu Bundesstaat unterschiedliche Informationen enthalten – von der Wahlregistrierung bis hin zum tatsächlichen Wahlverhalten in Verbindung mit verschieden umfangreichen Informationen zu soziodemographischen Merkmalen. Diese Akten sind je nach Bundesstaat in unterschiedlichem Maße öffentlich und frei zugänglich (z. B. NYT, 31.12.2015; WP, 9.7.2015).

Viele dieser Beiträge thematisieren dabei die unterschiedlichen **Datenschutzbedingungen** in den USA gegenüber Europa allgemein, beziehungsweise Deutschland und Großbritannien im Speziellen, die in europäischen Ländern sowohl das Sammeln und Speichern als auch die Weitergabe und den Erwerb von Daten rechtlich beschränken.

„A key to its [Cambridge Analyticas] strategy is the fact that U.S. laws regarding the release of personal information are more lax than in some other countries, enabling it to mine far more data here. For example, European Union countries require much more personal data to be kept private unless a person 'opts in' to its release. The United States, by contrast, allows much data to be released unless a person 'opts out.' Thus, companies such as Cambridge Analytica can collect thousands of pieces of information about voters.” (WP, 16.10.2016)

So sagt beispielsweise Julius van de Laar, Kampagnenberater u. a. im Obama-Wahlkampf 2012, dass in Deutschland ein Big-Data-Wahlkampf „völlig unvorstellbar“ sei: „Die Grundlage dafür ist der direkte Zugriff auf ein umfassendes Wählerregister, das Parteien in Deutschland nicht haben.“ (FAZ, 14.5.2017)

Neben Positionen, die europäische Länder als Hochburgen des Datenschutzes darstellen, finden sich diejenigen, die darauf verweisen, dass auch in Ländern mit vergleichsweise strikten Datenschutzbestimmungen Big-Data-Analysen möglich seien. So betont Alexander Nix, CEO von Cambridge Analytica: „Der gesetzliche Rahmen in Europa ist komplexer. Große Datensätze zu kaufen oder lizenzieren wie in den USA ist hier nicht so einfach möglich. Aber man kann mit Datensätzen arbeiten, die unseren Kunden gehören. Parteien haben oft ihre eigene Wählerforschung, die 20 oder 30 Jahre zurückreicht. Wir könnten Wähler auch Ocean-Fragebögen ausfüllen lassen und auf dieser Basis unser Modell bilden.“ (Handelsblatt online, 14.12.2016) Jan Philipp Albrecht, Europaabgeordneter der Grünen und Verhandlungsführer der europäischen Datenschutznovelle (ab 2018), wird in der *Frankfurter Allgemeinen Zeitung* mit der Aussage zitiert, dass die alte Datenschutz-Richtlinie Praktiken der Wahlbeeinflussung nicht ausreichend erfassen könne (17.12.2016). Maximilian Schrems, ein österreichischer Jurist und Datenschutzaktivist, sagt im gleichen Artikel, dass auch in Europa Parteien problemlos Personendaten einkaufen und österreichische Parteien auch Wählerlisten mit persönlichen Kontaktdaten erhalten könnten. Ein vom britischen *Guardian* angeführter Bericht der London School of Economics vom April 2017 kommt zu dem Schluss, dass auch die britische Gesetzgebung zum Wahlkampf dringend überarbeitet werden müsse, diese sei „weak and helpless“ (Guardian, 7.5.2017).

Insgesamt vermittelt der Diskurs, dass der Einsatz von Microtargeting-Methoden in allen modernen Wahlkämpfen zum Standard gehöre: „Big Data ist im Wahlkampf Standard.“ (FAZ, 15.12.2016), „The idea that a single company [Cambridge Analytica] influenced an entire election is also difficult to maintain because every single candidate used some form of profiling and micro-targeting to persuade voters—including Hillary Clinton and Trump’s competitors in the primaries. Not every campaign used personality profiles but that doesn’t make it any less invasive or creepy!“ (Medium, 13.4.2017)

7.3.2 DATENBASIERTE REPRÄSENTATIONSBEZIEHUNGEN

Die datenbasierten Repräsentationsbeziehungen werden im Diskurs auf drei Ebenen verhandelt. Einmal mit Bezug auf den **Nutzen datenintensiver Wahlkampfmethoden** des Micro- oder Psychotargeting, dann mit Blick auf **Big Data-basierte Wahlprognosen** und schließlich mit einem Schwerpunkt auf den diskutierten **Gefahren im Zusammenhang mit Big Data-Anwendungen in Wahlkämpfen**.

a) Big-Data-basiertes Microtargeting und Psychotargeting

Zur „epistemischen Autorität“ von Daten in Wahlkampagnen lassen sich **drei Positionen** im Diskurs finden: Optimisten, Skeptiker und Pessimisten.

Die **Optimisten** vertreten eine positive und positivistische Sicht auf datenintensive Praktiken. Sie versprechen sich durch sie Erkenntnisgewinne, die mit Zunahme der vorliegenden Daten steigen. Sie gehen davon aus, dass Daten valide Vorhersagen ermöglichen und dass eine Zunahme der Daten die Genauigkeit der Modelle verbessert. Hierdurch würden politische Kampagnen effizienter und kostengünstiger. Die Gruppe der Optimisten setzt sich hauptsächlich aus Datenfirmen-Mitarbeiter*innen und Analyst*innen zusammen, die besonders in den Monaten rund um den Wahlsieg Trumps den Diskurs bestimmen. Sie besitzen qua Expertise eine Deutungsmacht über die Wirkung datengetriebener Methoden, die von Journalist*innen, oft informationstechnischen Laien, häufig übernommen wird. So z. B. hier: „Donald Trump's U.S. election victory threw polling experts – from the New York Times to Nate Silver's FiveThirtyEight – into serious soul-searching after their forecast of a Hillary Clinton win turned out to be wrong. But for a little-known British company working for the Trump campaign it was confirmation that their data analysis – based on a hyper-targeted psychological approach – was working.“ (Mashable, 10.11.2016). Besonders unkritisch gehen Blogs von IT-Unternehmen und Branchenmagazine bzw. -websites mit dem behaupteten Einfluss von datenintensiven Wahlkampfmethoden um, was sich bereits in den Titeln der Beiträge widerspiegelt: „Der große Trump-Trick mit Big Data“ (Gründerszene, 5.12.2016), „How a little-known data firm helped Trump become president“ (Mashable, 10.11.2016), „How a Silicon Valley Campaign Strategy Won Trump the Election“ (Business 2 Community, 29.1.2017), „How the Trump Campaign Built an Identity Da-

tabase and Used Facebook Ads to Win the Election“ (Startup Grind, 18.11.2016), “How Political Campaigns Are Putting People Data To Work“ (Adexchanger, 22.3.2016).

Ein Zitat Michal Kosinskis, des Erfinders der Facebook-Persönlichkeitstests, verdeutlicht den epistemischen Anspruch an Daten: „Ja, mehr Daten sind immer besser. Aber letztendlich reichen bei vielen Menschen dann doch zehn Likes, um schon recht genaue Vorhersagen über den Charakter zu treffen.“ (taz, 17.12.2016) Laut Kosinski genügten zehn Likes, um über eine Person mehr zu wissen als ein Arbeitskollege, 70 überböten Freund*innen, 130 die Eltern und 300 Partner*innen (FAZ, 11.12.2016). Big Data wird mit Attributen versehen, die deren Einsatz als entscheidenden Vorsprung gegenüber politische Konkurrenz markieren – im Wahlkampf oder in Wahlvorhersagen: „secret weapon in building an insurmountable delegate lead“, „crucial tactical advantage“ (Politico, 7.9.2016), „essential to victory“ (WP, 19.10.2015), „‘significant impact’ on the election“ (The Guardian, 3.5.2017), „key role in winning the referendum“ (BuzzFeed, 30.11.2016), „key element of the successful Obama for America campaign“ (Forbes, 29.6.2016), „responsible for the win“ (The Huffington Post, 8.12.2016), „political weapon“ (WP, 28.10.2016), „razorsharp operations“ (WP, 24.11.2015).

Mit der Überzeugung datenbasierter Erkenntnisgewinne gehen teilweise auch Verheißungen auf demokratischen Fortschritt einher. Entgegen seiner im *Magazin*-Artikel geäußerten dystopischen Vorstellungen über den Einfluss psychographischer Microtargeting-Methoden auf den Wahlausgang, preist Michal Kosinski einige Monate später die demokratischen Chancen dieser Methode:

„‘One can enter the political race at much lower cost, and we could see that with Brexit, but also Trump and Bernie Sanders,’ Kosinski says. ‘It completely changes the economics of political targeting.’ And it could allow politicians to reach groups who have previously been ignored and may not have bothered voting as a result. ‘With both Donald Trump and Bernie Sanders, they attracted huge followings among groups of people who previously were not politically active,’ Kosinski says. ‘It’s creating turbulence as we speak in different countries. But I think in the long term, if we can survive the turbulence now, it’s ac-

tually great news for democracy that more people enter the dialogue and are being engaged.” (The Register, 7.3.2017)

Der demokratische Nutzen bestehe also in der Informationsvermittlung an politisch wenig interessierte Bürger*innen, denen das Hintergrundwissen geliefert werde, welches sie zu einer informierten Wahlentscheidung benötigten. In den Worten des Cambridge Analytica CEOs Alexander Nix: „Was Ihnen die Technologie erlaubt: Bei einem heißen Match sicherzustellen, dass jeder Wähler die Information bekommt, die er für eine aufgeklärte Entscheidung braucht.“ (FAZ, 13.3.2017)

Der Journalist Adrian Lobe schreibt ebenfalls in dieser Linie in der *Frankfurter Allgemeinen Zeitung* einen polemischen Artikel, der gegen basisdemokratische Entscheidungen argumentiert, die auf vermeintlich in Big Data abgebildeten Wählerwillen basieren. Befürworter direkter Demokratie können in seiner Argumentation Argumente für datengetriebene Entscheidungen finden, nämlich gerechtere Wahlergebnisse:

„Wählen kann zeitaufwendig sein. Man sollte Parteiprogramme studieren, sich über Kandidaten informieren, aktuelle Entwicklungen beobachten und am Ende weiß man immer noch nicht, wen man wählen soll. Vielleicht kann eine künstliche Intelligenz anhand unserer aggregierten Suchanfragen und Klicks präziser ermitteln, welche Partei oder welcher Kandidat am ehesten mit unseren politischen Überzeugungen übereinstimmt. Etwas zugespitzt: Vielleicht kennt eine Maschine unsere Positionen besser als wir selbst. Dann kann sie auch gleich für uns wählen gehen. Und vielleicht bildet der Chor der virtuellen Assistenten dann ja auch am besten den Wählerwillen ab und führt zu einem Wahlergebnis, das niemand anfechten mag. Den Bedarf dafür scheint es zu geben: In der Woche vor der Wahl in den Vereinigten Staaten registrierte Google Trends einen signifikanten Anstieg der Suchanfrage ‚Vote for‘.“ (FAZ, 15.12.2016)

Die **Skeptiker** hingegen betonen, dass Microtargeting auf Basis von Big-Data-Prognosen durchaus hilfreich sei(n) (könne), aber nicht allein wahlentscheidend. Beispielsweise stellt Robby Mook, Kampagnenmanager von Hillary Clinton, fest: „You don’t win campaigns because of data. You run your campaign more efficiently and effectively with the data.’ Clinton, he said, ‘has to get out there every day and make her case.’“ (Politico, 7.9.2016) Auch andere relativieren den Einfluss von Microtargeting-

Maßnahmen, indem sie auf weitere Einflussfaktoren für Wahlsiege verweisen und Microtargeting damit die wahlentscheidende Rolle absprechen bzw. es im Spektrum der Wahlkommunikation kontextualisieren: „Dabei gab es noch andere Datenanalyse-Firmen, die Trumps und Clintons Lager beschäftigt haben, sowie viele nicht-technische Gründe, die letztlich zum Sieg des republikanischen Kandidaten führten. Die Demokraten, der FBI-Direktor, das Wahlsystem, die Russen, die Medien, die Liberalen – sie alle wurden schon mitverantwortlich für den Wahlausgang gemacht. Die Liste ließe sich noch lange fortsetzen. Den einen Grund für Trumps Sieg gibt es nicht.“ (Zeit online, 6.12.2016) Auch Cambridge Analytica selbst werden mit der Zeit vorsichtiger in ihren Aussagen, hier in einem Zitat Alexander Nix: „‘We bake a cake, it’s got 10 ingredients in it. Psychographics is one of them,’ he said. ‘It’s very difficult to isolate exactly what the impact of that ingredient is.’“ (NYT, 6.3.2017)

In dieser Lesart bleibt der Ausgang von Wahlen ein Resultat zahlreicher Faktoren, die gerade in ihrem komplexen Zusammenspiel schwer instrumentell komplett manipuliert oder dirigistisch gesteuert werden können.

Andere argumentieren in dieser Hinsicht auch vom Ergebnis her, betonen aber, dass Wählerbeeinflussung schon immer stattgefunden habe, nur mit anderen Mitteln:

„Auch ohne Targeting wurden in der Vergangenheit schon zweifelhafte Persönlichkeiten gewählt und fragwürdige Abstimmungen gewonnen. Manipulation ist auch ohne Targeting möglich, und Targeting muss nicht zwangsläufig zur Manipulation eingesetzt werden. Aber natürlich macht Targeting auch Manipulation effektiver, und es ist eine beängstigende Vorstellung, dass in Zukunft nicht mehr Argumente, sondern Datenanalysten den Ausgang unserer Wahlen und Abstimmungen entscheiden.“ (Martin Sauter, 4.12.2016)

Michal Kosinski argumentiert – entgegen seiner einst warnenden Äußerungen – später ähnlich in Richtung einer Kontinuität der Wähler*innenbeeinflussung:

„Politicians can use it to bad ends, he agrees, but adds: ‘Lying to people and trying to scare them about others, or misinforming people using propaganda or discouraging people from voting – those are bad examples of political communication. Those bad and good examples of political communication are completely independent of the channel. You don’t need digital targeting of your

political messages to lie to people or discourage them voting. You can use traditional TV or a newspaper.” (The Register, 7.5.2017)

Selbst wenn auf Basis von Microtargeting die Wähler*innen identifiziert würden, die eine Kampagne ansprechen will, führt Hersh bereits zu Beginn des Jahres 2016 aus, bedeute dies noch lange nicht, dass diese in der intendierten Weise auf die Botschaften reagierten und aktiv entsprechend wählen würden:

„And even when the message is delivered, a lot of experts are skeptical that individual personalities can be captured by big-picture data and modeling. ‚Persuadability is finicky,‘ said Yale political scientist Eitan Hersh, who studies microtargeting and recently wrote a book about it. ‚You’re persuadable sometimes but not other times,‘ Hersh said. ‚You’re persuadable from a certain kind of appeal at a certain moment by a certain candidate. But what persuaded you then might not persuade you a week later.‘” (npr, 19.2.2016)

Die **Pessimisten** dagegen stellen den Einfluss von Microtargeting und Psycho-Targeting generell in Abrede. Ihr Pessimismus liegt also darin begründet, dass sie Big Data-basierte Effekte bezweifeln, zurückweisen bzw. stark moderieren; sie könnten somit auch, quasi als Negierer, als radikalere Variante oder Unterkategorie der Skeptiker geführt werden.

Ein Argument, das die direkte und ausschlaggebende Wirkung von neuen Microtargeting-Maßnahmen entlarven soll, ist der Verweis auf nicht erfolgreiche Big-Data-Kampagnen, etwa von Ted Cruz, Hillary Clinton oder Bernie Sanders: „Wenn die Dienstleistung von Cambridge Analytica aber so gut ist, wie das Unternehmen behauptet – warum ist Ted Cruz dann nicht der neue Präsident? Schließlich konnte er zu einer Zeit auf die Analyseergebnisse zurückgreifen, zu der Donald Trump es noch nicht tat.“ (WDR, 5.12.2016).

„So gerne sich Nix mit dem erfolgreichen Trump-Wahlkampf rühmt, so zweifelhaft ist dieser Ruhm auch. Denn Belege dafür, dass so eine Werbung tatsächlich wirkt, gibt es nicht. Die Methode hat ihre Grenzen. Kritiker der Cambridge Analytica ziehen gerne den ehemaligen Präsidentschaftskandidaten und schärfsten Trump-Gegner Ted Cruz heran. Diesen hat die Analyse-Firma im Wahlkampf ebenfalls unterstützt. Da der Herausforderer am Ende nicht als

Wahlsieger dasteht, wird aber klar, dass Nix wohl keinen Königsweg zum Erfolg gefunden hat.“ (MEEDIA, 3.3.2017)

Andere weisen auf die fehlende wissenschaftliche Evidenz der Wirksamkeit der Methoden hin: „Campaigns spend millions on data science to understand their own potential supporters—to whom they’re likely already credible messengers—but here Trump is speaking to his opponent’s. Furthermore, there’s no scientific basis for thinking this ploy will convince these voters to stay home. It could just as easily end up motivating them.“ (Bloomberg Businessweek, 27.10.2016), „Das zugrunde liegende Persönlichkeitsmodell OCEAN ist wissenschaftlich umstritten und längst nicht auf alle Menschen anwendbar“ (Zeit online, 6.12.2016), „Die vermeintlich passenden Psychogramme, die für Trump eine Rolle gespielt haben sollen – was allerdings auch bestritten und als PR-Angeberei der Firma Cambridge Analytica und deren Chefs Alexander Nix kritisiert wird –, machen Dispositionen erkennbar, mehr nicht.“ (FAZ, 7.12.2016)

Eithan Hersh (Tufts University, Massachusetts) tritt mehrfach im Diskurs als Zweifler am Nutzen von Big Data-Wahlkämpfen in Erscheinung. Er erklärt in einem Interview, dass Targeting durchaus funktioniere, jedoch böten Konsumentendaten oder andere Online-Big-Data-Quellen keinen epistemischen Mehrwert gegenüber soziodemographischen Daten in Verbindung mit den in den USA zugänglichen Wählerakten:

„Knowing what car you drive might have value, but it probably doesn’t have much added value. Here’s why. For many voters, it’s straightforward to determine whether they are likely to be Democrats or Republicans, or voters or likely abstainers – and without any help from consumer information. Before knowing what car you drive, campaigns already have obtained data from the Census Bureau about all sorts of characteristics of the neighborhood and town that you live in. From your voter registration, they also know your gender, age, and maybe race and party affiliation. On top of this information, the car you drive doesn’t provide much new insight. If a campaign knows that you live in a highly educated neighborhood in the center of a college town and that you voted in the last five Democratic primaries, it doesn’t need to know that you own a Prius to peg you as a Democrat. Moreover, for the more complicated dispositions that campaigns care about – like how persuadable you are or who you will support in a primary – consumer data are generally irrelevant. Ameri-

cans are politically divided, but those divisions are much more apparent when voters are segmented by geography, race, age or gender than when they are segmented by the cars they drive or the food they eat.” (Politico, 9.7.2015)

Ähnlich argumentiert ein anderer republikanischer Datenanalyst: „A Republican data scientist for a rival firm said he did not use psychographics. ‘If you get a voter on the phone, why are you asking them what their favourite ice cream is or what their favourite colour is – why don’t you just ask them who they’re going to vote for?’ He added: ‘They’ve [Cambridge Analytica] got a smooth-talking Brit wearing Savile Row suits who gives you a great pitch and wows you a little bit; they’ve got a great PR operation, but with psychographic profiling, there’s nothing there. They’re really, really smart people. It’s like they’re a bunch of board-certified doctors who decided to make a lot more money selling snake oil.’” (The Spectator, 3.12.2016)

Einmal wird im Diskurs auch der Verdacht artikuliert, dass datenintensive Kampagnen mehr der Imagepflege einer Partei als fortschrittlich und dynamisch nütze als tatsächlich der Persuasion von Wähler*innen mittels individuell adressierter Botschaften: „But some Republican operatives with digital experience privately wonder how much the emphasis on digital and analytics is simply for public image purposes. Will the digital team, they worry, still be the first item cut when the budget gets crunched?” (NYT, 11.6.2015)

b) Big-Data-basierte Wahlprognosen

Auch im thematischen Bereich der Wahlprognosen finden sich die drei epistemischen Grundhaltungen im Diskurs wieder:

Optimisten sprechen Big Data das Potential zu, Standardfehler und Irrtumswahrscheinlichkeit der Prognosen zu minimieren, d. h. die Voraussagen genauer und zuverlässiger zu machen: „Pollsters would be able to access a large pool of social media data, giving them a far richer demographic background of the audience than is currently possible with phone polling.“ (Venturebeat, 10.12.2016) Kampagnenberater Patrick Ruffini (Echelon Insights) weist zudem auf den Echtzeitvorsprung durch Big-Data-Analysen gegenüber traditionellen Umfragen hin:

„So much of any campaign happens in real time, but polls are a reflection of what’s happened days ago. There’s going to be a shift toward more real-time analysis and a much clearer, shorter feedback loop between action and reaction. Traditional research isn’t good at that. If there’s a Trump gaffe or big Hillary news, campaigns need a way to understand and respond to that in real time with more rigor.” (Digitalist Magazine, 15.9.2016)

Ein weiteres Argument für datenbasierte Prognosen versucht Alexander Nix zu etablieren, allerdings ist es im Diskurs nicht zu einer prominenten Referenz geworden. Er stigmatisiert traditionell soziodemographisch vorgehende Targeting-Maßnahmen, also areal-basierter Ansprache, gegenüber der individualisierten Ansprache des Microtargeting als diskriminierend, weil wenig differenziert in Bezug auf die involvierten Datenarten: „Your decision-making is based on your personality,’ he says. ‘And not on your gender. Nor is it based on your age or your wealth or any other demographic or geographic factor.’” (npr, 19.2.2016). Nix versucht also, seinen Standpunkt im Diskurs zu stärken, indem er die Methoden seiner Firma als demokratischen Schritt des Empowerment von Individuen darstellt. In den Worten des *Magazin*-Textes:

„Bisher, so Nix, seien Wahlkampagnen nach demografischen Konzepten geführt worden, eine lächerliche Idee, wenn Sie drüber nachdenken: Alle Frauen erhalten die gleiche Nachricht, bloss [sic] weil sie das gleiche Geschlecht haben – oder alle Afroamerikaner, wegen ihrer Rasse?’ So dilettantisch arbeitet das Kampagnenteam von Hillary Clinton, das braucht Nix hier gar nicht zu erwähnen, es unterteilt die Bevölkerung in vermeintlich homogene Gruppen – genauso wie all die Meinungsforschungsinstitute es taten, die Clinton bis zuletzt als Gewinnerin sahen.” (Das Magazin, 3.12.2016)

Skeptiker verweisen ebenfalls darauf, dass die traditionellen Methoden der Wähler*innenbefragung (*polling*) veraltet seien, betonen aber, dass daher gerade die etablierten Werkzeuge der Meinungsforscher*innen der Datenlage angepasst werden müssten:

„Civis’ theory also points to another larger problem in the industry: Today’s polls just aren’t as reliable as ones a few decades ago, particularly as pollsters

struggle to adapt to changing technological habits. Polls now need to include not just landlines but cell phones and online sampling – and, given the low response rates, they might still be as accurate as a generation before. ‘People are in the middle of a technological transition,’ Wagner [CEO Civis Analytics] says. ‘It’s compromising the use of phones as a tool for political measurement.’” (Wired, 15.11.2016)

In Hinblick auf die britischen Unterhauswahlen rät ein *Economist*-Artikel ebenfalls zu neuen, umfassenderen Sampling-Methoden:

„And dealing with declining response rates will probably require new ways to contact prospective voters. During the early days of internet polling, many feared that online samples were bound to be unrepresentative, mainly because they would include too few older people. But Britain’s online pollsters silenced their critics in the Brexit vote, where they came two percentage points closer than telephone pollsters to the result. Some startups are now testing what they call ‘programmatic sampling’: advertising very short surveys to smartphone users. Google, which runs bespoke market surveys for companies, tries to ensure representative samples by using browsing history to guess respondents’ demographics. Finally, pollsters will have to become more statistically sophisticated. Sampling 1,000-2,000 people and massaging their responses to correct for past errors looks increasingly antiquated. YouGov’s recent success was based on rolling questionnaires administered daily to 7,000 people from a 50,000-strong online panel, with the results combined using advanced number-crunching known as ‘multilevel regression and post-stratification’.” (The Economist, 17.7.2017)

Das hier vorgebrachte Argument erinnert an Wired-Chefredakteur Chris Andersons These vom „End of Theory“ (Wired, 23.6.2008), die die Überflüssigkeit von Modellen und schließender Statistik prophezeit, weil durch Big Data bald Vollerhebungen Standard würden, aus denen alle gewünschten Informationen abzulesen wären, ohne dass hierfür konzeptuelle Vorüberlegungen oder Auswahlstrategien notwendig seien.

Andere betonen, dass prognostische Modelle für Wahlen unter Normalbedingungen gut funktionierten, dass sie aber durch einen atypischen Kandidaten wie Trump ihre Vorhersagekraft verloren hätten:

„Abramovitz [Wahlforscher an der Emory University, Georgia] just thinks his model is wrong this year. ‚The model is based on the assumption that the parties are going to nominate mainstream candidates who will be able to unite the party, and that the outcome will be similar to a generic vote, a generic presidential vote for a generic Democrat versus a generic Republican,‘ he told me. ‚That’s usually a pretty reasonable assumption and produces pretty accurate predictions.‘ [...] Election models are valuable so long as the election is broadly similar in terms of candidate quality, campaign tactics, and party coalitions to the elections that have happened before. That is, election models are good at predicting elections that are like past elections. They are bad at predicting elections that are not like past elections.“ (vox.com, 14.6.2016)

Pessimisten verweisen wiederum darauf, dass Wahlen und Referenden niemals akkurat zu prognostizieren seien. Sie stellen damit die grundsätzliche Ambition, zukünftige Wahlausgänge abschätzen und Wahlhandlungen bestimmen zu können, in Zweifel.

„In Fällen wie dem Brexit oder der US-Wahl gebe es nicht genug Daten, schreibt der Kolumnist und Analyst Barry Ritholtz: ‚Man kann Emotionen nicht modellieren. Modellierung funktioniert gut, wenn man eine lange Serie wiederholbarer Ereignisse hat, so was wie Transaktionen mit der Kreditkarte.‘ Ein Brexit hingegen sei eine einmalige Sache. Es gebe keine Daten eines früheren Brexit. Das Gleiche gilt seiner Meinung nach für Präsidentschaftswahlen. Sie finden nur alle vier Jahre statt. ‚Die Daten sind einfach nicht alt genug‘, schreibt er. In den vier Jahren dazwischen haben sich die Einstellungen der Wähler so gewandelt, dass die Wahlen nur bedingt miteinander vergleichbar sind, auch wenn die Wahlprognosen etwas anderes suggerieren. Was wie Big Data aussieht, ist eigentlich Small Data. Und der Wissenschaftler und Big-Data-Experte Viktor Mayer-Schönberger sagt: Small Data ist tot.“ (Zeit online, 9.11.2016)

Als weiteres Argument gegen die robuste und erwartungssichere Prognostizierbarkeit von Wahlverhalten wird eine beschleunigte Gegenwart sozialen Geschehens, kommunikativer Vorgänge und (politischer) Meinungsbildung angeführt, in der Stabilitäten und Gewissheiten abhandengekommen seien:

„Das Problem ist, dass diese Unsicherheit wohl zunehmen wird. ‚Faktoren, die Prognosen früher Sicherheit gegeben haben, nehmen ab‘, sagt der Politikwissenschaftler Eric Linhart von der TU Chemnitz. Beispielsweise sinke die Identifikation mit einer bestimmten Partei in vielen Ländern, darunter Deutschland. Zudem sind Wahlentscheidungen zunehmend von Stimmungen abhängig, Ereignisse wie ein Terroranschlag können Emotionen schnell kippen lassen. Weiß ein Wähler fünf Minuten vor Stimmabgabe selbst nicht, wo er sein Kreuz macht, ‚dann kann man das auch nicht prognostizieren‘, so Linhart.“ (SZ, 6.2.2017)

Andere unterstreichen die Unmöglichkeit per se, menschliches Verhalten genau zu prognostizieren: „Wahlumfragen liegen oft daneben – und selbst Facebook und Co. Bringen wohl keine Besserung“ (FAZ, 22.4.2017). „Die Adressaten selbst der perfidesten Onlinewerbung und Fake News sind immer noch autonome Bürgerinnen und Bürger, die erst mit ihren eigenen Beschäftigungen und Daten die Botschaften anlocken. Und die natürlich auch selbst ihre Wahlentscheidung treffen.“ (Tages-Anzeiger, 10.12.2016)

Der populäre Datenanalyst Nate Silver, der aufgrund der Treffsicherheit seiner Prognosen seit den US-Wahlen 2008 bekannt geworden war, reagiert mit einer Reihe Veröffentlichungen auf seinem Blog *FiveThirtyEight* relativierend auf die pauschale Kritik, dass Prognosen trotz Big Data bei den US-Wahlen falsch gelegen hätten. Auch in Interviews argumentiert er, dass nicht die Daten und Prognosen größtenteils schlecht gewesen seien, sondern die Interpretationen:

„Ironically, the polls in the primaries were actually pretty good on Donald Trump. They correctly showed him leading and were correct in most states. They went high on Trump in a couple of states, like Iowa, but for the most part they did a pretty good job on the Republican primaries. So that was another case where people, including us, were ignoring data, or at least one type of data. Obviously, Trump himself is not a very data-driven president, but the notion that this was some big failure of data doesn't really match with the evidence. It's a giant, enormous, gaping failing for conventional wisdom. But people are often afraid to admit that their perspective on the world is sometimes wrong.“ (Harvard Gazette, 30.3.2017)

Er erklärt, dass mit der exponentiellen Zunahme von verfügbaren Daten die Komplexität der Analysen steige und teilweise Zusammenhänge gesehen würden, die nichts bedeuteten: „Data is not interesting for its own sake, but for how it relates to everything else. It’s like a map, you need to see how the data relates to everything else. A coordinate itself doesn’t tell you anything. But as your data set gets larger, you go from five variables to ten, you begin to see an exponential increase in complexity. The widespread availability of data in some fields adds to this complexity.” (The New Stack, 20.1.2017) Die größte Gefahr sieht Silver in Fehlinterpretation statistischer Daten durch die Medien:

„To be clear, if the polls themselves have gotten too much blame, then misinterpretation and misreporting of the polls is a major part of the story. Throughout the campaign, the polls had hallmarks of high uncertainty, indicating a volatile election with large numbers of undecided voters. And at several key moments they’d also shown a close race. In the week leading up to Election Day, Clinton was only barely ahead in the states she’d need to secure 270 electoral votes. Traditional journalists, as I’ll argue in this series of articles, mostly interpreted the polls as indicating extreme confidence in Clinton’s chances, however.” (FiveThirtyEight, 19.1.2017)

Als Beispiel führt er die Berichterstattung der *New York Times* an, die Clintons Sieg bei den Präsidentschaftswahlen als sicherer dargestellt hatten, als es die Prognosen zuließen. Dadurch, so Silver, seien Wähler*innen bei der Wahl zuhause geblieben, die bei weniger sicher vermuteten Vorsprüngen wählen gegangen wären, um einen Sieg Trumps zu verhindern. Dabei einzurechnen wäre die Rolle der *New York Times* als Leitmedium, dass zum einen von den politischen und gesellschaftlichen Eliten gelesen würde und dass zum anderen auch die Deutungslinien für den weiteren Diskurs informiert.

Im Ländervergleich fällt auf, dass in Deutschland eher die Big-Data-Praktiken in den USA im Vordergrund der Debatten stehen, während sich die britischen und US-amerikanischen Kommunikate hauptsächlich auf die Praktiken im eigenen Land konzentrieren. Insgesamt liegt der Diskursschwerpunkt auf US-amerikanischen Datenpraktiken, was auch im Diskurs selbst auf die dort am weitesten fortgeschrittene Elaboration und Verbreitung der Methoden zurückgeführt wird: „Rasmus Kleis Nielsen of

Oxford University, who has written a book on political campaigns in America, thinks that such targeting will remain largely confined to that country. Nowhere else have party organisations access to so much money, data and technical talent. Moreover, America's political system lends itself well to analytics because once voters get to the polling station they often have only two options." (The Economist, 26.3.2016) Das de facto Zweiparteiensystem der USA trage dazu bei, dass sich sowohl finanzielle Mittel als auch persuasive Macht auf wenige Akteure konzentrierte. Hinzu kommt die Etablierung einer institutionellen Wahlkampfökonomie, in der sich auf Datenpraktiken spezialisierte Firmen häufig aus Kampagnen heraus ausgründen und ihre Dienste anderen Wahlkampfteams anbieten.

c) Potentielle Gefahren in Zusammenhang mit Big-Data-Anwendungen

Ein Bereich, der in den Kommunikaten kritisch betrachtet wird, sind **Datenschutzfragen**. Sie gehen davon aus, dass die Sammlung von Daten und die Möglichkeiten der Zusammenführung aus verschiedenen Datenquellen die Identifikation von Individuen ermöglichen. Selbst wenn die Daten auf legalem Wege erworben worden seien, könne mit dem heutigen Stand der Technik nicht ausreichend gewährleistet werden, dass die Daten nicht an Unbefugte gelangen:

„But while most big data analytics technology groups insist they do not know individuals' names or identities, some people are concerned that increasingly sophisticated data techniques could infringe privacy. Dave Maass of the Electronic Frontier Foundation, a non-profit organisation defending people's rights online, says people should worry about whether their political preferences are being sold or traded. Sometimes, candidates will pass lists of voters on to other politicians when they drop out of a race, or use them in the next election, Mr Maass warns. A greater concern for many is whether political campaigns can be trusted to protect personal data from cyber attack. Last week, James Clapper, US director of national intelligence, warned that hackers – possibly working for foreign governments – were trying to infiltrate both the Democratic and Republican presidential campaigns.“ (Financial Times, 27.5.2016)

Weiterhin wird kritisiert, dass Nutzer*innen häufig gar nicht die Wahl hätten zu entscheiden, ob ihre Daten gespeichert und verwertet werden sollen, wenn sie bestimmte Angebote nutzen wollen:

„Bangor's McStay says that informed consent for online psychological profiling is currently the exception, given that most web browsers accept and retain cookies by default. Martin Sykora, a lecturer in information management at Loughborough University whose research using Twitter predicted Trump's success, points out that many apps demand wide-ranging access to personal data as a condition of use. They don't explain why, so it can't be called informed consent. 'Companies often don't seem to really bother with that too much and maybe should give it a bit more thought,' Sykora says.” (The Register, 7.5.2017)

Gerade in den USA wird darüber nachgedacht, dass denjenigen, die Zugang zu Daten haben, unfaire Vorteile gegenüber datenarmen Kandidat*innen oder Kampagnen entstehen. So sei die **politische Chancengleichheit** gefährdet. In den USA haben Parteien oftmals mehr Rechte und leichteren Zugang zu *voter files* als parteilose Kandidat*innen, was diese benachteilige: „The high cost of acquiring big election data effectively subverts democracy by privileging the major political parties and the wealthy over the public.” (WP, 19.10.2015)

7.3.3 DATENBASIERTE REGULIERUNGSABSICHTEN

Mit der Annahme einer breiten Wirkung von Microtargeting-Maßnahmen geht häufig die **Warnung vor Regulierungsabsichten politischer** (z. B. Trump, Leave-Kampagne) und **wirtschaftlicher Akteure** (SCL, Robert Mercer) einher, die mithilfe datenintensiver Maßnahmen Meinungen und Verhalten zu manipulieren und in ihrem Sinne zu steuern in der Lage seien.

Diese kritische Perspektive wird hauptsächlich von zu Wort kommenden Journalist*innen und Wissenschaftler*innen vertreten und geht einher mit einer negative Assoziationen andeutenden bzw. insinuirenden Wortwahl, die den dystopischen Charakter der vorgestellten Szenarien und Implikationen unterstreicht:

„Trump’s presidential election victory is the most successful digital voter suppression operation in American history. The secret weapons in Trump’s digital arsenal were Project Alamo, his database of 220 million people in the United States, and the Facebook Advertising Platform. By leveraging Facebook’s sophisticated advertising tools, including Facebook Dark Posts, Facebook Audience-Targeting, and Facebook Custom Audiences from Customer Lists, the Trump campaign was able to secretly target Hillary Clinton’s supporters and covertly discourage them from going to the polls to vote.” (Startup Grind, 18.11.2016).

Die Maßnahmen werden mit kriegerischen Vokabeln charakterisiert: *suppression operation*, *secret weapons*, *targeting*. Anhand zweier Artikel der britischen Journalistin Carole Cadwalladr, gegen deren Verbreitung Cambridge Analytica und die SCL-Gruppe einen bis heute andauernden Rechtsstreit führen, kann die Argumentation der **Warner** vor datenintensiven Methoden in politischen Wahlkämpfen nachgezeichnet werden. Carole Cadwalladr’s Artikel können in diesem Sinn beispielhaft für die Aufladung des Diskurses mit kriegerischen und bedrohlichen Metaphern und sprachlichen Bildern angesehen werden. In dem ersten Artikel mit dem Titel „Robert Mercer: the big data billionaire waging war on mainstream media“ (The Guardian, 26.2.2017) beschreibt die Autorin Robert Mercers Verbindungen in die Politik (Cruz- und Trump-Wahlkampf, Leave-Kampagne) und seine Großinvestitionen in rechtspopulistische Medienunternehmen (z. B. Breitbart News). Sie zitiert u.a. Andy Wigmore von Leave.EU: „A Facebook ‘like’, he said, was their most ‘potent weapon’.“ Professor Jonathan Rust, Leiter des Cambridge University’s Psychometric Centre, wo die Facebook-Persönlichkeitsprofile erfunden und gesammelt wurden, warnt im selben Artikel vor *psychologischer Kriegsführung*, *Gehirnwäsche* und *Verhaltenssteuerung*, der hier in einer längeren Passage zitiert wird:

„The danger of not having regulation around the sort of data you can get from Facebook and elsewhere is clear. With this, a computer can actually do psychology, it can predict and potentially control human behaviour. It’s what the scientologists try to do but much more powerful. It’s how you brainwash someone. It’s incredibly dangerous. It’s no exaggeration to say that minds can be changed. Behaviour can be predicted and controlled. I find it incredibly

scary. I really do. Because nobody has really followed through on the possible consequences of all this. People don't know it's happening to them. Their attitudes are being changed behind their backs.' Mercer invested in Cambridge Analytica, the Washington Post reported, 'driven in part by an assessment that the right was lacking sophisticated technology capabilities'. But in many ways, it's what Cambridge Analytica's parent company does that raises even more questions. Emma Briant, a propaganda specialist at the University of Sheffield, wrote about SCL Group in her 2015 book, *Propaganda and Counter-Terrorism: Strategies for Global Change*. Cambridge Analytica has the technological tools to effect behavioural and psychological change, she said, but it's SCL that strategises it. It has specialised, at the highest level – for Nato, the MoD, the US state department and others – in changing the behaviour of large groups. It models mass populations and then it changes their beliefs. SCL was founded by someone called Nigel Oakes, who worked for Saatchi & Saatchi on Margaret Thatcher's image, says Briant, and the company had been 'making money out of the propaganda side of the war on terrorism over a long period of time. There are different arms of SCL but it's all about reach and the ability to shape the discourse. They are trying to amplify particular political narratives. And they are selective in who they go for: they are not doing this for the left. [...] Later, when Trump picked up Mercer and Cambridge Analytica, the game changed again. 'It's all about the emotions. This is the big difference with what we did. They call it bio-psycho-social profiling. It takes your physical, mental and lifestyle attributes and works out how people work, how they react emotionally.' [Andy Wigmore] Bio-psycho-social profiling, I read later, is one offensive in what is called 'cognitive warfare'. Though there are many others: 'recoding the mass consciousness to turn patriotism into collaborationism,' explains a Nato briefing document on countering Russian disinformation written by an SCL employee. 'Time-sensitive professional use of media to propagate narratives,' says one US state department white paper. 'Of particular importance to psyop personnel may be publicly and commercially available data from social media platforms.'" (The Guardian, 26.2.2017)

Hier werden Cambridge Analyticas Verstrickungen in militärische Operationen nachgezeichnet, die als Legitimierung der Übertragung des kriegerischen Vokabulars auf zi-

vile Einsätze der Firma dienen. Kriegerische Metaphern und Bilder werden sowohl von Beobachtern als auch Mitarbeitern und Auftraggebern des Unternehmens aufgenommen: „'Politics is war,' said Steve Bannon last year in the Wall Street Journal." (The Guardian, 26.2.2017) Trotz Klage wiederholt die Autorin in einem drei Monate später erschienen Artikel ihre Vorwürfe und führt sie weiter aus und ordnet die Maßnahmen des Unternehmens semantisch dem Feld des Verbrechens zu: „The great British Brexit Robbery: how our democracy was hijacked“ (The Observer, 7.5.2017). Ein ehemaliger Mitarbeiter von Cambridge Analytica bezeichnet darin die Maßnahmen der Firma als psychologische Kriegsführung: „It was back when we were still just a psychological warfare firm.' Was that really what you called it, I ask him. Psychological warfare? 'Totally. That's what it is. Psyops. Psychological operations – the same methods the military use to effect mass sentiment change. It's what they mean by winning 'hearts and minds'. We were just doing it to win elections in the kind of developing countries that don't have many rules.'“ Explizit wird vor den **Folgen der datenintensiven Maßnahmen für die Demokratie** gewarnt:

„This is not just a story about social psychology and data analytics. It has to be understood in terms of a military contractor using military strategies on a civilian population. Us. David Miller, a professor of sociology at Bath University and an authority in psyops and propaganda, says it is 'an extraordinary scandal that this should be anywhere near a democracy. It should be clear to voters where information is coming from, and if it's not transparent or open where it's coming from, it raises the question of whether we are actually living in a democracy or not.'“ (The Observer, 7.5.2017).

Unbewusste Manipulation und eine dadurch entstehende Untergrabung demokratischer Willensbildung werden insgesamt am häufigsten als Gefahr von Big-Data-Anwendungen angeführt. Ein *Times*-Artikel mit dem Titel: „Political advertising is the real fake news. A secretive world of online marketing is distorting our politics and urgently needs to be made accountable“ (The Times, 28.2.2017) zeigt exemplarisch die Argumentation. Sie folgt der Idee, dass die Form politischer Persuasion unterschwellig und undurchsichtig sei. Sie würde zudem nicht von einer neutralen Instanz kontrolliert und somit blieben etwa mögliche Falschinformationen unberücksichtigt. Schließlich fördere die datenbasierte Individualansprache eine Zersplitterung politischer Mei-

nungsbildung und der Rolle von politischer Debatte als gemeinschaftsstiftender Veranstaltung.

“Now, many will recognise this as precisely how internet advertising normally works, selling cars, or toothpaste, or even newspapers. The big question is whether it is acceptable for politics to be sold in the same way. It isn't. First, it's opaque. This weekend, The Observer reported that Robert Mercer, the owner of Cambridge Analytica, which would later assist Trump, also weighed in to help Leave.eu free of charge. This was not reported as a donation. It might never have been revealed at all had the Leave.eu communications director Andy Wigmore not let it slip. Second, there are no checks against dishonesty, even of the sort you get when advertising cars or toothpaste. I tried Ofcom, the Advertising Standards Agency and the Electoral Commission, all of whom sent me to each other. None regards verification of political claims as its remit, as you may remember from the fuss over the '£350 million a week' figure on the Brexit bus. Cambridge Analytica, initially chatty, stopped replying when I asked who checked its communications were actually true. The real answer, almost unbelievably, is nobody. The big problem, though, is that it fractures our politics. This is why we cannot just sit back and wait for all sides to adopt the same tactics, as they eventually did with billboards and broadcast.”

Neben **mangelnder Kontrolle der Datensammel-, Datenauswertungs- und Datenanwendungspraktiken** wird hier vor allem vor einer **Parzellierung der Wählerschaft** gewarnt, wie sie durch gezielte Ansprache von Individuen entstehe.

Auffallend ist, dass im diskursiven Material bezüglich der Nutzungsmotive von datenbasierten Praktiken durch Kampagnen unterschiedliche Deutungsmuster aktiviert werden, je nachdem über wessen datenintensive Wahlkampfpraktiken gesprochen wird. Joshua Green und Sasha Issenberg betonen in ihrem Bericht in der *Bloomberg Businessweek* (27.10.2016) mit dem Kriegs-Assoziationen weckenden Titel „Inside the Trump Bunker, With Days to Go“, in dem die Autoren Einblicke in die Digitalkampagne Trumps geben, dass diese hauptsächlich einer *get out the vote*-Strategie folge, die potentielle Unterstützer*innen der Demokraten davon abbringen solle, zur Wahl zu gehen. „We have three major voter suppression operations under way,” says a senior official. They're aimed at three groups Clinton needs to win overwhelmingly: idealistic

white liberals, young women, and African Americans.” Beispielsweise wurden Afroamerikaner*innen Clips auf Facebook gezeigt, in denen Hillary Clinton Afroamerikaner als “super predators” bezeichnet – eine Animation, die Versatzstücke originaler Clinton-Aussagen über eine Comic-Animation legt. Es handelte sich um einen sogenannten „dark post“, den nur die durch Microtargeting identifizierten Personen sahen. Auch über den Einsatz von Microtargeting auf Seiten des Clinton-Lagers wird berichtet, allerdings liegt der Fokus in den Kommunikaten auf der Identifikation potentieller Wähler, die erreicht werden sollen: „By contrast [to Trump’s amateurish campaign], Clinton has run a highly effective campaign that has raised massive amounts of money, stayed relentlessly on message, and used data analytics to target and convert voters. Her ads have been clever, well-timed and carefully targeted.” (WP, 28.10.2016) Clintons Microtargeting-Aktivitäten werden demzufolge eher positiv gerahmt, indem auf das aktivierende Potential der Maßnahmen verwiesen wird, während Trumps Microtargeting-Strategie ein destruktives, blockierendes Potential zugeschrieben wird, indem Personen demotiviert werden, sich an Wahlen zu beteiligen.

Über das dominante Deutungsmuster von Microtargeting-Maßnahmen herrscht ein diskursiver Machtkampf zwischen Datenfirmen beziehungsweise Nutzer*innen von Big Data-Analysen, meist Politiker*innen, und Diskursakteuren, die für sich beanspruchen, für die Gruppe der Bürger*innen beziehungsweise Wähler*innen insgesamt zu sprechen wie beispielsweise Carole Cadwalladr, die oftmals mit Bedrohungsszenarien arbeiten, um ihren Argumenten Gewicht zu verleihen. Ein Interviewzitat von Alexander Nix verdeutlicht die Strategie der Big-Data-Befürworter:

„Nicht nur Gesetze schränken Sie hier ein. Die Deutschen sind weltbekannt für ihr Misstrauen gegenüber Datensammelei. Die Bevölkerung wird sich in dieser Frage weiterentwickeln. Wir müssen immer wieder deutlich machen, dass keiner dieser Datenpunkte die Privatsphäre verletzt. Wir verwenden keine Gesundheits- oder Finanzdaten. Wenn die Werbeindustrie weiß, welche Cornflakes Sie zum Frühstück essen, werden die wenigsten damit ein Problem haben.“ (Handelsblatt online, 14.12.2016)

Kritik an der Praxis des Datensammelns und -auswertens wird hier durch Diskreditierung der Kritiker*innen als übervorsichtige, antiquierte Bedenkenträger*innen und die rhetorische Mobilisierung von Euphemismus, einer sprachlichen Verharmlosung bezie-

ungsweise einer Strategie des Abwertens und Kleinredens der Bedeutung von Erkenntnissen aus Datensammlungen zu entkräften versucht. Indessen reagieren Kritiker*innen wie Michael Zimmer, Associate Professor an der University of Wisconsin, USA, auch auf diese Einwände skeptisch: „It's one thing for a marketer to try to predict if people like Coke or Pepsi,' he said, 'but it's another thing for them to predict things that are much more central to our identity and what's more personal in how I interact with the world in terms of social and cultural issues.'” (Guardian, 24.11.2017) Es komme auf die Verwendung der Daten an, nicht auf deren Inhalt.

Einige wenige Kommunikate thematisieren Big-Data-Anwendungen auch explizit als **politische Machtinstrumente**. Dabei sind es vor allem deutsche Beiträge, die ausführlich vor Datennutzung als Herrschaftsinstrumenten warnen. Ein markanter Artikel von Harald Staun in der *Frankfurter Allgemeinen Sonntagszeitung* interpretiert den Wahlsieg Trumps als Aufstand der Bürger*innen gegen eine Kontrolle durch Big Data. Hier werden prognostische Methoden der Datenauswertungen mit Überwachung und Kontrolle assoziiert, denen sich Bürger*innen entziehen können und sollen:

„Die Stimmabgabe ist ein Zeichen der Selbstbehauptung gegen das eigene Datenprofil – dabei muss es sich nicht einmal um einen bewussten Akt handeln. Wer sich die Gesellschaft als Verschwörung der Eliten vorstellt, entwickelt instinktiv Strategien gegen alles, was er für ein Instrument ihrer Herrschaft hält. Die Utopie der perfekten Datengesellschaft besteht in einer kybernetischen Regierung, die politischen Dissens durch effiziente algorithmisierte Prozesse ersetzen will. Man muss sich nicht mit der Technik solcher Verfahren auskennen, um die Alternativlosigkeit abzulehnen, mit der eine solche Rationalität präsentiert wird. Die Wahl von Trump ist nicht nur Ausdruck ästhetischer und moralischer Ressentiments, sie ist auch ein Akt der Sabotage. Das ist, wenn man so will, die gute Nachricht dieser Katastrophe: Die Menschen haben gegen ihre digitalen Schatten gewonnen. Dass es Menschen mit eher dampfbetriebenen Gehirnen und vormodernen Ansichten waren, ist womöglich mehr als ein bedauernswerter Zufall: Es macht wieder einmal sichtbar, wie groß das Defizit einer fortschrittlichen Kritik an einem Regime der Daten ist.“ (FAS, 13.11.2016)

Auch Carole Cadwalladr zieht in ihrem Artikel Parallelen zwischen Microtargeting-Maßnahmen und der Entwicklung hin zu einem Überwachungsstaat: „In the US, the

government is bound by strict laws about what data it can collect on individuals. But, for private companies anything goes. Is it unreasonable to see in this the possible beginnings of an authoritarian surveillance state? A state that is bringing corporate interests into the heart of the administration.” (The Observer, 7.5.2017) Big Data als Kontrollwissen wird hier als Gefahr in den Händen einzelner Akteure oder des Staates skizziert.

Johann Schloemann (SZ, 10.12.2016) ordnet die Diskussion um Big-Data-basierte Manipulation in einen kulturkritischen Diskurs ein. In diesem reihen sich – vor dem Hintergrund politischen Wettbewerbs – die neuen Formen datenbasierter Kommunikation in die Instrumente der Manipulation und propagandistischen Unterwanderung ein.

„Immer schon gehörte es zum demokratischen Wettstreit, die Wähler der Gegenseite nicht als von falschen Meinungen überzeugt, sondern als Opfer demagogischer Techniken darzustellen, von der Rhetorik bis zur Bestechung. Das eigene Lager behauptet hingegen, kommunikativ arglos zu sein, und will dies als Aufrichtigkeit im Dienst des Gemeinwohls verstanden wissen. Aber mit dem Aufkommen von professioneller Produktwerbung und politischer Massenpsychologie im 19. und 20. Jahrhundert hat das Misstrauen andere Dimensionen angenommen. Die Diskussion über das politische Psycho-Targeting erinnert etwa an literarische Überwachungsszenarien, aber auch an den Schock, den der amerikanische Wissenschaftsautor Vance Packard 1957 mit seinem Buch ‚Die geheimen Verführer‘ (‚The Hidden Persuaders‘) über die Tricks der Werbeindustrie auslöste [...] Aus der Verbindung von Werbung und Medien führt denn auch historisch der Weg zur heutigen politischen Manipulation im Internet. Der amerikanische Jurist und Netztheoretiker Tim Wu vergleicht in seinem gerade erschienenen Buch ‚The Attention Merchants. The Epic Scramble to Get Inside Our Heads‘ die ‚Aufmerksamkeitshändler‘ seit dem 19. Jahrhundert mit den kalifornischen Internetgiganten von heute. Das überwiegende Geschäftsmodell aller privaten Medien sei eigentlich immer dasselbe geblieben, nur werde es von Facebook & Co. heute auf die Spitze getrieben: Aufmerksamkeit gegen Werbung.“

Lediglich zwei Artikel thematisieren indessen „**representative claims**“, die sich auf Big Data stützen. Beide beziehen sich auf die Folgen von Big Data-getriebener Politik, die

auf vermeintliche Mehrheitsmeinungen reagiert: „Die Datenwissenschaftlerin Cathy O’Neil stellte in einem Beitrag für die *Huffington Post* die These auf, dass Donald Trump operiere wie ein verzerrter, maschinell lernender Algorithmus: Er werde mit falschen Daten – denen seiner Anhänger – gefüttert, handle danach und habe gar keine eigenen politischen Überzeugungen. Das wäre die Steigerung des auf den Ada-Algorithmus gestützten Wahlkampfes von Hillary Clinton.¹ Ihr wurde eine politische Stimmungslage vorgemacht, bei Trump, dessen Personalentscheidungen zu größten Befürchtungen Anlass geben, gibt es die vermeintlich den Wünschen der Wähler folgende Politik. Deren Folgen jedoch rechnet keine Maschine aus.“ (FAZ, 15.12.2016) Der Autor, Adrian Lobe, macht hier auf eine besondere Art der Instrumentalisierung von „representative claims“ aufmerksam: Während Politiker vorgeben, den in Daten abgebildeten Wähler*innenwillen zu erfüllen, bilde dieser in Wahrheit nur den eines Teils der Wählerschaft ab, die in den Daten erfasst werde. Auch in dem *Huffington Post*-Beitrag wird in Hinblick auf die US-Politik davor gewarnt, kurzsichtige datenbasierte Entscheidungen zu treffen: „Especially since the predictions are carried everywhere by social media, analytic models have become powerful information shortcuts in American politics with intense capacity to shape our perceptions of reality. The algorithms also lie to us. We do not have a good understanding of the full range of distortions that are introduced, but we know that predictive models simplify political reality and collective behavior.“ (The Huffington Post, 23.11.2016)

8 DISKUSSION

Im Folgenden werden die Einsichten der Untersuchung in Hinblick auf die eingangs formulierten Fragen und Thesen zusammengefasst und erläutert.

¹ Ada-Algorithmus: Der Name eines von Clintons Wahlkampfteam genutzten Algorithmus bezieht sich auf Ada Lovelace (1815-1852). Sie schrieb 1843 den ersten Algorithmus.

Frage 1: Wie wird das Verhältnis von Daten und gesellschaftlicher Realität in öffentlichen Diskursen verhandelt?

Ziel der Analyse war es, die interpretative Dimension von Big Data im Rahmen ihrer diskursiven Verhandlung aufzuzeigen. Der Begriff 'Big Data' wird im untersuchten Zeitraum mit drei relevanten Bedeutungskomplexen verknüpft:

- **Big Data als Technologie:** Insbesondere die Ubiquität von Daten und Datenpraktiken wird im Diskurs betont. Die Verwendung von Online-Daten (z. B. Cookies, Social-Media-Profile, Internetkäufe) und Offline-Daten (z. B. soziodemographische Merkmale, Konsumentendaten, Mitgliedschaften) sowie deren Zusammenführung, das sogenannte Onboarding, werden als technologischer Standard moderner politischer Wahlkämpfe dargestellt. Dabei werden verschiedene Möglichkeiten der Datengewinnung (z. B. Facebook-Quizze, Wahlregistereinträge, Einkauf von Konsumentendaten) dargestellt.
- **Big Data als analytische Ambition:** Methoden der Datenauswertung zur Identifikation von Wählerinnen und Wählern, die gezielt angesprochen werden sollen, sowie der Einsatz datenintensiver Methoden für Wahlprognosen werden ebenfalls diskursiv verhandelt.
- **Big Data als regulatorische Herausforderung:** Status und zukünftige Maßnahmen der Kontrolle und Überwachung datenintensiver Wahlkampfpraktiken werden mit Blick auf Datenschutz, Sicherheit, informationelle Selbstbestimmung und bürgerliche Freiheiten diskutiert.

Die Verhandlung der verschiedenen Bedeutungen von Big Data in Wahlkämpfen erfolgte wie erwartet aus **affirmativer** genauso wie aus **kritischer Perspektive**. Genauer wurden **drei grundlegende Diskurspositionen** bestimmt: Optimisten, Skeptiker und Pessimisten. Diese lassen sich einzelnen Sprecher*innen zuordnen, grundsätzlich aber bestimmen sie den Standpunkt und die Haltung, welche Akteure einnehmen können bzw. ihnen zugewiesen wird.

- **Optimisten** vertreten eine zugleich positive und positivistische Sicht auf datenintensive Wahlkampfpraktiken. Sie propagieren Erkenntnisgewinne durch Big Data-Analysen, die Wahlprognosen genauer und schneller sowie Wahlkampagnen planba-

rer, effektiver und effizienter machten, indem potentielle Spender*innen, Unterstützer*innen oder Wähler*innen mit auf ihre Bedürfnisse hin zugeschnittenen Botschaften gezielt angesprochen und auf diese Weise entweder zu einer bestimmten unterstützenden Handlung beziehungsweise Wahlentscheidung motiviert oder in sogenannten ‚get out the vote‘-Strategien davon abgehalten werden könnten, sich an Wahlen zu beteiligen. Sowohl in Bezug auf Microtargeting-Maßnahmen als auch im Hinblick auf Wahlprognosen versprechen sie sich mit einer Zunahme verfügbarer Daten eine Präzisierung der Ergebnisse und verlässlichere Vorhersagen. Diese Position wird hauptsächlich von Datenanalysten, die kommerziell arbeiten, sowie deren Kunden, also Parteien und Kampagnen, vertreten. Ihr Hauptargument besteht in der Annahme, dass sich Unsicherheit von Vorhersagen durch größere Datenmengen reduzieren ließe. Neben Effizienzsteigerung und Kostenminimierung in politischen Wahlkämpfen, wovon auch Kandidat*innen kleinerer Parteien und Parteilose profitierten, wird auf die Steigerung demokratischer Beteiligung verwiesen, da potentielle Nicht-Wähler*innen gezielt angesprochen werden und mit für sie relevanten Informationen versorgt werden könnten, die sie zur Stimmabgabe motivieren würden. Alexander Nix, CEO von Cambridge Analytica, führt zudem in einer Einzelposition bleibenden Überlegung aus, dass areal-basierte Targeting-Maßnahmen diskriminierend seien, weil sie anhand soziodemographischer Merkmale wie Wohnort, Geschlecht oder Hautfarbe Rückschlüsse auf politische Präferenzen zögen, während individualisierte Targeting-Maßnahmen die Persönlichkeit jedes einzelnen Wählers bzw. jeder Wählerin berücksichtigten.

- **Skeptiker** halten Erkenntnisvorteile durch Big Data-Auswertungen und darauf basierende Optimierungsmöglichkeiten von Wahlkampagnen für möglich, sehen die Expansion von Erkenntnis und Wirkung durch Big Data aber als begrenzt. Sie betonen zum Beispiel, dass datenintensive Methoden nur eine von vielen Komponenten eines Wahlkampfes seien, die zu dessen Erfolg beitragen können, sodass Wähleransprachen auf vielen Kanälen funktionieren müssten, nicht nur in maßgeschneiderten Online-Botschaften. Weiterhin weisen sie auf die Schwierigkeiten hin, wahlprognostische Methoden der neuen Datenlage anzupassen. Darüber hinaus gebe es immer Unsicherheitsfaktoren, die zuverlässige Prognosen störten – wie das Auftreten eines untypischen Präsidentschaftskandidaten wie Donald Trump. Diese Position wird hauptsächlich von Journalist*innen und Wissenschaftler*innen vertreten, wenn diese mit Kommentaren oder Stellungnahmen auftreten.

- **Pessimisten** stellen die erkenntniserweiternde Autorität von Daten generell infrage. Sie verweisen auf datenbasierte Wahlkämpfe, die nicht erfolgreich waren, wie die von Ted Cruz, Bernie Sanders, Marco Rubio oder Hillary Clinton, und die daher als Zeugnisse der Unwirksamkeit der Maßnahmen dienen. Sie betonen die fehlende wissenschaftliche Evidenz der Wirksamkeit der Methoden und heben hervor, dass diese keinen Mehrwert gegenüber traditionellen Formen der Umfrageforschung und Targeting-Maßnahmen böten. Neben Zweifeln an der Methodik datenintensiver Prognosen und des Microtargeting betonen sie eine generelle Unvorhersehbarkeit des menschlichen Handelns, indem sie sich auf das aufklärerische Konzept des freien Willens berufen. Menschen seien zu eigenständigen Entscheidungen fähig, bei denen sie sich von Informationen, Interessen, Neigungen, aber auch Stimmungen leiten ließen, weshalb die planmäßige Wirksamkeit individueller persuasiver Botschaften zweifelhaft sei. Auch wird als Argument angeführt, dass datenintensive Kampagnen eher der Imagepflege eines Kandidaten oder einer Partei als besonders fortschrittlich und dynamisch zuträglich sei statt dem tatsächlichen Erkenntnisvorsprung. Diese Position wird hauptsächlich von Journalist*innen vertreten, wenn diese sich meinungsbildend am Diskurs beteiligen.

Eine kritische Perspektive im Diskurs nimmt zudem ein weiterer Typ der **Warnenden** ein. Sie problematisieren die Anwendung von Big-Data-Analysen aus **datenschutzrechtlicher, persönlichkeitsrechtlicher** oder **demokratiethoretischer Sicht**. Vertreten wird diese Position hauptsächlich von Journalist*innen und Wissenschaftler*innen, die ihre Funktion als Diskursteilnehmende im Sinne einer Informierung und Stärkung der Position der Bürger*innen sehen, die von Datensammel- und Auswertungspraktiken und darauf basierender Wahlkampfmaßnahmen betroffen sind. Als konkrete Problemfelder werden benannt:

- **Datenschutz:** Neben der Infragestellung der Rechtmäßigkeit einzelner Datensammelpraktiken betonen sie, dass selbst bei rechtmäßigem Besitz von aggregierten Daten nicht ausreichend sichergestellt werden könne, dass die Daten nicht in die Hände Unbefugter gelangten. Weiterhin wird kritisiert, dass Nutzer*innen heute vielfach gar keine Wahlmöglichkeit hätten, die Speicherung und Verwendung ihrer Daten zu unterbinden, wenn sie auf bestimmte mediale Inhalte und Informationsangebote zugreifen wollen.

- **Datenklufft:** Warnende befürchten eine Kluft (als *data divide*) zwischen 'Datenreichen' und 'Datenarmen', die neue Hürden in der politischen Partizipation installiert. So seien in den USA Wähler*innendaten für Parteien leichter und kostengünstiger zugänglich als für Parteilose. Auch seien die Beschaffung und Auswertung von Datensammlungen mit Kosten verbunden, die sich nicht jeder Akteur leisten könne, sodass es zu richer-get-rich-Effekten kommen könne, wobei etablierte Parteien mit solider finanzieller Ausstattung bzw. finanzstarke Einzelkandidat*innen ihre privilegierte Position durch datenintensive Maßnahmen noch ausbauen und weiter festigen könnten, während neue Kandidat*innen und kleine Parteien aus dem politischen Wettbewerb verdrängt würden.
- **Manipulation:** Die Befürchtung besteht, dass durch Microtargeting-Maßnahmen Wähler*innen ohne jede Form von Kontrolle manipuliert werden könnten, weil jede*r von ihnen andere Botschaften sehe und die Auslieferung persuasiver Botschaften auf Social-Media-Kanälen aufgrund der Opazität der Strukturen und der Geschäftsgeheimnisse der Betreiber nicht kontrolliert werden könne. Dabei steht besonders die Angst im Vordergrund, dass Wähler*innen dazu gebracht werden könnten, populistische Positionen, Kandidat*innen und Parteien zu unterstützen.
- **Überwachung:** Big Data könnten als Herrschafts- und Kontrollwissen von Firmen und Regierungen eingesetzt werden, die ihre 'gläsernen' Bürger*innen bei von der gewünschten Norm abweichendem Verhalten gezielt steuern können – durch persuasive Anreize oder sogar durch Exekutivgewalt.
- **Parzellierung der Wählerschaft und Fragmentierung politischer Öffentlichkeit:** Durch die gezielte Ansprache einzelner Wähler*innen mit individualisierten Botschaften werde die Herausbildung einer politischen Öffentlichkeit im Habermas'schen Sinne verhindert und damit die demokratische Legitimation einer Gesellschaft untergraben. Demokratischer Konsens werde torpediert, weil jeder Bürger und jede Bürgerin andere Themen und andere Deutungsmuster dieser Themen zugespielt bekomme, sodass keine gemeinsame Diskursgrundlage mehr vorhanden sei.

Anhand des untersuchten Materials lässt sich eine diskursive Auseinandersetzung um das dominante Deutungsmuster der Rolle von Big Data in politischen Prozessen feststellen. Gleich ob Big Data in kritischer oder affirmativer Perspektive verhandelt wird, in den meinungsmäßig pointiertesten Stellungnahmen versuchen die Sprecher*innen,

die jeweils anderen Positionen zu diskreditieren: Die Optimisten stellen die Warnenden als fortschrittsängstliche Bedenkenträger dar. Die Warner wiederum bezichtigen die Optimisten, manipulative, demokratiezersetzende und Persönlichkeitsrechte verletzende Praktiken durch das Vorantreiben datenintensiven Methoden zu ermöglichen.

Jedoch teilen Optimisten und Warnende einen entscheidenden Punkt: Sie sind überzeugt von der tiefgreifenden Wirkung von Big Data. Der Unterschied liegt darin, dass die einen diese Effekte positiv oder gar utopisch auslegen, während die anderen sie negativ bis dystopisch deuten. Das zeigt sich etwa mit Blick auf gesellschaftliche Machtverhältnisse: Die Optimisten meinen, Big Data würde marginalisierten Gesellschaftsgruppen Chancen bieten, während die Warner auf die ungleiche Verteilung beim Zugang zu Daten (bzw. den Auswertungsmöglichkeiten) verweisen. Beide gehen aber von gravierenden Effekten durch Big Data aus, womit sie sich von den Positionen der Pessimisten und Skeptiker unterscheiden.

Frage 2: Inwiefern wird die Repräsentativität der in Daten abgebildeten und aus diesen resultierenden Willensbekundungen, Aktivitäten oder Relationen thematisiert bzw. problematisiert?

Die **Repräsentativität der Datenlage** wird im untersuchten Diskursmaterial kaum kritisch hinterfragt. Meist gehen die Kommunikate implizit von einer Ubiquität datenintensiver Praktiken aus, von der jeder Mensch gleichermaßen betroffen sei – von Praktiken des Datensammelns und -auswertens genauso wie auf darauf basierenden Microtargeting-Maßnahmen. Nur in Ausnahmefällen wird die Repräsentativität von Online-Daten im Rahmen von Wahlprognosen kritisch diskutiert. Insgesamt entsteht sowohl eher der Eindruck, dass Daten in ausreichender, valider Form über einen Großteil der Bevölkerung vorlägen beziehungsweise diese in umfassender Weise durch Microtargeting-Maßnahmen erreicht werden könnten als auch, dass mit Ausweitung der Möglichkeiten der Generierung, Speicherung und Auswertung der Daten eine vollständige datenbasierte Durchdringung der Gesellschaft einhergehen werde.

Obwohl neue Akteure in prominenten Sprecherpositionen im Diskurs auftreten (z. B. Datenanalysten, Statistiker, Leiter von Digitalkampagnen) werden von ihnen kaum 're-

presentative claims' artikuliert, also Ansprüche aus in Daten abgebildeten Mehrheitsmeinungen. Lediglich zwei Artikel warnen davor, Politik am algorithmisierten, vermeintlichen Wählerwillen auszurichten. Eine Erklärung für die im Diskurs fehlende Artikulation von Repräsentationsansprüchen kann im Deutungsmuster gefunden werden, mit dem das Ziel datenintensiver Maßnahmen vornehmlich artikuliert wird: Politische Kampagnen seien nicht mehr an der Beeinflussung der öffentlichen Meinung interessiert, sondern an der Einflussnahme auf individuelle Meinungen durch maßgeschneiderte Botschaften. Auch **'data activism'**, also Bürgerinitiativen, die Big Data ihrerseits nutzen, um etwa Themen auf die politische Agenda zu setzen oder Einfluss auf die Gestaltung von Politikfeldern zu nehmen, spielt im Diskurs keine Rolle. Bürger*innen werden vielmehr als passive Opfer von Datensammlungs- und Datenauswertungspraktiken und darauf basierenden Beeinflussungsversuchen dargestellt.

Frage 3: Inwiefern werden im öffentlichen Diskurs die analytischen Praktiken in Bezug gesetzt zum politischen System, insbesondere Wahlsystem, sowie der politischen Kultur, vor allem im Blick auf das Parteienverhältnis und die Formen des Wahlkampfes?

Explizit thematisiert werden im Diskurs unterschiedliche datenschutzrechtliche Regulierungsansätze in den USA im Gegensatz zu Europa allgemein beziehungsweise zu Deutschland oder Großbritannien im Speziellen. Die USA werden als Vorreiter datenbasierter Wahlkampfmethoden präsentiert, die auf eine längere und ausgeprägtere Tradition der Sammlung von Wählerdaten zurückblickten ('voter files') als europäische Staaten, welche die Speicherung personenbezogener Daten zum Wahlverhalten aus Datenschutzgründen verböten bzw. einschränkten. Auch seien in Europa die Möglichkeiten rechtlich begrenzt, aggregierte Daten bei kommerziellen Datenanbietern (z. B. Acxiom, Experian, Epsilon, Datalogix, Infogroup, Affinity Answers, Affinio) zu kaufen. Weiterhin wird darauf verwiesen, dass die USA mit ihrem de facto Zweiparteiensystem und der letztlichen Konzentration auf die beiden Kandidat*innen für das Amt des Präsidenten in der Abschlussphase des Wahlkampfes größere Wirkungen durch Microtargeting erzielen könnten als in europäischen Ländern mit Mehrparteiensystemen und mehrheitlich partei- statt personengeführten Wahlkämpfen. Nichtsdestotrotz werden Big-Data-Anwendungen in politischen, zusehends auch personalisierten

Wahlkämpfe auch in Europa für möglich gehalten oder als bereits praktiziert dargestellt.

Der **Diskursschwerpunkt** liegt sowohl in als auch auf den USA: Die Mehrheit der Kommunikate (52,9%) stammt aus den USA. Diskursiv problematisiert werden vorrangig Status, Implikationen und Beschränkungen datenintensiver Wahlkampf- und Prognosemethoden, wobei britische und deutsche bzw. deutschsprachige Kommunikate mehrheitlich (zu 70%) auf Entwicklungen und Methoden in den USA rekurrieren. US-amerikanische Artikel und Posts verweisen – wenn überhaupt – auf die Rolle der Datenanalysefirma Cambridge Analytica beim britischen Brexit-Referendum, weil dieselbe Firma auch für Donald Trumps Wahlkampf tätig war. Überhaupt erreicht der Diskurs die deutschsprachige Diskussion mit deutlicher Verzögerung, nämlich erst nach dem Wahlsieg Trumps am 8. November 2016. Insbesondere die Veröffentlichung des *Magazin*-Artikels „Ich habe nur gezeigt, dass es die Bombe gibt“ am 3. Dezember 2016 setzt eine Diskursintensivierung in Gang.

Insgesamt werden die an den Diskurs vorab herangetragenen Thesen durch die Analyseergebnisse gestützt:

Die herausgearbeiteten Diskurspositionen – Optimisten, Skeptiker, Pessimisten und Warnende – verdeutlichen, dass sich bisher keine solitäre hegemoniale Vorstellung darüber herausgebildet hat, welche Bedeutung Big Data in und für Wahlkämpfe hat (**These 1**). Vielmehr werden der Begriff und die damit verknüpften datenbasierten Praktiken mit variablen Bedeutungen aufgeladen, die aus technischen Sachbezügen stammen, mit analytischen Erwartungen einhergehen, epistemischer Vorteile herausstreichen oder reelle bzw. potentielle Bedrohungen thematisieren.

Auch eine fundamentale Unsicherheit hinsichtlich probabilistischer Schlüsse aus datenintensiven Handlungen kann aus den verschiedenen Diskurspositionen abgelesen werden. Vor allem die Gruppe der Skeptiker betonen die diskursive Spannung zwischen dem Definieren retrospektiver Gewissheiten, dem Prognostizieren prospektiver Wahrscheinlichkeiten und dem Diskutieren ihres Eintretens bzw. Nicht-Eintretens (**These 2**).

Zentrale Sprecher- und Subjektpositionen, die von teilweise neuen Akteuren im politischen Diskurs eingenommen werden, konnten ebenfalls identifiziert werden (**These 3**):

a. Politiker*innen: Insbesondere die Kampagnen von Donald Trump, Hillary Clinton, aber auch von Ted Cruz und Marco Rubio werden als Pioniervorhaben datenbasierter Wahlkampfpraktiken diskutiert. Daneben widmen sich viele Kommunikate populistischen Bewegungen wie den Leave.EU-Kampagnen. **b. Datenanalysefirmen:** Im Zentrum der Aufmerksamkeit steht Cambridge Analytica, deren (mögliche) Rolle in allen drei untersuchten Wahlkämpfen thematisiert wird. Neben Cambridge Analytica werden zahlreiche andere, weniger bekannte Firmen in Verbindung mit datenbasierten Wahlprognosen und Kommunikationsmaßnahmen genannt, teilweise wird auch deren Rolle und Einfluss rekonstruiert (u.a. Proofpoint, Aggregate IQ, Civis Analytics, Tableau, ClearStory Data, Targeted Victory, Blue State Digital, Harris Media, Optimus, Push Digital). **c. Wissenschaftler*innen & Analyst*innen:** Neben Alexander Nix, CEO von Cambridge Analytica, stehen besonders Michal Kosinski, der Erfinder der Facebook-Persönlichkeitstests, und Nate Silver als prominenter Statistiker und Wahlforscher im Zentrum des Diskurses bzw. kommen selbst zu Wort. **d. Bürger*innen/Wähler*innen/Social-Media-Nutzer*innen:** Auffällig ist, dass die Kommunikate mehrheitlich die Auswirkungen von Big Data-Praktiken auf die Gesellschaft thematisieren. Damit stellen sie keine Individuen ins Zentrum, sondern Kollektive wie die Wählerschaft, deren Wahlentscheidung möglicherweise durch gezielte, personalisierte Wahlwerbung beeinflusst sein könnte, oder das Kollektiv der Social-Media-Nutzer*innen, deren Daten ohne deren Wissen oder Einwilligung für politische Zwecke gebraucht werden.

9 HANDLUNGSEMPFEHLUNGEN

Auf Basis der Ergebnisse, wie datenbasierte Wahlkämpfe in der Öffentlichkeit aufgenommen und verhandelt werden, sollen im Folgenden Handlungsempfehlungen abgeleitet werden. Sie richten sich an die im Diskurs identifizierten Akteure bzw. Akteurskonstellationen und beziehen weitere Interessengruppen und zu beteiligende Institutionen ein. Im Kern aller Empfehlungen steht die Herausforderung, **Repräsentationsbeziehungen als grundsätzlich wechselseitige und kommunikative Prozesse zu verstehen**, in denen politische Kandidat*innen und gewählte Entscheidungsträger*innen mit Bürger*innen in Austausch bleiben, um, aus Sicht der politischen Reprä-

sentant*innen, Initiativen und Entscheidungen zu vermitteln, abzugleichen und zu erläutern und um, aus Sicht der Bürger*innen, Interessen und Bedenken weiterzugeben, mitzuteilen und einzuordnen. Diese grundlegenden Bedingungen demokratischer Repräsentationsbeziehungen gelten auch für datenbasierte Politik und sollen, so der Anspruch, auch unter veränderten technischen und informationellen Bedingungen hergestellt werden.

1. Erweiterung der politischen Bildung von Bürgerinnen und Bürgern über die Möglichkeiten und Formen datenbasierter Politik.

Bürger*innen werden zunehmend Fokus und Bezugsgruppe datenbasierter Politik. Die Big Data-geführten Wahlkämpfe sind eine intensive, zeitlich begrenzte und öffentlich besonders diskutierte Form dieses Wandels. Um die datenanalytischen Verfahren in ihren Ansätzen, ihren Orientierungen und ihren Konsequenzen für den Einzelnen, für Gruppen und für die Gesellschaft besser verstehen und einschätzen zu können, sind entsprechende Bildungsangebote her- und bereitzustellen. Orte dieser Medien- und 'Daten'-Bildung sind Schulen und Universitäten ebenso wie Stiftungen und Einrichtungen der politischen Bildung. Auftrag ist die Befähigung der Bürger*innen, die Rolle von aggregierten Daten und darauf bezogenen Operationen auch für Fragen der demokratischen Willensbildung und der Gestaltung des Gemeinwohls zu diskutieren und bewerten zu können. Damit einher geht die Erwartung, dadurch einen Beitrag zu leisten, die Optionen persönlichen Datenschutzes bzw. selbstverantwortlicher datenbezogener Reaktionen (etwa als *data obfuscation*) aufzuzeigen und zu erproben. Einen Schritt weiter noch geht die Befähigung, selbst datenbezogene politische Aktivitäten im Sinne des *data activism* durchzuführen.

2. Erweiterung und Vertiefung journalistischer Expertise

Damit Journalismus als sozusagen Fourth Estate seine gesellschaftliche Aufgabe einer kritischen Begleitung, Reflexion und Beurteilung von politischen Prozessen und deren Auswirkungen auf das Gemeinwesen und die Chancen persönlicher Lebensführung erfüllen kann, müssen Journalist*innen über das entsprechende Fach- und Bewertungswissen verfügen. Die neue Realität datenbasierter Kampagnen stellt in diesem Sinn die Journalist*innen vor die Aufgabe, sich die notwendige Expertise anzueignen, um die

komplexen Vorgänge zu überblicken, um sie in die Öffentlichkeit zu bringen und um sie zugleich in ihren Bedingungen oder Implikationen zu problematisieren. Der Erwerb der erforderlichen *data/code literacy* ist nicht auf die publizistische Begleitung von Wahlkämpfen oder politischer Praxis im Allgemeinen beschränkt. Weil Daten und datenbasierte Anwendungen zusehends mehr Gesellschaftsbereiche und Ressorts durchdringen wird sie eine journalistische Qualifikation per se und sollte Teil des beruflichen Profils sein. Auf diese Weise kann es auch möglich werden, die mediale Behandlung und Deutung von Big Data über dichotome Szenarien und Zustimmung oder Ablehnung hinaus zu führen und eine (ebenfalls sensibilisierten und kritikfähigen) Öffentlichkeit in eine differenzierte Auseinandersetzung einzubeziehen.

3. Auskunftspflicht für datenbasierte Aktivitäten von Parteien

Im Rahmen politischer Aktivitäten erkennen nicht nur staatliche Agenturen und Verwaltungen die Möglichkeiten einer wirksamen, datenbasierten Verhaltenssteuerung (*nudging*). Auch Parteien greifen in Wahlkämpfen auf datenbasierte Auswertungen zurück, um ihre Botschaften zu kommunizieren, Wähler*innen zu adressieren und Wahlentscheidungen zu beeinflussen. Dabei reicht das Spektrum von etablierten Formen der Ansprache bis zu neuen Arten des individualisierten und zum Teil subkutanen Priming von Botschaften. Nicht immer scheinen diese Verfahren dazu zu dienen, Informationen und Aussagen im Sinn einer Überzeugungsabsicht zu verbreiten, sondern es umfasst auch Schattierung von Persuasion und Propaganda und kann dazu beitragen, die Wählerschaft zusehends zu parzellieren, die Wahlchancen kleiner Parteien zu marginalisieren und die Wahlbeteiligung kritischer Gruppen zu verringern. Aus diesem Grund scheint es geboten, eine Auskunftspflicht zu installieren. Dadurch wären, ähnlich wie im Falle der Parteienfinanzierung praktiziert, alle in einer Wahl antretenden Parteien verpflichtet, Auskunft zu geben über die genutzten Verfahren, die dabei herangezogenen Daten und die eventuell in Anspruch genommenen Dienstleistungen von Dritten. Eine Prüfung könnte durch eine unabhängige Instanz in Form einer einzurichtenden Prüfkommision bzw. einer Ombudsperson erfolgen.

4. Rechtliche Bedingungen für Datenbesitz, Datenweitergabe und Datenzusammenführung prüfen

Im Ländervergleich zeigen sich die deutlichen Unterschiede in den rechtlichen Rahmenbedingungen für Datenbesitz, Datenweitergabe und Datenzusammenführung. Darauf gründen die unterschiedlich ausgeprägten Möglichkeiten einer datenbasierten Meinungsforschung und Ansprache. Zugleich aber zeigt der Vergleich die Expansion der vornehmlich in den USA geprägten Formen und die Ambitionen, diese auch in andere Wahlsysteme und Jurisdiktionen zu transferieren. Vor diesem Hintergrund ist es geboten, die rechtlichen Bedingungen für Datenbesitz, Datenweitergabe und Datenzusammenführung mit Hinblick auf die aktuell zirkulierenden Datenarten und die Innovationen in Erhebungs- und Auswertungstechniken zu überprüfen und gegebenenfalls anzupassen. Leitidee für diese Aufgabe in bundesdeutschem und europäischem Rahmen sollte die Gewährleistung von Datenintegrität und informationeller Selbstbestimmung der Bürger*innen sein. Teil der rechtlichen Prüfung sollte ebenso die Gewährleistung der Datensicherheit durch die datenverarbeitenden Stellen sein, insbesondere mit Auftreten neuer politisch relevanter Akteure wie Plattformen und Kampagnenagenturen, aber auch staatliche und Partei-Server.

10 LITERATUR

- Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*. Abgerufen von http://www.wired.com/science/discoveries/magazine/16-07/pb_theory
- Andrejevic, M. (2014). The Big Data Divide. *International Journal of Communication*, 8, 1673-1689.
- Anstead, N. (2008). The Internet and Campaign Finance in the U.S. and the UK: An Institutional Comparison. *Journal of Information Technology & Politics*, 5(3), 285-302. doi:10.1080/19331680802425214
- Anstead, N. (2016). A Different Beast? Televised Election Debates in Parliamentary Democracies. *The International Journal of Press/Politics*, 21(4), 508-526.
- Anstead, N. (2017). Data-Driven Campaigning in the 2015 United Kingdom General Election. *The International Journal of Press/Politics*, 22(3). doi:1940161217706163
- Anstead, N. & Chadwick, A. (2008). Parties, Election Campaigning, and the Internet: Toward a Comparative Institutional Approach. In A. Chadwick. & P. Howard (Hrsg.), *The Handbook of Internet Politics* (S. 56–71). London: Routledge.
- Arceneaux, K. & Nickerson, D. W. (2009). Comparing Negative and Positive Campaign Messages: Evidence from Two Field Experiments. *American Politics Research*, 38(1), 54-83. doi:10.1177/1532673x09331613
- Barassi, V. (2016). Contested Visions: Digital Discourses as Empty Signifiers from the 'Network' to 'Big Data'. *Communication and the Public*, 1(3).
- Beer, D. (2016). How Should We Do the History of Big Data? *Big Data & Society*, 1(3), 1-10.
- Behnke, J., Blätte, A., Schnapp, K.-U. & Wagemann, C. (Hrsg.). (in Druck). *Big Data? New Data!* Baden-Baden: Nomos.
- Bimber, B. (2014). Digital Media in the Obama Campaigns of 2008 and 2012: Adaptation to the Personalized Political Communication Environment. *Journal of Information Technology & Politics*, 11(2), 130-150. doi:10.1080/19331681.2014.895691
- Blaemire, R. (2012). An Explosion of Innovation: The Voter-Data Revolution. In N. G. Pearlman (Hrsg.), *Margin of Victory: How Technologists Help Politicians Win Elections* (S. 107-120). Santa Barbara, CA: ABC-CLIO.
- Boellstorff, T. (2015). Making Big Data, in Theory. In T. Boellstorff & B. Maurer (Hrsg.) *Data. Now Bigger and Better* (S.87-108). Chicago: Prickly Paradigm Press.
- Bowker, G. (2013). Data Flakes. In L. Gitelman (Hrsg.) "Raw Data" is an Oxymoron (S. 167-172). Cambridge, MA: MIT Press.
- Boyd, D & Crawford, K. (2013). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), 662–679.
- Broockman, D. & Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science*, 352(6282), 220-224. doi:10.1126/science.aad9713
- Burrows, R. & Savage, M. (2014). After the Crisis? Big Data and the Methodological Challenges of Empirical Sociology. *Big Data & Society* 1(1), 1-6.

- Callon, M. (2007). What Does It Mean to Say That Economics Is Performative? In D. MacKenzie, F. Muniesa & L. Siu (Hrsg.), *Do Economists Make Markets? On the Performativity of Economics*. (S. 311-357) Princeton: Princeton University Press.
- Castleman, D. (2016). *Essentials of Modeling and Microtargeting*. In A. Therriault (Hrsg.), *Data and Democracy: How Political Data Science Is Shaping the 2016 Elections* (S.1-6). Sebastopol, CA: O'Reilly Media.
- Couldry, N. & Hepp, A. (2017). *The Mediated Construction of Reality*. Cambridge: Polity.
- Couldry, N. & Powell, A. (2014). Big data from the bottom up. *Big Data & Society*, 1(2). Abgerufen von <http://journals.sagepub.com/doi/abs/10.1177/2053951714539277>
- Cowley P. & Kavanagh, D. (2015). *The British General Election of 2015*. Basingstoke: Palgrave Macmillan.
- Crabtree, J. (.2010). David Cameron's Battle for Connect. *Wired Magazine*. Abgerufen von <http://www.wired.co.uk/article/david-cameros-battle-to-connect>
- Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197-208.
- Disch, L. (2008). The People as "Presupposition" of Representative Democracy - An Essay on the Political Theory of Pierre Rosanvallon. *Redescriptions: Political Thought, Conceptual History and Feminist Theory*, 12(1), 47-71.
- Dorschel, J. (Hrsg.). (2015): *Praxishandbuch Big Data*. Wiesbaden: Springer.
- Esser, F. & Strömbäck, J. (2012). Comparing Election Campaign Communication. In F. Esser & T. Hanitzsch (Hrsg.), *The Handbook of Comparative Communication Research* (S. 289-307). New York: Routledge.
- Falguera E., Jones S. & Ohman M. (2014). *Funding of Political Parties and Election Campaign: A Handbook of Political Finance*. Stockholm: Institute for Democracy and Electoral Assistance.
- Ferguson, A. G. (2017). *The Rise of Big Data Policing*. New York: NYU Press.
- Ferree M. M., Gamson W. A., Gerhards J. & Rucht D. (06.2002). *Theory and Society*, 31(3), 289-324.
- Gerber, A. S. & Green, D. P. (2012). *Field Experiments: Design, Analysis, and Interpretation*. New York, NY: W. W. Norton & Company.
- Gerber, A. S., Kessler, D. P. & Meredith, M. (2011). The Persuasive Effects of Direct Mail: A Regression Discontinuity Based Approach. *The Journal of Politics*, 73(1), 140-155. doi:10.1017/s0022381610000927
- Geiselberger, H. & Moorstedt, T. (2013). Vorwort. In Diess. (Hrsg.), *Big Data. Das neue Versprechen der Allwissenheit* (S. 7-20). Berlin, Suhrkamp.
- Gitelman, L. & Jackson, V. (2013). Introduction. In L. Gitelman (Hrsg.) "Raw Data" is an Oxymoron (S. 1-14). Cambridge, MA: MIT Press.
- Green, D. P. & Gerber, A. S. (2015). *Get Out the Vote: How to Increase Voter Turnout* (3. Auflage). Washington, DC: Brookings Institution Press.
- Gülden-zopf, R. (2011). Die Evolution eines Wahlkampfinstrumentes - von Ventura bis Obama. In R. Grünewald, R. Gülden-zopf & P. Melanie (Hrsg.), *Politische Kommunikation: Beiträge zur Politischen Bildung* (S. 301-308). Berlin: Lit Verlag.
- Hajer, M. (1995). *The Politics of Environmental Discourse*. Oxford: Oxford University Press.

- Halavais, A. (2015). Bigger sociological imaginations: framing big social data theory and methods. *Information, Communication & Society* 18(5), 583-594.
- Hallin, D.C. & Mancini, P. (2004). *Comparing media systems: three models of media and politics*. New York: Cambridge University Press.
- Herschinger, E. & Nonhoff, M. (2014). Diskursforschung in der Politikwissenschaft. In J. Angermüller, M. Nonhoff, E. Herschinger, F. Magilchrist, M. Reisigl, J. Wedl, D. Wrana & A. Ziem (Hrsg.), *Diskursforschung. Ein interdisziplinäres Handbuch. Band I: Theorien, Methodologien und Kontroversen* (S. 192-207). Bielefeld: transcript.
- Hersh, E. D. (2015). *Hacking the Electorate: How Campaigns Perceive Voters*. New York, NY: Cambridge University Press.
- Hilbert, M. & López, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science* 332(60), 60-65.
- Hoeren, T. (2017). Big data and the legal framework for data quality. *International Journal of Law and Information Technology*, 25(1), 26-37.
- Howard, P. N. (2006). *New Media Campaigns and the Managed Citizen*. Cambridge: Cambridge University Press.
- Issenberg, S. (2012). *The Victory Lab: The Secret Science of Winning Campaigns*. New York: Crown.
- Jungherr, A. (2012). Online Campaigning in Germany: The CDU Online Campaign for the General Election 2009 in Germany. *German Politics*, 21(3), 317-340. doi:10.1080/09644008.2012.716043
- Jungherr, A. (2013). Die Rolle des Internets in deutschen Wahlkämpfen. *ZPB Zeitschrift für Politikberatung*, 2013(2). 91-95. doi:10.5771/1865-4789-2013-2-89
- Jungherr, A. (2015). *Analyzing political communication with digital trace data*. Cham, Switzerland: Springer.
- Jungherr, A. (2016). Four Functions of Digital Tools in Election Campaigns: The German Case. *International Journal of Press/Politics*, 21(3), 358-377. doi:10.1177/1940161216642597
- Jungherr, A. (2017). Das Internet in der politischen Kommunikation: Forschungsstand und Perspektiven. *Politische Vierteljahresschrift* 58(2), 285-316. doi: 10.5771/0032-3470-2017-2-285
- Jungherr, A. & Schoen, H. (2013). *Das Internet in Wahlkämpfen: Konzepte, Wirkungen und Kampagnenfunktionen*. Wiesbaden: Springer VS.
- Karpf, D. (2016). *Preparing for the Campaign Tech Bullshit Season*. Civicist: Civic Tech News & Analysis. Abgerufen von <http://civichall.org/civist/preparing-campaign-tech-bullshit-season>
- Karpf, D. (2017). *Analytical Activism: Digital Listening and the New Political Strategy*. New York: Oxford University Press.
- Keller, R., Knoblauch, H., & Reichertz, J. (Hrsg.). (2013). *Kommunikativer Konstruktivismus*. Wiesbaden: VS.
- Kennedy, H. & Moss, G. (2015). Known or knowing publics? Social media data mining and the question of public agency. *Big Data & Society*. Abgerufen von <http://journals.sagepub.com/doi/abs/10.1177/2053951715611145>
- Kenski, K. & Jamieson, K (Hrsg.). (2014). *The Oxford Handbook of Political Communication*. Oxford: Oxford University Press.

- Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures & their Consequences*. London: Sage.
- Kitchin, R. (2016). Thinking critically about and research algorithms. *Information, Communication & Society*, 20(1), 14-29.
- Kreiss, D. (2012). *Taking Our Country Back: The Crafting of Networked Politics from Howard Dean to Barack Obama*. New York: Oxford University Press.
- Kreiss, D. (2016). *Prototype Politics: Technology-Intensive Campaigning and the Data of Democracy*. New York: Oxford University Press.
- Kreiss D. & Howard P. N. (2010). New Challenges to Political Privacy: Lessons from the First US Presidential Race in the Web 2.0 Era. *International Journal of Communication*, 19(4). Abgerufen von <http://ijoc.org/index.php/ijoc/article/view/870/473>
- Laclau, E. (1996). *Emancipations*. London: Verso.
- Lane, J., Stodden, V., Bender, S. & Nissenbaum, H. (Hrsg.). (2014). *Privacy, Big Data, and the Public Good*. Cambridge: Cambridge University Press.
- Law, J. (2009). Seeing like a survey. *Cultural Sociology*, 3(2), 239-256.
- Law, J. (2012). Collateral realities. In F.D. Rubio & P. Baert (Hrsg.), *The Politics of Knowledge* (S.156-178). London: Routledge.
- Lazer, D., Pentland, A. & Adamic, L., et al. (2009). *Computation Social Science*. *Science* 323(5915), 721-723.
- Lundry, A. (2016). Data-Driven Media Optimization. In A. Therriault (Hrsg.), *Data and Democracy: How Political Data Science Is Shaping the 2016 Elections* (S. 19-24). Sebastopol, CA: O'Reilly Media.
- Lyon, D. (2014). Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique. *Big Data & Society* 1(2), 1-13.
- Madrigal, A. C. (16.11.2012). When the Nerds Go Marching. *The Atlantic*. Abgerufen von http://www.theatlantic.com/technology/archive/2012/11/when-the-nerds-go-marching-in/265325/?single_page=true
- Mayer-Schönberger, V. & Cukier, K. (2013). *Big Data. A Revolution That Will Transform How We Live, Work and Think*. New York: John Murray.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse*. Weinheim: Beltz.
- McKibbin, R. (2016). A Brief Supremacy: The Fragmentation of the Two-Party System in British Politics, c. 1950–2015. *Twentieth Century British History* 27(3), 450–469.
- Meier, S., & Pentzold, C. (2010). Theoretical sampling als auswahlstrategie für Online-Inhaltsanalysen. *Die Online-Inhaltsanalyse: Forschungsobjekt Internet*. Köln: Herbert von Halem Verlag, 103-123.
- Michael, M. & Lupton, D. (2015). Toward a manifesto for the “public understanding of big data”. *Public Understanding of Science*. Abgerufen von <http://journals.sagepub.com/doi/abs/10.1177/0963662515609005>
- Milan, S. (2013). *Social Movements and Their Technologies: Wiring Social Change*. Basingstoke: Palgrave Macmillan.
- Nickerson, D. W. & Rogers, T. (2014). Political Campaigns and Big Data. *The Journal of Economic Perspectives* 28(2), 51-74. doi:10.1257/jep.28.2.51
- Nielsen, R. K. (2012). *Ground Wars: Personalized Communication in Political Campaigns*. Princeton, NJ: Princeton University Press O’Neil.

- Pearlman, N. G. (2012). Bootstrapping an Enterprise: NGP and the Evolution of Campaign Software. In N. G. Pearlman (Hrsg.), *Margin of Victory: How Technologists Help Politicians Win Elections* (S. 189-202). Santa Barbara, CA: ABC-CLIO.
- Pfetsch, B. (2014). The idea of political communication cultures and its empirical correlates. In B. Pfetsch (Hrsg.) *Political communication cultures in Europe: attitudes of political actors and journalists in nine countries* (S. 13-30). Houndmills, Basingstoke: Palgrave Macmillan.
- Portmess, L. & Tower, S. (2015). Data Barns, Ambient Intelligence and Cloud Computing: The Tacit Epistemology and Linguistic Representation of Big Data. *Ethics and Information Technology*, 17(1). 1-9.
- Puschmann, C. & Burgess, J. (2014). Metaphors of Big Data. *International Journal of Communication*, 8, 1690-1709.
- Ribes, D. & Jackson, S. J. (2013). Data Bite Man: The Work of Sustaining a Long-Term Study. In L. Gitelman (Hrsg.), "Raw Data" is an Oxymoron (S.147-166). Cambridge, MA: MIT Press.
- Richel, M. (4.12.2016). Nach Cambridge Analytica: Vom kleinen Big Data deutscher Parteien. Zielgruppenfernes Verhalten: Politik & Kampagnen. Abgerufen von <http://www.zielgruppenfernes-verhalten.de/2016/12/04/nach-cambridge-analytica-vom-kleinen-big-data-deutscher-parteien>
- Römmele, A. & Von Schneidmesser, D. (2016). Election Campaigning Enters a Fourth Phase: The Mediatized Campaign. *Zeitschrift für Politikwissenschaft*, 2016(4), 184-197.
- Rosanvallon, P. (2006). *Democracy past and future*. New York: Columbia University Press.
- Ross, T. (16.5.2015). Secrets of the Tories' Election 'War Room'. *The Daily Telegraph*.
- Ruppert, E., Law, J. & Savage, M. (2013). Reassembling the Social Science Methods. *Theory, Culture & Society* 30(4), 22-46.
- Saatchi, E. (2012). Actionable Data: Using Social Technology to Change Organizations. In N. G. Pearlman (Hrsg.), *Margin of Victory: How Technologists Help Politicians Win Elections* (S. 203-214). Santa Barbara, CA: ABC-CLIO.
- Saward, M. (2010). *The Representative Claim*. Oxford: Oxford University Press.
- Schnapp, K.-U., Behnke, N. & Behnke, J. (Hrsg.). (2009). *Datenwelten. Datenerhebung und Datenbestände in der Politikwissenschaft*. Baden-Baden: Nomos.
- Schreier, M. (2014). Qualitative content analysis. *The SAGE Handbook of Qualitative Data Analysis*, 170-183.
- Schrock, A. (2016). Civic Hacking as Data Activism and Advocacy: A History from Publicity to Open Government Data. *New Media & Society*, 18(4), 581-599.
- Scott, J. C. (1999). *Seeing Like a State*. New Haven, CT: Yale University Press.
- Sides, J. & Vavreck, L. (21.1.2014). Obama's Not-So-Big Data. *Pacific Standard*. Abgerufen von <http://www.psmag.com/navigation/politics-and-law/obamas-big-data-inconclusive-results-political-campaigns-72687/>
- Shorey, S. & Howard, P. (2016). Automation, Algorithms, and Politics| Automation, Big Data and Politics: A Research Review. *International Journal of Communication*, 10. Abgerufen von <http://ijoc.org/index.php/ijoc/article/view/6233>
- Star, S.L. & Bowker, G. (1999). *Sorting Things Out. Classification and Its Consequences*. Cambridge, MA: MIT Press.

- Steinke, Ines (2008). Gütekriterien qualitativer Forschung. In Uwe Flick, Ernst von Kardorff & Ines Steinke (Hrsg.), *Qualitative Forschung – Ein Handbuch* (6. Aufl., S. 319–331). Reinbek: Rowohlt.
- Stier, S. (2016). Strukturbedingungen im Online-Wahlkampf: USA und Deutschland im Vergleich. In C. Bieber & K. Kamps (Hrsg.), *Die US-Präsidentschaftswahl 2012: Analysen der Politik- und Kommunikationswissenschaft* (S. 363-382). Wiesbaden, DE: Springer VS.
- Strauss, A. (2012). Acting Intelligently: A Brief History of Political Analytics. In N. G. Pearlman (Hrsg.), *Margin of Victory: How Technologists Help Politicians Win Elections* (S. 175-188). Santa Barbara, CA: ABC-CLIO.
- Strauss, A. L., Corbin, J. M., & Niewiarra, S. (1996). *Grounded Theory: Grundlagen qualitativer Sozialforschung*. Weinheim: Beltz.
- Stromer-Galley, J. (2014). *Presidential Campaigning in the Internet Age*. New York, NY: Oxford University Press.
- Sullivan, M. L. (2012). A New Model: VAN and the Challenge of the Voter-File Interface. In N. G. Pearlman (Hrsg.), *Margin of Victory: How Technologists Help Politicians Win Elections* (S. 133-146). Santa Barbara, CA: ABC-CLIO.
- Thompson, N. (2012). International Campaign Finance: How Do Countries Compare? CNN. Abgerufen von <http://edition.cnn.com/2012/01/24/world/global-campaign-finance/>
- Tufekci, Z. (2014). Engineering the Public: Big Data, Surveillance and Computational Politics. *First Monday*, 19. Abgerufen von <http://firstmonday.org/article/view/4901/4097>
- Ulbricht, L. & Grafenstein, M. (2016). Big Data: Big Power Shifts? *Internet Policy Review* 5(1). Abgerufen von: <http://policyreview.info/articles/analysis/big-data-big-power-shifts>
- Ulbricht, L., Hofmann J., Haunss, S., Klinger, U., Passoth, J.-H., Pentzold, C., Schneider, I., Strassheim, H. & Voß, J.-P. (2017). Dimensionen von Big Data: eine politikwissenschaftliche Systematisierung Basisgutachten des Arbeitskreises Politikwissenschaft. ABIDA – Assessing Big Data-Projekt.
- Zittel, T. (2010). Mehr Responsivität durch neue digitale Medien? Die elektronische Wählerkommunikation von Abgeordneten in Deutschland, Schweden und den USA. Baden-Baden: Nomos.

11 ANHANG

Im Anhang führen wir das Codebuch, die Quellenübersicht und wesentliche Beiträge des Diskurses auf.

| Code | Anzahl |
|---|-------------|
| Codesystem | 4784 |
| Microtargeting | 148 |
| Big Data | 152 |
| Akteure | 0 |
| Arten von Akteuren | 0 |
| CEO / Leiter / Mitarbeiter / Eigentümer Dat... | 169 |
| Forscher / Analysten | 98 |
| Datenfirma | 211 |
| Datenhändler | 16 |
| Big-Data-Endverbraucher | 289 |
| andere Akteure | 51 |
| Aussage Akteure zu Big Data | 426 |
| Infoquelle | 111 |
| personale Verbindungen | 121 |
| Abbildungen in Daten | 0 |
| Offline Big Data | 65 |
| Online Big Data | 86 |
| Datenanalyse | 297 |
| Identifikation von Personen durch Big Data | 168 |
| Methoden Big Data Analysen | 359 |
| Handlungen aufgrund von Ergebnissen von Big-D... | 290 |
| Auswirkungen von Big-Data-Anwendungen | 0 |
| Zweifel an Wirksamkeit | 214 |
| unspezifisches Potential/Wirkweisen von Big-... | 215 |
| Wahlsieg durch Big-Data-Anwendungen | 143 |
| andere konkrete Wirkweisen von Big-Data-An... | 114 |
| Manipulation von Verhalten | 86 |
| Mittlere Position: Targeting wichtig, aber nur e... | 34 |
| Demokratischer Nutzen von Big Data | 22 |
| Gefahren von Big-Data-Anwendungen | 273 |

| | | |
|--|---|-----|
| • Einsatz von Big Data im Wahlkampf in Ländern/Ka... | ■ | 131 |
| • historischer Vergleich - vor Big Data | ■ | 72 |
| • Länderspezifik | ■ | 63 |
| • Datenschutz | ■ | 96 |
| • Wahlprognosen | ■ | 226 |
| • Änderungsvorschläge Umgang mit Big Data | ■ | 38 |

Quelle: Screenshot des Codesystems in MAXQDA

| Obercode | Subcode Ebene 1 | Subcode Ebene 2 | Codieranweisung |
|---|------------------------------------|---|---|
| Microtargeting | | | |
| Big Data | | | |
| Akteure | | | Akteure, die in Zusammenhang mit Big-Data-Sammlung, -Auswertung, -Verwertung im Bereich politischer Kampagnen gebracht werden. |
| | Arten von Akteuren | | Nur einmal pro Artikel codieren, wenn Person darin mehrmals genannt wird. Alle Infos zur Person werden mit codiert. Wenn neue Infos zur Person genannt werden, kann Code erneut vergeben werden. |
| | | CEO / Leiter / Mitarbeiter / Eigentümer Datenfirma | z.B. Steve Bannon, Robert Mercer |
| | | Forscher / Analysten | Alle Akteure, die Big Data direkt „verarbeiten“ und nicht explizit in ihrer Verwaltungsposition innerhalb einer Firma angesprochen werden (Alexander Nix kann beispielsweise als Analyst vorgestellt werden - in diesem Fall würde er mit dem Code „Forscher/Analysten“ versehen, er kann aber auch als CEO von Cambridge Analytica benannt werden, dann wäre entsprechend „CEO/Leiter/Mitarbeiter/Eigentümer Datenfirma“ zu codieren.). z.B. Michal Kosinski, Nate Silver |
| | | Datenfirma | z.B. Cambridge Analytica, Proofpoint, Targeted Victory |
| | | Datenhändler | z.B. Acxiom |
| | | Big-Data-Endverbraucher | z.B. Parteien, Politiker, Kampagnenmanager, Wahlforscher |
| | | andere Akteure | |
| | Aussage Akteure zu Big Data | | Codiert werden direkte und indirekte Zitate von Akteuren, die sich zum Einsatz von Big Data in Wahlkampfperioden äußern. |
| | Infoquelle | | Sekundärquellen wie Zeitschriftenartikel, die zum Belegen von Aussagen herangezogen werden; alle Quellen mit denen der/die Verfasser*in nicht selbst gesprochen hat, sondern Infos aus zweiter Hand |
| | personale Verbindungen | | Verbindungen von Big-Data-Personen (Mitarbeiter*innen, Vorsitzende von Datenanalysefirmen, Forscher*innen) zu Politiker*innen und/oder Spender*innen: z.B. Steve Bannon im Executive Board von Cambridge Analytica und Berater von Trump |
| Abbildungen in Daten | | | Hier wird codiert, welche Art von Phänomenen in quantifizierter Form erhoben werden. z.B. Einkäufe offline und online |
| | Offline Big Data | | Daten, die auf Phänomenen in der Offline-Welt beruhen. z.B. Wahlverhalten, Alter, Wohnort |
| | Online Big Data | | Daten, die auf Online-Aktivitäten beruhen. z.B. Facebook-Persönlichkeitsquiz |
| Datenanalyse | | | Auswertungen von Datenmengen/ Verhalten, das nicht näher spezifiziert wird. Z.B. "Datenanalysen sind im Wahlkampf wichtig". |
| Identifikation von Personen durch Big Data | | | Hier wird konkret beschrieben, dass und wie Personen durch Big-Data-Analysen identifiziert werden können. z.B. „Man kann nicht nur aus Daten psychologische Profile erstellen, man kann auch umgekehrt nach bestimmten Profilen suchen – etwa: alle besorgten Familienväter, alle wütenden Introvertierten. Oder auch: alle unentschlossenen Demokraten. Was Kosinski genau genommen erfunden hat, ist eine Menschensuchmaschine.“ |
| Methoden Big Data Analysen | | | Es werden gezielte Erhebungsverfahren von Big Data genannt oder erläutert (z.B. Facebook-Quizze). Der Hinweis, dass Big Data in einer Online- oder Offline-Umgebung entsteht (z.B. in Social Media), genügt nicht, sondern es muss konkret die Methodik der Erhebung angesprochen werden. z.B. „Psychometrie, manchmal auch Psychografie genannt, ist der wissenschaftliche Versuch, die Persönlichkeit eines Menschen zu vermessen. In der modernen |

| | | |
|--|---|--|
| | | Psychologie ist da für die sogenannte Ocean-Methode zum Standard geworden.“ |
| Handlungen aufgrund von Ergebnissen von Big-Data-Analysen | | z.B. gezielte, personalisierte Wahlwerbung: per Telefon, Post, Social Media etc. |
| Auswirkungen von Big-Data-Anwendungen | | |
| | Zweifel an Wirksamkeit | Entweder werden die Big-Data-Sammlungs- oder Auswertungsverfahren infrage gestellt oder die Maßnahmen, die auf diesen beruhen. z.B. „I think there’s a big question about whether we think psychographic profiling even works.“ |
| | unspezifisches Potential/Wirkweisen von Big-Data-Anwendungen | Verheißungen von Big Data werden angesprochen, ohne die Auswirkungen konkret zu benennen. z.B. „Standing before political and business leaders in New York last fall, Alexander Nix promised a revolution.“ |
| | Wahlsieg durch Big-Data-Anwendungen | Eine spezielle Wirkweise von Big-Data-Anwendungen ist ein diesen zugeschriebener Wahlsieg. z.B.: "How Big Data helped Trump win the election" |
| | andere konkrete Wirkweisen von Big-Data-Anwendung | Hier werden Auswirkungen von Maßnahmen angesprochen, die auf Big-Data-Analysen beruhen. Der Analyseprozess selbst steht also nicht im Vordergrund. Z.B.: „Psychologisches Targeting, wie Cambridge Analytica es verwendete, steigert die Clickraten von Facebook-Anzeigen um über 60 Prozent.“ |
| | Manipulation von Verhalten | Nur wenn gezielt von Manipulation, Persuasion oder übermäßiger Beeinflussung die Rede ist. Die bloße Ansprache von potentiellen Wählern oder das Versenden personalisierter Botschaften wird unter dem allgemeineren Code „Handlungen aufgrund von Big-Data-Anwendungen“ codiert. z.B. „But the company that can perfect psychological targeting could offer far more potent tools: the ability to manipulate behavior.“ |
| | Mittlere Position: Targeting wichtig, aber nur ein Faktor | Die Bedeutung von Big Data wird anerkannt, aber relativiert. z.B. “We bake a cake, it’s got 10 ingredients in it. Psychographics is one of them,” he said. “It’s very difficult to isolate exactly what the impact of that ingredient is.” |
| | Demokratischer Nutzen von Big Data | Gerechtigkeit, besseres Repräsentationsverhältnis etc. |
| | Gefahren von Big-Data-Anwendungen | Nicht nur Hinweise auf Auswirkungen, sondern klare Bewertung der Auswirkung als potentiell negativ. Schlüsselwörter: warnen, gefährlich, schlecht z.B. - nicht jeder Kandidat kann sich Datensätze leisten: „Who is hurt by this? Independent candidates and nonprofit organizations that want to run national campaigns but don't have deep pockets. If someone like Donald Trump launched an independent presidential run, he could buy the necessary data without much difficulty. But a nonprofit focused on mobilizing low-income voters could be stretched thin.“ - "The great British Brexit Robbery. How our democracy was hijacked" |

| | |
|---|--|
| Einsatz von Big Data im Wahlkampf in Ländern/Kampagnen | <p>Es wird festgestellt oder darüber spekuliert, in welchen Ländern / Kampagnen Big-Data-Anwendungen zum Einsatz gekommen sind, kommen werden oder im Gegenteil: gerade nicht eingesetzt werden.</p> <p>z.B. „The spokesman also said neither Cambridge nor SCL had done any work, paid or unpaid, with the pro-“Brexit” Leave.eu campaign last year, although Mr. Nix once claimed that Cambridge had helped “supercharge” Leave.eu’s social media campaign. British authorities are now investigating the company’s exact role with Leave.eu and whether Cambridge’s techniques violated British and European privacy laws.“</p> |
| historischer Vergleich - vor Big Data | <p>Wahlkampagnen oder Wahlprognosen heute werden mit jenen verglichen, als noch keine Big-Data-Anwendungen zum Einsatz kamen.</p> <p>z.B. „Bisher, so Nix, seien Wahlkampagnen nach demografischen Konzepten geführt worden, «eine lächerliche Idee, wenn Sie drüber nachdenken: Alle Frauen erhalten die gleiche Nachricht, bloss weil sie das gleiche Geschlecht haben – oder alle Afroamerikaner, wegen ihrer Rasse?»</p> |
| Länderspezifik | <p>Die Besonderheiten und Unterschiede datenbasierter Methoden im Wahlkampf werden betont.</p> <p>Z.B. „The United States represented a critical new market. Europe has strict privacy protections that limit the use of personal information, but America is more lightly regulated, allowing the sale of huge troves of consumer data to any company or candidate who can afford them.“</p> |
| Datenschutz | <p>Datenschutz wird explizit problematisiert. Die Thematisierung der Verwendung von personenbezogenen Daten allein reicht nicht aus für eine Codierung.</p> |
| Wahlprognosen | <p>Thematisierung von Big Data für Wahlprognosen.</p> |
| Änderungsvorschläge Umgang mit Big Data | <p>Maßnahmen, die vor negativen Auswirkungen von Big-Data-Anwendungen in Wahlkämpfen schützen sollen.</p> |

| Name der Quelle | Häufigkeit | Prozent |
|--------------------------------|------------|---------|
| New York Times | 21 | 13,73 |
| Frankfurter Allgemeine Zeitung | 14 | 9,15 |
| Guardian | 11 | 7,19 |
| Washington Post | 8 | 5,23 |
| Politico | 5 | 3,27 |
| Huffington Post | 5 | 3,27 |
| Zeit Online | 4 | 2,61 |
| Bloomberg | 4 | 2,61 |
| BuzzFeed | 4 | 2,61 |
| The Times | 4 | 2,61 |
| FiveThirtyEight | 4 | 2,61 |
| Twitter | 3 | 1,96 |
| Süddeutsche Zeitung | 3 | 1,96 |
| Forbes | 3 | 1,96 |
| Wired | 3 | 1,96 |
| The Economist | 2 | 1,31 |
| Spectator | 2 | 1,31 |
| MEEDIA | 1 | 0,65 |
| Buisness 2 Community | 1 | 0,65 |
| The Data Scientist | 1 | 0,65 |
| Heinrich Böll Stiftung | 1 | 0,65 |
| Startup Grind | 1 | 0,65 |
| CNN | 1 | 0,65 |
| Campaigns & Elections | 1 | 0,65 |

| | | |
|-----------------------|---|------|
| taz | 1 | 0,65 |
| PersonalData.IO | 1 | 0,65 |
| Harvard gazette | 1 | 0,65 |
| Futurezone | 1 | 0,65 |
| The New Stack | 1 | 0,65 |
| AdExchanger | 1 | 0,65 |
| The Independant | 1 | 0,65 |
| Financial Times | 1 | 0,65 |
| DigidayUK | 1 | 0,65 |
| Venturebeat | 1 | 0,65 |
| facebook buisness | 1 | 0,65 |
| npr | 1 | 0,65 |
| LSE USAPP blog | 1 | 0,65 |
| Medium | 1 | 0,65 |
| ayasdi.com | 1 | 0,65 |
| Facebook | 1 | 0,65 |
| MIT Technology Review | 1 | 0,65 |
| Vice | 1 | 0,65 |
| Open Democracy UK | 1 | 0,65 |
| Mashable | 1 | 0,65 |
| Scout | 1 | 0,65 |
| zscaler | 1 | 0,65 |
| SparkCognition Inc | 1 | 0,65 |
| NBC News | 1 | 0,65 |
| Tages-Anzeiger | 1 | 0,65 |
| The Register | 1 | 0,65 |

| | | |
|---------------------|-----|--------|
| The Telegraph | 1 | 0,65 |
| D!gitalist | 1 | 0,65 |
| proofpoint | 1 | 0,65 |
| The TYEE | 1 | 0,65 |
| BBC | 1 | 0,65 |
| USA today | 1 | 0,65 |
| Vox Media | 1 | 0,65 |
| ZeroHedge | 1 | 0,65 |
| Handelsblatt Online | 1 | 0,65 |
| TechRepublic | 1 | 0,65 |
| FORTUNE | 1 | 0,65 |
| InfoWorld | 1 | 0,65 |
| Civis Analytics | 1 | 0,65 |
| Inc. | 1 | 0,65 |
| ADZINE | 1 | 0,65 |
| WDR | 1 | 0,65 |
| SpiegelOnline | 1 | 0,65 |
| Martin Sauter | 1 | 0,65 |
| GRÜNDERSZENE | 1 | 0,65 |
| Das Magazin | 1 | 0,65 |
| Gesamt | 153 | 100,00 |